Shaping of Adaptive Immunity and Celiac Disease Autoimmunity in a Population based Childhood Cohort

The Generation R Study

Michelle Anne Elisa Jansen

Shaping of Adaptive Immunity and Celiac Disease Autoimmunity in a Population based Childhood Cohort

The Generation R Study

Michelle Anne Elisa Jansen

Colofon

The first phase of The Generation R Study was made possible by financial support from the Erasmus Medical Center, Rotterdam, The Erasmus University Rotterdam, the Netherlands Organization for Health Research and Development (ZonMW), the Netherlands Organization for Scientific Research (NWO), the ministry of Health, Welfare and Sport, and the Ministry of Youth and Families, the Netherlands. Additionally, grant support was provided for the studies described in this thesis by European Container Terminal (ECT).

The printing of this thesis has been financially supported by the Erasmus University Rotterdam and the Generation R Study Group, Rotterdam.

ISBN: 978-94-6295-762-6.

Lay-out and print by: ProefschriftMaken // www.proefschriftmaken.nl

Shaping of Adaptive Immunity and
Celiac Disease Autoimmunity
in a Population based Childhood Cohort
The Generation R Study

De vorming van specifieke afweer en coeliakie autoimmuniteit
in een populatie gebaseerd kinder cohort
Het Generation R Onderzoek

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
Rector Magnificus

Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
vrijdag 8 december 2017 om 11:30 uur

door

Michelle Anne Elisa Jansen

geboren te Deventer

Erasmus University Rotterdam

Erasmus

PROMOTIECOMMISSIE

Promotor	Prof.dr. H.A. Moll
Co promotor	Dr. M.C. van Zelm
Overige leden	Prof.dr. J.C. Escher Prof.dr. H. Raat Prof.dr. F. Koning

TABLE OF CONTENTS

MANUSCRIPTS THAT FORM THE BASIS OF THIS THESIS

Part I

1. Van den Heuvel D, Jansen MAE, K Nasserinejad, Dik WA, van Lochem EG, Bakker-Jonges LE, Bouallouch-Charif H, Jaddoe VWV, Hooijkaas H, van Dongen JJM, Moll HA, van Zelm MC. Effects of nongenetic factors on immune cell dynamics in early childhood: The Generation R Study. J Allergy Clin Immunol. 2017;139(6): 1923-1934.e17.

2. Jansen MAE, van den Heuvel D, van Zelm C, Jaddoe VWV, Hofman A, de Jongste JC, Hooijkaas H, Moll HA. Decreased memory B cells and increased CD8 memory T cells in Blood of Breastfed Children. The Generation R Study. PLoS One 2015; May 18;10(5).

3. Van den Heuvel D, Jansen MAE, Dik WA, Bouallouch-Charif H, Zhao D, van Kester K, Smits-te Nijenhuis MAW, Kolijn-Couwenberg MJ, Jaddoe VWV, Arens R, van Dongen JJM, Moll HA, van Zelm MC. CMV- and EBV-related T-cell expansions in young children do not impair naive T-cells nor vaccination responses. The Generation R Study. J Infect Dis. 2016 Jan 15;213(2):233-42.

4. Jansen MAE, Van den Heuvel D, Jaddoe VWV, Moll HA, van Zelm MC. No Interactive Effects of Sex and Persistent Cytomegalovirus on Immune Phenotypes in Young Children: The Generation R Study. J. Infect Dis 2017 March 15;215(6):883-888.

5. Jansen MAE, van den Heuvel D, Jaddoe VWV, van Zelm MC, Moll HA. Abnormalities in CD57+ cytotoxic T cells and Vδ1+ TCRγδ cells in subclinical celiac disease in childhood are affected by cytomegalovirus. The Generation R Study. Clin Immunol. 2017 Apr 26. [Epub ahead of print].

Part II

6. Jansen MAE, Tromp IIM, Kiefte-de Jong JC, Jaddoe VWV, Hofman A, Escher JC, Hooijkaas H, Moll HA. Infant feeding and anti-tissue transglutaminase antibody concentrations in the Generation R Study. Am J Clin Nutr 2014; Oct;100(4):1095-101.

7. Jansen MAE, van den Heuvel D, van der Zwet KVM, Jaddoe VWV, Hofman A, Escher JC, Fraaij PLA, Hooijkaas H, van Zelm MC, Moll HA. Herpesvirus infections and Transglutaminase type 2 Antibody positivity in Childhood. The Generation R Study. J Pediatr Gastroenterol Nutr. 2016 Oct;63(4):423-30.

8. Jansen MAE, Beth SA, van den Heuvel D, Kiefte-de Jong JC, Raat H, Jaddoe VWV, van Zelm MC, Moll HA. Ethnic differences in coeliac disease autoimmunity in Childhood. The Generation R Study. Arch Dis Child. 2017 Jun; 102(6):529-534

9. Jansen MAE, van den Heuvel D, Bouthoorn SH, Jaddoe VWV, Hooijkaas H, Raat H, Fraaij PLA, van Zelm MC, Moll HA. Determinants of Ethnic differences in Cytomegalovirus, Epstein-Barr Virus, and Herpes Simplex Virus Type 1 Seroprevalence in Childhood. J Pediatr. 2016 Mar;170:126-34.e1-6.

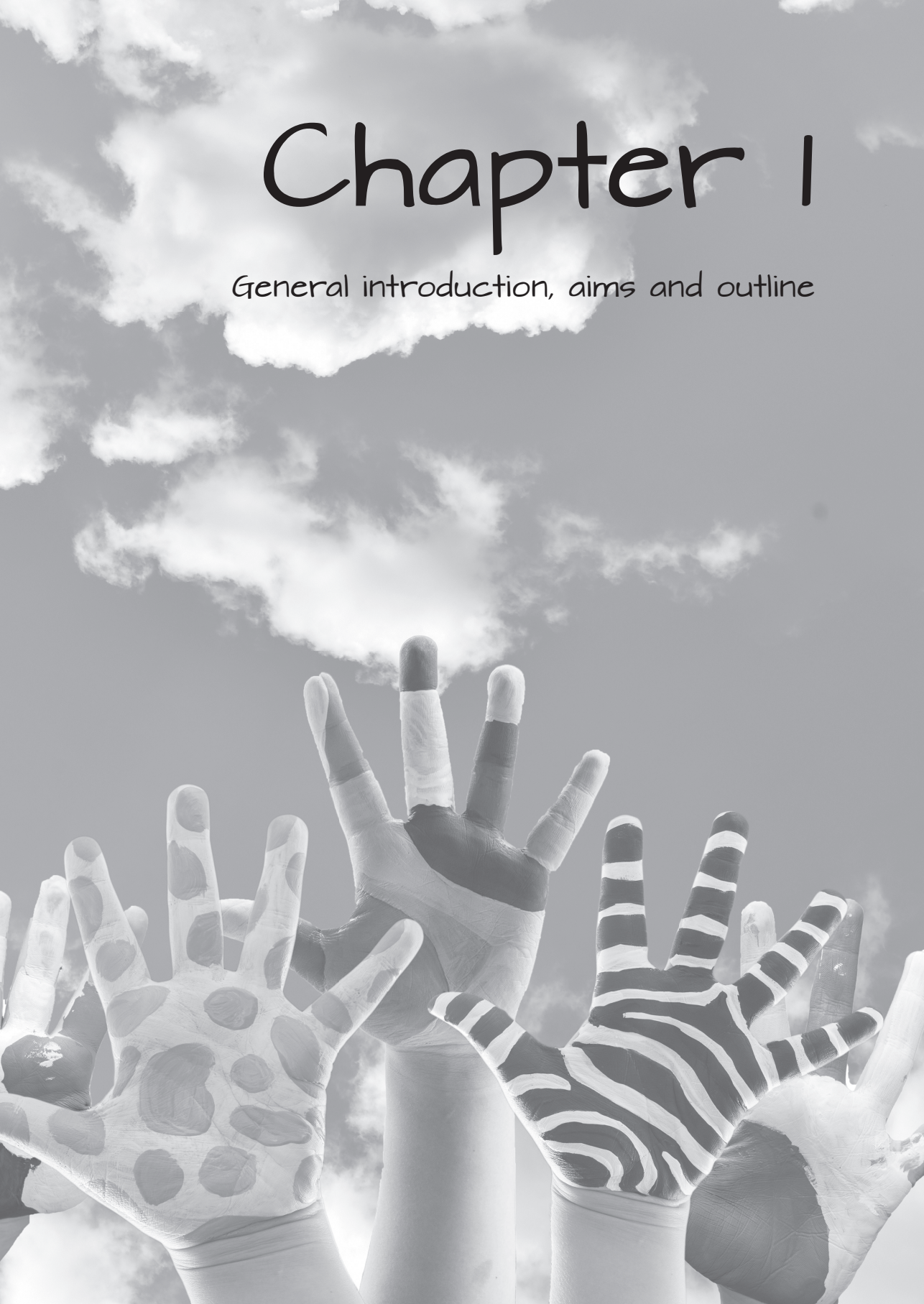
Part III

10. Jansen MAE, van Zelm MC, Groeneweg M, Jaddoe VWV, Dik WA, Schreurs MWJ, Hooijkaas H, Moll HA, Escher JC. The identification of celiac disease in asymptomatic children: the Generation R Study. J. Gastroenterol. 2017 June 6. [Epub ahead of print].

11. Jansen MAE, Kiefte-de Jong JC, Gaillard R, Escher JC, Hofman A, Jaddoe VWV, Hooijkaas H, Moll HA. Growth Trajectories and Bone Mineral Density in Anti-Tissue Transglutaminase Antibody-positive Children: The Generation R Study. Clin Gastroenterol Hepatol 2015; May;13(5):913-20.

Chapter 1

General introduction, aims and outline



DELICATE BALANCE OF IMMUNE RESPONSES IN EARLY CHILDHOOD

The human body is continuously exposed to various pathogens. To protect itself from these potential harmful agents, an immune system is in place with mechanisms to inhibit the entrance of pathogens, and if this fails, to recognize, neutralize and quickly remove these pathogens. Aberrant immune protection, however, results in clinical disease, such as autoimmune diseases with an immune response directed against self-antigens, or allergies/hypersensitivities with an excessive reaction to harmless foreign particles. Immune competence, therefore, reflects a delicate balance between 'too little and too much' tolerance, which develops in early childhood, or possibly even in utero.[1, 2] This process may be prone to various external factors, especially during the first 5 years of life when antigens encounter the immature immune system for the first time. Immune mediated diseases such as celiac disease, diabetes mellitus type 1, allergic rhinitis, atopic dermatitis and allergic asthma are prevalent and have reached even epidemic proportions nowadays, resulting in a substantial global burden of disease in western countries.[3-8] The incidence of these immune mediated disorders began to increase in the 1950s, and have doubled or tripled during the last 3 decades worldwide, particularly in young children in industrialized countries.[4, 8] Concomitantly, an obvious decline in infectious diseases was observed, which together, led to the hypothesis that early exposure to potential pathogens may protect against the development of autoimmune and allergic diseases.[8] The early childhood period, therefore may reflect a window of vulnerability, during which various determinants 'shape' the developing immune system and mediate the origin of immune mediated diseases.

THE SHAPING OF THE ADAPTIVE IMMUNE SYSTEM

The human immune system is formed during fetal development and "innate" responses can be generated after birth to protect the newborn against pathogens. However, at this stage the immune system is still immature. Following interactions with pathogens, the immune system "adapts" and memory is formed to quickly eliminate foreign substances without damaging the body's own healthy tissue. However, when the delicately balanced regulation between self- and non-self-antigens is disturbed, autoimmunity might develop.

Memory B and T cell populations

B and T cells are the key cells of adaptive immunity. Mature B and T cells circulate in the blood and in secondary lymphoid tissues after they have differentiated from precursor B and T cells in the bone marrow and thymus, respectively. These lymphocytes carry specialized receptors, i.e. a B-cell receptor (BCR) or T cell receptor (TCR). Each type of receptor is unique for each newly generated B and T cell. While B cells recognize complete antigens, T

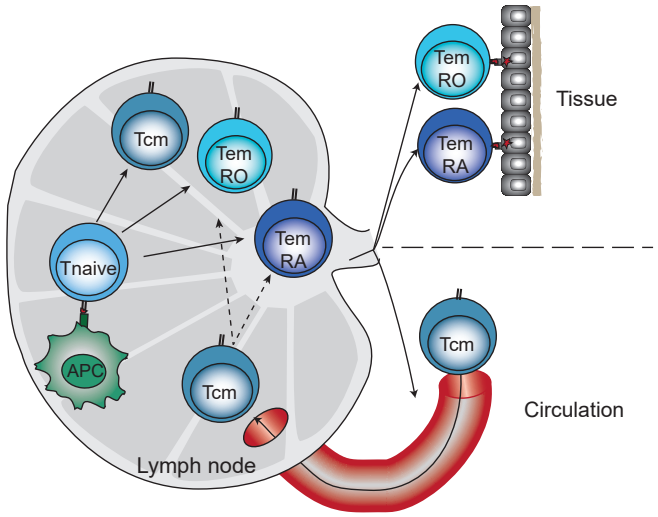
cells only recognize an antigen when it is presented on specialized protein complexes, i.e. the human leukocyte antigens (HLA; also known as the major histocompatibility complex (MHC)). These antigens are presented by dendritic cells (DCs) in the lymph node, after these have migrated there from local tissue. Upon antigen recognition, a B or T cell becomes activated, resulting in extensive clonal proliferation and differentiation into effector cells.

T cells comprise of 2 major lineages: cytotoxic (CD8+) and T-helper (CD4+) cells. The TCR of cytotoxic T cells recognizes antigens presented by HLA class I proteins that are expressed on all cells in the body, except on erythrocytes. Recognition of virus-infected cells by activated T cells induces cytokine production, resulting in lysis of the infected cell.[9] The TCR of T helper cells recognizes antigens presented by HLA class II proteins. Class II proteins present peptides taken up from the cells exterior, e.g. extracellular bacterial antigens. HLA class II proteins are only presented by immune cells, especially DCs and other antigen presenting cells (APCs). Therefore, helper T cell interactions are limited to the cells of the immune system. Activated helper T cells stimulate the activation of surrounding antigen specific B and cytotoxic T lymphocytes.[9]

Within the CD4+ and CD8+ T cell lineages, distinct T cell subsets can be distinguished depending on their function and differentiation pathway, reflected by differential expression of their cell-surface markers, such as costimulatory and adhesion molecules, and chemokines receptors (Figure 1). The term 'differentiation' relates to expression levels of genes that control cell cycle, survival, migration and effector functions of the cells,[10, 11] and is therefore generally accepted to describe lymphocyte maturation or development.[12] Although the exact differentiation pathways through which the various memory T cells populations are generated remain actively debated,[10, 13] general agreement includes the characterization of naive, central memory (TCM), and effector memory T cells (TEM) within both the CD4+ and CD8+ lineage.[12, 13] *Naive T cells (T_{naive})* express CD45RA, CCR7 and the lymph node homing receptor L-selectin (CD62L), which together, enables migration to T-cell areas in secondary lymphoid organs to search for antigen presented by dendritic cells.[12, 14] Upon antigen recognition, the cell will replace CD45RA with alternative splice variant CD45RO. The CD45-isotype-switch is associated with increased proliferation and T cell activation.[15] *Central memory T cells (TCM)* are part of the CD45RO+ memory cells, but express high levels of CCR7 and L-selection as well.[12, 14] Due to the maintained expression of these migration markers, these cells circulate through the blood, lymphoid system and secondary lymphoid tissues, but they lack potent effector functions.[12] It has been proposed that T_{CM} cells are responsible for maintaining life-long memory: Upon re-exposure to antigens, T_{CM} cells proliferate and differentiate into effector memory cells (TEM). *Effector memory T (TEM) cells* express either CD45RO+ *(TemRO),* or re-express CD45RA+ *(TemRA)* (the latter being considered as having more effector cell properties),[16] but in contrast to TCM, they lack expression of CCR7 and L-selectin. Instead, T_{EM} express receptors (CCR4 and CCR5) for migration to inflamed tissues, such as the skin or intestine. Additionally, they express increased levels of effector functions, e.g. they carry perforin, and rapidly produce

IL-2, IL-4, IL5, IFN-γ and TNF-α following antigen encounter. The Tem are thus a mixture of T_H1, as well as T_H2, and cytotoxic T cells. [12]

A



B

		Tnaive	Tcm	Tem RO early	Tem RO interm	Tem RO late	Tem RA early	Tem RA interm	Tem RA late
co-stimulation cell survival regulation of activation	CD45RA								
	CD45RO								
	CD27								
	CD28								
	CD127 (IL-7R)								
	CD57								
	KLRG1								
	PD-1								
cell migration and adhesion	CCR7								
	CXCR4								
	CD62L								
	CCR4								
	CCR5								
	CXCR1								
	CX3CR1								
cytokine production and cytotoxic function	IL-2								
	IFNγ								
	TNFα								
	Perforin/Granzyme B								
	CD95L (FASL)								

Figure 1. A, CD8+ T-cell responses Activated naive T cells differentiate into central memory T cells (T_{CM}) or effector memory T cells (T_{EM}), the latter being either CD45RO+ (TemRO) or CD45RO- (TemRA). T_{CM} circulate between the blood and lymphatic system. TemRO and TemRA cells migrate into tissue and kill virus-infected cells. Figure adapted from van den Heuvel et al.[146] **B,** T-cell subset expression levels of proteins involved in co-stimulation, cell survival and regulation of activation; cell migration and adhesion, cytokine production and cytotoxic function. White, not expressed, light blue, intermediate expression, dark-blue high expression. Adapted from van den Heuvel et al.[146]

Based on the expression of the B cell surface markers IgD, CD27, and CD38, different stages of B-cell maturation can be identified (Figure 2). The majority of circulatory B cells comprise *naive B cells* (IgD+CD27-).[17] Of these, *'transitional B cells'* (IgD+ CD24+CD38^{hi}) are new bone marrow emigrants and less mature than the *naive mature B cells* (IgD+ CD27-IgM+CD24^{dim}CD38^{dim}).[17, 18] Upon antigen recognition, naive mature B cells proliferate and differentiate either into memory B cells or plasma cells. Plasma cells directly produce antibodies to mediate protective memory, whereas memory B cells are necessary for reactive memory, as reflected by the proliferation and differentiation into plasma cells upon secondary antigen encounters.[12] The majority of memory B cells are T cell dependent (TD): They need a second signal from T-helper cells in the germinal center to become activated, resulting in extensive B cell proliferation and IgG class switch recombination.[19] Hence, high affinity memory B cells and IgG producing plasma cells are formed.[19] Five distinct memory B cell subsets are thought to develop via T cell dependent germinal center responses,

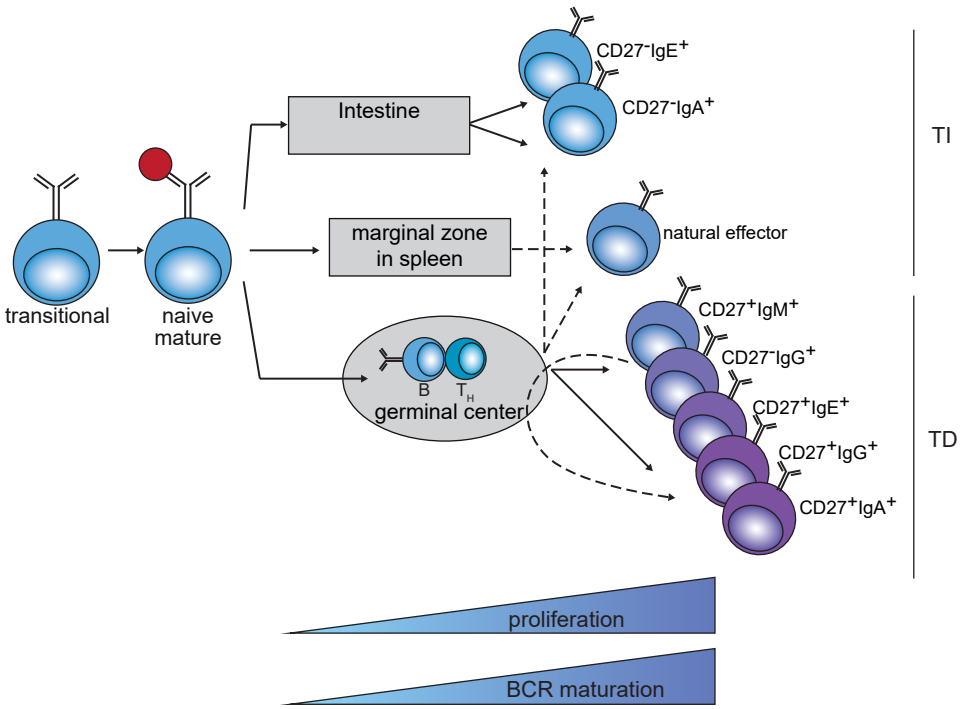


Figure 2. Memory B cell formation Naive B cells can undergo T cell dependent (TD) or T cell independent (TI) immune responses. TI responses occur locally in the mucosal tissue, inducing the formation of CD27-IgE+ and CD27-IgA+ cells, or systematically in the splenic marginal zone, generating natural effector cells. TI responses are characterized by low proliferation and B-cell receptor (BCR) maturation levels as compared with TD responses. TD responses occur in germinal center structures. Primary germinal center responses induce the formation of CD27+IgM+, CD27-IgG+, and CD27+IgE+ memory B cells with intermediate proliferation and BCR maturation. Secondary TD germinal center responses induce the formation of CD27+IgG+ and CD27+IgA+ memory B cells that underwent even stronger proliferation and BCR maturation. Adapted from Berkowska et al,[19] and van den Heuvel et al.[146]

and show subsequent increased proliferation: CD27+IgM+IgD- "*IgM-only", CD27-IgG+, CD27+IgE+, CD27+IgG+* and *CD27+IgA+*.[19, 20] Additionally, B cells can respond to antigens independent of T cells (TI). These antigens are mostly blood-borne pathogens in the splenic marginal zone and non-protein antigens in mucosal tissues. CD27+IgM+ memory B cells that still co-express IgD+ (natural effectors) develop in the splenic marginal zone,[21] and are thought to proliferate outside of the germinal centers of the spleen and lymph nodes. The other two TI B cell subsets *(CD27-IgA+* and *CD27-IgE+)* originate from mucosal tissues.[19, 20]

During childhood, the composition and numbers of immune cells considerably fluctuate as a consequence of ongoing antigen exposure, of whom the most obvious changes occur in the first 5 to 6 years of age.[18, 22-24] Concurrent with an expansion of memory cells with increasing age, the repertoire of naive cells becomes smaller due to involution of the thymus and decreased bone marrow output.[18, 25] Immunity in adulthood, and subsequently protection against recurrence of disease in elderly, therefore, is largely depended on the memory pool that is built-up in the first years of life.[25] Despite several cross-sectional studies presenting consistent observations of normal, i.e. age-dependent reference, values during childhood,[18, 23, 24, 26-32] none of the previous studies assessed these kinetics longitudinally. Moreover, many external factors are thought to affect these patterns, but are previously largely unaddressed. Although differentiation pathways are still not fully elucidated, and the usefulness of various expression markers are still debated, the identification of various cell populations as described above is general accepted, which enables studying kinetics and determinants of these lymphocytes during early childhood. Understanding which factors contribute to the maturation of a delicately balanced and mature immune system will be of benefit for the assessment of normal and disturbed immune development.

Immune shaping through external determinants

Pre– and perinatal determinants

During pregnancy, the fetal immune system is largely immature, and predominantly composed of innate leukocytes and naive lymphocytes.[27, 32] Maternal IgG can be transferred through the placenta though, providing passive protection to the unborn child.[33] Still, environmental factors, such as prenatal lifestyle, parity, fetal growth, and prematurity might affect programming of the fetal immune system in utero, and contribute to later development of immune-mediated diseases.[1, 2, 34] Maternal and fetal anti-inflammatory cytokines have been shown to strongly correlate,[2] suggesting 'transplacental immune regulation', i.e. a significant alignment or crosstalk between maternal and fetal immunity, which may have long term effects for both mother and child.[2] The prenatal period may therefore reflect a possible sensitive window of vulnerability during which intrauterine exposures can modulate the programming of the child's immune system, and affect the course of immune development and tolerance mechanisms.[2, 34] Additionally, perinatal determinants, such as mode of delivery, duration of labor, and birth weight have been shown to affect innate leukocyte and memory lymphocyte populations, possibly partly through perinatal

stress and differences in microbial population of the gut.[35-39] These immune alterations at birth have been associated with 'T_H2'-type immune mediated diseases, including food allergies and eczema.[38-40] Still, it remains unclear whether a sensitive window of immune vulnerability exists, and whether innate and adaptive immune cell kinetics are predominantly influenced by pre-, peri- and/or postnatal determinants. Furthermore, none of the previous studies assessed the development of detailed innate leukocyte and memory lymphocyte populations longitudinally, covering intrapersonal correlations between repeated measures.

Postnatal determinants

Breastfeeding

After birth, breastfeeding is known to protect the child from respiratory and gastrointestinal infections.[41] Colostrum and breastmilk contain numerous beneficially factors, including maternal IgA, IgG and cytokines, providing the child passive protection during the period of breastfeeding, and possibly shortly after weaning.[42] During this period in early life, the child has to build up its own active immunity, when the immune system is still functionally immature and encountering various, mainly viral antigens.[43] Many previous studies have focused on the beneficial effects of breastfeeding on innate immunity and protection towards infectious diseases in childhood.[41, 42, 44] However, studies focusing on adaptive immunity are scarce, and often not adjusted for potential confounding factors. Studying determinants of adaptive immunity is of particular importance, as adaptive immunity provides an extraordinary specific immune response against pathogens, and confers life-long immunity: principles that form the basis of vaccination.

Herpesviruses

Most viruses that children encounter in the first few years of life are respiratory viruses, which are effectively cleared by the immune system, and memory B and T cells are formed. Several viruses, however, escape from this clearance. To date, 8 herpesviruses have been identified to infect humans, including Cytomegalovirus (CMV), Epstein-Barr virus (EBV), Herpes Simplex virus type 1 (HSV-1) and Varicella Zoster virus (VZV). Primary infection with these viruses predominantly occurs during childhood, and is often mild or asymptomatic.[43] Herpesviruses are effectively suppressed by the immune system, but establish lifelong latency, with intermittent bouts of reactivation, causing for example scores or herpes zoster in cases of immune suppression. To prevent reactivation, the immune system, therefore, should constantly suppress these viruses, which may adversely affect the control of other encountering allergens or pathogens. Moreover, the constant viral pressure of latent infection might cause 'immune exhaustion' on the long run. It is not known, however, to what extent herpesviruses affect the balance and diversity between specific and preexisting memory B and T cells in early childhood. For example, newly formed CMV specific CD8+ T cells may expand the preexisting CD8+ repertoire, but may also 'compete' with the preexist-

ing memory CD8+ T cells for 'immune-space', resulting in substantial depletion of preexisting memory. In elderly, adverse effects have been described on the diversity of the B and T cell memory repertoire, resulting in overcrowding (accumulation of late differentiated CD8 memory, concurrent with reduced naive cells) and eventually loss of immunity against other pathogens.[45-47] Still, these observations have not been persistently replicated.[48-50] Nevertheless, the flexibility of the immune system, viral pathogenesis and clinical outcome may be differentially affected by age. For example, beneficial protective effects of EBV have been observed in relation to atopic and allergic diseases in childhood,[51-53] but seemingly depended on age, as early (<2y) and late (>2y) EBV infection resulted in opposite effects.[52]

The seroprevalence of herpesviruses in adulthood reaches 30-100%,[54-62] but these seroprevalences depend largely on geographic location, race/ethnic background, socioeconomic disparities, population dynamics and various other factors that facilitate transmission.[56-58, 61, 63] Cumulative population prevalence are difficult to estimate because of the asymptomatic course of primary infection in childhood. Thus, estimates had to be based on specific IgG antibodies against the virus.[43] Most prevalence studies are from the 1980s-1990's,[57] but social and demographic factors have changed over time, limiting up-to-date conclusions on trends in childhood herpesvirus epidemiology nowadays. Moreover, most studies have not incorporated the relative distribution of all combined environmental factors. Insight into these infection dynamics and their determinants is important to understand herpesvirus associated disease epidemiology, and in the light of prevention, to advise on vaccination strategies on the long run.

Conspicuously in industrialized countries, a shift towards an older age of primary infections, concurrent with an increase in atopic and autoimmune diseases was observed, which led to the 'hygiene' or 'old friends' hypothesis.[64] This hypothesis suggests a causal relationship between the decreasing incidence of infectious diseases in countries with a Western lifestyle and urbanization, and the increasing incidence and severity of atopic and autoimmune diseases.[64-66] In this context, a diverse microenvironment in early life might protect against these diseases. In contrast, other studies observed that an increase in infections, provokes autoimmunity.[67] Herpesviruses primarily, have been suggested to initiate and progress numerous autoimmune syndromes,[68, 69] whereas others suggested rather protective effects.[51-53, 70] Hence, evidence on this highly debated topic is conflicting, and simple mechanisms to describe pathways in autoimmunity do not capture the probably multifaceted and multiplicative signals/pathways involved in autoimmune development. Still, all autoimmune diseases have in common that tolerance mechanisms 'fail', resulting in an aberrant immune response against harmless food or allergens.

THE EPIDEMIOLOGY, DETERMINANTS AND CONSEQUENCES OF UNRECOGNIZED CELIAC DISEASE

Celiac disease (CeD) is a common chronic multisystem immune-mediated disorder, characterized by an aberrant immune response against the body's own cells and tissues in response to gluten ingestion in genetic susceptible individuals. Ingestion of gluten by persons with CeD causes immune-mediated inflammatory damage to the small intestine, which can cause gastrointestinal and nongastrointestinal illness. The local CD4+ T cell dependent immune reaction involves both TCR's responses against gluten peptides, as well as B cell antibody formation (IgA or IgG) against the gliadin deaminating enzyme tissue transglutaminase type 2 (TG2), endomysium (EmA) or deamidated gliadin peptide (DGP).[71] The presence of TG2-specific antibodies in serum is the most specific (98%) and sensitive (98%) diagnostic marker of active CeD.[71-74] A cascade of inflammatory responses results in flattening of the small intestinal mucosa and a variable degree of small intestinal mucosal atrophy. It is the most prevalent (1-3%) and chronic autoimmune disease in childhood, though severely underdiagnosed: Only 1:3 to 1:5 cases are recognized due to often atypical, mild or even asymptomatic modes of presentation in childhood.[73, 75-77] Classical clinical features include weight loss, abdominal pain, chronic or intermittent diarrhea or constipation.[73, 77] However, CeD can also manifest with extra-intestinal manifestations, such as dermatitis herpetiformis or gluten ataxia, that are direct consequences of autoimmunity.[78] Other extra-intestinal symptoms are indirectly related to inflammation and/or malabsorption, such as anemia, delayed or stunted growth, delayed puberty and osteoporosis.[78] Several organs including the central nervous system, joints, liver and teeth can be affected,[78, 79] reflecting its broad and variable clinical presentation. In most patients with CeD, strict adherence to a life-long gluten free diet (GFD), especially if started in early life, leads to a complete recovery of enteropathy and disappearance of CeD-associated symptoms and complications.[80-82]

DEFINITIONS OF CELIAC DISEASE

The spectrum of CeD has broadened the past three decades, and also includes children with unrecognized forms of CeD, such as subclinical or potential forms of disease that are below the threshold of clinical detection (Figure 3, Table 1).[80, 81] Children with subclinical and potential CeD, are per definition identified by serology screening. Depending on the presence or absence of villous atrophy, they are considered as having either subclinical or potential CeD. However, when the status of biopsy is not known, the term 'CD autoimmunity' (CDA) or 'TG2A positivity' should be used. Although the definition of CD autoimmunity requires TG2A measurements on at least 2 occasions, we use the term "*CD autoimmunity*" (CDA) in the present thesis defined as *single TG2A positive* children, when the status of

biopsy is not known. Unrecognized CeD may include all forms of CeD, but as symptomatic and classical CeD are often diagnosed by routine clinical care, the majority concerns subclinical or potential disease.

Table 1.

Definitions of Celiac Disease		Equivalent discouraged term
Symptomatic	Clinically evident gastrointestinal and/or extra-intestinal symptoms attributable to gluten intake in accordance with enteropathy (defined as Marsh-Oberhuber grade ≥2).	Overt
Classical	Signs and symptoms of malabsorption. Diarrhoea, steatorrhoea, weight loss or growth failure is required in accordance with enteropathy (defined as Marsh-Oberhuber grade ≥2).	Typical
Subclinical‡	Positive serology (TG2A, EmA, DGPA (IgA/IgG)) and mucosal atrophy, without signs or symptoms sufficient to trigger CeD testing in clinical practice.	Silent/ asymptomatic†
Potential‡	Positive CeD serology, but without symptoms and mucosal atrophy. At increased risk for CeD. If the biopsy is positive, then this is CeD, if the biopsy is negative than it is potential CeD.	Latent
Refractory	Persistent or recurrent malabsorptive symptoms and signs with villous atrophy despite a strict GFD for more than 12 months.	
Celiac disease autoimmunity	Positive TG2A or EmA, on at least 2 occasions when the status of biopsy is not known.	
TG2A positivity*	When TG2A has only been measured on one occasion	

According to the Oslo definitions [80]

‡ Only identified by serology screening.

*In the present thesis TG2A positivity was defined by using the cutoff for clinical practice of 7 U/mL.

†Asymptomatic CeD is not accompanied by symptoms even in response to direct questioning at initial diagnosis. Many of these patients suffer from decreased quality of life. Sometimes minor symptoms (eg, fatigue) are only recognized after the introduction of a GFD; such patients do not suffer from true asymptomatic CeD and should be reclassified as having subclinical CeD.

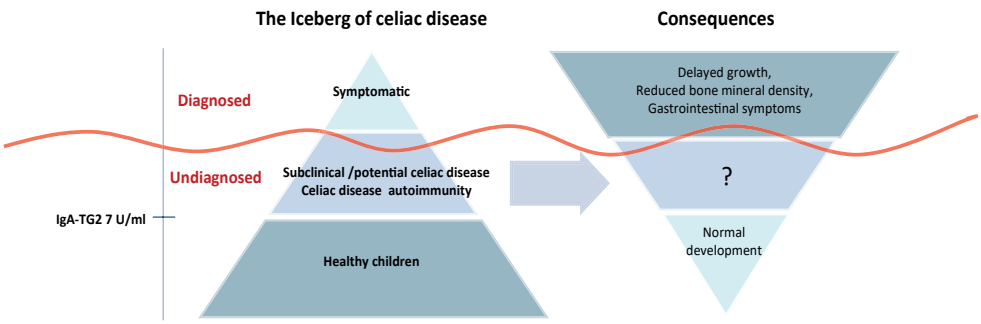


Figure 3. The iceberg of celiac disease The tip of the iceberg reflects symptomatic and diagnosed celiac disease (above the red line). However, subclinical and potential forms of the disease are often not recognized (below the red line). Children with a positive serology (indicated by positive IgA-TG2) are at increased risk of developing symptomatic CeD. The consequences of these unrecognized forms of CeD remain largely unknown.

Multifactorial etiology

Carriership of either the HLA-DQ 2.2, HLA-DQ 2.5 or HLA-DQ8 haplotypes is a prerequisite for CeD.[71] However, >25% of the general population carries one of the necessary risk alleles of CeD, whereas only a small fraction (1-3%) develops CeD.[83-85] Hence, >90% of genetically predisposed individuals do not develop CeD. It remains unknown why only a minority of genetic susceptible individuals develop CeD. Besides gluten and a genetic prerequisite, other determinants must contribute to disease initiation.[71, 75] Though, the search for environmental factors that contribute to the development of CeD has reached limited success.[71] Swedish birth cohort studies observed a difference in CeD incidence after a change in feeding recommendations in the 1980s and 1990s,[86-88] which gave rise to the concept that breastfeeding and the timing of gluten introduction may affect CeD development in predisposed persons. Introducing gluten between 4-6 months of age while breastfeeding was thought to prevent, or to delay, onset of CeD by inducing desensitization, or 'oral tolerance'.[88-92] I.e. if gradually small amounts of food substances are administered during a critical time period, the small intestinal barrier is more mature, and the local immature immune system 'learns' not to respond to this substance. Following the Scandinavian observations,[86, 87, 90, 91] some observational studies, however, could not confirm these findings.[93-95] The lack of consistency between these observational studies may be partly due to the majority of studies being retrospective or ecological. Hence, selection or recall bias of feeding habits may be an important concern. Because of the ongoing discussion on the possible role of gluten introduction and breastfeeding in the pathogenesis of CeD in 2013-2014, we studied these associations in a population of children screened for CeD. We conducted and published our study prior to the publication of 2 randomized clinical trials in 2014, including one from the 'PreventCD' Study,[84] that convincingly showed that neither early introduction (4 months of age) nor late (1 year of age) of gluten while breastfeeding affects the development of CeD.[84, 96] Several other determinants including a variety of (viral) infections and epigenetic factors could affect the intestinal microbiome, disrupt tolerance mechanisms in the intestine, affect the developing immune system and contribute to CeD development.[75, 97, 98] Substantial geographical differences in CeD prevalence have been observed between European countries, ranging from 0.3% in Germany to 2.4% in Finland, despite using common criteria for CeD diagnosis.[76] Therefore, environmental factors need further investigation.[97] CeD has been reported to be particularly on the rise in Western countries,[99-102] where a simultaneous decrease in herpesvirus infections over time was observed.[103, 104] Additionally, ethnic/racial disparities in CeD prevalence have been observed within single regions: CeD is more prevalent in non-Hispanic whites than in blacks, [105] which is in line with the delayed age of primary infection in non-Hispanic whites relative to blacks.[103] Hence, on top of differences in genetic predisposition, a link between herpesviruses and CeD may exist. Although previous studies have addressed associations between 'infections' and CeD in early life,[94, 98, 106, 107] these studies are inconclusive and difficult to compare due to often retrospective designs, various differences in pathogens and

port d'entrée's, differences in timing of infection, data acquisition (questionnaire-derived or serology proven) and limited statistical power. On top of this, determinants of symptomatic diagnosed CeD may be different than for subclinical or potential CeD.

Consequences of Celiac Disease Autoimmunity

CeD has been associated with a wide range of adverse effects on health and development. One of the most common complications of CeD are changes in bone mineralization, osteopenia and osteoporosis, which affect up to 75% of adult patients.[108, 109] Both osteoporosis and osteopenia increase osteoporotic fracture risk (i.e fracture risk is 2-7 times higher in CeD patients), leading to lower quality of life.[108, 110, 111] Other severe adverse consequences include cognitive impairment, infertility (although evidence is not consistent),[112] intestinal lymphomas, adenocarcinomas, and other malignancies including non-Hodgkin lymphomas.[81, 82, 113-122] Considerable evidence shows that a GFD can reverse the increased risk for associated malignancies,[116, 123-125] but the increased risk for non-Hodgkin lymphomas seems to persist.[116] Overall mortality of CeD patients has not been found increased in the majority of studies,[121, 126] neither in symptomatic CeD, nor in subclinical CeD.[113, 127] The strength of this association is, however, conflicting and varies widely among different studies, ranging from no associations to a 3,9 fold increased risk.[101, 113, 121, 125, 127-131] Nevertheless, it should be noted that several factors might dilute mortality rates and complication risks, such as differences in age at diagnosis and definitions of seropositivity, the adherence to a GFD, follow-up time, and type of clinical presentation.[132-134] Importantly, complications may be different for symptomatic than for subclinical CeD.[131] However, data on consequences of unrecognized CeD in children are scarce, complicating conclusions on whether these children should be identified and treated earlier.

In newly diagnosed children with both symptomatic and asymptomatic CeD, delayed growth and decreased BMD have been observed, which completely recover upon gluten avoidance.[135-137] Reduced bone mineralization have been shown to be present even before the appearance of any other symptom.[137, 138] Once the GFD starts, most of these children catch up their height and weight growth and bone mineralization rate within 1-2 years.[108, 136, 139] The younger the children start, especially before puberty, the better the bone mineral density (BMD).[108, 137, 138] In contrast, by the time CeD is diagnosed in adulthood, complete bone recovery is not always achieved.[140] In fact, delayed diagnosis or lack of treatment in childhood, induces severe osteoporosis in up to one third of adult patients.[137, 141] It is estimated that by the time childhood CeD is diagnosed, one third of patients have osteoporosis, one third have osteopenia and only the remaining one third have a normal BMD.[131] Consequently, part of the idiopathic adult osteoporosis may be explained by unrecognized CeD in childhood.[109] Importantly, adverse effects on bone health have also been observed in adult patients with subclinical and potential CeD, which improved upon introduction of a GFD,[108, 109, 138, 142, 143] indicating that merely the presence of positive serology is sufficient to have adverse effects. This highlights the importance of CeD diagnosis as early as possible,

and it suggests that these patients would benefit from early treatment. Still, it is currently unknown whether adverse effects on growth and development are present in screening identified, TG2A positive children, prior to the onset of other complications, and irrespective of the presence of villous atrophy (Figure 3).

AIMS

The first aim of this thesis was to identify determinants of the developing immune system in childhood, focusing predominantly on the adaptive immune compartment. The second aim was to study determinants and the consequences of celiac disease autoimmunity (CDA) within a population based prospective cohort.

THE GENERATION R STUDY

The studies in this thesis were embedded within the Generation R Study, a population based prospective multi-ethnic birth cohort, following pregnant mothers and their children from fetal life until young adulthood, in Rotterdam, The Netherlands.[144, 145] In total, 9,778 pregnant women were included in the cohort, all residing in Rotterdam at their delivery date. All children were born between April 2002 and January 2006. During the two postnatal phases of the study (0-4 years and 6 years), information was obtained from 7,893 and 8,305 children, respectively. In the pre-school period, information was collected by parental reported questionnaires at the ages of 2,6,12,14,24, 36 and 48 months, and regular routine visits to the child health centers. At 6 years of age, blood samples were drawn from 4,593 children. The overall blood sample response rate at 6 years of age was 69%. The 31% missing were due to lack of consent by the parents or non-successful vena punctures.[144] The majority of studies presented in Part I "Shaping of adaptive immunity", were performed in a subgroup of 1,232 Dutch pregnant women and their children born between February 2003 and August 2005, and include more detailed measurements.[144] This subgroup is ethnically homogeneous to exclude confounding or effect modification by ethnicity. From a total of 1,247 children, 1,079 children were available for postnatal analyses. After birth, the children were invited to visit the research center at 1.5, 6, 14, 24 and 48 months of age. Measurements during these visits included, among others, collection of venous blood samples (both serum and plasma). Detailed immune phenotyping was available from cord blood (n=230), 6 months (n=377), 14 months (n=243), and 24 months (n=259). At 6 years of age, n=1,121 Dutch and non-Dutch children of the cohort provided data on detailed immune phenotyping. Data on total T, B and NK cell numbers (TruCounts) at 6 years of age were available in n=3,465 children. Information on determinants such as breastfeeding, and many other environmental determinants (e.g. socioeconomic status, maternal

smoking habits, number of siblings, daycare attendance and so on) was obtained from a combination of parental reported pre- and postnatal questionnaires at the ages of 1.5 and 6 months, 1,2,3,4, and 6 years of age. Celiac testing at 6 years of age was performed in 4,442 children.

OUTLINE

The objectives are addressed in several studies presented in this thesis. **Part I** is devoted to the shaping of the adaptive immune system. In **Chapter 2**, we aim to identify developmental age-related patterns of blood innate leukocyte, naive and memory B and T cells numbers, and to study external determinants which may influence these patterns. Next, we describe to what extent breastfeeding (**Chapter 3**), and persistent herpesvirus infections (**Chapter 4**) influence the build-up of adaptive immune cell numbers, and whether the effects of cytomegalovirus (CMV) on the immune system are differentially affected by sex (**Chapter 5**). Additionally, we study whether CDA is associated with abnormalities in blood T cell numbers (**Chapter 6**). **Part II** focuses on determinants of CDA. We study the associations between infant feeding and anti-tissue transglutaminase positivity (**Chapter 7**), and between herpesvirus infections and CDA (**Chapter 8**). Additionally, we describe whether ethnic differences are present in CDA and to what extent these can be explained by environmental factors (**Chapter 9**). **Chapter 10** describes to what extent ethnic differences in herpesvirus infections can be explained by social and environmental factors. **Part III** covers the consequences of CDA. First, we study consequences of CDA on childhood growth and bone development (**Chapter 11**). Second, we describe the identification of subclinical CeD by screening in asymptomatic children (**Chapter 12**). Finally, **Chapter 13** provides an overall discussion, including a description of the main findings of this thesis. The results are described in a broader context, including methodological considerations, and directions for future research. A summary of this thesis is presented in **Chapter 14.**

REFERENCES

1. Kragh M, Larsen JM, Thysen AH, Rasmussen MA, Wolsk HM, Bisgaard H, et al. Divergent response profile in activated cord blood T cells from first-born child implies birth-order-associated in utero immune programming. Allergy. 2016;71:323-32.
2. Santner-Nanan B, Straubinger K, Hsu P, Parnell G, Tang B, Xu B, et al. Fetal-maternal alignment of regulatory T cells correlates with IL-10 and Bcl-2 upregulation in pregnancy. J Immunol. 2013;191:145-53.
3. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee. Lancet. 1998;351:1225-32.
4. Harjutsalo V, Sjoberg L, Tuomilehto J. Time trends in the incidence of type 1 diabetes in Finnish children: a cohort study. Lancet. 2008;371:1777-82.
5. Wheatley LM, Togias A. Clinical practice. Allergic rhinitis. N Engl J Med. 2015;372:456-63.
6. Sicherer SH, Sampson HA. Food allergy: Epidemiology, pathogenesis, diagnosis, and treatment. J Allergy Clin Immunol. 2014;133:291-307; quiz 8.
7. WHO. Prevalence of asthma and allergies in children. Fact sheet no 3.1; may 2007. (http://www.euro.who.int/__data/assets/pdf_file/0012/96996/3.1.pdf). . World Health Organization; 2007.
8. Bach JF. The effect of infections on susceptibility to autoimmune and allergic diseases. N Engl J Med. 2002;347:911-20.
9. Murphy KP. Janeway's immunobiology 8ed. New York: Garland Science, Taylor & Francis Group, LLC; 2012.
10. Appay V, Rowland-Jones SL. Lessons from the study of T-cell differentiation in persistent human virus infection. Semin Immunol. 2004;16:205-12.
11. Lanzavecchia A, Sallusto F. Progressive differentiation and selection of the fittest in the immune response. Nat Rev Immunol. 2002;2:982-7.
12. Sallusto F, Geginat J, Lanzavecchia A. Central memory and effector memory T cell subsets: function, generation, and maintenance. Annu Rev Immunol. 2004;22:745-63.
13. Appay V, van Lier RA, Sallusto F, Roederer M. Phenotype and function of human T lymphocyte subsets: consensus and issues. Cytometry A. 2008;73:975-83.
14. Sallusto F, Lenig D, Forster R, Lipp M, Lanzavecchia A. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. Nature. 1999;401:708-12.
15. Johannisson A, Festin R. Phenotype transition of CD4+ T cells from CD45RA to CD45R0 is accompanied by cell activation and proliferation. Cytometry. 1995;19:343-52.
16. Henson SM, Riddell NE, Akbar AN. Properties of end-stage human T cells defined by CD45RA re-expression. Curr Opin Immunol. 2012;24:476-81.
17. Perez-Andres M, Paiva B, Nieto WG, Caraux A, Schmitz A, Almeida J, et al. Human peripheral blood B-cell compartments: a crossroad in B-cell traffic. Cytometry B Clin Cytom. 2010;78 Suppl 1:S47-60.
18. Morbach H, Eichhorn EM, Liese JG, Girschick HJ. Reference values for B cell subpopulations from infancy to adulthood. Clin Exp Immunol. 2010;162:271-9.
19. Berkowska MA, Driessen GJ, Bikos V, Grosserichter-Wagener C, Stamatopoulos K, Cerutti A, et al. Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways. Blood. 2011;118:2150-8.
20. Berkowska MA, Heeringa JJ, Hajdarbegovic E, van der Burg M, Thio HB, van Hagen PM, et al. Human IgE(+) B cells are derived from T cell-dependent and T cell-independent pathways. J Allergy Clin Immunol. 2014;134:688-97 e6.
21. Weller S, Braun MC, Tan BK, Rosenwald A, Cordier C, Conley ME, et al. Human blood IgM "memory" B cells are circulating splenic marginal zone B cells harboring a prediversified immunoglobulin repertoire. Blood. 2004;104:3647-54.

22. Piatosa B, Wolska-Kusnierz B, Pac M, Siewiera K, Galkowska E, Bernatowska E. B cell subsets in healthy children: reference values for evaluation of B cell maturation process in peripheral blood. Cytometry B Clin Cytom. 2010;78:372-81.
23. Duchamp M, Sterlin D, Diabate A, Uring-Lambert B, Guerin-El Khourouj V, Le Mauff B, et al. B-cell subpopulations in children: National reference values. Immun Inflamm Dis. 2014;2:131-40.
24. Tosato F, Bucciol G, Pantano G, Putti MC, Sanzari MC, Basso G, et al. Lymphocytes subsets reference values in childhood. Cytometry A. 2015;87:81-5.
25. Qi Q, Zhang DW, Weyand CM, Goronzy JJ. Mechanisms shaping the naive T cell repertoire in the elderly - thymic involution or peripheral homeostatic proliferation? Exp Gerontol. 2014;54:71-4.
26. Tsao PN, Chiang BL, Yang YH, Tsai MJ, Lu FL, Chou HC, et al. Longitudinal follow-up of lymphocyte subsets during the first year of life. Asian Pac J Allergy Immunol. 2002;20:147-53.
27. Comans-Bitter WM, de Groot R, van den Beemd R, Neijens HJ, Hop WC, Groeneveld K, et al. Immunophenotyping of blood lymphocytes in childhood. Reference values for lymphocyte subpopulations. J Pediatr. 1997;130:388-93.
28. van Gent R, van Tilburg CM, Nibbelke EE, Otto SA, Gaiser JF, Janssens-Korpela PL, et al. Refined characterization and reference values of the pediatric T- and B-cell compartments. Clin Immunol. 2009;133:95-107.
29. Shearer WT, Rosenblatt HM, Gelman RS, Oyomopito R, Plaeger S, Stiehm ER, et al. Lymphocyte subsets in healthy children from birth through 18 years of age: the Pediatric AIDS Clinical Trials Group P1009 study. J Allergy Clin Immunol. 2003;112:973-80.
30. de Vries E, de Bruin-Versteeg S, Comans-Bitter WM, de Groot R, Hop WC, Boerma GJ, et al. Longitudinal survey of lymphocyte subpopulations in the first year of life. Pediatr Res. 2000;47:528-37.
31. Driessen GJ, Dalm VA, van Hagen PM, Grashoff HA, Hartwig NG, van Rossum AM, et al. Common variable immunodeficiency and idiopathic primary hypogammaglobulinemia: two different conditions within the same disease spectrum. Haematologica. 2013;98:1617-23.
32. Sagnia B, Ateba Ndongo F, Ndiang Moyo Tetang S, Ndongo Torimiro J, Cairo C, Domkam I, et al. Reference values of lymphocyte subsets in healthy, HIV-negative children in Cameroon. Clin Vaccine Immunol. 2011;18:790-5.
33. Virella G, Silveira Nunes MA, Tamagnini G. Placental transfer of human IgG subclasses. Clin Exp Immunol. 1972;10:475-8.
34. Marques AH, O'Connor TG, Roth C, Susser E, Bjorke-Monsen AL. The influence of maternal prenatal and early childhood nutrition and maternal prenatal stress on offspring immune system development and neurodevelopmental disorders. Front Neurosci. 2013;7:120.
35. Duijts L, Bakker-Jonges LE, Labout JA, Jaddoe VW, Hofman A, Steegers EA, et al. Perinatal stress influences lymphocyte subset counts in neonates. The generation R study. Pediatr Res. 2008;63:292-8.
36. Yildiran A, Yurdakul E, Guloglu D, Dogu F, Arsan S, Arikan M, et al. The effect of mode of delivery on T regulatory (Treg) cells of cord blood. Indian J Pediatr. 2011;78:1234-8.
37. Duijts L, Bakker-Jonges LE, Labout JA, Jaddoe VW, Hofman A, Steegers EA, et al. Fetal growth influences lymphocyte subset counts at birth: the Generation R Study. Neonatology. 2009;95:149-56.
38. Zhang Y, Collier F, Naselli G, Saffery R, Tang ML, Allen KJ, et al. Cord blood monocyte-derived inflammatory cytokines suppress IL-2 and induce nonclassic "TH2-type" immunity associated with development of food allergy. Sci Transl Med. 2016;8:321ra8.
39. Ismail IH, Boyle RJ, Mah LJ, Licciardi PV, Tang ML. Reduced neonatal regulatory T cell response to microbial stimuli associates with subsequent eczema in high-risk infants. Pediatr Allergy Immunol. 2014;25:674-84.
40. Thysen AH, Rasmussen MA, Kreiner-Moller E, Larsen JM, Folsgaard NV, Bonnelykke K, et al. Season of birth shapes neonatal immune function. J Allergy Clin Immunol. 2015.
41. Duijts L, Ramadhani MK, Moll HA. Breastfeeding protects against infectious diseases during infancy in industrialized countries. A systematic review. Matern Child Nutr. 2009;5:199-210.

42. Hanson LA, Korotkova M, Lundin S, Haversen L, Silfverdal SA, Mattsby-Baltzer I, et al. The transfer of immunity from mother to child. Ann N Y Acad Sci. 2003;987:199-206.
43. Alter SJ, Bennett JS, Koranyi K, Kreppel A, Simon R. Common childhood viral infections. Curr Probl Pediatr Adolesc Health Care. 2015;45:21-53.
44. Victora CG, Bahl R, Barros AJ, Franca GV, Horton S, Krasevec J, et al. Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. Lancet. 2016;387:475-90.
45. Derhovanessian E, Maier AB, Hahnel K, McElhaney JE, Slagboom EP, Pawelec G. Latent infection with cytomegalovirus is associated with poor memory CD4 responses to influenza A core proteins in the elderly. J Immunol. 2014;193:3624-31.
46. Frasca D, Diaz A, Romero M, Landin AM, Blomberg BB. Cytomegalovirus (CMV) seropositivity decreases B cell responses to the influenza vaccine. Vaccine. 2015;33:1433-9.
47. Khan N, Hislop A, Gudgeon N, Cobbold M, Khanna R, Nayak L, et al. Herpesvirus-specific CD8 T cell immunity in old age: cytomegalovirus impairs the response to a coresident EBV infection. J Immunol. 2004;173:7481-9.
48. Remmerswaal EB, Havenith SH, Idu MM, van Leeuwen EM, van Donselaar KA, Ten Brinke A, et al. Human virus-specific effector-type T cells accumulate in blood but not in lymph nodes. Blood. 2012;119:1702-12.
49. van Leeuwen EM, Koning JJ, Remmerswaal EB, van Baarle D, van Lier RA, ten Berge IJ. Differential usage of cellular niches by cytomegalovirus versus EBV- and influenza virus-specific CD8+ T cells. J Immunol. 2006;177:4998-5005.
50. Odumade OA, Knight JA, Schmeling DO, Masopust D, Balfour HH, Jr., Hogquist KA. Primary Epstein-Barr virus infection does not erode preexisting CD8(+) T cell memory in humans. J Exp Med. 2012;209:471-8.
51. Nilsson C, Larsson Sigfrinius AK, Montgomery SM, Sverremark-Ekstrom E, Linde A, Lilja G, et al. Epstein-Barr virus and cytomegalovirus are differentially associated with numbers of cytokine-producing cells and early atopy. Clin Exp Allergy. 2009;39:509-17.
52. Saghafian-Hedengren S, Sverremark-Ekstrom E, Linde A, Lilja G, Nilsson C. Early-life EBV infection protects against persistent IgE sensitization. J Allergy Clin Immunol. 2010;125:433-8.
53. Alcantara-Neves NM, Veiga RV, Dattoli VC, Fiaccone RL, Esquivel R, Cruz AA, et al. The effect of single and multiple infections on atopy and wheezing in children. J Allergy Clin Immunol. 2012;129:359-67, 67 e1-3.
54. Smith JS, Robinson NJ. Age-specific prevalence of infection with herpes simplex virus types 2 and 1: a global review. J Infect Dis. 2002;186 Suppl 1:S3-28.
55. Korndewal MJ, Mollema L, Tcherniaeva I, van der Klis F, Kroes AC, Oudesluys-Murphy AM, et al. Cytomegalovirus infection in the Netherlands: seroprevalence, risk factors, and implications. J Clin Virol. 2015;63:53-8.
56. Cannon MJ, Schmid DS, Hyde TB. Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. Rev Med Virol. 2010;20:202-13.
57. Staras SA, Dollard SC, Radford KW, Flanders WD, Pass RF, Cannon MJ. Seroprevalence of cytomegalovirus infection in the United States, 1988-1994. Clin Infect Dis. 2006;43:1143-51.
58. Bradley H, Markowitz LE, Gibson T, McQuillan GM. Seroprevalence of herpes simplex virus types 1 and 2--United States, 1999-2010. J Infect Dis. 2014;209:325-33.
59. Gaytant MA, Galama JM, Semmekrot BA, Melchers WJ, Sporken JM, Oosterbaan HP, et al. The incidence of congenital cytomegalovirus infections in The Netherlands. J Med Virol. 2005;76:71-5.
60. Dowd JB, Palermo T, Brite J, McDade TW, Aiello A. Seroprevalence of Epstein-Barr virus infection in U.S. children ages 6-19, 2003-2010. PLoS One. 2013;8:e64921.
61. Pembrey L, Raynor P, Griffiths P, Chaytor S, Wright J, Hall AJ. Seroprevalence of cytomegalovirus, Epstein Barr virus and varicella zoster virus among pregnant women in Bradford: a cohort study. PLoS One. 2013;8:e81881.

62. van Lier A, Smits G, Mollema L, Waaijenborg S, Berbers G, van der Klis F, et al. Varicella zoster virus infection occurs at a relatively young age in The Netherlands. Vaccine. 2013;31:5127-33.
63. Dowd JB, Aiello AE, Alley DE. Socioeconomic disparities in the seroprevalence of cytomegalovirus infection in the US population: NHANES III. Epidemiol Infect. 2009;137:58-65.
64. Rook GA. Hygiene hypothesis and autoimmune diseases. Clin Rev Allergy Immunol. 2012;42:5-15.
65. Dreyfus DH. Herpesviruses and the microbiome. J Allergy Clin Immunol. 2013;132:1278-86.
66. Janson C, Asbjornsdottir H, Birgisdottir A, Sigurjonsdottir RB, Gunnbjornsdottir M, Gislason D, et al. The effect of infectious burden on the prevalence of atopy and respiratory allergies in Iceland, Estonia, and Sweden. J Allergy Clin Immunol. 2007;120:673-9.
67. Benoist C, Mathis D. Autoimmunity provoked by infection: how good is the case for T cell epitope mimicry? Nat Immunol. 2001;2:797-801.
68. Dreyfus DH. Autoimmune disease: A role for new anti-viral therapies? Autoimmun Rev. 2011;11:88-97.
69. Draborg AH, Duus K, Houen G. Epstein-Barr virus in systemic autoimmune diseases. Clin Dev Immunol. 2013;2013:535738.
70. Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. Nature. 2007;448:470-3.
71. van Bergen J, Mulder CJ, Mearin ML, Koning F. Local communication among mucosal immune cells in patients with celiac disease. Gastroenterology. 2015;148:1187-94.
72. Lebwohl B, Green PH. Antitissue Transglutaminase IgA for Celiac Disease Testing. JAMA. 2016;315:81-2.
73. Husby S, Koletzko S, Korponay-Szabo IR, Mearin ML, Phillips A, Shamir R, et al. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr. 2012;54:136-60.
74. Leffler DA, Schuppan D. Update on serologic testing in celiac disease. Am J Gastroenterol. 2010;105:2520-4.
75. Dieli-Crimi R, Cenit MC, Nunez C. The genetics of celiac disease: A comprehensive review of clinical implications. J Autoimmun. 2015;64:26-41.
76. Mustalahti K, Catassi C, Reunanen A, Fabiani E, Heier M, McMillan S, et al. The prevalence of celiac disease in Europe: results of a centralized, international mass screening project. Ann Med. 2010;42:587-95.
77. Guandalini S, Assiri A. Celiac disease: a review. JAMA Pediatr. 2014;168:272-8.
78. Leffler DA, Green PH, Fasano A. Extraintestinal manifestations of coeliac disease. Nat Rev Gastroenterol Hepatol. 2015;12:561-71.
79. Lundin KE, Wijmenga C. Coeliac disease and autoimmune disease-genetic overlap and screening. Nat Rev Gastroenterol Hepatol. 2015;12:507-15.
80. Ludvigsson JF, Leffler DA, Bai JC, Biagi F, Fasano A, Green PH, et al. The Oslo definitions for coeliac disease and related terms. Gut. 2013;62:43-52.
81. Maki M, Collin P. Coeliac disease. Lancet. 1997;349:1755-9.
82. Catassi C, Bearzi I, Holmes GK. Association of celiac disease and intestinal lymphomas and other cancers. Gastroenterology. 2005;128:S79-86.
83. Romanos J, Rosen A, Kumar V, Trynka G, Franke L, Szperl A, et al. Improving coeliac disease risk prediction by testing non-HLA variants additional to HLA variants. Gut. 2014;63:415-22.
84. Vriezinga SL, Auricchio R, Bravi E, Castillejo G, Chmielewska A, Crespo Escobar P, et al. Randomized feeding intervention in infants at high risk for celiac disease. N Engl J Med. 2014;371:1304-15.
85. Monsuur AJ, de Bakker PI, Zhernakova A, Pinto D, Verduijn W, Romanos J, et al. Effective detection of human leukocyte antigen risk alleles in celiac disease using tag single nucleotide polymorphisms. PLoS One. 2008;3:e2270.
86. Ivarsson A, Myleus A, Norstrom F, van der Pals M, Rosen A, Hogberg L, et al. Prevalence of childhood celiac disease and changes in infant feeding. Pediatrics. 2013;131:e687-94.

87. Ivarsson A, Persson LA, Nystrom L, Ascher H, Cavell B, Danielsson L, et al. Epidemic of coeliac disease in Swedish children. Acta Paediatr. 2000;89:165-71.
88. Carlsson A, Agardh D, Borulf S, Grodzinsky E, Axelsson I, Ivarsson SA. Prevalence of celiac disease: before and after a national change in feeding recommendations. Scand J Gastroenterol. 2006;41:553-8.
89. Norris JM, Barriga K, Hoffenberg EJ, Taki I, Miao D, Haas JE, et al. Risk of celiac disease autoimmunity and timing of gluten introduction in the diet of infants at increased risk of disease. JAMA. 2005;293:2343-51.
90. Stordal K, White RA, Eggesbo M. Early feeding and risk of celiac disease in a prospective birth cohort. Pediatrics. 2013;132:e1202-9.
91. Ivarsson A, Hernell O, Stenlund H, Persson LA. Breast-feeding protects against celiac disease. Am J Clin Nutr. 2002;75:914-21.
92. D'Amico MA, Holmes J, Stavropoulos SN, Frederick M, Levy J, DeFelice AR, et al. Presentation of pediatric celiac disease in the United States: prominent effect of breastfeeding. Clin Pediatr (Phila). 2005;44:249-58.
93. Peters U, Schneeweiss S, Trautwein EA, Erbersdobler HF. A case-control study of the effect of infant feeding on celiac disease. Ann Nutr Metab. 2001;45:135-42.
94. Welander A, Tjernberg AR, Montgomery SM, Ludvigsson J, Ludvigsson JF. Infectious disease and risk of later celiac disease in childhood. Pediatrics. 2010;125:e530-6.
95. Ascher H, Krantz I, Rydberg L, Nordin P, Kristiansson B. Influence of infant feeding and gluten intake on coeliac disease. Arch Dis Child. 1997;76:113-7.
96. Lionetti E, Castellaneta S, Francavilla R, Pulvirenti A, Tonutti E, Amarri S, et al. Introduction of gluten, HLA status, and the risk of celiac disease in children. N Engl J Med. 2014;371:1295-303.
97. Liu E, Lee HS, Aronsson CA, Hagopian WA, Koletzko S, Rewers MJ, et al. Risk of pediatric celiac disease according to HLA haplotype and country. N Engl J Med. 2014;371:42-9.
98. Stene LC, Honeyman MC, Hoffenberg EJ, Haas JE, Sokol RJ, Emery L, et al. Rotavirus infection frequency and risk of celiac disease autoimmunity in early childhood: a longitudinal study. Am J Gastroenterol. 2006;101:2333-40.
99. Kang JY, Kang AH, Green A, Gwee KA, Ho KY. Systematic review: worldwide variation in the frequency of coeliac disease and changes over time. Aliment Pharmacol Ther. 2013;38:226-45.
100. Lohi S, Mustalahti K, Kaukinen K, Laurila K, Collin P, Rissanen H, et al. Increasing prevalence of coeliac disease over time. Aliment Pharmacol Ther. 2007;26:1217-25.
101. Rubio-Tapia A, Kyle RA, Kaplan EL, Johnson DR, Page W, Erdtmann F, et al. Increased prevalence and mortality in undiagnosed celiac disease. Gastroenterology. 2009;137:88-93.
102. Riddle MS, Murray JA, Porter CK. The incidence and risk of celiac disease in a healthy US adult population. Am J Gastroenterol. 2012;107:1248-55.
103. Xu F, Sternberg MR, Kottiri BJ, McQuillan GM, Lee FK, Nahmias AJ, et al. Trends in herpes simplex virus type 1 and type 2 seroprevalence in the United States. JAMA. 2006;296:964-73.
104. Aarnisalo J, Ilonen J, Vainionpaa R, Volanen I, Kaitosaari T, Simell O. Development of antibodies against cytomegalovirus, varicella-zoster virus and herpes simplex virus in Finland during the first eight years of life: a prospective study. Scand J Infect Dis. 2003;35:750-3.
105. Choung RS, Ditah IC, Nadeau AM, Rubio-Tapia A, Marietta EV, Brantner TL, et al. Trends and racial/ethnic disparities in gluten-sensitive problems in the United States: findings from the National Health and Nutrition Examination Surveys from 1988 to 2012. Am J Gastroenterol. 2015;110:455-61.
106. Sandberg-Bennich S, Dahlquist G, Kallen B. Coeliac disease is associated with intrauterine growth and neonatal infections. Acta Paediatr. 2002;91:30-3.
107. Marild K, Kahrs CR, Tapia G, Stene LC, Stordal K. Infections and risk of celiac disease in childhood: a prospective nationwide cohort study. Am J Gastroenterol. 2015;110:1475-84.
108. Lucendo AJ, Garcia-Manzanares A. Bone mineral density in adult coeliac disease: an updated review. Rev Esp Enferm Dig. 2013;105:154-62.

109. Corazza GR, Di Stefano M, Maurino E, Bai JC. Bones in coeliac disease: diagnosis and treatment. Best Pract Res Clin Gastroenterol. 2005;19:453-65.
110. Ludvigsson JF, Michaelsson K, Ekbom A, Montgomery SM. Coeliac disease and the risk of fractures - a general population-based cohort study. Aliment Pharmacol Ther. 2007;25:273-85.
111. Ludvigsson JF, Card TR, Kaukinen K, Bai J, Zingone F, Sanders DS, et al. Screening for celiac disease in the general population and in high-risk groups. United European Gastroenterol J. 2015;3:106-20.
112. Dhalwani NN, West J, Sultan AA, Ban L, Tata LJ. Women with celiac disease present with fertility problems no more often than women in the general population. Gastroenterology. 2014;147:1267-74 e1; quiz e13-4.
113. Godfrey JD, Brantner TL, Brinjikji W, Christensen KN, Brogan DL, Van Dyke CT, et al. Morbidity and mortality among older individuals with undiagnosed celiac disease. Gastroenterology. 2010;139:763-9.
114. Hu WT, Murray JA, Greenaway MC, Parisi JE, Josephs KA. Cognitive impairment and celiac disease. Arch Neurol. 2006;63:1440-6.
115. Lichtwark IT, Newnham ED, Robinson SR, Shepherd SJ, Hosking P, Gibson PR, et al. Cognitive impairment in coeliac disease improves on a gluten-free diet and correlates with histological and serological indices of disease severity. Aliment Pharmacol Ther. 2014;40:160-70.
116. Green PH, Fleischauer AT, Bhagat G, Goyal R, Jabri B, Neugut AI. Risk of malignancy in patients with celiac disease. Am J Med. 2003;115:191-5.
117. Elfstrom P, Granath F, Ye W, Ludvigsson JF. Low risk of gastrointestinal cancer among patients with celiac disease, inflammation, or latent celiac disease. Clin Gastroenterol Hepatol. 2012;10:30-6.
118. Jatla M, Zemel BS, Bierly P, Verma R. Bone mineral content deficits of the spine and whole body in children at time of diagnosis with celiac disease. J Pediatr Gastroenterol Nutr. 2009;48:175-80.
119. Matysiak-Budnik T, Malamut G, de Serre NP, Grosdidier E, Seguier S, Brousse N, et al. Long-term follow-up of 61 coeliac patients diagnosed in childhood: evolution toward latency is possible on a normal diet. Gut. 2007;56:1379-86.
120. Singh P, Arora S, Lal S, Strand TA, Makharia GK. Celiac Disease in Women With Infertility: A Meta-Analysis. J Clin Gastroenterol. 2016;50:33-9.
121. Abdul Sultan A, Crooks CJ, Card T, Tata LJ, Fleming KM, West J. Causes of death in people with coeliac disease in England compared with the general population: a competing risk analysis. Gut. 2015;64:1220-6.
122. Green PH, Cellier C. Celiac disease. N Engl J Med. 2007;357:1731-43.
123. Corrao G, Corazza GR, Bagnardi V, Brusco G, Ciacci C, Cottone M, et al. Mortality in patients with coeliac disease and their relatives: a cohort study. Lancet. 2001;358:356-61.
124. Holmes GK, Prior P, Lane MR, Pope D, Allan RN. Malignancy in coeliac disease--effect of a gluten free diet. Gut. 1989;30:333-8.
125. Logan RF, Rifkind EA, Turner ID, Ferguson A. Mortality in celiac disease. Gastroenterology. 1989;97:265-71.
126. Elfstrom P, Granath F, Ekstrom Smedby K, Montgomery SM, Askling J, Ekbom A, et al. Risk of lymphoproliferative malignancy in relation to small intestinal histopathology among patients with celiac disease. J Natl Cancer Inst. 2011;103:436-44.
127. Canavan C, Logan RF, Khaw KT, West J. No difference in mortality in undetected coeliac disease compared with the general population: a UK cohort study. Aliment Pharmacol Ther. 2011;34:1012-9.
128. Ludvigsson JF, Montgomery SM, Ekbom A, Brandt L, Granath F. Small-intestinal histopathology and mortality risk in celiac disease. JAMA. 2009;302:1171-8.
129. Tio M, Cox MR, Eslick GD. Meta-analysis: coeliac disease and the risk of all-cause mortality, any malignancy and lymphoid malignancy. Aliment Pharmacol Ther. 2012;35:540-51.
130. Peters U, Askling J, Gridley G, Ekbom A, Linet M. Causes of death in patients with celiac disease in a population-based Swedish cohort. Arch Intern Med. 2003;163:1566-72.
131. Goddard CJ, Gillett HR. Complications of coeliac disease: are all patients at risk? Postgrad Med J. 2006;82:705-12.

132. Biagi F, Corazza GR. Do different patients with coeliac disease have different mortality rates? Gut. 2015;64:1187-8.
133. Biagi F, Corazza GR. Mortality in celiac disease. Nat Rev Gastroenterol Hepatol. 2010;7:158-62.
134. Lebwohl B, Rubio-Tapia A, Assiri A, Newland C, Guandalini S. Diagnosis of celiac disease. Gastrointest Endosc Clin N Am. 2012;22:661-77.
135. Mora S, Barera G, Beccio S, Proverbio MC, Weber G, Bianchi C, et al. Bone density and bone metabolism are normal after long-term gluten-free diet in young celiac patients. Am J Gastroenterol. 1999;94:398-403.
136. Kavak US, Yuce A, Kocak N, Demir H, Saltik IN, Gurakan F, et al. Bone mineral density in children with untreated and treated celiac disease. J Pediatr Gastroenterol Nutr. 2003;37:434-6.
137. Turner J, Pellerin G, Mager D. Prevalence of metabolic bone disease in children with celiac disease is independent of symptoms at diagnosis. J Pediatr Gastroenterol Nutr. 2009;49:589-93.
138. Mazure R, Vazquez H, Gonzalez D, Mautalen C, Pedreira S, Boerr L, et al. Bone mineral affection in asymptomatic adult patients with celiac disease. Am J Gastroenterol. 1994;89:2130-4.
139. Mora S, Barera G, Ricotti A, Weber G, Bianchi C, Chiumello G. Reversal of low bone density with a gluten-free diet in children and adolescents with celiac disease. Am J Clin Nutr. 1998;67:477-81.
140. Mautalen C, Gonzalez D, Mazure R, Vazquez H, Lorenzetti MP, Maurino E, et al. Effect of treatment on bone mass, mineral metabolism, and body composition in untreated celiac disease patients. Am J Gastroenterol. 1997;92:313-8.
141. Cellier C, Flobert C, Cormier C, Roux C, Schmitz J. Severe osteopenia in symptom-free adults with a childhood diagnosis of coeliac disease. Lancet. 2000;355:806.
142. Corazza GR, Di Sario A, Cecchetti L, Jorizzo RA, Di Stefano M, Minguzzi L, et al. Influence of pattern of clinical presentation and of gluten-free diet on bone mass and metabolism in adult coeliac disease. Bone. 1996;18:525-30.
143. Mustalahti K, Collin P, Sievanen H, Salmi J, Maki M. Osteopenia in patients with clinically silent coeliac disease warrants screening. Lancet. 1999;354:744-5.
144. Kruithof CJ, Kooijman MN, van Duijn CM, Franco OH, de Jongste JC, Klaver CC, et al. The Generation R Study: Biobank update 2015. Eur J Epidemiol. 2014;29:911-27.
145. Kooijman MN, Kruithof CJ, van Duijn CM, Duijts L, Franco OH, van IMH, et al. The Generation R Study: design and cohort update 2017. Eur J Epidemiol. 2016;31:1243-64.
146. Van den Heuvel D. The Immune System out of Shape? Shaping of adaptive immunity by persistent viral infections in young children. Rotterdam: Erasmus University Rotterdam; 2015.

Part I

The shaping of adaptive immunity

Chapter 2

Effects of nongenetic factors on immune cell dynamics in early childhood: The Generation R Study

Diana van den Heuvel, Michelle A.E. Jansen, Kazem Nasserinejad, Wim A Dik, Ellen G van Lochem, Liesbeth E. Bakker-Jonges, Halima Bouallouch-Charif, Vincent W.V. Jaddoe, Herbert Hooijkaas , Jacques J.M. van Dongen, Henriette A. Moll , Menno C. van Zelm MC

J Allergy Clin Immunol. 2016 November 29

ABSTRACT

Background

Numbers of blood leukocyte subsets are highly dynamic in childhood, and differ greatly between subjects. Interindividual variation is only partly accounted for by genetic factors.

Objective

We sought to determine which nongenetic factors affect the dynamics of innate leukocytes and naive and memory lymphocyte subsets.

Methods

We performed 6-color flow cytometry and linear mixed-effects modeling to define the dynamics of 62 leukocyte subsets from birth to 6 years of age in 1182 children, with 1 to 5 measurements per subject. Subsequently, we defined the effect of prenatal maternal lifestyle-related or immunemediated determinants, birth characteristics, and bacterial/viral exposure–related determinants on leukocyte subset dynamics.

Results

Functionally similar leukocyte populations were grouped by using unbiased hierarchical clustering of patterns of age-related leukocyte dynamics. Innate leukocyte numbers were high at birth and predominantly affected by maternal low education level. Naive lymphocyte counts peaked around 1 year, whereas most memory lymphocyte subsets more gradually increased during the first 4 years of life. Dynamics of $CD4^+$ T cells were predominantly associated with sex, birth characteristics, and persistent infections with cytomegalovirus (CMV) or EBV. $CD8^+$ T cells were predominantly associated with CMV and EBV infections, and T-cell receptor $\gamma\delta^+$ T cells were predominantly associated with premature rupture of membranes and CMV infection. B-cell subsets were predominantly associated with sex, breast-feeding, and Helicobacter pylori carriership.

Conclusions

Our study identifies specific dynamic patterns of leukocyte subset numbers, as well as nongenetic determinants that affect these patterns, thereby providing new insights into the shaping of the childhood immune system.

INTRODUCTION

The human immune system shows high diversity in cellular composition and functional responses between subjects. Blood leukocytes in young children are highly dynamic in number and composition.[1-6] Numbers of innate cells, such as neutrophils and natural killer (NK) cells, are higher in neonates than in children or adults[1,5,7] and already display dynamic changes within the first few days after birth.[1,5,6] B and T cells are mostly naive in infants, whereas protective immunity is gradually built up in the form of increasing numbers of memory B (Bmem) and memory T (Tmem) cells during the first 5 years of life, after which these numbers stabilize.[2-4,8-11] Recent studies have found that 25% to 50% of interindividual variation in cellular composition and functional responses were accounted for by genetic factors,[12,13] indicating that nongenetic determinants underlie a large part of immune trait variance,[14] which will potentially have long-term effects and underlie part of the immune trait variation in adults. To date, correlations of patterns between different immune cells are incompletely studied because of either the restricted numbers of analyzed subsets, the short-term follow-up, or the lack of longitudinal analyses correcting for intrapersonal correlations between repeated measures. Furthermore, the nongenetic determinants that drive the kinetics of each type of immune cell remain less well studied. It is likely that these involve various environmental determinants, such as maternal lifestyle, maternal immune-mediated diseases, birth characteristics, and bacterial and viral exposure in childhood.[15-24] Studying the effect of these nongenetic factors on the dynamics of childhood immune development requires large cohorts of healthy young children, with multiple measurements per subject. We investigated which nongenetic factors related to the prenatal maternal lifestyle, prenatal maternal immune-mediated diseases, and birth characteristics or bacterial/viral exposure- related characteristics influence the dynamics of blood leukocyte populations from birth until 6 years of age. This concerned a total of 62 leukocyte populations, including innate leukocyte subsets and naive T- and B-cell and Bmem and Tmem subsets.

PATIENTS AND METHODS

Study subjects

This study was embedded in the Generation R Study, a prospective population-based cohort study from fetal life until young adulthood.[25,26] The current study focused on a subgroup of 1182 two-generation Dutch children born between August 2003 and August 2006. Peripheral blood was obtained at birth and a median age of 6, 14, 25, and 72 months. Detailed immunophenotyping was performed at 1 to 5 time points per child, resulting in a total of 2010 data points (details are provided in supplemental material). Written informed

consent was obtained from parents according to the Medical Ethics Committee guidelines of Erasmus MC.

Immunophenotyping

Absolute numbers of leukocytes, NK cells, T cells and B cells were obtained with a diagnostic lyse-no-wash protocol, using a BD FACSCalibur (BD Biosciences, San Jose, Calif). Six-color flow cytometry was performed to quantify 62 well-defined leukocyte populations (Table 1), which included NK-cell, monocyte, granulocyte, naive and memory B-cell, naive and memory TCRαβ+ T-cell and TCRγδ+ T-cell subsets (Supplemental Table 1). Flowcytometric analyses were performed on an LSRII (BD Biosciences) by using standardized measurement settings (details in supplemental material).27

Table 1. Definition of leukocyte subsets

n=62	n=31	Name population	Phenotype definition
		Innate leukocytes	
1		**Granulocytes**	$SSC^{high}CD45^{+}$
2		CD15+ granulocytes	$SSC^{high}CD45^{+}CD15^{+}$
3	+	Neutrophils	$SSC^{high}CD45^{+}CD16^{+}$
4	+	Eosinophils	$SSC^{high}CD45^{+}CD16^{-}$
5		**Monocytes**	$FSC^{inter}SSC^{inter}CD45^{+}$
6	+	Classical monocytes	$FSC^{inter}SSC^{inter}CD45^{+}CD14^{+}CD16^{-}$
7	+	Intermediate monocytes	$FSC^{inter}SSC^{inter}CD45^{+}CD14^{+}CD16^{+}$
8	+	Non-classical monocytes	$FSC^{inter}SSC^{inter}CD45^{+}CD14^{-}CD16^{+}$
9	+	**NK cells**	$SSC^{low}CD45^{+}CD3^{-}CD16^{+}$ or $CD56^{+}$
		Lymphocytes	
10		**T cells**	$SSC^{low}CD45^{+}CD3^{+}$
11		TCRαβ+ T cells	$SSC^{low}CD3^{+}TCR\gamma\delta^{-}TCR\alpha\beta^{+}$
12		CD4+ TCRαβ+ T cells	$SSC^{low}CD3^{+}TCR\gamma\delta^{-}TCR\alpha\beta^{+}CD8^{-}CD4^{+}$
13		CD8+ TCRαβ+ T cells	$SSC^{low}CD3^{+}TCR\gamma\delta^{-}TCR\alpha\beta^{+}CD8^{+}CD4^{-}$
14		TCRγδ+ T cells	$SSC^{low}CD3^{+}TCR\gamma\delta^{+}TCR\alpha\beta^{-}$
15		CD4+ TCRγδ+ T cells	$SSC^{low}CD3^{+}TCR\gamma\delta^{+}TCR\alpha\beta^{-}CD8^{-}CD4^{+}$
16		CD8+ TCRγδ+ T cells	$SSC^{low}CD3^{+}TCR\gamma\delta^{+}TCR\alpha\beta^{-}CD8^{+}CD4^{-}$
17		Vδ1-Vδ2- TCRαβ- T cells	$SSC^{low}CD3^{+}TCR\alpha\beta^{-}V\delta1^{-}V\delta2^{-}$
18		Vδ2+ T cells	$SSC^{low}CD3^{+}TCR\alpha\beta^{-}V\delta1^{-}V\delta2^{+}$
19	+	Vδ1+ T cells	$SSC^{low}CD3^{+}TCR\alpha\beta^{-}V\delta1^{+}V\delta2^{-}$
20		Vγ9+ T cells	$FSC^{low}SSC^{low}CD3^{+}TCR\alpha\beta^{-}V\gamma9^{+}$
21	+	Vδ2+Vγ9+ T cells	$FSC^{low}SSC^{low}CD3^{+}TCR\alpha\beta^{-}V\delta2^{+}V\gamma9^{+}$
22		CD4+ T cells	$SSC^{low}CD45^{+}CD3^{+}CD4^{+}CD8^{-}$
23	+	CD4+ Tnaive	$FSC^{low}SSC^{low}CD3^{+}CD8^{-}CD197^{+}CD45RO^{-}CD27^{+}CD28^{+}$
24		CD4+ Tmem	$FSC^{low}SSC^{low}CD3^{+}CD8^{-}$ $CD197^{-}$ and $CD197^{+}CD45RO^{+}$

Table 1. **Definition of leukocyte subsets (continued)**

n=62	n=31	Name population	Phenotype definition
25	+	CD4+ Tcm	$FSC^{low}SSC^{low}CD3^{+}CD8^{-}CD197^{+}CD45RO^{+}CD27^{+}CD28^{+}$
26		CD4+ TemRO	$FSC^{low}SSC^{low}CD3^{+}CD8^{-}CD197^{-}CD45RO^{+}$
27	+	CD4+ early TemRO	$FSC^{low}SSC^{low}CD3^{+}CD8^{-}CD197^{-}CD45RO^{+}CD27^{+}CD28^{+}$
28	+	CD4+ interm TemRO	$FSC^{low}SSC^{low}CD3^{+}CD8^{-}CD197^{-}CD45RO^{+}CD27^{-}CD28^{+}$
29	+	CD4+ late TemRO	$FSC^{low}SSC^{low}CD3^{+}CD8^{-}CD197^{-}CD45RO^{+}CD27^{-}CD28^{-}$
30		CD4+ TemRA	$FSC^{low}SSC^{low}CD3^{+}CD8^{-}CD197^{-}CD45RO^{-}$
31	+	CD4+ early TemRA	$FSC^{low}SSC^{low}CD3^{+}CD8^{-}CD197^{-}CD45RO^{-}CD27^{+}CD28^{+}$
32	+	CD4+ interm TemRA	$FSC^{low}SSC^{low}CD3^{+}CD8^{-}CD197^{-}CD45RO^{-}CD27^{-}CD28^{+}$
33	+	CD4+ late TemRA	$FSC^{low}SSC^{low}CD3^{+}CD8^{-}CD197^{-}CD45RO^{-}CD27^{-}CD28^{-}$
34		CD8+ T cells	$SSC^{low}CD45^{+}CD3^{+}CD4^{-}CD8^{+}$
35	+	CD8+ Tnaive	$FSC^{low}SSC^{low}CD3^{+}CD8^{+}CD197^{+}CD45RO^{-}CD27^{+}CD28^{+}$
36		CD8+ Tmem	$FSC^{low}SSC^{low}CD3^{+}CD8^{+}CD197^{-}$ and $CD197^{+}CD45RO^{+}$
37	+	CD8+ Tcm	$FSC^{low}SSC^{low}CD3^{+}CD8^{+}CD197^{+}CD45RO^{+}CD27^{+}CD28^{+}$
38		CD8+ TemRO	$FSC^{low}SSC^{low}CD3^{+}CD8^{+}CD197^{-}CD45RO^{+}$
39	+	CD8+ early TemRO	$FSC^{low}SSC^{low}CD3^{+}CD8^{+}CD197^{-}CD45RO^{+}CD27^{+}CD28^{+}$
40	+	CD8+ interm TemRO	$FSC^{low}SSC^{low}CD3^{+}CD8^{+}CD197^{-}CD45RO^{+}CD27^{+}CD28^{-}$
41	+	CD8+ late TemRO	$FSC^{low}SSC^{low}CD3^{+}CD8^{+}CD197^{-}CD45RO^{+}CD27^{-}CD28^{-}$
42		CD8+ TemRA	$FSC^{low}SSC^{low}CD3^{+}CD8^{+}CD197^{-}CD45RO^{-}$
43	+	CD8+ early TemRA	$FSC^{low}SSC^{low}CD3^{+}CD8^{+}CD197^{-}CD45RO^{-}CD27^{+}CD28^{+}$
44	+	CD8+ interm TemRA	$FSC^{low}SSC^{low}CD3^{+}CD8^{+}CD197^{-}CD45RO^{-}CD27^{+}CD28^{-}$
45	+	CD8+ late TemRA	$FSC^{low}SSC^{low}CD3^{+}CD8^{+}CD197^{-}CD45RO^{-}CD27^{-}CD28^{-}$
46		**B cells**	$SSC^{low}CD45^{+}CD19^{+}$
47	+	Bnaive	$SSC^{low}CD19^{+}CD27^{-}IgD^{+}$
48		Bmem	$SSC^{low}CD19^{+}IgD^{-}$ and $CD27^{+}IgD^{+}$
49		CD27- Bmem	$SSC^{low}CD19^{+}CD27^{-}IgD^{-}$
50		CD27+ Bmem	$SSC^{low}CD19^{+}CD27^{+}IgD^{-}$
51		IgM+ Bmem	$SSC^{low}CD19^{+}CD27^{+}IgM^{+}$
52	+	Natural effector	$SSC^{low}CD19^{+}CD27^{+}IgM^{+}IgD^{+}$
53	+	IgMonly	$SSC^{low}CD19^{+}CD27^{+}IgM^{+}IgD^{-}$
54		IgA+ Bmem	$SSC^{low}CD19^{+}IgA^{+}$
55	+	CD27-IgA+	$SSC^{low}CD19^{+}CD27^{-}IgA^{+}$
56	+	CD27+IgA+	$SSC^{low}CD19^{+}CD27^{+}IgA^{+}$
57		IgG+ Bmem	$SSC^{low}CD19^{+}IgG^{+}$
58	+	CD27-IgG+	$SSC^{low}CD19^{+}CD27^{-}IgG^{+}$
59	+	CD27+IgG+	$SSC^{low}CD19^{+}CD27^{+}IgG^{+}$
60		CD21low B cells	$SSC^{low}CD19^{+}CD38^{low}CD21^{low}$
61		Igλ+ B cells	$SSC^{low}CD19^{+}Igκ^{-}Igλ^{+}$
62		Igκ+ B cells	$SSC^{low}CD19^{+}Igκ^{+}Igλ^{-}$

Determinants

Information on 22 dichotomized and 4 continuous determinants was obtained (Table II).[25,26] Six determinants were related to prenatal maternal lifestyle and evaluated by using questionnaires in the first, second, and third trimesters of pregnancy[26]: maternal age and body mass index before pregnancy, education, net household income, and smoking or alcohol use during pregnancy. Three determinants were related to prenatal maternal immunemediated diseases: data on maternal atopy were obtained by means of maternal reported questionnaires during pregnancy,[26] and serum anti–thyroid peroxidase IgG (anti-TPO) and anti–tissue transglutaminase IgA antibody levels were measured in the second trimester of pregnancy (mean 6 SD: 13.4 ± 2.0 and 20.6 ± 1.2 weeks, respectively).[28,29] Information on 6 birthrelated determinants (ie, sex, gestational age [preterm birth, <37 weeks], birth weight [low birth weight, <2500 g], premature ruptures of the membranes, mode of delivery, and birth season) was obtained from obstetric records assessed in midwife practices and hospitals.[25,26] Eleven bacterial or viral infection- related determinants were included: breast-feeding and breast-feeding duration, having siblings, antibiotics use, and presence of upper or lower respiratory tract infections in the first year of life (obtained by using questionnaires at child's age of 2, 3, 6, and 12 months),[26] Helicobacter pylori carriership at 6 years of age, and seropositivity (IgG) for cytomegalovirus (CMV), EBV, herpes simplex virus 1 (HSV-1), and varicella zoster virus at the age of 6 years.[22,30,31]

Statistical modeling

Linear mixed-effects analyses were performed to model leukocyte dynamics between birth and age 6 years. By including random effects in the model, this approach enabled modeling of cross-sectional data, with further improvement in accuracy by incorporating longitudinal follow-up data from individual children. To capture the trend in the data more precisely, the age of the children was included as a natural spline with different knots (0-3 knots) into the models. Basically, the number of knots is inversely related to the smoothness of the curve. Positions of the knots were defined as the 50th percentile (25.5 months) in the 1-knot model and the 33rd and 66th percentiles (14.1 and 70 months) in the 2-knot model. The knots in the 3-knot model were defined manually at 6, 14, and 24 months, focusing around the time points of data measurement. Model selection was performed by using the likelihood ratio test. Next, for each leukocyte population, the effect of all 26 determinants on the models was assessed by first adding each determinant univariably into the model and analyzing fixed-effects estimates. To correct for potential multiple testing errors, a correction for the 4 groups of determinants was performed, and consequently, only determinants with an effect P value of less than .0125 were defined as significant. Subsequently, for each leukocyte population, the determinants with a significant effect (<3 determinants were significant per model) were combined in a multivariable model to correct for possible confounding effects. Finally, to test whether changes in leukocyte dynamics over time (from birth until 6 years) induced by the determinant of interest was

most noticeable at a specific age interval, the individual determinants were tested in relation to the age of the children in a multivariable model. The P value for significance for the multivariable analyses was less than .05. Statistical analyses were performed with R software (version R-3.2.1; details are provided in Supplemental Material and Methods.[32]

Clustering analyses

The modeled data of each population were normalized into zero means and unit SDs (z-score) to cluster patterns of leukocyte subset kinetics by using the following calculation:

$$\text{z-score} = \frac{\textit{population size (at month x)} - \textit{average population size (at months 0–76)}}{\textit{standard deviation of population size (at months 0–76)}}$$

The normalized leukocyte models were subsequently clustered using agglomerative (bottom-up) Ward hierarchical clustering at each step, clustering two clusters with minimum between-cluster distance by using the Euclidean distance measure.

RESULTS

Linear mixed effect modeling of leukocyte–subset cell numbers versus child's age

To study the dynamics of blood immune cells in young children, we quantified cell numbers of 11 leukocyte subsets in 1,182 children between birth and 76 months (6y) of age. Statistical modeling showed that the numbers of four innate leukocyte subsets, i.e. monocytes, neutrophils, eosinophils and NK cells were high at birth, quickly declined within the first 6 months of age and subsequently remained stable (Figure 1A). Naive B-cell and T-cell numbers strongly increased after birth, peaked between 6 and 14 months of age and subsequently decreased to stable levels around the age of 2 to 6 years (Figure 1B). Memory B and memory T cell numbers slowly increased within the first 6 to 14 months of life, after which numbers decreased marginally and stabilized from approximately 3 years onward (Figure 1B). Total TCRγδ^{+} T-cell numbers increased until 6 months and remained quite stable at these levels (Figure 1C).

Distinct dynamics of innate leukocyte, and naive and memory B and T cells

To evaluate whether functionally related immune populations would follow similar dynamics with increasing age, we extended our analysis in the same children to 62 leukocyte subsetsbased on recent insights into naive and memory B-cell and T-cell subsets (Table 1 and Supplemental Figures 1-5; reference values per age category are presented in Supplemental Table 2).[24, 33]

Hierarchical clustering of all 62 leukocyte populations resulted in 4 distinct clusters of leukocyte dynamics (Figure 2A). Cell numbers of all populations in the first cluster were high at birth, followed by a sharp decrease within the first 6 months of life, after which

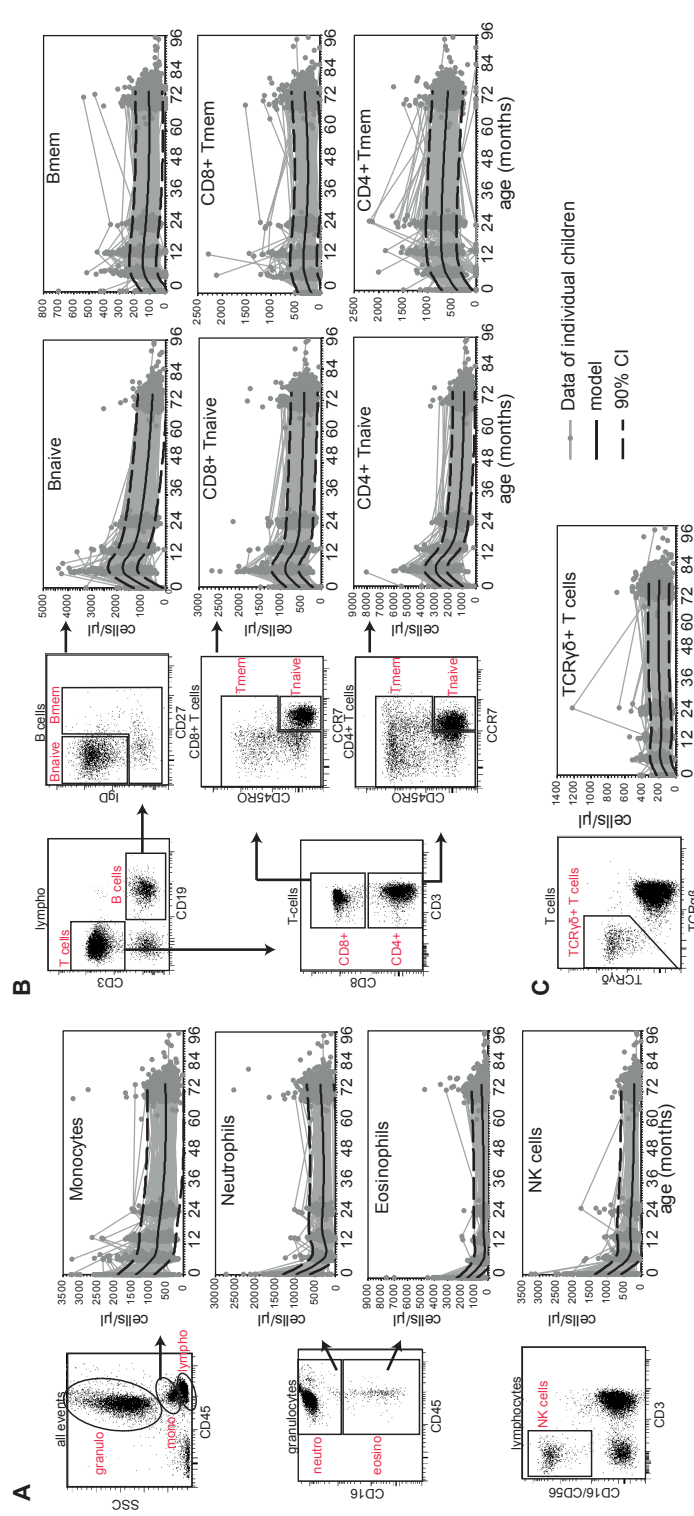


Figure 1. Dynamics of innate leukocyte and naive and memory lymphocyte populations in children between birth and 6 years of age.
Dynamics of **A)** monocytes, neutrophils, eosinophils and NK cells, **B)** naive and memory B-cell and T-cell subsets, and **C)** TCRγδ$^+$ T cells. Flow cytometry plots depict population definitions in one representative 6-year-old individual. Graphs depict absolute numbers of cells in blood of 1,182 children with in total 2,010 measurements (gray dots). Linear mixed effect models are depicted by a solid black line for each population and the 90% confidence interval (CI) of the model with dashed black lines. For clarity, only direct follow-up time points of an individual were connected with gray lines; i.e. 0-6m, 6-14m, 14-25m or 25-76m.

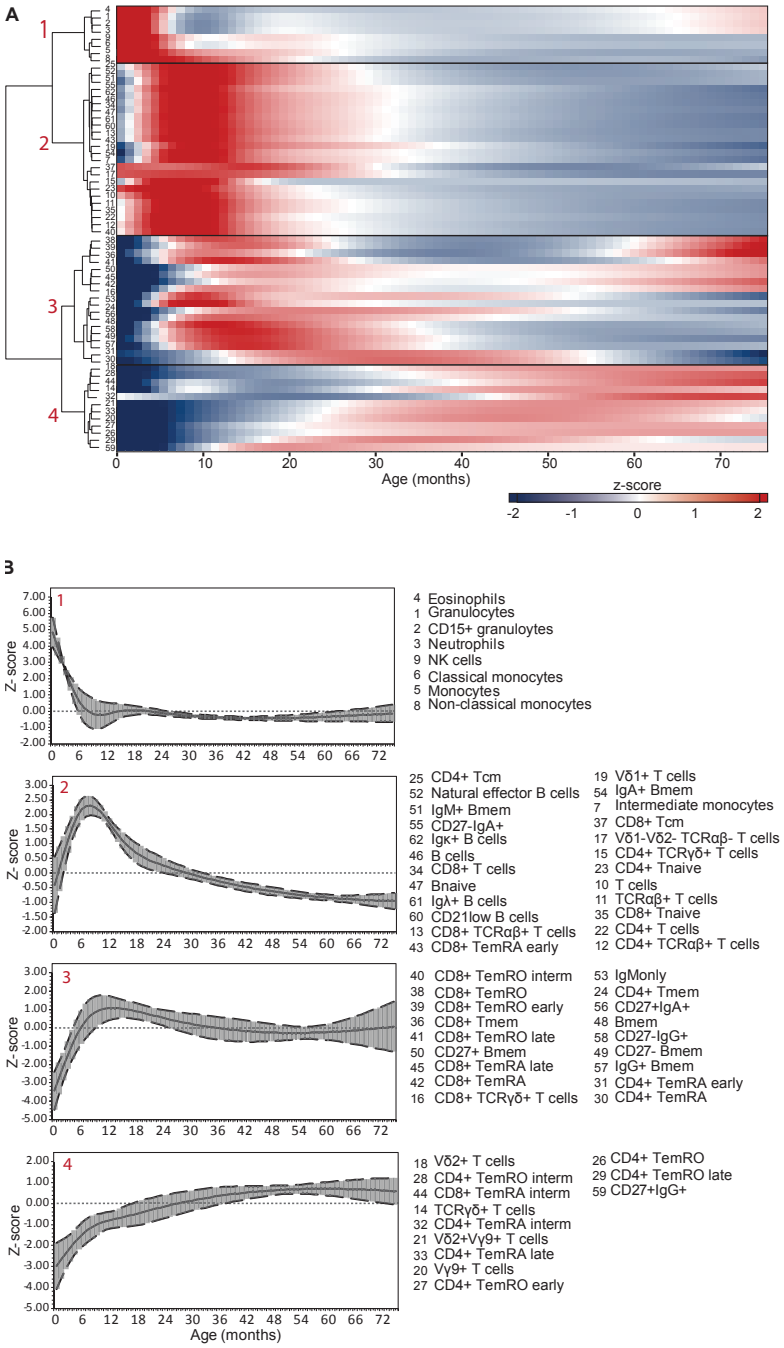


Figure 2. Hierarchical clustering of leukocyte subset dynamics in early childhood.

A) Ward's hierarchical clustering of the cellularity of the 62 populations derived from the linear mixed models. Cellularity between 0-76 months of age was converted to zero mean and unit standard deviation (z score). **B)** Average patterns +/- 1 standard deviation of the subsets in each of the 4 major clusters. Populations are numbered according to their order in Table 1.

they stabilized (Figure 2B). This cluster exclusively contained innate leukocyte subsets. All defined innate leukocyte subsets clustered within cluster 1, except for the intermediate monocytes.

The populations in cluster 2 strongly increased shortly after birth and peaked before 14 months of age, followed by a long-term gradual decrease. Cluster 2 included the 3 major naive lymphocyte subsets ($CD4^+$ Tnaive, $CD8^+$ Tnaive and Bnaive), as well as 5 memory populations ($CD4^+$ central memory T cells (Tcm), $CD8^+$ Tcm cells, early differentiated $CD8^+$ $CD45RO^-$ effector memory T cells ($CD8^+$ early TemRA), $CD27^-IgA^+$ memory B cells and natural effector memory B cells). Also the total B-cell, $CD8^+$ T-cells, $CD4^+$ T-cells and total T-cell subsets clustered in cluster 2, as did the intermediate monocytes and the $V\delta1^+$ T cells.

The populations in clusters 3 and 4 gradually increased in cell number and peaked either at 14 months (cluster 3) or after approximately 4 years (cluster 4). Clusters 3 and 4 contained only memory B-cell and T-cell subsets and the $TCR\gamma\delta^+$ T-cell subsets. Cluster 3 contained the total memory lymphocyte populations Bmem, $CD4^+$ Tmem and $CD8^+$ Tmem. In particular, these memory populations included the early, intermediate and late $CD8^+$ TemRO subpopulations, late $CD8^+$ TemRA cells, early CD4+ TemRA cells, and the IgMonly, $CD27^+IgA^+$ and $CD27^-IgG^+$ memory B-cell subsets. Cluster 4 contained total $TCR\gamma\delta^+$ T cells and the $V\delta2^+$ and the $V\gamma9^+$ T-cell subsets, as well as the early, intermediate and late $CD4^+$ TemRO subpopulations, late $CD4^+$ TemRA cells, intermediate $CD8^+$ TemRA cells, and $CD27^+IgG^+$ memory B cells.

An additional clustering was performed on a selection of 31 non-overlapping populations (Table 1) to test the effect of including populations with overlapping population definitions on clustering. The resulting clusters showed similar dynamics (Figure 2B and Supplemental Figure 6B) with only 3 populations (non-classical monocytes, $CD8^+$ late TemRO and $CD8^+$ late TemRA) assigned to a different cluster, indicating that the overlapping populations did not overtly skew our clustering approach. Thus, functionally related populations, as well as populations with phenotypic overlapping definitions, displayed similar dynamics in early childhood.

Associations between nongenetic factors and leukocyte subset dynamics in the 4 clusters

The memory B- and T-cell subset showed large interindividual variation in dynamics of cell numbers (Figure 1). To study the effects of external factors, we analyzed 6 maternal life style- and 3 maternal immune-related determinants, 6 birth characteristics, and 11 bacterial/viral exposure-related determinants (Table 2). Twelve of these 26 determinants showed a significant association with the dynamics of one or more of the 62 leukocyte populations after multivariable correction (Table 2). Bacterial/viral exposure-related determinants were more frequently found to affect cell numbers in clusters 2, 3 and 4 than in cluster 1 (Figure 3A-B).

Table 2. Characteristics of nongenetic factors

	negative individuals (%)	positive individuals (%)	missing (%)	significant effect on cluster [a]
Prenatal maternal life style				
Maternal age (years)	(continuous determinant)		0 (0)	-
Low maternal educational level	400 (33.8)	760 (64.3)	22 (1.9)	1,2,3,4
Net household income per month >€ 2200	242 (20.5)	847 (71.7)	93 (7.9)	-
Smoking during pregnancy	815 (69)	256 (21.7)	111 (9.4)	-
Alcohol use continued during pregnancy	299 (25.3)	750 (63.5)	133 (11.3)	-
Body Mass Index before pregnancy (kg/m$^{2)}$	(continuous determinant)		181 (15.3)	-
Prenatal maternal immune-mediated diseases				
Maternal atopy (eczema, allergy HDM, hay-fever	683 (57.8)	383 (32.4)	116 (9.8)	-
Anti-TPO (before 18 weeks of pregnancy) (mU/ml)	(continuous determinant)		278 (23.5)	2
Maternal TG2A during pregnancy U/ml	(continuous determinant)		208 (17.6)	-
Birth characteristics				
Gender (girl yes/no)	600 (50.8)	582 (49.2)	0 (0)	1,2,3,4
Preterm birth <37 weeks	76 (6.4)	1106 (93.6)	0 (0)	-
Low birth weight <2500g	1119 (94.7)	63 (5.3)	0 (0)	-
Premature rupture of membranes	1089 (92.1)	53 (4.5)	40 (3.4)	4
Caesarian section versus vaginal/ forceps/vacuum assisted	903 (76.4)	154 (13)	125 (10.6)	3,4
Birth season (born in Fall/Winter)	682 (57.7)	500 (42.3)	0 (0)	-
Bacterial/viral exposure-related characteristics				
Breastfeeding ever	109 (9.2)	969 (82)	104 (8.8)	-
Breastfeeding at 6 months of age	754 (63.8)	299 (25.3)	129 (10.9)	2,4
Siblings >1	687 (58.1)	480 (40.6)	15 (1.3)	1
Antibiotics/Penicillin use in 1st y	636 (53.8)	342 (28.9)	204 (17.3)	1,3
Upper respiratory tract infections in 1st y	547 (46.3)	475 (40.2)	160 (13.5)	-
Lower respiratory tract infections in 1st y (doctor attended)	819 (69.3)	131 (11.1)	232 (19.6)	-
Carrier of Helicobacter pylori within 6yrs	887 (75)	49 (4.1)	246 (20.8)	2,3
Seropositivity for Cytomegalovirus (CMV) at 6y	665 (56.3)	269 (22.8)	248 (21)	1,2,3,4
Seropositivity for Epstein Barr virus (EBV) at 6y	534 (45.2)	400 (33.8)	248 (21)	2,3,4
Seropositivity for Herpes simplex virus-1 (HSV-1) at 6y	810 (68.5)	124 (10.5)	248 (21)	2
Seropositivity for Varicella zoster virus (VZV) at 6y	73 (6.2)	861 (72.8)	248 (21)	-

[a] First, each determinant was added to the linear mixed effect model univariately to test for a significant effect (Significance was defined as p<0.0125); Subsequently, per leukocyte population, all determinants with a significant individual effect were added in a multivariable model to correct for confounding effects. Up to 3 determinants were added per multivariable model. Significance was defined as p<0.05. IU, international units. No confounding effect was observed

In cluster 1, a low maternal education level was associated with a reduction in the patterns of eosinophils and classical monocytes (Figure 4). Female sex was associated with a significant increase in the pattern of neutrophils and the phenotypically overlapping CD15^{+}

granulocyte population. Having more than 1 sibling was associated with a reduction in NK cell numbers, and CMV infection with an increase in NK cell numbers. Antibiotic use in the first year of life was associated with an increase in non-classical monocytes. Overall effect estimates and associations of effects with specific age-intervals are presented in Supplemental Table 3.

In cluster 2, none of the determinants included in our study affected the patterns of more than 25% of the populations (Figure 3C-D). However, female sex was associated with an increase in the patterns of $CD4^+$ naive and Tcm cells and, consequently, in the phenotypically overlapping $CD4^+TCR\alpha\beta^+$ and total $CD4^+$ T cells. Although breastfeeding for more than 6 months was associated with a reduction in the pattern of total IgA^+ and $CD27^-IgA^+$ memory B cells, H. pylori carriership was associated with an increase in these populations. Furthermore, H. pylori carriership was associated with an increase in total $Ig\kappa^+$ B cells and $TCR\alpha\beta^+$ T cell counts. CMV and EBV were both associated with an increase in the strongly related total and $TCR\alpha\beta^+$ $CD8^+$ T cell numbers, and CMV was associated with an increase in $V\delta1^+$ T cell numbers. HSV-1 seropositivity was associated with a decrease in the large population of naive B cells and, consequently, with total B cells and $Ig\kappa^+$ and $Ig\lambda^+$ B-cell subsets.

In cluster 3 CMV and EBV seropositivity were significantly associated with changes in dynamics of greater than 40% of subsets (Figure 3C; Suppl Table 3). Both viruses were associated with an increase in $CD8^+$ intermediate and late TemRO cell numbers, and the phenotypically-related total $CD8^+$ TemRO and total $CD8^+$ Tmem populations. In addition, CMV was associated with an increase in $CD8^+TCR\gamma\delta^+$ T cells and $CD8^+$ late TemRA cells and the phenotypically-related total $CD8^+$ TemRA population. EBV was associated with an increase in $CD8^+$ early TemRO cells and a decrease in numbers of total Bmem cells and the $CD27^-$ and $CD27^+$ Bmem subsets. The effects on $CD8^+$ late TemRO and TemRA cells were still present in the analysis of 31 non-overlapping population, even though these were now included in cluster 4 (Figure 3D).

In cluster 4 sex, premature rupture of membranes, CMV and EBV were significantly associated with changes in dynamics of 40% of subsets (Figure 3C and D; Suppl Table 3). Female sex was associated with an increase in the patterns of $CD27^+IgG^+$ memory B cells, and $CD4^+$ early TemRO cells and with a reduction in total $TCR\gamma\delta^+$ T cells and $CD4^+$ intermediate and late TemRA populations. Premature rupture of membranes was associated with an increase in numbers of $CD4^+$ late TemRO cells, $CD4^+$ intermediate TemRA cells and $V\delta2^+$ and $V\gamma9^+$ T-cell subsets. CMV and EBV seropositivity were both associated with an increase in numbers of total $CD4^+$ TemRO and $CD4^+$ late TemRO and TemRA cells, as well as $CD8^+$ intermediate TemRA cells. In addition, CMV was associated with increased numbers of $CD4^+$ intermediate TemRO cells, and EBV with an increase in $CD4^+$ early TemRO cell numbers.

Altogether, we identified multiple determinants that had a significant effect on the dynamics of leukocyte subset numbers within 1 or more of the 4 distinct age-related patterns. Still, considerable variation of effects could be observed within individual clusters.

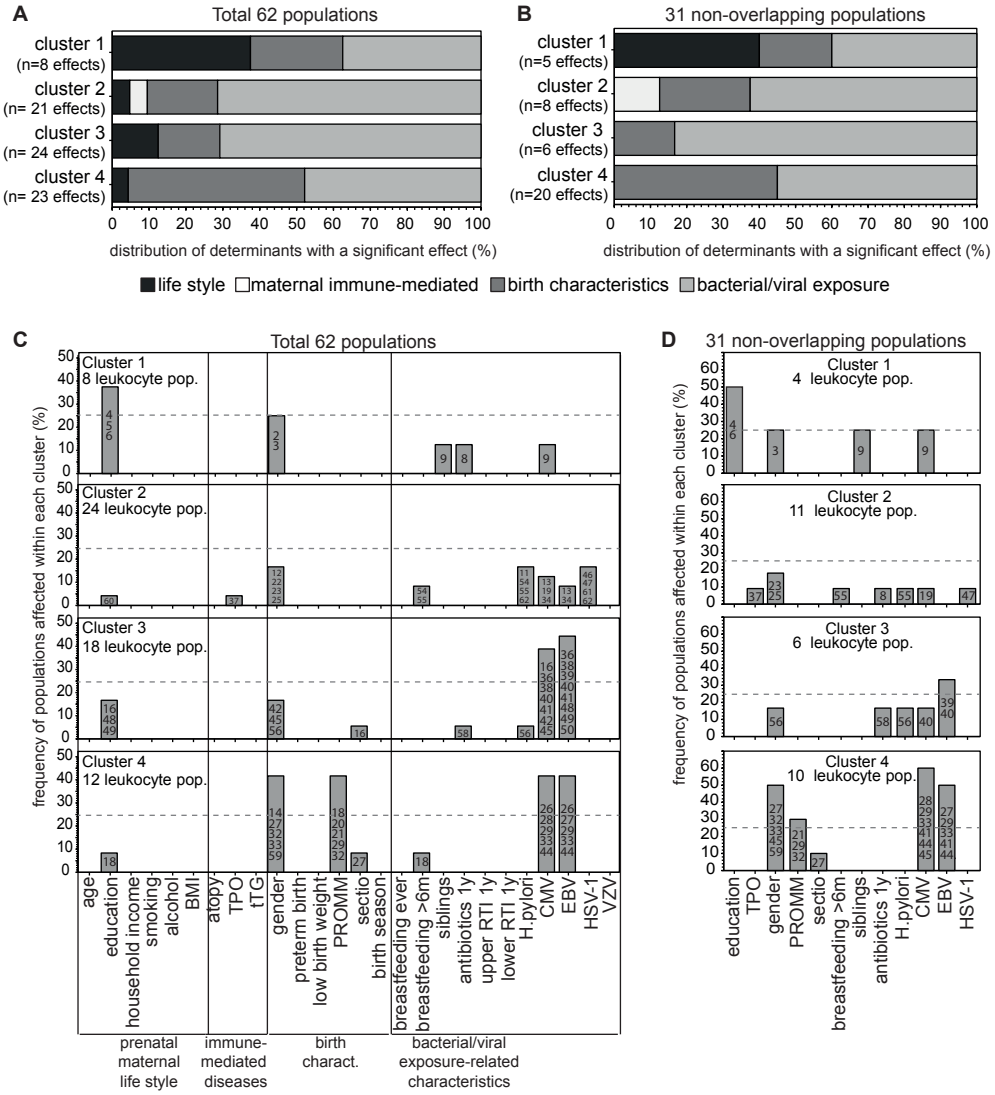


Figure 3. Nongenetic factors associated with leukocyte dynamics in the four clusters.

Significant associations of the 26 determinants with leukocyte subsets in clusters 1-4 based on the clustering of 62 leukocyte populations as in Figure 2 (**A** and **C**) or of the 31 non-overlapping subsets as in Suppl Figure 6 (**B** and **D**) corrected for confounding effects. **A and B)** Relative distribution of types of determinants that significantly affect leukocyte kinetics. **C and D)** Frequencies of populations within a cluster that were affected by individual determinants. Abbreviations: PROMM, premature rupture of membranes; RTI, respiratory tract infection. Numbers in each bar refer to the populations as defined in Table 1.

Effects of nongenetic factors on leukocyte dynamics within distinct leukocyte lineages

Because clusters 2, 3 and 4 each contained various B-cell and T-cell populations, each with distinct humoral or cellular functions, we next studied the effects of the 26 determinants

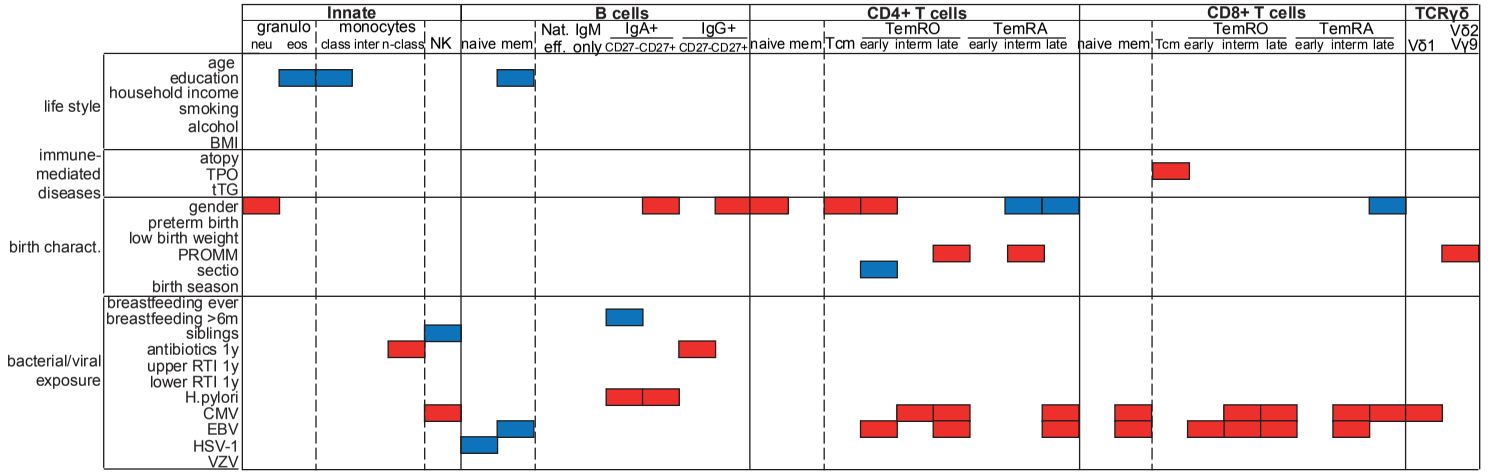


Figure 4. Effects of nongenetic factors on the dynamics of innate leukocytes, and B-cell and T-cell subsets.

Significant effects are shown derived from multivariable linear mixed effect analysis of all 26 determinants on the 31 non-overlapping populations, and the total memory subsets ("mem") within the B-cell, $CD4^+$ T-cell and $CD8^+$ T-cell lineages. Blue boxes represent a reduction, red boxes an increase in the pattern of the indicated population. Effect sizes are presented in Supplemental Table 3, and details of each determinant in Table 2. Additional abbreviations: PROMM, premature rupture of membranes; RTI, respiratory tract infection; neu, neutrophils; eos, eosinophils; class, classical monocytes; inter, intermediate monocytes; n-class, non-classical monocytes.

on the patterns of individual leukocyte lineages.

Within the B-cell lineage, low maternal educational level was associated with a reduction in the patterns of total memory B cells (Figure 4). Female sex was associated with an increase in the patterns of $CD27^+IgA^+$ and $CD27^+IgG^+$ memory B cells, although these effects were not reflected in the total memory B-cell populations. Breastfeeding for more than 6 months was associated with a selective reduction in the pattern of $CD27^-IgA^+$ memory B cells, antibiotics use in the first year of life was associated with a selective increase in the pattern of $CD27^-IgG^+$ memory B cells, and H. pylori carriership was associated with an increase in the patterns of both $CD27^-IgA^+$ and $CD27^+IgA^+$ memory B cells. EBV infection was associated with a reduction in the pattern of total memory B cells, although not with a specific memory B-cell subset, and HSV-1 infection was associated with a selective reduction in the pattern of naive B cells.

Within the $CD4^+$ and $CD8^+$ T-cell lineages, increased serum anti-TPO levels were associated with an increase in $CD8^+$ Tcm cells. Female sex was associated with an increase in the patterns of $CD4^+$ naive, Tcm and early TemRO cells, although these effects were not reflected in the total $CD4^+$ memory T-cell populations. In contrast, female sex was associated with a decrease in the patterns of $CD4^+$ intermediate and late TemRA and $CD8^+$ late TemRA cells. Whereas premature rupture of membranes was associated with an increase in $CD4^+$ late TemRO and $CD4^+$ intermediate TemRA cells, child birth through caesarian section was associated with a reduction in the pattern of $CD4^+$ early TemRO cells. Persistent viral infection with CMV and EBV significantly affected $CD4^+$ and $CD8^+$ T-cell dynamics in contrast to breastfeeding, antibiotics usage, H. pylori, HSV-1 and VZV carriership. CMV infection associated exclusively with an increase in the patterns of intermediate and late differentiated $CD4^+$ and $CD8^+$ TemRO and TemRA cells. Associations with EBV were slightly more variable, also including early differentiated $CD4^+$ and $CD8^+$ TemRO cells.

Within the $TCR\gamma\delta^+$ T-cell lineage, premature rupture of membranes was associated with an increase in $V\delta2^+V\gamma9^+$ T cell counts and CMV infection associated with an increase in the pattern of $V\delta1^+$ T cells.

DISCUSSION

Here we modeled the kinetics of 62 leukocyte subsets and identified distinct patterns of cell numbers in the first 6 years of life for innate leukocytes, naive B and T cells, and memory B and T cells. Unsupervised clustering revealed that leukocyte dynamics between birth and age 6 years could be summarized into 4 major profiles, with either (1) early predominance and fast decrease to stable numbers, (2) gradual increased in first year followed by a gradual decline, or (3 and 4) a slow increase in the first 1 to 2 years, followed by stabilization.

Consistent with their early predominance, innate leukocyte kinetics were affected by maternal education level, which might already influence the fetus prenatally. Low mater-

nal education level is strongly related to a less healthy life.[34] These observations suggest that maternal life style is especially important for shaping of innate leukocyte populations, although the exact mediator remains to be determined.

The non-classical and intermediate monocytes did not consistently cluster with the innate leukocyte subsets because of a later peak in numbers or a slower decrease. This altered kinetics could be the result of these being derived from classical monocytes, and not directly from precursors in bone marrow.[35] Still, the various subsets of monocytes express distinct levels of proteins involved in HLA-class II-dependent antigen presentation and CD40-CD40L co-stimulation,[35] and differ in parasite pattern recognition.[35-38] Thus our data support the need to discriminate intermediate and non-classical monocytes from the dominant population of classical monocytes in kinetics studies.

All naive B- and T-cell populations followed profile 2, with a gradual increase in the first 14 months followed by a more gradual decrease. After the first 1 to 2 years of life, B-lymphocyte production in bone marrow decreases,[39, 40] and the thymus starts to involute.[41] These processes are likely causes of the decrease we observed.

Memory B-cell populations showed different patterns. Natural effector and $CD27^-IgA^+$ memory B-cell kinetics were similar to naive B cells. Both memory B-cell populations are suggested to derive, at least in part, from germinal center-independent responses without T-cell help in the splenic marginal zone and the intestinal lamina propria.[33, 42, 43] The relatively fast increase in these cell numbers within the first 6 months of life contrasts with the kinetics of the other T-cell dependent memory B-cell subsets. IgM-only, $CD27^-IgG^+$ and $CD27^+IgA^+$ memory B cells followed profile 3, representing gradual memory formation. The $CD27^+IgG^+$ population followed profile 4 and developed slightly slower than the other memory B-cell populations. These observations supports the concept of rapid generation of T-cell independent memory B cells in the absence of the extensive proliferation and selection processes of the germinal center,[33] whereas $CD27^+IgG^+$ memory B cells might represent the most mature memory B-cell population. In adults, this subset shows a higher level of affinity maturation, a more extensive replication history, and more frequent development through consecutive class-switching than $CD27^-IgG^+$ memory B cells.[33, 44] It is perceivable that many of these $CD27^+IgG^+$ cells are generated from $CD27^-IgG^+$ memory B cells in secondary responses, which could explain the gradual decrease in $CD27^-IgG^+$ B cell numbers after 14 months of age. Together, the different clustering of the memory B-cell populations thereby follows their increasing functional maturity (Supplemental Figure 5 I).

Several external determinants significantly affected the B-cell kinetics. First, we confirmed our previous observations regarding lower numbers of memory B cells in breastfed children.[20, 22] However, in the current analysis we could not reproduce the previously reported association between breastfeeding duration and germinal center-dependent memory B-cell numbers at 6 months of age.[22] The difference is likely due to our current analysis in which duration of breastfeeding was not included and the overall pattern between birth and 6 years of age was studied.

Numbers of IgA^+ memory B cells were significantly increased in H. pylori positive children. The colonization of the gastric mucus by H. pylori was found to correlate with an increase in total blood B-cell counts,[45] on top of the local expansion of antibody-secreting IgA^+ cells,[46] which are important for the protection against H. pylori infection.[47] Our observed expansion of circulating IgA^+ memory B cells included both the $CD27^+IgA^+$ and mucosa-derived $CD27^-IgA^+$ subsets. This would suggest that the presence of H. pylori does not only result in a local expansion of plasma cells, but also a systemic expansion of memory B cells in otherwise asymptomatic carriers. It remains to be determined whether this expansion is beneficial for the host as the bacterial protein CagA inhibits B-cell apoptosis and thereby increases the risk for mucosa-associated B-cell malignancies.[48, 49]

Memory T-cell subsets were found in clusters 2, 3 and 4. $CD4^+$ and $CD8^+$ Tcm (cluster 2) express lymph node homing markers and are the presumed precursors of effector memory T cells.[24, 50, 51] The early peak in Tcm numbers prior to those of effector memory T cells in young children would fit with this function. Similarly, $CD8^+$ early TemRA cells might be precursors for further differentiated TemRA subsets. $V\delta1^+$ T-cell numbers (cluster 2) peaked prior to the $V\delta2^+$ and $V\gamma9^+$ subsets (cluster 4), thereby confirming the previously observed early shift from $V\delta1^+$ to $V\delta2^+V\gamma9^+$ predominance in children.[52, 53]

In line with previous observations in both children and elderly,[22, 54-58] both $CD4^+$ and $CD8^+$ memory T-cell dynamics were predominantly affected by CMV and/or EBV seropositivity. This concerned mostly late effector memory T-cell numbers for CMV and early effector memory T-cell subsets for EBV.[24, 59] HSV-1 seropositivity did not affect $CD4^+$ and $CD8^+$ T-cell populations,[22, 58] but was associated with a loss of naive B cells. This association has not been described before. EBV infection was associated with a reduction in memory B cells, likely because of the selective EBV persistence in these cells.[60, 61] Our large-scale analysis allowed us to separate the herpesvirus-associated effects, with CMV and EBV being associated with memory T-cell expansions, and HSV-1 and EBV being associated with a decrease in naive or memory B-cell numbers, respectively.

Sex had a widespread effect on 14 subsets within multiple leukocyte lineages. Interestingly, girls showed a skewing of humoral and early differentiated $CD4^+$ T-cell responses over cellular cytotoxic responses in contrast to boys. These effects might be associated with differences in sex-hormone levels (testosterone, estradiol) which are already detectable during early infancy, as well as with genetic differences between females and males.[62] These insights can especially be important for dissection of auto-immune diseases, which are much more prevalent in females than in males, although predisposition for allergic diseases seems to be opposite in infancy.[62, 63]

Methodological considerations

The strength of this study is its prospective longitudinal population-based design with more than 1000 children and the possibility to study 26 external determinants in 4 subgroups with adjustments for major confounders. The linear mixed model approach enabled mod-

eling of cross-sectional data, with further improvement of the accuracy by incorporating additionally available longitudinal follow-up data. Furthermore, we included only children with a two-generation Dutch ancestry, which prevented interference of our analyses by strong ethnic and cultural influences. However, extrapolation of our findings to different ethnic and cultural populations might be limited, and would require additional analysis of ethnically-different population cohorts.

The inclusion of 62 leukocyte populations allowed for the large-scale analysis of the effect of external determinants on both total cell lineages and small subsets defined by extensive and detailed phenotypic definition. The overlap in some populations could have skewed the hierarchical clustering. However, our selection of 31 phenotypically non-overlapping populations resulted in clusters with similar patterns, indicating the robustness of the 4 major patterns of leukocyte dynamics and the observed effects of external determinants.

Our study was primarily explorative with a focus on the identification of determinants that affected leukocyte dynamics between birth and 6 years of age. Although we defined whether determinants had a positive or negative association with leukocyte numbers, we were unable to identify the exact nature of the effect (i.e. exactly when these effects presented and whether these effects will be transient, persistent or potentially even increasing over time). Consequently, more research into individual determinants will be needed to extend our observations by specifying these effects.

Finally, special consideration needs to be taken for serology of infectious agents that were measured at the age of 6 years. We cannot determine the exact timing of primary infections, and this can consequently be in the whole period preceding the age of 6 years. Still, most H. pylori infections already occur in early childhood, and IgG seropositivity to herpesviruses only appears several weeks to 3 months after infection. Thus, these determinants were present already during or before the sixth year of life.

Conclusions

With our unbiased approach, for the first time, we classified leukocyte populations according to their dynamics between birth and 6 years of age. Moreover, we identified nongenetic factors that are associated with the dynamics of cell lineages or specific leukocyte subsets and that underlie, at least in part, human immunological diversity. These newly identified determinants can provide new targets for studies on the molecular processes that regulate leukocyte development and immune responses and that together underlie formation of long-lasting immunity without inducing destructive, excessive or insufficient immune responses.

REFERENCES

1. Comans-Bitter WM, de Groot R, van den Beemd R, Neijens HJ, Hop WC, Groeneveld K, et al. Immunophenotyping of blood lymphocytes in childhood. Reference values for lymphocyte subpopulations. J Pediatr 1997; 130:388-93.
2. van Gent R, van Tilburg CM, Nibbelke EE, Otto SA, Gaiser JF, Janssens-Korpela PL, et al. Refined characterization and reference values of the pediatric T- and B-cell compartments. Clin Immunol 2009; 133:95-107.
3. Tosato F, Bucciol G, Pantano G, Putti MC, Sanzari MC, Basso G, et al. Lymphocytes subsets reference values in childhood. Cytometry A 2015; 87:81-5.
4. Duchamp M, Sterlin D, Diabate A, Uring-Lambert B, Guerin-El Khourouj V, Le Mauff B, et al. B-cell subpopulations in children: National reference values. Immun Inflamm Dis 2014; 2:131-40.
5. Xanthou M. Leucocyte blood picture in healthy full-term and premature babies during neonatal period. Arch Dis Child 1970; 45:242-9.
6. Manroe BL, Weinberg AG, Rosenfeld CR, Browne R. The neonatal blood count in health and disease. I. Reference values for neutrophilic cells. J Pediatr 1979; 95:89-98.
7. Sagnia B, Ateba Ndongo F, Ndiang Moyo Tetang S, Ndongo Torimiro J, Cairo C, Domkam I, et al. Reference values of lymphocyte subsets in healthy, HIV-negative children in Cameroon. Clin Vaccine Immunol 2011; 18:790-5.
8. Huck K, Feyen O, Ghosh S, Beltz K, Bellert S, Niehues T. Memory B-cells in healthy and antibody-deficient children. Clin Immunol 2009; 131:50-9.
9. Morbach H, Eichhorn EM, Liese JG, Girschick HJ. Reference values for B cell subpopulations from infancy to adulthood. Clin Exp Immunol 2010; 162:271-9.
10. Piatosa B, Wolska-Kusnierz B, Pac M, Siewiera K, Galkowska E, Bernatowska E. B cell subsets in healthy children: reference values for evaluation of B cell maturation process in peripheral blood. Cytometry B Clin Cytom 2010; 78:372-81.
11. Driessen GJ, Dalm VA, van Hagen PM, Grashoff HA, Hartwig NG, van Rossum AM, et al. Common variable immunodeficiency and idiopathic primary hypogammaglobulinemia: two different conditions within the same disease spectrum. Haematologica 2013; 98:1617-23.
12. Brodin P, Jojic V, Gao T, Bhattacharya S, Angel CJ, Furman D, et al. Variation in the human immune system is largely driven by non-heritable influences. Cell 2015; 160:37-47.
13. Orru V, Steri M, Sole G, Sidore C, Virdis F, Dei M, et al. Genetic variants regulating immune cell levels in health and disease. Cell 2013; 155:242-56.
14. Carr EJ, Dooley J, Garcia-Perez JE, Lagou V, Lee JC, Wouters C, et al. The cellular composition of the human immune system is shaped by age and cohabitation. Nat Immunol 2016; 17:461-8.
15. Pabst HF, Spady DW, Pilarski LM, Carson MM, Beeler JA, Krezolek MP. Differential modulation of the immune response by breast- or formula-feeding of infants. Acta Paediatr 1997; 86:1291-7.
16. Winkler B, Aulenbach J, Meyer T, Wiegering A, Eyrich M, Schlegel PG, et al. Formula-feeding is associated with shift towards Th1 cytokines. Eur J Nutr 2015; 54:129-38.
17. Hanson L, Silfverdal SA, Stromback L, Erling V, Zaman S, Olcen P, et al. The immunological role of breast feeding. Pediatr Allergy Immunol 2001; 12 Suppl 14:15-9.
18. Hanson LA, Korotkova M, Lundin S, Haversen L, Silfverdal SA, Mattsby-Baltzer I, et al. The transfer of immunity from mother to child. Ann N Y Acad Sci 2003; 987:199-206.
19. Duijts L, Ramadhani MK, Moll HA. Breastfeeding protects against infectious diseases during infancy in industrialized countries. A systematic review. Matern Child Nutr 2009; 5:199-210.
20. Hawkes JS, Neumann MA, Gibson RA. The effect of breast feeding on lymphocyte subpopulations in healthy term infants at 6 months of age. Pediatr Res 1999; 45:648-51.

21. M'Rabet L, Vos AP, Boehm G, Garssen J. Breast-feeding and its role in early development of the immune system in infants: consequences for health later in life. J Nutr 2008; 138:1782S-90S.
22. Jansen MA, van den Heuvel D, van Zelm MC, Jaddoe VW, Hofman A, de Jongste JC, et al. Decreased memory B cells and increased CD8 memory T cells in blood of breastfed children: the generation R study. PLoS One 2015; 10:e0126019.
23. Bekker V, Scherpbier H, Pajkrt D, Jurriaans S, Zaaijer H, Kuijpers TW. Persistent humoral immune defect in highly active antiretroviral therapy-treated children with HIV-1 infection: loss of specific antibodies against attenuated vaccine strains and natural viral infection. Pediatrics 2006; 118:e315-22.
24. Appay V, van Lier RA, Sallusto F, Roederer M. Phenotype and function of human T lymphocyte subsets: consensus and issues. Cytometry A 2008; 73:975-83.
25. Jaddoe VW, van Duijn CM, Franco OH, van der Heijden AJ, van Iizendoorn MH, de Jongste JC, et al. The Generation R Study: design and cohort update 2012. Eur J Epidemiol 2012; 27:739-56.
26. Kruithof CJ, Kooijman MN, van Duijn CM, Franco OH, de Jongste JC, Klaver CC, et al. The Generation R Study: Biobank update 2015. Eur J Epidemiol 2014; 29:911-27.
27. Kalina T, Flores-Montero J, van der Velden VH, Martin-Ayuso M, Bottcher S, Ritgen M, et al. EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. Leukemia 2012; 26:1986-2010.
28. Korevaar TI, Medici M, de Rijke YB, Visser W, de Muinck Keizer-Schrama SM, Jaddoe VW, et al. Ethnic differences in maternal thyroid parameters during pregnancy: the Generation R study. J Clin Endocrinol Metab 2013; 98:3678-86.
29. Kiefte-de Jong JC, Jaddoe VW, Uitterlinden AG, Steegers EA, Willemsen SP, Hofman A, et al. Levels of antibodies against tissue transglutaminase during pregnancy are associated with reduced fetal weight and birth weight. Gastroenterology 2013; 144:726-35 e2.
30. Jansen MA, Tromp, II, Kiefte-de Jong JC, Jaddoe VW, Hofman A, Escher JC, et al. Infant feeding and anti-tissue transglutaminase antibody concentrations in the Generation R Study. Am J Clin Nutr 2014; 100:1095-101.
31. den Hollander WJ, Holster IL, van Gilst B, van Vuuren AJ, Jaddoe VW, Hofman A, et al. Intergenerational reduction in Helicobacter pylori prevalence is similar between different ethnic groups living in a Western city. Gut 2014.
32. R Core Team; R Foundation for Statistical Computing V, Austria. R: A language and environment for statistical computing. . 2015.
33. Berkowska MA, Driessen GJ, Bikos V, Grosserichter-Wagener C, Stamatopoulos K, Cerutti A, et al. Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways. Blood 2011; 118:2150-8.
34. Bouthoorn SH, Silva LM, Murray SE, Steegers EA, Jaddoe VW, Moll H, et al. Low-educated women have an increased risk of gestational diabetes mellitus: the Generation R Study. Acta Diabetol 2015; 52:445-52.
35. Wong KL, Tai JJ, Wong WC, Han H, Sem X, Yeap WH, et al. Gene expression profiling reveals the defining features of the classical, intermediate, and nonclassical human monocyte subsets. Blood 2011; 118:e16-31.
36. Turner JD, Bourke CD, Meurs L, Mbow M, Dieye TN, Mboup S, et al. Circulating CD14brightCD16+ 'intermediate' monocytes exhibit enhanced parasite pattern recognition in human helminth infection. PLoS Negl Trop Dis 2014; 8:e2817.
37. Hijdra D, Vorselaars AD, Grutters JC, Claessen AM, Rijkers GT. Phenotypic characterization of human intermediate monocytes. Front Immunol 2013; 4:339.
38. Ziegler-Heitbrock L, Hofer TP. Toward a refined definition of monocyte subsets. Front Immunol 2013; 4:23.
39. Kogut I, Scholz JL, Cancro MP, Cambier JC. B cell maintenance and function in aging. Semin Immunol 2012; 24:342-9.

40. Jensen K, Schaffer L, Olstad OK, Bechensteen AG, Hellebostad M, Tjonnfjord GE, et al. Striking decrease in the total precursor B-cell compartment during early childhood as evidenced by flow cytometry and gene expression changes. Pediatr Hematol Oncol 2010; 27:31-45.
41. Flores KG, Li J, Sempowski GD, Haynes BF, Hale LP. Analysis of the human thymic perivascular space during aging. J Clin Invest 1999; 104:1031-9.
42. Weller S, Braun MC, Tan BK, Rosenwald A, Cordier C, Conley ME, et al. Human blood IgM "memory" B cells are circulating splenic marginal zone B cells harboring a prediversified immunoglobulin repertoire. Blood 2004; 104:3647-54.
43. Weller S, Faili A, Garcia C, Braun MC, Le Deist FF, de Saint Basile GG, et al. CD40-CD40L independent Ig gene hypermutation suggests a second B cell diversification pathway in humans. Proc Natl Acad Sci U S A 2001; 98:1166-70.
44. Jackson KJ, Wang Y, Collins AM. Human immunoglobulin classes and subclasses show variability in VDJ gene mutation levels. Immunol Cell Biol 2014; 92:729-33.
45. Helmin-Basa A, Michalkiewicz J, Gackowska L, Kubiszewska I, Eljaszewicz A, Mierzwa G, et al. Pediatric Helicobacter pylori infection and circulating T-lymphocyte activation and differentiation. Helicobacter 2011; 16:27-35.
46. Mattsson A, Quiding-Jarbrink M, Lonroth H, Hamlet A, Ahlstedt I, Svennerholm A. Antibody-secreting cells in the stomachs of symptomatic and asymptomatic Helicobacter pylori-infected subjects. Infect Immun 1998; 66:2705-12.
47. Lee CK, Weltzin R, Thomas WD, Jr., Kleanthous H, Ermak TH, Soman G, et al. Oral immunization with recombinant Helicobacter pylori urease induces secretory IgA antibodies and protects mice from challenge with Helicobacter felis. J Infect Dis 1995; 172:161-72.
48. Umehara S, Higashi H, Ohnishi N, Asaka M, Hatakeyama M. Effects of Helicobacter pylori CagA protein on the growth and survival of B lymphocytes, the origin of MALT lymphoma. Oncogene 2003; 22:8337-42.
49. Lin WC, Tsai HF, Kuo SH, Wu MS, Lin CW, Hsu PI, et al. Translocation of Helicobacter pylori CagA into Human B lymphocytes, the origin of mucosa-associated lymphoid tissue lymphoma. Cancer Res 2010; 70:5740-8.
50. Sallusto F, Geginat J, Lanzavecchia A. Central memory and effector memory T cell subsets: function, generation, and maintenance. Annu Rev Immunol 2004; 22:745-63.
51. Sallusto F, Lenig D, Forster R, Lipp M, Lanzavecchia A. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. Nature 1999; 401:708-12.
52. Sandberg Y, Almeida J, Gonzalez M, Lima M, Barcena P, Szczepanski T, et al. TCRgammadelta+ large granular lymphocyte leukemias reflect the spectrum of normal antigen-selected TCRgammadelta+ T-cells. Leukemia 2006; 20:505-13.
53. Breit TM, Wolvers-Tettero IL, van Dongen JJ. Unique selection determinant in polyclonal V delta 2-J delta 1 junctional regions of human peripheral gamma delta T lymphocytes. J Immunol 1994; 152:2860-4.
54. Bekker V, Bronke C, Scherpbier HJ, Weel JF, Jurriaans S, Wertheim-van Dillen PM, et al. Cytomegalovirus rather than HIV triggers the outgrowth of effector CD8+CD45RA+CD27- T cells in HIV-1-infected children. Aids 2005; 19:1025-34.
55. Weinberger B, Lazuardi L, Weiskirchner I, Keller M, Neuner C, Fischer KH, et al. Healthy aging and latent infection with CMV lead to distinct changes in CD8+ and CD4+ T-cell subsets in the elderly. Hum Immunol 2007; 68:86-90.
56. van Leeuwen EM, Remmerswaal EB, Vossen MT, Rowshani AT, Wertheim-van Dillen PM, van Lier RA, et al. Emergence of a CD4+CD28- granzyme B+, cytomegalovirus-specific T cell subset after recovery of primary cytomegalovirus infection. J Immunol 2004; 173:1834-41.

57. Chidrawar S, Khan N, Wei W, McLarnon A, Smith N, Nayak L, et al. Cytomegalovirus-seropositivity has a profound influence on the magnitude of major lymphoid subsets within healthy individuals. Clin Exp Immunol 2009; 155:423-32.
58. Derhovanessian E, Maier AB, Hahnel K, Beck R, de Craen AJ, Slagboom EP, et al. Infection with cytomegalovirus but not herpes simplex virus induces the accumulation of late-differentiated CD4+ and CD8+ T-cells in humans. J Gen Virol 2011; 92:2746-56.
59. Appay V, Dunbar PR, Callan M, Klenerman P, Gillespie GM, Papagno L, et al. Memory CD8+ T cells vary in differentiation phenotype in different persistent virus infections. Nat Med 2002; 8:379-85.
60. Babcock GJ, Decker LL, Volk M, Thorley-Lawson DA. EBV persistence in memory B cells in vivo. Immunity 1998; 9:395-404.
61. Chaganti S, Heath EM, Bergler W, Kuo M, Buettner M, Niedobitek G, et al. Epstein-Barr virus colonization of tonsillar and peripheral blood B-cell subsets in primary infection and persistence. Blood 2009; 113:6372-81.
62. Ober C, Loisel DA, Gilad Y. Sex-specific genetic architecture of human disease. Nat Rev Genet 2008; 9:911-22.
63. Whitacre CC, Reingold SC, O'Looney PA. A gender gap in autoimmunity. Science 1999; 283:1277-8.

SUPPLEMENTAL MATERIALS AND METHODS

Study subjects

Included in our study are 1182 two-generation Dutch children. Peripheral blood was obtained from 220 children at birth, from 376 children at the age of 5-9 months, from 241 children at the age of 13-17 months, from 257 children at the age of 22-30 months and from 916 children at the age of 61-95 months. Of 237 children, detailed immunophenotyping data was available at 3 or more follow-up time points, of 90 children at 4 or more time points. The complete follow-up at 5 time points was available for 12 children.

Sample preparation

Approximately 4ml of blood was collected in EDTA-anticoagulant tubes by venous puncture. Blood was kept at room temperature and processed within 24 hours after sampling. 50µl Full blood was mixed with BD Trucount beads (BD Biosciences; San Jose, CA) and stained for CD3, CD16, CD56, CD19 and CD45 in a diagnostic lyse-no-wash approach to obtain absolute cell counts. Subsequently, red blood cells were lysed by adding NH_4Cl. Absolute numbers of leukocytes, NK cells, T cells and B cells were obtained without a wash step on a BD FACSCalibur (BD Biosciences; San Jose, CA).

2ml Full blood was subjected to bulk lysis of red blood cells by 10 min incubation with NH_4Cl. Following wash steps, the remaining white blood cells were divided over 10 tubes for 6-color flowcytometric analyses with the antibodies listed in Supplemental Table 1. Following 10 min incubation at room temperature, cells were washed and immunophenotypic measurements were performed on a 3-laser BD LSRII (BD Biosciences; San Jose, CA). The 488nm laser was used to excite the FITC (530/30nm filter), PE (575/26nm), PerCP-Cy5.5 (695/40nm) and PE-Cy7 (780/60nm) fluorophores, and the 633nm laser for APC (660/20nm) and APC-Cy7 (780/60nm). PMT voltages were set using 8-peak rainbow beads based on target values for the 7^{th} peak as established within the EuroFlow consortium.[27] Absolute cells per µl blood of leukocyte subpopulations were recalculated using the total leukocyte, NK-cell, T-cell and B-cell numbers obtained from the Trucount analysis.

Statistical modeling

Our dataset included 62 leukocyte populations and 26 determinants for 1182 children, with for each leukocyte population (defined as population "i"), a total of 2010 data points.

Linear mixed effect modeling of leukocyte kinetics

For each leukocyte population "i" in our dataset, the following 4 linear mixed effect models were tested:

lmR_0knots<-
lmer(dataset[,i] ~ ns(Age,1)+(1|Id),REML=FALSE,data=data)

lmR_1knots<-
lmer(dataset[,i] ~ ns(Age,2, knots = c(25.5))+(1|Id),REML=FALSE,data=dataset)

lmR_2knots<-
lmer(dataset[,i] ~ ns(Age,3, knots = c(14.1 , 70))+(1|Id),REML=FALSE,data=dataset)

lmR_3knots<-
lmer(dataset[,i] ~ ns(Age,4, knots = c(6, 14, 24))+(1|Id),REML=FALSE,data=dataset)

 # ns(Age,...) defines the natural spline of the age of the children, with indicated knots;
 # (1|Id) defines the child's Id-number as a random effect in the model, to include longitudinal measurements.

The optimal model was selected by likelyhood ratio test:

anov[1]<-ifelse(anova(lmR_0knots,lmR_1knots)[[8]][2]<0.05,2,0)
anov[2]<-ifelse(anova(lmR_1knots,lmR_2knots)[[8]][2]<0.05,3,0)
anov[3]<-ifelse(anova(lmR_2knots,lmR_3knots)[[8]][2]<0.05,4,0)
selected.lmR<-max(anov)

Univariate analyses of the effect of nongenetic determinants on leukocyte kinetics

Each determinant "j" in our dataset was included independently as a fixed effect to the optimal linear mixed effect model of individual leukocyte populations (below an example for a leukocyte population modelled by model lmR_3knots)

*lmer(dataset[,i] ~ ns(Age,4, knots = c(6,14,24))**+determinant[,j]**+(1|Id),REML=FALSE,data=dataset)*

Determinants with an effect of p<0.0125 were defined as significant.

Multivariable analyses of the effect of nongenetic determinants on leukocyte kinetics

All determinants with a significant univariate effect on a leukocyte population (up to three determinants per leukocyte population), were combined and included as fixed effects to the optimal linear mixed effect model of individual leukocyte populations to correct for confounding effects (below an example for a leukocyte population modelled by model lmR_3knots)

*lmer(dataset[,i] ~ ns(Age,4, knots = c(6,14,24))**+significant determinant 1+ significant determinant 2+ etc**+(1|Id),REML=FALSE,data=dataset)*

Determinants with an effect of $p<0.05$ were defined as significant after multivariable correction.

Define whether a determinant has an age–associated affect on a leukocyte population

For determinants that still affected leukocyte kinetics after multivariable correction, we subsequently defined whether this effect was significantly stronger within a selective age-period (either between the age of 0months – knot1, knot1 – knot2 , knot2 – knot3 or knot3 – 76 months). To test this, the models for multivariable analyses were adjusted, testing individual determinants in relation to the age in the children:

e.g. Test the age-associated effect of determinant 1: lmer(dataset[,i] ~ ns(Age,4, knots = c(6,14,24)) **** significant determinant 1 + significant determinant 2+ etc*** *+(1|Id),REML=FALSE,data=dataset)*

 # "" indicates a test of the natural spline of the age of the children in relation to determinant 1*
 # the other determinants significantly affecting this leukocyte population were still included as fixed effects to correct for potential confounding effects

SUPPLEMENTAL TABLES

Supplemental Table 1. Antibody details

Tube		Fluorochrome					
		FITC	PE	PerCP-Cy5.5	PE-Cy7	APC	APC-Cy7
	Antibody	**CD3**	**CD16+CD56**	**CD45**	**CD4**	**CD19**	**CD8**
1	Clone	SK7	B73.1 + C5.9	2D1	SK3	SJ25C1	SK1
	Manufacturer	BD	BD+Dako	BD	BD	BD	BD
	Antibody	**CD15**	**-**	**CD45**	**CD16**	**-**	**CD14**
2	Clone	MMA	-	2D1	3G8	-	MO-P9
	Manufacturer	BD	-	BD	BD	-	BD
	Antibody	**Igκ**	**Igλ**	**CD19**	**CD21**	**-**	**-**
3	Clone	polyclonal	polyclonal	SJ25C1	B-ly-4	-	-
	Manufacturer	Dako	Southern Biotech	BD	BD	-	-
	Antibody	**CD38**	**-**	**CD19**	**CD21**	**-**	**-**
4	Clone	HB7	-	SJ25C1	B-ly-4	-	-
	Manufacturer	BD	-	BD	BD	-	-
	Antibody	**IgD**	**-**	**CD19**	**-**	**IgM**	**CD27**
5	Clone	polyclonal	-	SJ25C1	-	polyclonal	L128
	Manufacturer	Southern Biotech	-	BD	-	BD	BD
	Antibody	**IgA**	**IgG**	**CD19**	**-**	**IgM**	**CD27**
6	Clone	polyclonal	polyclonal	SJ25C1	-	polyclonal	L128
	Manufacturer	Kallestad	Southern Biotech	BD	-	BD	BD
	Antibody	**TCRαβ**	**TCRγδ**	**CD3**	**CD4**	**CD8**	**-**
7	Clone	WT31	11F2	SK7	SK3	SK1	-
	Manufacturer	BD	BD	BD	BD	BD	-
	Antibody	**Vδ2**	**Vδ1 ***	**CD3**	**CD4**	**TCRαβ**	**CD8**
8	Clone	B6.1	R9.12	SK7	SK3	IP26	SK1
	Manufacturer	BD	Beckman Coulter	BD	BD	eBiosciences	BD
	Antibody	**CD28**	**CD197**	**CD3**	**CD8**	**CD45RO**	**CD27**
9	Clone	CD28.2	3D13	SK7	SK1	UCHL-1	L128
	Manufacturer	BD	eBiosciences	BD	BD	BD	BD
	Antibody	**Vδ2**	**Vγ9**	**CD3**	**CD4**	**TCRαβ**	**CD8**
10	Clone	B6.1	B3.1	SK7	SK3	IP26	SK1
	Manufacturer	BD	BD	BD	BD	eBiosciences	BD

Manufactures listed are BD Biosciences, San Jose, Calif; DakoCytomation, Glostrup, Denmark; Southern Biotech, Birmingham, Ala; Kallestad Diagnostics, Chaska, Minn; Beckman Coulter, Indianapolis, Ind; and eBioscience, San Diego, Calif.

APC, Allophycocyanin; ***FITC,*** fluorescin isothiocyanate; ***PE,*** phycoerythrin; ***PerCP,*** peridin-chlorophyll-protein complex.

* Unconjugated antibody detected with Goat anti-Mouse IgG PE (polyclonal; Invitrogen, Waltham, Mass).

Supplemental Table 2. Age-associated reference values of leukocyte populations

		Age category (months)				
	percentile	0	6	14	25	72
Number of children	-	220	376	241	257	916
Age (months)	5th	0.0	5.5	13.6	23.6	69.3
	25th	0.0	6.0	14.1	24.4	70.6
	50th	0.0	6.2	14.4	25.1	72.4
	75th	0.0	6.7	14.8	25.8	75.0
	95th	0.0	7.7	16.1	27.3	79.0
Innate Leukocytes						
Granulocytes	5th	4884.2	1026.1	1196.3	1352.5	1562.4
	25th	8263.0	2055.0	2250.4	1990.9	2840.2
	50th	10572.7	2850.0	3074.1	2822.9	3715.0
	75th	13727.9	4167.8	4143.0	3832.6	5041.0
	95th	18784.8	6519.8	6056.8	6467.5	7597.8
CD15+ granulocytes	5th	5161.2	871.6	1138.2	1310.8	1521.0
	25th	8010.4	1735.8	2072.7	1931.9	2736.1
	50th	10397.8	2617.3	2973.2	2690.5	3608.8
	75th	13564.7	3789.1	4066.6	3746.3	4896.7
	95th	17933.3	5913.7	5633.9	6194.1	7527.4
Neutrophils	5th	4433.3	824.4	1026.1	1116.3	1321.7
	25th	7064.5	1668.6	1926.4	1840.5	2414.2
	50th	9091.8	2497.6	2750.9	2533.5	3274.0
	75th	11634.0	3671.1	3795.5	3665.9	4451.2
	95th	16483.1	5653.1	5506.1	5983.6	6810.9
Eosinophils	5th	489.5	155.2	84.5	83.9	109.4
	25th	891.3	253.1	178.1	142.2	199.1
	50th	1340.0	351.6	248.0	203.8	356.4
	75th	1999.2	486.7	373.2	315.6	598.1
	95th	3721.9	856.0	763.6	617.5	1261.5
Monocytes	5th	496.1	360.1	320.6	272.8	233.9
	25th	1002.2	582.8	504.0	433.1	370.4
	50th	1346.5	789.6	691.6	577.4	490.9
	75th	1749.6	1065.8	962.7	729.6	623.7
	95th	2429.0	1755.0	1510.6	1066.6	924.1
Classical monocytes	5th	378.1	220.4	192.7	190.0	168.8
	25th	829.5	338.6	301.7	304.0	267.4
	50th	1114.3	461.1	425.3	402.6	349.6
	75th	1425.4	623.9	572.6	522.9	455.9
	95th	2125.5	992.3	849.1	714.1	676.4

Supplemental Table 2. Age-associated reference values of leukocyte populations (continued)

		Age category (months)				
	percentile	0	6	14	25	72
Intermediate monocytes	5th	20.6	19.4	26.1	11.6	10.1
	25th	46.6	36.0	42.4	27.9	21.0
	50th	72.3	55.1	70.3	43.7	31.0
	75th	112.1	87.0	116.7	72.0	44.8
	95th	170.0	202.7	243.3	124.3	89.8
Non-classical monocytes	5th	24.0	16.2	13.4	4.2	4.8
	25th	50.6	39.3	24.3	12.6	11.7
	50th	77.3	60.8	41.3	26.8	22.0
	75th	116.0	92.6	66.1	48.9	41.2
	95th	207.5	162.2	139.2	93.4	80.3
NK cells	5th	229.5	160.0	130.0	120.0	90.0
	25th	460.0	250.0	210.0	180.0	140.0
	50th	840.0	330.0	280.0	240.0	190.0
	75th	1350.0	450.0	400.0	320.0	250.3
	95th	2100.5	752.5	690.0	536.0	420.0
Lymphocytes						
T cells	5th	1478.5	2587.5	2180.0	1586.0	1329.0
	25th	2187.5	3740.0	3080.0	2220.0	1770.0
	50th	2695.0	4620.0	3700.0	2830.0	2102.0
	75th	3457.5	5690.0	4610.0	3650.0	2571.8
	95th	4745.5	7722.5	6010.0	4956.0	3560.0
TCRαβ+ T cells	5th	1403.3	2509.1	2063.6	1461.9	1170.2
	25th	2130.3	3680.3	2934.0	2064.0	1593.1
	50th	2607.5	4477.9	3534.7	2691.3	1916.9
	75th	3318.6	5426.5	4401.4	3345.1	2328.9
	95th	4598.5	7552.0	5697.5	4565.6	3295.1
CD4+TCRαβ+ T cells	5th	957.5	1788.3	1388.1	896.6	685.2
	25th	1446.6	2679.3	1926.2	1338.0	949.5
	50th	1848.8	3295.1	2401.3	1739.1	1177.0
	75th	2338.7	3994.7	3023.6	2233.5	1450.3
	95th	3085.2	5597.1	4019.3	3167.0	2067.5
CD8+TCRαβ+ T cells	5th	351.4	575.1	503.3	411.9	364.1
	25th	542.7	856.5	796.1	630.5	537.4
	50th	736.4	1101.7	1013.3	813.2	669.9
	75th	957.6	1427.4	1291.6	1053.7	857.0
	95th	1454.5	2112.0	1894.5	1582.7	1227.1

Supplemental Table 2. Age-associated reference values of leukocyte populations (continued)

		Age category (months)				
	percentile	**0**	**6**	**14**	**25**	**72**
TCRγδ+ T cells	5th	31.5	69.1	68.8	63.7	82.3
	25th	63.1	123.3	106.5	109.3	132.2
	50th	87.7	164.0	162.8	163.1	178.2
	75th	126.0	226.5	213.2	219.8	248.2
	95th	225.1	333.2	332.6	352.8	377.7
CD4+ TCRγδ+ T cells	5th	5.4	6.6	4.8	3.4	2.1
	25th	11.5	13.7	9.1	6.6	4.4
	50th	17.6	20.4	13.3	10.4	6.9
	75th	26.0	31.6	19.3	14.7	10.4
	95th	43.5	50.8	36.0	25.2	18.5
CD8+ TCRγδ+ T cells	5th	3.2	8.5	7.2	6.7	6.9
	25th	7.6	18.2	15.7	14.3	14.3
	50th	12.6	27.7	28.3	25.7	22.3
	75th	19.4	44.4	42.6	41.2	33.7
	95th	39.8	81.0	92.5	75.4	62.2
Vδ1-Vδ2- TCRαβ- T cells	5th	22.9	69.1	28.0	30.6	23.6
	25th	43.1	113.5	53.6	49.0	43.9
	50th	62.5	173.8	89.9	70.7	65.9
	75th	89.7	274.8	130.6	95.9	93.6
	95th	164.1	448.2	288.3	151.5	164.7
Vδ2+ T cells	5th	4.0	13.1	11.4	12.9	22.8
	25th	8.1	30.4	28.8	30.8	50.9
	50th	12.2	43.5	47.5	53.5	77.1
	75th	19.7	65.5	74.1	90.9	118.4
	95th	35.0	124.4	127.8	177.5	212.9
Vδ1+ T cells	5th	7.6	29.0	23.4	18.5	13.6
	25th	18.6	48.2	42.2	33.2	25.4
	50th	30.7	65.1	62.8	49.7	37.7
	75th	43.8	96.4	87.4	67.8	54.3
	95th	69.2	163.8	135.5	134.4	94.4
Vγ9+ T cells	5th	15.1	25.0	28.0	24.7	25.5
	25th	20.1	39.9	43.5	45.1	46.8
	50th	29.7	56.7	62.8	67.1	69.5
	75th	38.7	80.8	100.3	99.2	104.0
	95th	53.6	119.0	181.9	174.8	187.0

Supplemental Table 2. Age-associated reference values of leukocyte populations (continued)

		Age category (months)				
	percentile	**0**	**6**	**14**	**25**	**72**
	5th	1.5	5.2	9.0	8.7	13.3
	25th	3.6	11.8	18.3	20.9	31.0
Vδ2+Vγ9+ T cells	50th	5.7	22.9	31.7	39.2	49.9
	75th	12.5	38.4	58.0	66.0	83.9
	95th	24.3	71.0	116.5	135.7	163.1
	5th	1060.7	1820.5	1430.2	922.3	737.1
	25th	1552.3	2706.7	1991.4	1362.5	988.2
CD4+ T cells	50th	1972.5	3366.2	2457.8	1782.7	1240.0
	75th	2498.9	4101.9	3062.0	2288.7	1525.3
	95th	3266.2	5553.1	4123.0	3217.7	2134.0
	5th	848.7	1519.9	941.0	575.3	392.9
	25th	1321.1	2233.6	1469.5	916.6	621.4
CD4+ Tnaive	50th	1695.9	2849.1	1972.0	1256.0	821.9
	75th	2159.7	3460.9	2461.5	1662.9	1069.7
	95th	2898.2	4841.4	3373.4	2336.6	1538.0
	5th	161.0	354.1	331.6	325.1	344.5
	25th	259.6	470.8	526.9	504.6	461.7
CD4+ Tmem	50th	355.9	580.2	669.0	651.4	559.2
	75th	465.1	713.2	834.9	850.9	679.5
	95th	754.6	953.0	1227.9	1274.9	878.5
	5th	39.5	134.9	111.0	86.7	74.2
	25th	81.4	202.7	176.0	129.5	112.5
CD4+ Tcm	50th	121.3	261.9	242.8	173.1	149.0
	75th	198.3	322.1	313.2	223.6	198.9
	95th	395.6	449.3	427.4	349.3	285.0
	5th	23.4	61.2	31.8	76.9	118.3
	25th	51.5	93.0	123.7	152.4	181.5
CD4+ TemRO	50th	73.8	138.3	184.3	205.4	235.2
	75th	107.0	177.5	254.7	260.3	295.5
	95th	172.6	260.8	387.9	387.4	408.5
	5th	21.0	50.8	20.6	54.6	82.5
	25th	47.6	79.0	95.6	116.7	127.1
CD4+ early TemRO	50th	65.0	110.3	148.8	154.8	168.0
	75th	98.9	144.8	199.8	197.5	210.1
	95th	155.8	211.9	301.6	276.4	292.4

Supplemental Table 2. Age-associated reference values of leukocyte populations (continued)

		Age category (months)				
	percentile	0	6	14	25	72
	5th	0.0	2.7	2.6	9.7	18.2
	25th	1.8	10.2	14.0	17.7	29.7
CD4+ intermediate TemRO	50th	3.8	17.3	23.0	26.2	43.9
	75th	7.0	24.7	33.2	36.4	61.5
	95th	19.6	38.5	53.4	64.0	105.2
	5th	0.0	0.0	0.0	1.4	1.1
	25th	0.0	0.0	2.5	3.8	3.0
CD4+ late TemRO	50th	0.0	0.0	5.1	7.7	6.3
	75th	0.0	3.8	11.2	17.6	12.9
	95th	3.3	14.3	29.0	43.4	30.1
	5th	35.9	79.9	50.3	58.1	55.8
	25th	89.2	122.0	115.1	144.0	102.4
CD4+ TemRA	50th	126.8	178.0	221.2	245.8	153.6
	75th	181.1	241.2	343.4	415.3	227.8
	95th	312.1	376.8	570.5	812.5	371.7
	5th	27.9	66.7	33.1	40.3	36.3
	25th	74.2	104.3	91.1	114.3	67.6
CD4+ early TemRA	50th	113.7	149.9	189.8	213.6	110.7
	75th	162.7	204.6	304.4	356.7	168.4
	95th	287.3	357.8	542.8	782.6	317.4
	5th	0.0	0.0	0.0	0.0	0.8
	25th	0.0	0.0	0.0	1.4	1.8
CD4+ intermediate TemRA	50th	2.7	3.6	2.7	2.7	3.6
	75th	5.1	5.7	4.6	4.7	6.9
	95th	13.5	11.0	9.0	7.8	18.2
	5th	0.0	0.0	0.0	1.7	1.7
	25th	0.0	0.0	3.3	4.4	4.7
CD4+ late TemRA	50th	1.5	3.6	6.9	8.8	9.4
	75th	2.9	6.6	15.2	16.0	17.4
	95th	10.4	17.8	44.5	42.0	42.6
	5th	316.6	527.3	447.2	389.3	335.8
	25th	493.3	836.6	776.2	619.4	509.0
CD8+ T cells	50th	655.5	1093.1	998.8	805.1	649.7
	75th	886.4	1389.5	1277.1	1038.8	841.6
	95th	1319.3	2060.9	1831.8	1571.6	1197.8

Supplemental Table 2. Age-associated reference values of leukocyte populations (continued)

		Age category (months)				
	percentile	**0**	**6**	**14**	**25**	**72**
CD8+ Tnaive	5th	182.4	380.4	269.6	205.6	166.1
	25th	351.1	591.2	453.9	363.8	287.2
	50th	487.2	835.1	610.3	465.1	383.6
	75th	695.2	1081.6	809.5	640.3	513.2
	95th	1037.3	1632.3	1238.7	1033.7	768.1
CD8+ Tmem	5th	57.2	98.1	102.0	71.6	122.4
	25th	102.8	171.4	181.4	152.3	199.0
	50th	149.1	236.3	268.9	224.3	261.0
	75th	214.9	343.6	432.3	331.4	365.3
	95th	351.2	674.6	888.7	550.0	597.5
CD8+ Tcm	5th	10.2	16.7	13.5	9.5	5.1
	25th	27.7	32.1	26.1	18.6	9.3
	50th	47.7	49.2	39.8	31.0	14.4
	75th	71.7	72.6	68.8	56.3	21.7
	95th	147.6	128.0	140.4	109.2	43.8
CD8+ TemRO	5th	7.9	16.7	17.8	14.0	40.5
	25th	17.3	39.3	43.5	36.0	75.5
	50th	29.6	62.6	74.9	67.5	114.9
	75th	48.4	106.4	151.2	108.1	166.3
	95th	98.7	284.8	362.8	244.3	281.8
CD8+ early TemRO	5th	5.5	10.7	7.4	6.4	19.4
	25th	12.2	22.7	22.7	19.6	38.1
	50th	22.2	36.5	38.3	30.5	56.6
	75th	32.9	55.4	68.0	48.7	79.9
	95th	64.3	106.8	145.0	88.0	132.0
CD8+ intermediate TemRO	5th	0.3	0.9	0.1	0.5	3.1
	25th	2.1	4.7	4.7	3.4	8.5
	50th	5.8	10.2	11.6	8.2	17.1
	75th	11.3	21.5	29.9	18.1	30.2
	95th	27.7	64.9	98.6	56.1	64.0
CD8+ late TemRO	5th	0.0	0.0	0.0	0.8	2.2
	25th	0.0	1.2	3.2	3.7	6.5
	50th	0.0	3.1	8.3	9.2	14.0
	75th	0.0	9.4	22.0	24.5	30.4
	95th	0.7	85.0	94.8	73.9	72.6

Supplemental Table 2. Age-associated reference values of leukocyte populations (continued)

		Age category (months)				
	percentile	0	6	14	25	72
CD8+ TemRA	5th	18.9	34.4	33.0	14.8	44.5
	25th	43.1	76.8	69.9	53.0	86.0
	50th	62.5	109.4	122.2	102.8	129.4
	75th	96.3	168.9	224.6	180.9	194.2
	95th	165.1	345.1	467.4	367.6	346.9
CD8+ early TemRA	5th	11.7	22.3	13.7	5.5	17.9
	25th	29.2	44.5	36.4	28.8	33.5
	50th	48.0	68.3	62.1	54.1	52.1
	75th	75.7	106.0	101.0	89.9	76.4
	95th	123.8	167.0	211.4	169.2	139.8
CD8+ intermediate TemRA	5th	1.4	2.6	2.9	1.1	8.9
	25th	6.7	9.8	11.0	6.8	20.4
	50th	11.5	20.8	22.8	17.5	33.8
	75th	18.8	38.9	46.8	34.7	53.4
	95th	40.9	91.3	114.4	79.7	108.0
CD8+ late TemRA	5th	0.0	0.0	1.2	1.2	4.4
	25th	0.0	2.8	5.5	5.4	11.9
	50th	0.4	6.3	11.8	12.9	24.6
	75th	1.0	15.6	36.6	33.0	52.2
	95th	2.1	90.4	187.3	151.6	162.6
B cells	5th	240.0	810.0	710.0	520.0	321.5
	25th	467.5	1310.0	1150.0	860.0	480.0
	50th	675.0	1720.0	1490.0	1130.0	630.0
	75th	930.0	2260.0	1950.0	1440.0	765.3
	95th	1663.5	3277.5	2590.0	1890.0	1110.0
Bnaive	5th	260.8	733.9	621.0	466.0	250.2
	25th	425.6	1190.7	1001.1	772.7	389.5
	50th	616.3	1609.4	1339.5	1019.3	513.6
	75th	854.8	2104.3	1757.7	1292.0	639.9
	95th	1500.9	3030.8	2310.1	1731.0	933.7
Bmem	5th	16.2	47.2	59.1	48.8	46.0
	25th	31.4	79.3	98.4	78.3	74.1
	50th	48.0	118.1	142.1	108.2	98.3
	75th	81.2	171.6	191.7	151.5	131.2
	95th	168.4	291.3	305.1	212.6	200.0

Supplemental Table 2. Age-associated reference values of leukocyte populations (continued)

		Age category (months)				
	percentile	0	6	14	25	72
	5th	7.7	20.7	27.4	29.4	20.8
	25th	19.3	38.7	52.6	44.2	35.8
CD27- Bmem	50th	29.6	62.1	72.2	62.8	48.8
	75th	51.7	88.6	98.6	90.5	65.9
	95th	116.2	153.7	157.5	130.0	103.8
	5th	0.0	4.1	6.2	6.3	9.8
	25th	1.2	10.0	18.4	12.0	17.0
CD27+ Bmem	50th	1.7	16.3	28.1	21.1	25.2
	75th	3.7	27.7	42.5	30.9	37.2
	95th	7.0	65.2	75.3	57.1	65.9
	5th	3.1	11.4	12.2	8.9	8.7
	25th	6.8	24.5	26.0	15.5	16.2
IgM+ Bmem	50th	13.0	41.8	41.2	24.7	25.0
	75th	24.7	65.9	68.0	38.1	36.8
	95th	67.3	131.2	133.8	70.2	61.7
	5th	2.7	8.8	9.6	6.9	6.9
	25th	6.3	19.3	20.1	12.3	13.0
Natural effector	50th	12.5	33.7	33.3	20.4	19.7
	75th	22.3	53.3	55.8	31.3	30.2
	95th	65.2	107.4	113.5	58.3	52.1
	5th	0.0	1.0	0.8	0.9	1.0
	25th	0.0	2.8	3.6	2.4	2.4
IgMonly	50th	0.6	5.7	6.0	4.0	4.1
	75th	1.3	10.8	10.1	6.5	7.1
	95th	3.6	27.2	25.2	15.6	15.7
	5th	5.8	16.9	16.4	15.0	11.6
	25th	11.8	41.3	38.0	30.7	20.4
IgA+ Bmem	50th	21.6	70.2	60.8	47.5	28.9
	75th	36.3	117.1	90.9	75.2	40.0
	95th	67.9	225.2	151.6	131.3	68.1
	5th	4.7	11.3	10.4	10.1	5.0
	25th	10.3	31.2	25.3	20.3	10.2
CD27-IgA+	50th	19.0	55.8	43.4	32.4	16.0
	75th	31.9	94.5	67.3	55.6	23.8
	95th	62.9	196.7	119.1	101.9	42.6

Supplemental Table 2. Age-associated reference values of leukocyte populations (continued)

		Age category (months)				
	percentile	0	6	14	25	72
	5th	0.0	3.0	3.3	3.5	4.5
	25th	0.9	7.4	9.8	8.1	8.4
CD27+IgA+	50th	1.7	12.0	16.3	13.2	12.2
	75th	3.1	20.1	26.7	21.6	17.3
	95th	7.9	42.2	44.2	37.1	27.9
	5th	4.2	30.9	33.1	29.7	31.3
	25th	8.5	61.0	74.3	61.6	55.7
IgG+ Bmem	50th	13.3	87.5	116.8	94.0	77.0
	75th	22.4	143.4	187.6	160.8	102.3
	95th	47.9	229.2	318.2	280.1	158.2
	5th	3.6	23.3	25.9	23.8	19.2
	25th	7.1	48.7	52.6	45.6	36.8
CD27-IgG+	50th	11.2	76.6	87.7	75.4	53.8
	75th	19.2	124.2	151.0	129.2	74.1
	95th	42.2	200.3	274.1	224.0	119.3
	5th	0.0	3.4	4.4	5.2	7.7
	25th	0.8	8.1	14.7	12.6	14.7
CD27+IgG+	50th	1.6	12.3	23.0	19.6	21.1
	75th	3.3	19.2	37.9	33.4	30.1
	95th	7.2	35.3	60.8	58.8	51.0
	5th	22.0	66.2	60.4	43.1	22.0
	25th	38.8	113.7	98.2	74.5	38.4
CD21low B cells	50th	57.8	157.2	142.8	100.3	55.2
	75th	91.1	222.4	193.7	134.8	78.1
	95th	166.2	351.1	282.9	220.0	124.0
	5th	115.0	323.7	290.3	203.9	127.0
	25th	220.7	545.5	500.0	351.5	188.8
Igλ+ B cells	50th	313.6	757.4	638.0	476.5	251.3
	75th	417.7	975.7	840.0	636.2	317.4
	95th	798.4	1486.1	1186.2	847.7	477.2
	5th	126.1	456.6	404.8	293.4	179.0
	25th	248.8	718.2	635.8	490.5	272.3
Igκ+ B cells	50th	352.2	967.0	813.6	634.5	357.0
	75th	493.0	1263.5	1101.9	821.5	439.7
	95th	841.4	1857.9	1429.7	1057.4	621.4

The presented reference values (in cells per microliter of blood) were obtained by using cross-sectional analyses of all 2010 data points on the indicated time points.

Supplemental Table 3. Overall effect estimates and association with specific age-periods

Innate leukocytes								
Models with 3 knots								
Leukocyte population	**Determinant**	**Significance**	**Effect estimate**	**Standard error**	**Significance age-intervals**			
					0-6m	**6-14m**	**14-24m**	**24-76m**
Granulocytes	-	-	-	-	-	-	-	-
CD15+ granulocytes	Gender (girl yes/no)	**	359.6	111.6	**	*	****	ns
Neutrophils	Gender (girl yes/no)	**	296.4	106.6	ns	ns	ns	ns
Eosinophils	Low maternal education level	**	-63.1	24.5	ns	ns	ns	ns
Monocytes	Low maternal education level	**	-51.5	18.1	ns	*	**	ns
Classical monocytes	Low maternal education level	**	-34.3	13.0	*	ns	**	ns
Intermediate monocytes	-	-	-	-	-	-	-	-
NK cells	Seropositivity for CMV at 6yr	**	34.7	12.9	ns	ns	ns	ns
	Siblings >1	****	-53.2	12.1	****	ns	****	ns

Models with 2 knots							
Leukocyte population	**Determinant**	**Significance**	**Effect estimate**	**Standard error**	**0-14.1m**	**14.1-70m**	**70-76m**
Non-classical monocytes	Antibiotics/Penicillin use in 1st yr	**	6.0	2.2	ns	ns	ns

Lymphocytes								
Models with 3 knots								
Leukocyte population	**Determinant**	**Significance**	**Effect estimate**	**Standard error**	**0-6m**	**6-14m**	**14-24m**	**24-76m**
T cells	-	-	-	-	-	-	-	-
TCRαβ+ T cells	Carrier of H. pylori within 6yrs	*	303.4	120.1	ns	ns	ns	*
CD4+ TCRαβ+ T cells	Gender (girl yes/no)	**	96.9	36.1	ns	ns	ns	ns

Lymphocytes								
Models with 3 knots								
Leukocyte population	**Determinant**	**Significance**	**Effect estimate**	**Standard error**	**0-6m**	**6-14m**	**14-24m**	**24-76m**
CD8+ TCRαβ+ T cells	Seropositivity for CMV at 6yr	***	79.2	22.1	ns	ns	*	ns
	Seropositivity for EBV at 6yr	*	51.4	20.4	ns	*	*	ns
TCRγδ+ T cells	Gender (girl yes/no)	*	-11.7	4.6	ns	ns	ns	ns
CD4+ TCRγδ+ T cells	-	-	-	-	-	-	-	-
CD8+ TCRγδ+ T cells	Low maternal education level	**	3.9	1.3	ns	ns	**	ns
	Seropositivity for CMV at 6yr	***	4.9	1.3	**	ns	ns	ns
	Caesarian section	**	-4.8	1.7	ns	ns	ns	ns
Vδ1-Vδ2- TCRαβ- T cells	-	-	-	-	-	-	-	-
Vδ2+ T cells	Low maternal education level	**	7.9	2.9	ns	ns	ns	ns
	Breastfeeding at 6 months of age	*	-6.2	3.1	ns	ns	ns	ns
	Premature rupture of membranes	**	19.5	7.2	ns	ns	ns	ns
Vδ1+ T cells	Seropositivity for CMV at 6yr	****	9.6	1.9	ns	ns	**	ns
CD4+ T cells	Gender (girl yes/no)	**	107.9	37.1	ns	ns	ns	ns
CD4+ Tnaive	Gender (girl yes/no)	*	81.3	32.4	ns	ns	ns	ns
CD4+ Tmem	-	-	-	-	-	-	-	-
CD4+ Tcm	Gender (girl yes/no)	**	12.1	4.6	ns	ns	**	ns
CD4+ TemRO	Seropositivity for CMV at 6yr	***	20.1	5.4	ns	ns	ns	ns
	Seropositivity for EBV at 6yr	**	15.5	5.0	ns	ns	*	ns
CD4+ early TemRO	Gender (girl yes/no)	**	11.2	3.8	ns	ns	ns	ns
	Seropositivity for EBV at 6y	*	9.8	3.8	ns	ns	*	ns
	Caesarian section	*	-13.1	5.3	ns	ns	ns	ns
CD4+ interm TemRO	Seropositivity for CMV at 6yr	***	5.4	1.4	ns	ns	ns	ns
CD8+ T cells	Seropositivity for CMV at 6yr	***	82.3	21.9	ns	ns	**	ns
	Seropositivity for EBV at 6yr	*	47.1	20.3	ns	*	**	ns

Lymphocytes								
Models with 3 knots								
Leukocyte population	**Determinant**	**Significance**	**Effect estimate**	**Standard error**	**0-6m**	**6-14m**	**14-24m**	**24-76m**
CD8+ Tnaive	-	-	-	-	-	-	-	-
CD8+ Tmem	Seropositivity for CMV at 6yr	****	104.3	11.0	*	ns	****	ns
	Seropositivity for EBV at 6yr	****	41.2	10.2	ns	ns	*	ns
CD8+ Tcm	Anti-TPO (before 18 weeks of pregnancy) (mU/ml)	***	0.02	0.007	****	****	**	****
CD8+ TemRO	Seropositivity for CMV at 6yr	****	41.0	6.2	ns	ns	*	ns
	Seropositivity for EBV at 6yr	****	31.3	5.8	ns	ns	*	ns
CD8+ early TemRO	Seropositivity for EBV at 6yr	****	14.1	2.6	ns	ns	*	ns
CD8+ interm TemRO	Seropositivity for CMV at 6yr	****	7.7	1.6	ns	ns	*	ns
	Seropositivity for EBV at 6yr	****	9.2	1.5	ns	ns	**	ns
CD8+ late TemRO	Seropositivity for CMV at 6yr	****	24.8	1.9	**	ns	***	ns
	Seropositivity for EBV at 6yr	**	4.8	1.8	ns	**	ns	ns
CD8+ TemRA	Gender (girl yes/no)	**	-18.1	6.0	**	ns	ns	ns
	Seropositivity for CMV at 6yr	****	64.9	6.5	ns	ns	****	ns
CD8+ early TemRA	-	-	-	-	-	-	-	-
CD8+ interm TemRA	Seropositivity for CMV at 6yr	****	8.0	2.0	ns	ns	ns	ns
	Seropositivity for EBV at 6yr	**	5.5	1.9	ns	ns	*	ns
CD8+ late TemRA	Gender (girl yes/no)	***	-13.7	3.6	**	*	ns	ns
	Seropositivity for CMV at 6yr	****	53.3	3.9	**	ns	****	ns

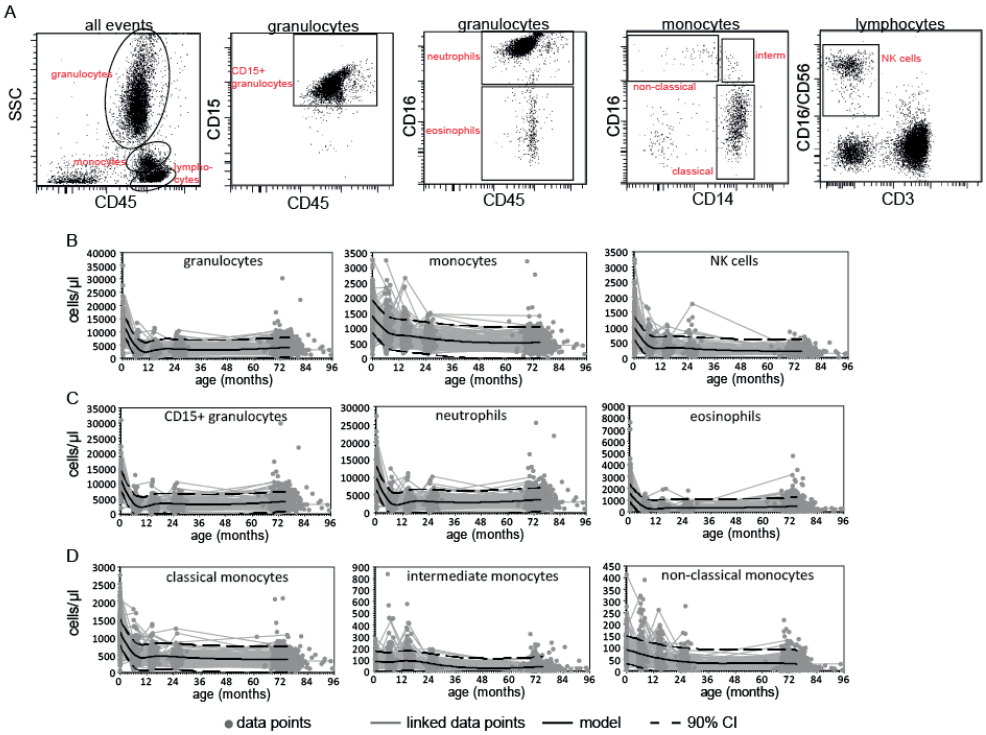
Supplemental Table 3 continued (Lymphocytes; models with 3 knots continued)

Leukocyte population	Determinant	Significance	Effect estimate	Standard error	Significance age-intervals			
					0-6m	6-14m	14-24m	24-76m
B cells	Seropositivity for HSV-1 at 6yr	**	-111.2	35.8	ns	ns	ns	ns
Bnaive	Seropositivity for HSV-1 at 6yr	***	-108.8	33.0	ns	ns	ns	ns
Bmem	Low maternal education level	*	-8.4	3.5	ns	ns	ns	ns
	Seropositivity for EBV at 6yr	**	-9.3	3.4	ns	ns	ns	ns
CD27- Bmem	Low maternal education level	*	-4.7	2.0	ns	ns	ns	ns
	Seropositivity for EBV at 6yr	**	-6.2	2.0	ns	ns	*	ns
CD27+ Bmem	Seropositivity for EBV at 6yr	*	-2.7	1.1	*	ns	ns	ns
IgM+ Bmem	-	-	-	-	-	-	-	-
Natural effector	-	-	-	-	-	-	-	-
IgMonly	-	-	-	-	-	-	-	-
IgA+ Bmem	Breastfeeding at 6 months of age	*	-5.8	2.3	ns	ns	ns	*
	Carrier of H. pylori within 6yrs	****	22.0	4.9	ns	ns	*	****
CD27-IgA+	Breastfeeding at 6 months of age	**	-4.9	1.9	ns	ns	ns	ns
	Carrier of H. pylori within 6yrs	****	17.8	4.1	ns	ns	*	****
CD27+IgA+	Gender (girl yes/no)	**	1.6	0.6	ns	ns	ns	ns
	Carrier of H. pylori within 6yrs	*	3.1	1.2	ns	ns	ns	*
IgG+ Bmem	-	-	-	-	-	-	-	-
CD27-IgG+	Antibiotics/Penicillin use in 1st yr	*	7.1	2.8	ns	ns	ns	ns
CD27+IgG+	Gender (girl yes/no)	**	2.1	0.7	ns	ns	ns	ns
CD21low B cells	Low maternal education level	***	-11.7	3.2	ns	ns	ns	ns
Igλ+ B cells	Seropositivity for HSV-1 at 6yr	**	-53.0	16.6	ns	ns	ns	ns
Igκ+ B cells	Seropositivity for HSV-1 at 6yr	**	-59.5	19.9	ns	ns	ns	ns
	Carrier of H. pylori within 6yrs	**	82.2	30.9	ns	ns	ns	**

Models with 2 knots							
Leukocyte population	**Determinant**	**Significance**	**Effect estimate**	**Standard error**	**0-14.1m**	**14.1-70m**	**70-76m**
CD4+ late TemRO	Seropositivity for CMV at 6yr	****	6.1	0.7	ns	ns	*
	Seropositivity for EBV at 6yr	***	2.7	0.7	ns	ns	ns
	Premature rupture of membranes	**	5.0	1.8	ns	****	ns
CD4+ TemRA	-	-	-	-	-	-	-
CD4+ early TemRA	-	-	-	-	-	-	-
CD4+ interm TemRA	Gender (girl yes/no)	**	-0.7	0.2	ns	ns	ns
	Premature rupture of membranes	**	1.9	0.6	ns	ns	ns

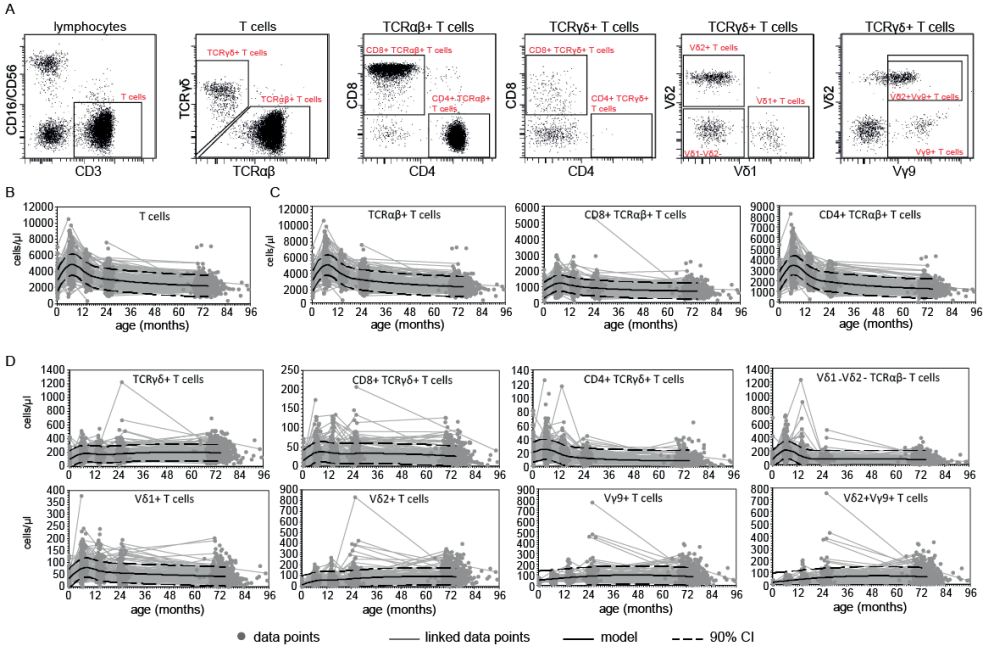
Models with 1 knot						
Leukocyte population	**Determinant**	**Significance**	**Effect estimate**	**Standard error**	**0-25.5m**	**25.5-76m**
CD4+ late TemRA	Gender (girl yes/no)	***	-3.0	0.9	ns	ns
	Seropositivity for CMV at 6yr	****	7.4	0.9	****	*
	Seropositivity for EBV at 6yr	*	1.8	0.9	ns	ns
Vγ9+ T cells	Premature rupture of membranes	**	23.1	7.4	*	ns
Vδ2+Vγ9+ T cells	Premature rupture of membranes	***	24.6	6.6	*	ns

NS, Not significant;
*P <0.05.
**P <0.01.
***P <0.001.
****P <0.0001.



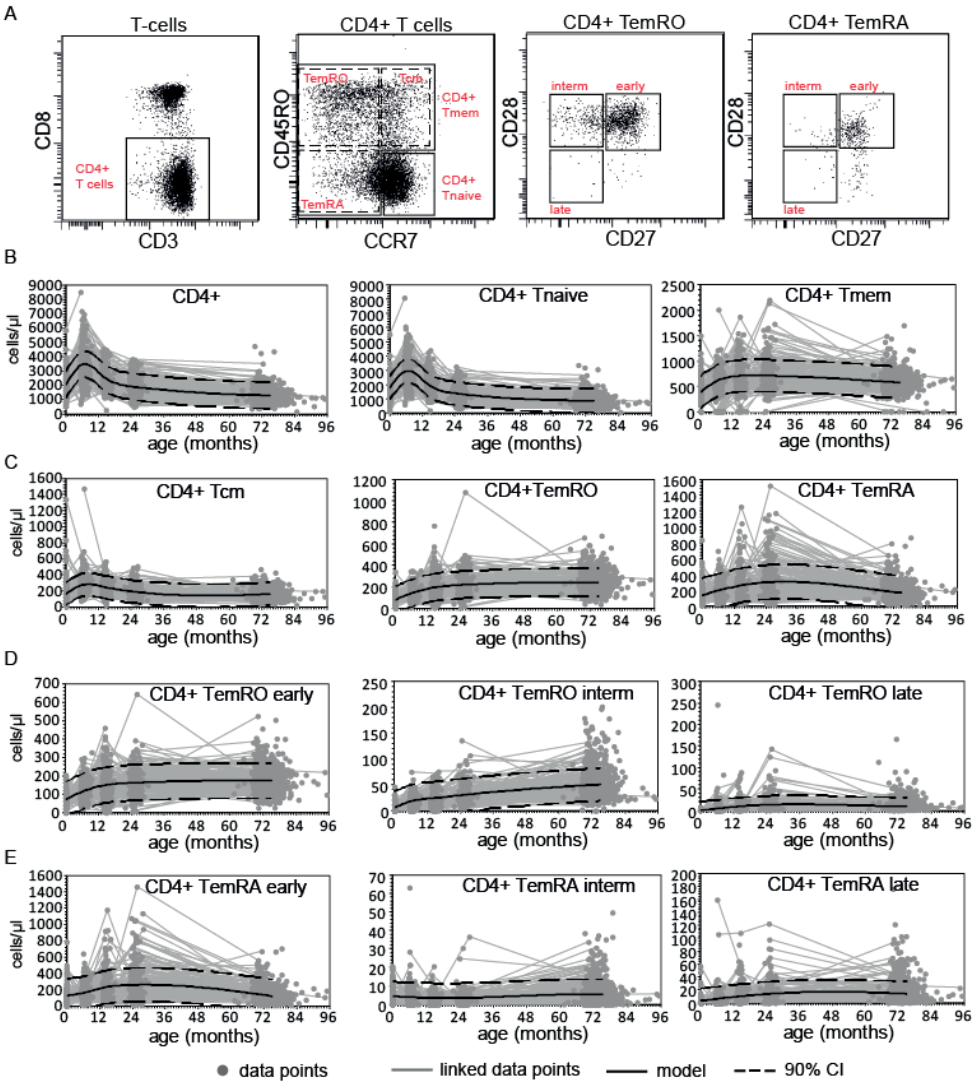
Supplemental Figure 1. Dynamics of innate leukocyte subsets in children between birth and 6 years of age.

A) Gating strategy used to identify innate leukocyte subsets. Plots depict data of a representative 61- to 95 months-old child. *SSC,* Side scatter. **B-D)** Absolute numbers and modeled dynamics of granulocytes, monocytes and NK cells **(B)**, $CD15^+$, neutrophilic and eosinophilic granulocytes **(C)**, and classical, intermediate and non-classical monocyte subsets **(D),** from birth until 6 years of age. Linear mixed effect models were generated for each population (***solid black line***) and represented with the 90% confidence interval (CI) of the model (***dashed black lines***). For clarity of the graphs, ***gray lines*** connect only consecutive time points within one individual; i.e. 0-6 months, 6-14 months, 14-25 months or 25-76 months.

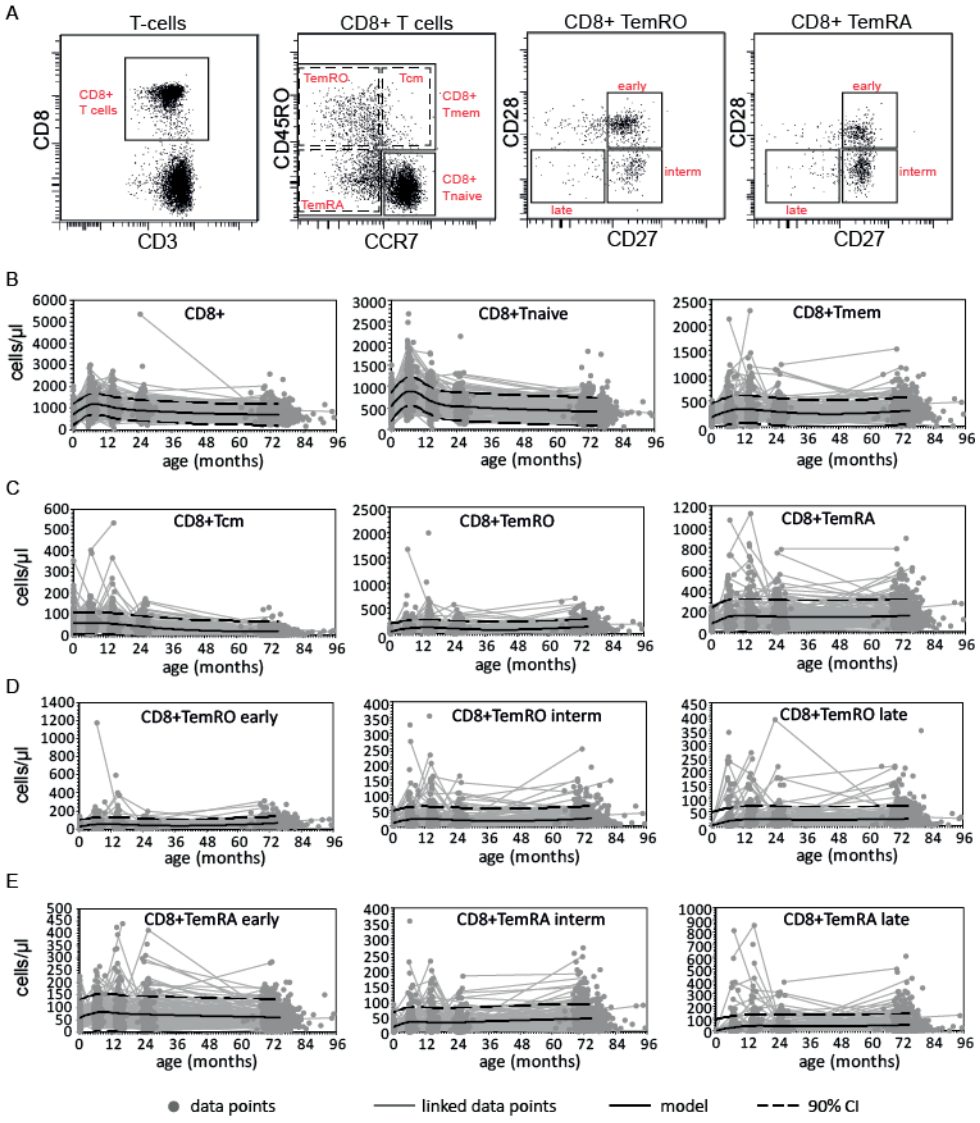


Supplemental Figure 2. Dynamics of TCR$\alpha\beta^+$ T-cell and TCR$\gamma\delta^+$ T-cell subsets in children between birth and 6 years of age.

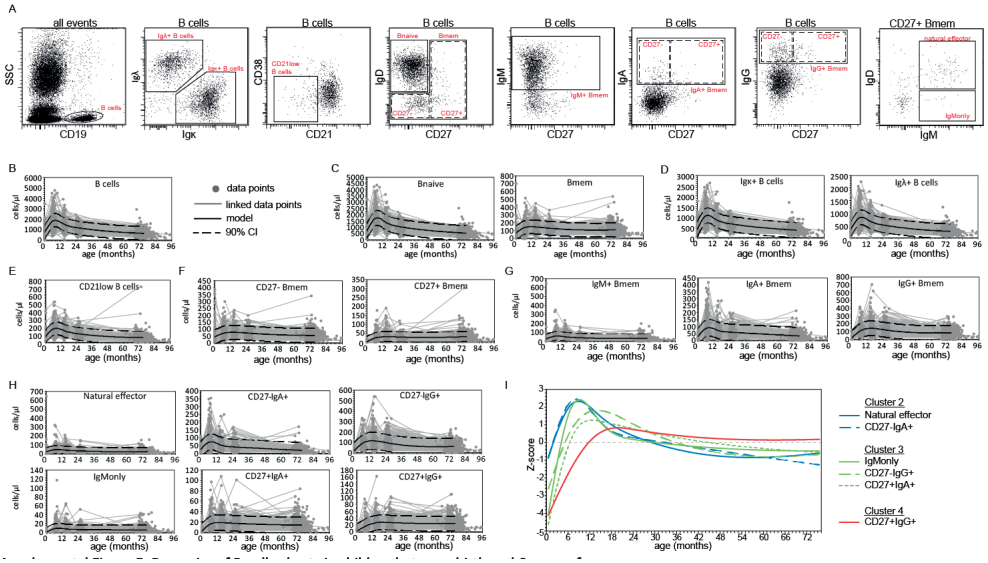
A) Gating strategy used to identify TCR$\alpha\beta^+$ T-cell and TCR$\gamma\delta^+$ T-cell subsets. Plots depict data of a representative 61- to 95 months-old child. **B-D)** Absolute numbers and modeled dynamics of T cells **(B)**, TCR$\alpha\beta^+$ T-cell populations **(C)**, and TCR$\gamma\delta^+$ T-cell populations **(D),** from birth until 6 years of age. Linear mixed effect models were generated for each population (***solid black line***) and represented with the 90% confidence interval (CI) of the model (***dashed black lines***). For clarity of the graphs, ***gray lines*** connect only consecutive time points within one individual; i.e. 0-6 months, 6-14 months, 14-25 months or 25-76 months.



Supplemental Figure 3. Dynamics of CD4$^+$ T-cell subsets in children between birth and 6 years of age.
A) Gating strategy used to identify CD4$^+$ T-cell subsets. Plots depict data of a representative 61- to 95 months-old child. **B-E)** Absolute numbers and modeled dynamics of total, naive and memory CD4$^+$ T cells **(B)**; CD4$^+$ central memory T cells (Tcm), CD45RO$^+$ effector memory (TemRO) and CD45RO$^-$ effector memory (TemRA) T-cell subsets **(C)**; early, intermediate or late differentiated TemRO subsets **(D)**; or early, intermediate or late differentiated TemRA subsets **(E)**, between birth and 6 years of age. Linear mixed effect models were generated for each population (***solid black line***) and represented with the 90% confidence interval (CI) of the model (***dashed black lines***). For clarity of the graphs, ***gray lines*** connect only consecutive time points within one individual; i.e. 0-6 months, 6-14 months, 14-25 months or 25-76 months.

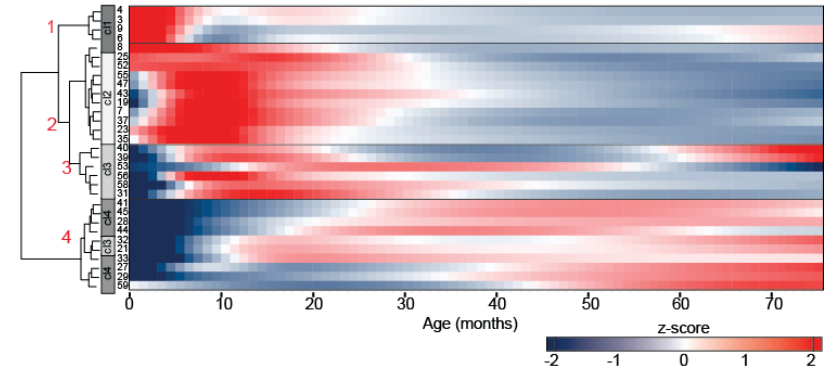


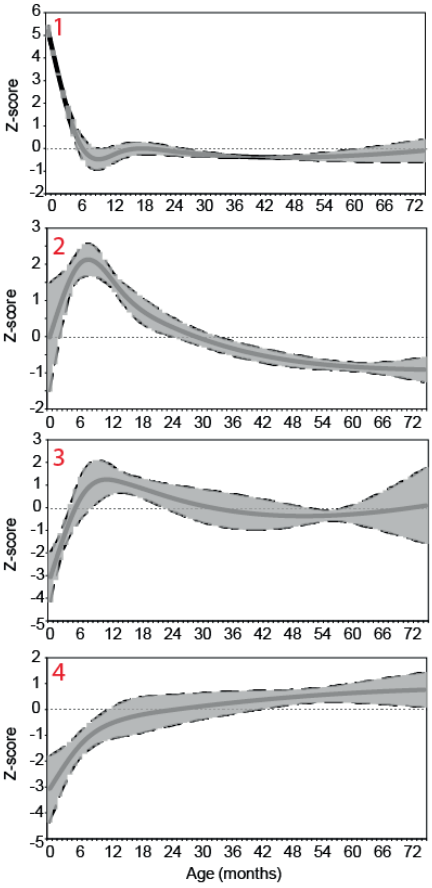
Supplemental Figure 4. Dynamics of CD8$^+$ T-cell subsets in children between birth and 6 years of age.
A) Gating strategy used to identify CD8$^+$ T-cell subsets. Plots depict data of a representative 61- to 95 months-old child. **B-E)** Absolute numbers and modeled dynamics of total, naive and memory CD8$^+$ T cells **(B)**; CD8$^+$ central memory T cells (Tcm), CD45RO$^+$ effector memory (TemRO) and CD45RO$^-$ effector memory (TemRA) T cell subsets **(C)**; early, intermediate or late differentiated TemRO subsets **(D)**; or early, intermediate or late differentiated TemRA subsets **(E)**, between birth and 6 years of age. Linear mixed effect models were generated for each population (*solid black line*) and represented with the 90% confidence interval (CI) of the model (*dashed black lines*). For clarity of the graphs, *gray lines* connect only consecutive time points within one individual; i.e. 0-6 months, 6-14 months, 14-25 months or 25-76 months.



Supplemental Figure 5. Dynamics of B-cell subsets in children between birth and 6 years of age.

A) Gating strategy used to identify B-cell subsets. Plots depict data of a representative 61-95 months-old child. **B-I)** Absolute numbers and modeled dynamics of total B cells **(B)**; naive and total memory B cells **(C)**; total Igκ^{+} and Igλ^{+} B-cells populations **(D)**; and CD21^{low} B cells **(E)**; CD27^{-} and CD27^{+} memory B cells **(F)**; total IgM^{+}, IgA^{+} or IgG^{+} memory B cells **(G)**; IgM^{+}, IgA^{+} or IgG^{+} memory B-cell subsets **(H)** between birth and 6 years of age. Linear mixed effect models were generated for each population (solid black line) and represented with the 90% confidence interval (CI) of the model (dashed black lines). For clarity of the graphs, gray lines connect only consecutive time points within one individual; i.e. 0-6 months, 6-14 months, 14-25 months or 25-76 months. **I)** Overlay of the normalized (zero mean; unit SD) linear mixed effects models of the memory B-cell populations as in panel **H**.





4 Eosinophils
3 Neutrophils
9 NK cells
6 Classical monocytes

8	Non-classical monocytes	cluster 1
25	CD4+ Tcm	
52	Natural effector B cells	
55	CD27-IgA+	
47	Bnaive	
43	CD8+ TemRA early	
19	Vδ1+ T cells	
7	Intermediate monocytes	
37	CD8+ Tcm	
23	CD4+ Tnaive	
35	CD8+ Tnaive	

40 CD8+ TemRO interm
39 CD8+ TemRO early
53 IgMonly
56 CD27+IgA+
58 CD27-IgG+
31 CD4+ TemRA early

41	CD8+ TemRO late	cluster 3
45	CD8+ TemRA late	
28	CD4+ TemRO interm	
44	CD8+ TemRA interm	
32	CD4+ TemRA interm	
21	Vδ2+Vγ9+ T cells	
33	CD4+ TemRA late	
27	CD4+ TemRO early	
29	CD4+ TemRO late	
59	CD27+IgG+	

Supplemental Figure 6. Hierarchical clustering of the dynamics of 31 non-overlapping leukocyte subsets in early childhood.

A) Ward hierarchical clustering was performed as in Figure 2, including only the 31 non-overlapping leukocyte subsets. Indicated in *gray squares* in front of the heat map is the cluster to which each population was assigned in the analysis of the total 62 leukocyte subsets in Figure 2. **B)** Average patterns ± 1 SD of the subsets in each of the 4 major clusters. Indicated in *black squares* are 3 subsets that were assigned to different clusters than in Figure 2.

Chapter 3

Decreased memory B cells and Increased CD8 memory T cells in Blood of Breastfed Children:

The Generation R Study

Michelle A.E. Jansen, Diana van den Heuvel, Menno C. van Zelm,
Vincent W.V. Jaddoe, Albert Hofman, Johan C. de Jongste, Herbert Hooijkaas , Henriette A. Moll

PLoS One 2015; May 18;10(5)

ABSTRACT

Background

Breastfeeding provides a protective effect against infectious diseases in infancy. Still, immunological evidence for enhanced adaptive immunity in breastfed children remains inconclusive.

Objective

To determine whether breastfeeding affects B- and T- cell memory in the first years of life.

Methods

We performed immunophenotypic analysis on blood samples within a population-based prospective cohort study. Participants included children at 6 months (n=258), 14 months (n=166), 25 months (n=112) and 6 years of age (n=332) with both data on breastfeeding and blood lymphocytes. Total B- and T-cell numbers and their memory subsets were determined with 6-color flow cytometry. Mothers completed questionnaires on breastfeeding when their children were aged 2, 6, and 12 months. Multiple linear regression models with adjustments for potential confounders were performed.

Results

Per month continuation of breastfeeding, a 3% (95% CI -6,-1) decrease in CD27+IgM+, a 2% (95 CI % -5,-1) decrease in CD27+IgA+ and a 2% (95% CI -4,-1) decrease in CD27-IgG+ memory B cell numbers were observed at 6 months of age. CD8 T-cell numbers at 6 months of age were 20% (95% CI 3,37) higher in breastfed than in non-breastfed infants. This was mainly found for central memory CD8 T cells and associated with exposure to breast milk, rather than duration. The same trend was observed at 14 months, but associations disappeared at older ages.

Conclusions

Longer breastfeeding is associated with increased CD8 T-cell memory, but not B-cell memory numbers in the first 6 months of life. This transient skewing towards T cell memory might contribute to the protective effect against infectious diseases in infancy.

INTRODUCTION

Breast milk contains factors that enhance nutrient absorption, stimulate growth and enhance the defense against pathogens.[1] Consequently, breastfeeding provides protection against infectious diseases during infancy.[2,3,4] The protective effect persists during childhood [5,6], and modulates vaccination responses.[7,8,9] Thus, it is likely that breastfeeding not only provides passive immunization, but also enhances adaptive immunity.[10]

B and T lymphocytes comprise the cellular components of adaptive immunity, and are generated throughout life. B cells contribute to humoral immunity through the production of immunoglobulins (Ig), whereas CD8+ cytotoxic T cells provide cellular immune responses. CD4+ helper T cells support both humoral and cellular immune responses. Each B and T cell generates a unique antigen receptor during precursor differentiation in bone marrow or thymus, respectively. Only those cells that specifically recognize antigen with their receptor will undergo clonal proliferation and are involved in the antigen response. Cells generated from the clonal expansion will remain present in the body as long-lived memory cells and will initiate a fast and quantitatively stronger response upon secondary antigen encounter. In addition to CD27- naive B cells, six memory B-cell subsets can be identified.[11] Four of these express CD27 and are either positive for IgM, IgM and IgD, IgA or IgG. In addition, CD27-IgA+ and CD27-IgG+ memory B cells can be identified. Within both the CD4 and CD8 T-cell lineages, central memory (CD45RO+CCR7+), CD45RO+CCR7- effector memory (TemRO) and CD45RO-CCR7- (TemRA) can be distinguished from naive T cells (CD45RO-CCR7+).[12] Central memory T cells are most efficient in generating a new immune response by proliferating extensively in response to an antigen upon secondary antigen encounter.[13,14,15] The diversity and composition of the B- and T-cell compartments are highly dynamic in the first years of life; blood cell counts are especially high up to 2 years of age, following which they slowly decline to reach adult levels between 6 and 10 years.[16-20] At birth, nearly all B and T cells are naive, and memory cells are gradually built up in the first 6 years of life.[19]

Several studies have addressed the effects of breastfeeding on adaptive immunity. Breastfeeding was found to be associated with decreased frequencies of blood CD4+ T cells.[21] This was mostly due to lower frequencies of naïve (CD45RA+) T cells in breastfed children. Still, these observations were not consistently reproduced with some studies showing increased, and some decreased numbers of CD4+ T cells.[21-24] Furthermore, long term breastfeeding was found to be associated with increasing numbers of CD4+ and CD8+ T cells.[24] Thus, although previous studies have addressed the influence of breastfeeding on blood lymphocyte populations,[21-24] the results remained inconclusive, mainly due to small samples sizes and limitations in the detection of memory cells.[21-24]

In the present study we used 6-color flow cytometric analysis of lymphocyte subsets in a population-based prospective cohort study to assess the impact of breastfeeding on build-up of memory B and T cells in infants and young children.

MATERIALS AND METHODS

Design and study population

This study was embedded in the Generation R Study, a prospective population-based cohort study that follows pregnant women and their children from fetal life onwards in the Netherlands.[25] The study has been approved by the Medical Ethics Committee of the Erasmus MC, University Medical Centre Rotterdam. Written informed consent was obtained from all parents of participants. We included 1,079 Dutch pregnant women and their children participating in a detailed subgroup study.[25] All children were born between February 2002 and August 2006. We excluded twins (n=27) in the present analysis to prevent bias due to correlation. Of these, data on both breastfeeding and immunophenotyping were available from 258 children at 6 months, 166 at 14 months, 112 at 25 months and 332 at 6 years of age. The main reasons for missing samples were due to non-consent of the parents (approximately 55% per visit) and technical or logistical failure (approximately 10% per visit).

Breastfeeding

Information regarding breastfeeding was obtained in postnatal questionnaires at the ages of 2, 6 and 12 months.[25] Mothers were asked whether they had ever breastfed their child (yes or no) and, if yes, at what age (months) they had stopped.[26,27] Breastfeeding duration was then categorized into four groups: never, ≤3 months, between 3 and 6 months and ≥6 months. An approximation of exclusive breastfeeding was performed according to whether the child received breastfeeding without any other bottle feeding, milk or solids.[27] Partial breastfeeding indicates infants receiving both breast-feeding, bottle feeding and/or solids in this period. Subsequently, the information on exclusiveness of breastfeeding was combined and categorized into the following breastfeeding categories: never; partial until 4 months and exclusive until 4 months.

Immunophenotyping of lymphocyte subsets

Flow cytometry was performed within 24 hours following sampling on fresh whole blood at the ages of 6 months (median 6.2; range 5.2; 8.2), 14 months (median 14.4; range 13.1-17.4), 25 months (median 25.2; range 23.3-29.8) and 6 years (median 5.9; range 5.1-7.2). Absolute counts of blood CD3+ T cells, CD16/56+ NK cells, and CD19+ B cells were obtained with a diagnostic lyse-no-wash protocol. Lymphocytes were gated on the basis of CD45, FSC and SSC characteristics. Gates were set based on cells that are known to lack expression of the indicated marker. Additionally, 6-color flow cytometry was performed on an LSRII (BD Biosciences) to distinguish naive and memory B- and T- lymphocyte subsets as defined previously (S3 Table).[11,12] All flow cytometry acquisition was performed on whole blood after red blood cell lysis with ammonium chloride.

Covariates

The covariates that were assessed in this study were obtained from midwife and hospital registries at birth (birth weight, gestational age and gender) or through measurements at the research center (child anthropometrics). Additionally, information on smoking and alcohol use during pregnancy and socioeconomic status was obtained by prenatal questionnaires sent during the first, second and third trimesters of pregnancy [28 29]. Information on day-care attendance was obtained from parent-reported questionnaires at the ages of 6 and 12 months.

Statistical methods

Because the distribution of lymphocyte numbers in different age groups was skewed, these values were normalized by transformation to a natural log-scale. Differences in maternal and infant characteristics between breastfed versus never breastfed children were tested using independent samples t-tests and Chi-Square tests. Differences in baseline characteristics among the groups with different duration of breastfeeding were assessed using ANOVA tests and Chi-Square tests. Additionally, the associations of breastfeeding, breastfeeding duration (measured in groups, and measured continuously per month continuation of breastfeeding) and breastfeeding exclusivity with the change in lymphocyte numbers were assessed using multiple linear regression models with adjustment for potential confounders. For all ages, the category with no breastfeeding was the reference. First, associations between breastfeeding and total B, T, CD4 and CD8 counts were assessed. Subsequently, associations for B, CD4 and CD8 subpopulations were studied to assess the effect of breastfeeding on memory cells specifically. Multivariable regression models were created with stepwise adjustment for potential confounders, which were selected based on previous literature. Potential confounders included: maternal age, socioeconomic status (SES), marital status, maternal BMI, maternal smoking and alcohol consumption during pregnancy, maternal reported autoimmune disease (including thyroid disease, multiple sclerosis, systemic lupus erythematosus, diabetes mellitus and arthritis), elevated anti-tTG level during pregnancy, maternal fever in the last trimester of pregnancy, family history of asthma or atopy (hay-fever, allergy, eczema), multiple parities, mode of delivery (caesarean section), gender, birth weight, gestational age, preterm birth, APGAR score, birth season, weight and age at focus visit, fever in the first 6 months (yes/no), frequency of upper and lower respiratory tract infections, and day-care attendance in the first year of life. Covariates were kept in the final multivariate model if the covariate resulted in an alteration in effect estimate of $\geq$ 10%,[30] or if the variables were associated with breastfeeding (determinant) and lymphocyte numbers (outcome) in our study. Because of the small numbers in the never breastfed group, final adjustment for potential confounders was restricted to those who attained the strongest alteration (%) in effect estimates. Because of the strong correlation between our outcomes (e.g. Pearson's correlation between total CD8 T cells and naive CD8 T cells r =0.82, and between total B cells and IgA r=0.57), and

unweighted calculations only hold if the tests are independent,[31,32] we did not perform adjustments for multiple testing. All measures of associations are presented with their 95% confidence interval. All statistical analyses were performed using the Statistical Package for the Social Sciences version 20.0 for Windows (SPSS Inc, Chicago, IL, USA). P values <0.05 were considered to be statistically significant.

RESULTS

Population characteristics

No major differences in characteristics between children included at 6, 14, 25 months and 6 years of age were observed (Table 1, Supplemental Table 1) Overall, more than 86% of mothers started breastfeeding. Mother's educational level was significantly associated with the start of breastfeeding at 6 and 25 months, and 6 years of age. Moreover, mother's educational level was significantly associated with both the duration of breastfeeding and with B cell memory subsets at 6 months of age [data not shown]. In addition, maternal alcohol use was related to both breastfeeding duration and total T, B, CD4 and CD8 cell numbers at 6 months [data not shown].

Both maternal educational level and alcohol use influenced the regression coefficients by more than 10%. Therefore, all subsequent analysis on breastfeeding duration and cell numbers were adjusted for both maternal education and maternal alcohol use during pregnancy.

Decrease in B–cell memory

Breastfeeding exposure (Fig. 1b) and duration (Fig. 1c) were not associated with total B cell numbers at 6, 14, 25 months and 6 years of age (Table 2, S2 Table). Furthermore, no associations were observed between breastfeeding exposure and duration and naive B cells, which constitute the majority of total B-cell numbers (Table 3, Figs. 1b and 1c) However, a longer duration of breastfeeding, was associated with changes in the memory B-cell compartment. Per month longer breastfeeding, a 3% decrease in absolute numbers of CD27+IgM+, and a 2% decrease in both CD27+IgA+ and CD27-IgG+ memory B cells at 6 months of age were observed (Table 3, Fig. 1c). CD27+IgG+ and CD27-IgA+ memory B cells at 6 months of age did not change with differences in breastfeeding duration. Stronger negative trends were observed for the associations between breastfeeding duration and frequencies of CD27+IgM+, CD27+IgA+ and CD27-IgG+ memory B cells within total B cells at 6 months of age, resp. -4% (95% CI -7,-1), -4% (95% CI -7, -1), and -3% (95% CI -5,-1) (data not shown).

At 14 months of age, similar trends for breastfeeding duration and total B-cell numbers were observed, although not significant (Table 2). At older ages the effects disappeared. (S2 Table). Thus, a longer breastfeeding duration seems to negatively impact B-cell numbers in infants at least until the age of 6 months, at older ages these effects disappeared.

Table 1. Maternal and infant characteristics of the study population at 6 months

	Not breastfed (n=35)	Breastfed (n=223)
Maternal characteristics (n=258)		
Age (Mean ± SD; years)	32 (3.6)	32 (3.8)
Educational level (n; %)		
Lower	24 (69%)	71 (32%)
Higher	11 (31%)	152 (68%)*
Net household income per month (n; %)		
< € 2400	2 (7%)	23 (11%)
≥ € 2400	29 (93%)	178 (89%)
Smoking continued during pregnancy (n; %)	4 (14%)	21 (13%)
Alcohol use continued during pregnancy (n; %)	8 (29%)	55 (35%)
Body Mass Index before pregnancy (Mean ± SD; kg/m^2)	23 (3)	23 (4)
Fever in third trimester of pregnancy (n; %)	3 (9)	14 (6%)
Maternal atopy (eczema, allergy HDM, hay-fever)(n; %)	8 (25%)	72 (35%)
Paternal atopy (eczema, allergy HDM, hay-fever) (n; %)	8 (28%)	55 (28%)
Family history of asthma / atopy (n; %)	13 (37%)	107 (49%)
Any reported autoimmune disease (diabetes mellitus, SLE, arthritis, MS, thyroid disorder, or celiac disease) (n; %)	0 (0%)	5 (0.02%)
Mode of delivery (n; %)		
Vaginal	21 (64%)	129 (61%)
Forceps or vacuum assisted	5 (15%)	44 (20%)
Caesarian section	7 (21%)	40 (19%)
Premature rupture of membranes (n; %)	3 (9%)	6 (3%)
Infant characteristics (n=258)		
Males (n; %)	19 (54%)	113 (51%)
Gestational age (Mean ± SD; weeks)	39.4 (2.4)	40.0 (1.7)
Preterm birth (<37 weeks) (n; %)	2 (6%)	11 (5%)
Birth weight (Mean ± SD; grams)	3439 (632)	3504 (524)
Apgar score at 5 min <7 (n; %)	0 (0%)	3 (1%)
Birth season (n; %)		
Winter (dec-jan-feb)	5 (14%)	36 (16%)
Spring (mar-apr-may)	10 (29%)	81 (36%)
Summer (jun-jul-aug)	11 (31%)	62 (28%)
Autumn (sept-oct-nov)	9 (26%)	44 (20%)
Siblings ≥1 (n; %)	4 (11%)	26 (12%)
Day-care >16 hours /week (n; %)	5 (33%)	77 (46%)
Fever in first 6 months (n; %)	12 (60%)	121 (62%)
Age at focus visit (Median ± range; months)	6.6 (6.2-8.2)	6.2 (5.2-7.9)*

Values are means (SD), absolute numbers (percentages) or #medians (90% range). *Significantly different between groups.

Data were missing on: Household income (n=26) Smoking during pregnancy (***n***=72), Alcohol during pregnancy (***n***=72) BMI before pregnancy (n=37), Fever in third trimester (n=9) Mode of delivery (n=12) Maternal atopy (n=22) Paternal atopy (n=29), Family history of asthma or atopy (n=3) Mode of delivery (n=12), maternal reported autoimmune disease (n=29), maternal reported any other chronic condition (n=31), premature rupture of membranes (n=5), Apgar (n=4), Day-care (n=42), weight at focus visit (n=1), fever in first 6 months (n=42). Any reported autoimmune disease included thyroid disease (n=4) and elevated anti-tTG level during pregnancy (n=1).

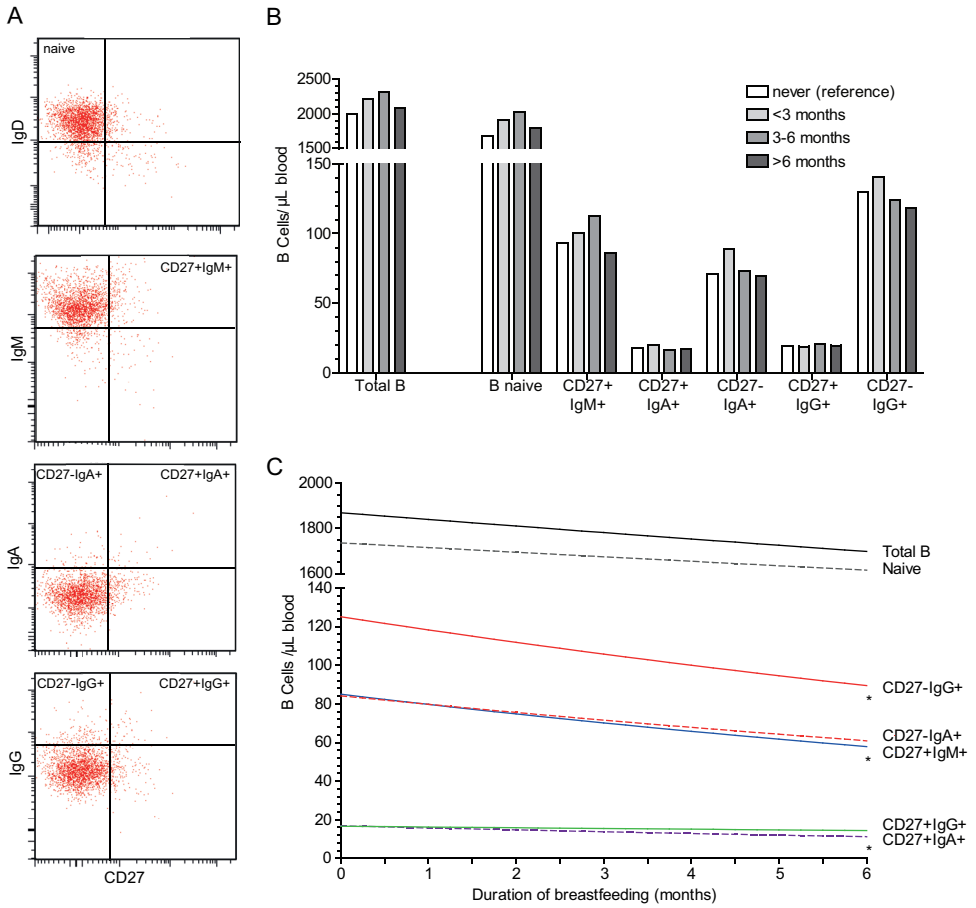


Figure 1. Impact of breastfeeding duration on B lymphocyte subsets at 6 months of age. Figure 1 vervangen door figuur met geode kwaliteit!

A) Gating strategy for dissection of CD19^{+} B cells into 1 naive and 5 memory B-cell subsets by flow cytometry. The reference plots depict density plots of total lymphocytes and were used to set the gates accordingly. **B)** Backtransformed B-cell counts (cells/µL) at 6 months of age according to different breastfeeding duration categories (reference category is never). Categories of breastfed children contain both partial and exclusively breasted children. **C)** The estimated backtransformed regression line reflects B-cell counts (cells/µL) at 6 months of age, per month increase in breastfeeding duration.*, P<0.05.

Increase in T–cell memory

Breastfeeding exposure and duration were not associated with total T-cell and CD4+ T-cell numbers at 6, 14, 25 months and 6 years of age (Table 2, Figs. 2b and 2c, and S2Table). However, the exposure to breastfeeding was associated with CD8+T cell numbers. CD8+ T-cell numbers were 19% (95% CI 3, 35) higher in 6-month old infants who were breastfed for less than 3 months and remained 20% (95% CI 3, 37) higher in children who were breastfed until 6 months (Table 2, Fig. 2b), than in children who were never breastfed. Comparable effect sizes were observed for exclusiveness of breastfeeding in relation to CD8 T cells (Table

Table 2. Adjusted associations between breastfeeding and cell numbers at 6 and 14 months

	Regression coefficients for logtransformed cell numbers (*10e9/L)										
	6 months (n=258)						14 months (n=166)				
	B	NK	T	CD4	CD8		B	NK	T	CD4	CD8
DURATION						**DURATION**					
Breastfeeding (n=223) increase per month	-1(-3,1)	0 (-2,2)	-1 (-2,1)	1(-2,1)	1 (-1,2)	Breastfeeding (n=150) increase per month	-2 (-4,0)	0 (-2,2)	0 (-2,1)	0 (-2,1)	0 (-2,2)
EXPOSURE						**EXPOSURE**					
Never (n= 35)	REF	REF	REF	REF	REF	Never (n=16)	REF	REF	REF	REF	REF
< 3 months (n= 86)	9 (-8,25)	4 (-16,24)	7 (-5,19)	6 (-7,19)	**19 (3,35)***	< 3 months (n=55)	15 (10,40)	21 (-5,47)	16(-3,35)	13 (-6,32)	19 (-5,44)
≥ 3 < 6 months (n= 64)	6 (-12,23)	0 (-21,22)	8 (-5,21)	7 (-7,22)	**20 (3,37)***	≥ 3 < 6 months (n=40)	26 (0,52)	-2 (-28,25)	20 (0,39)	18 (-2,38)	25 (0,50)
≥ 6 months (n= 73)	0 (-18,18)	9 (-13,31)	2 (-11,16)	0 (-14,15)	16 (-2,33)	≥ 6 months (n=55)	2 (-22,27)	9 (-16,34)	13 (-5,31)	13 (-6,31)	13 (-11,36)

Values are log transformed regression coefficients (95% confidence interval of cell numbers (*10^{e9}/L)) derived from linear regression models and reflect the increase or decrease (%) in log transformed cell numbers per month continuation of breastfeeding.(duration), and the increase or decrease (%) in log transformed cell numbers of groups exposed to breastfeeding relative to the reference group (never breastfeeding). Breastfed groups include both partially and exclusively breastfed children. *P-value <0.05.Adjusted for maternal education and maternal alcohol use during pregnancy.

Table 3. Adjusted associations between breastfeeding and memory B-cell and T-cell subsets at 6 months

	Regression coefficients for logtransformed cell numbers (*10e9/L)					
	B (n=258)					
	Naive B CD27- IgD+	CD27+ IgM+	CD27+ IgA+	CD27- IgA+	CD27+ IgG+	CD27- IgG+
DURATION						
Breastfeeding (n=223) increase per month	-1(-3,1)	**-3(-6,-1)***	**-2 (-5,-1)***	-2 (-5,1)	1 (-3,2)	**-2 (-4,-1)***
EXPOSURE						
Never (n= 35)	REF	REF	REF	REF	REF	REF
< 3 months (n= 86)	12 (-6,31)	9(-19,36)	1 (-28,30)	15 (-14,43)	-9 (-34,16)	-1 (-22,22)
≥ 3 < 6 months (n= 64)	18 (-1,38)	18(-11,48)	1 (-30,32)	14 (-17,44)	0 (-26,27)	1 (-22,25)
≥ 6 months (n= 73)	6 (-13,26)	-4(-34,25)	-11 (-40,19)	-2 (-31,28)	-7 (-33,19)	-13 (-36,10)

Values are log transformed regression coefficients (95% confidence interval of cell numbers (*10^{e9}/L)) derived from linear regression models and reflect the increase or decrease (%) in log transformed cell numbers per month continuation of breastfeeding.(duration), and the increase or decrease (%) in log transformed cell numbers of groups

4 and 5). Partial breastfeeding until 4 months was associated with a 20% increase in total CD8 T cells, and remained 21% higher in infants who were exclusively breastfed (Table 5). At 14 and 25 months of age, similar tendencies for breastfeeding exposure and CD8+ T cell numbers were observed, however the associations were not significant. (Table 2, S2 Table).

Within the CD8 T-cell compartment at 6 months of age, central memory cell numbers (CD45RO+CCR7+) were 30% (95% CI 4, 57) higher in children who were breastfed until 3 months than in non-breastfed children (Table 3, Fig. 2b). This effect was irrespective of

Table 4. Unadjusted associations between breastfeeding exclusivity and $CD8^+$ T cell numbers at 6 months

Regression coefficients for logtransformed cell numbers (*10e9/L)					
6 months (n=258)					
	CD8 Total	CD8 Tnaive	CD8 Tcm	CD8 TemRA	CD8 TemRO
EXCLUSIVENESS					
Never (n= 35)	REF	REF	REF	REF	REF
Partial until 4 months (n=157)	13 (-3,29)	9 (-11,29)	17 (-7,41)	-5 (-30,21)	3 (-27,32)
Exclusive until 4 months(n=80	14 (-3,31)	3 (-19,25)	**33 (7,59)***	12 (-16,40)	28 (-3,59)

Unadjusted values are log transformed regression coefficients (95% confidence interval of cell numbers (*10^{e9}/L)) derived from linear regression models and reflect the increase or decrease (%) in
log transformed cell numbers of partial and exclusive breastfed groups relative to the reference group (never breast-feeding). Partial breastfeeding indicates infants receiving both breast-feeding,
bottle feeding and/or solids in this period.*P-value <0.05.

T CD4+ (n=258)				T CD8+ (n=258)			
CD4+ naive	CD4+ cm	CD4+ TemRO	CD4+ TemRA	CD8+ T naive	CD8+ T cm	CD8+ TemRO	CD8+ emRA
-2 (-3,0)	-1(-2,1)	-1(-3,2)	0(-3,2)	-2 (-4,1)	0(-2,3)	0(-3,3)	2 (-1,5)
REF	REF	REF	REF	REF	REF	REF	REF
7(-10,24)	-3(-20,14)	17(-4,38)	14(-8,37)	23 (0,46)	**30 (4,57)***	15(-18,47)	20(-8,47)
4(-14,22)	-5(-23,13)	9(-14,31)	8(-16,32)	23 (-2,448)	17 (-11,45)	-4(-38,31)	9 (-20,39)
-5(-23,14)	-3(-21,15)	13(-10,35)	10(-14,34)	10(-15,34)	**29 (1,57)***	-5(-39,30)	21(-9,51)

exposed to breastfeeding relative to the reference group (never breastfeeding). Breastfed groups include both partially and exclusively breastfed children. *P-value <0.05.Adjusted for maternal education and maternal alcohol use during pregnancy.

duration; central memory cell numbers were still 29% (95% CI 1, 57) higher in infants who were breastfed for 6 months or longer than in infants who were never breastfed (Table 3, Fig 2b). In line with this, no significant associations were found between the duration of breastfeeding and CD8 central memory numbers (Table 3, Fig. 2c).

Thus, CD8 central memory T cells were associated with breastfeeding exposure, but not with the duration of breastfeeding.

Table 5. Adjusted associations between breastfeeding exclusivity and CD8$^+$ T cell numbers at 6 months

Regression coefficients for logtransformed cell numbers (*10e9/L)					
6 months (n=258)					
	CD8 Total	CD8 Tnaive	CD8 Tcm	CD8 TemRA	CD8 TemRO
EXCLUSIVENESS					
Never (n= 35)	REF	REF	REF	REF	REF
Partial until 4 months (n=143)	**20 (3, 38)***	20 (-5, 44)	12 (-17,40)	18 (-11, 47)	4 (-31, 39)
Exclusive until 4 months (80)	**21 (2, 40)***	14 (-14, 41)	16 (-15, 48)	27 (-5, 59)	12 (-28, 51)

Values are log transformed regression coefficients (95% confidence interval of cell numbers (*10^{e9}/L)) derived from linear regression models and reflect the increase or decrease (%) in
log transformed cell numbers of partial and exclusive breastfed groups relative to the reference group (never breastfeeding). Partial breastfeeding indicates infants receiving both breast-feeding,
bottle feeding and/or solids in this period.*P-value <0.05. Adjusted for maternal education and maternal alcohol use during pregnancy.

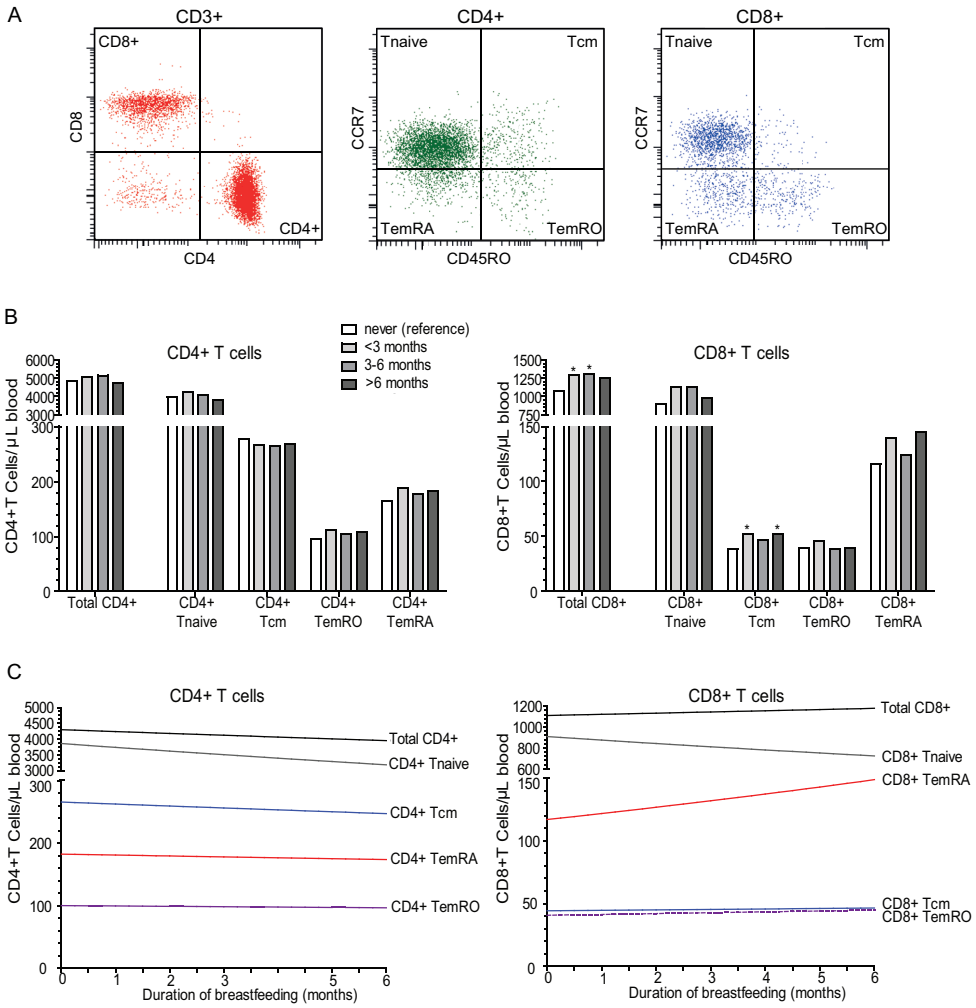


Figure 2. Impact of breastfeeding duration on T-lymphocyte subsets at 6 months of age.
A) Gating strategy for T-cell subset delineation. $CD4^+$ and $CD8^+$ T-cell subsets were defined within total $CD3^+$ T cells, and subsequently dissected using CD45RO and CCR7 into naive, central memory, $CD45RO^+$ effector memory (TemRO) and $CD45RO^-$ effector memory (TemRA) T cells. The reference depict density plots of total lymphocytes and were used to set the gates accordingly. **B)** Backtransformed CD4 and CD8 cell counts (cells/µL) at 6 months of age according to different breastfeeding duration categories (reference category is never). Categories of breastfed children contain both partial and exclusively breasted children. **C)** The estimated backtransformed regression line reflects CD4 and CD8 cell counts (cells/µL) at 6 months of age, per month increase in breastfeeding duration. *, P<0.05.

DISCUSSION

Overall we found that breastfeeding was associated with a decrease in B-cell memory and an increase in CD8 T-cell memory at 6 months of age. The same trend was observed at 14 months of age, and disappeared at older ages. The decrease in B-cell memory was associ-

ated with the duration of breastfeeding, while the increase in total CD8 T cells and central memory CD8 T cells did not depend on the duration of breastfeeding. This suggests that breastfeeding enhances T-cell maturation, but not B-cell maturation.

Total B-cell numbers nor frequencies were significantly related with breastfeeding. This is in line with previous studies.[21,23] Only one study reported higher frequencies of total B cells in breastfed children at 6 months of age, however sample sizes were small (n=7 breastfed infants).[22] We found that the duration of breastfeeding was associated with decreased numbers of CD27+IgA+, CD27+IgM+ and CD27-IgG+ memory B cells, which are mostly derived from systemic T-cell dependent responses.[11] This suggests that continuous breastfeeding inhibits memory B-cell development. One of the candidate factors in breast milk that is likely to mediate this, is secretory IgA (sIgA).[33,34] Continuous breastfeeding will provide a constant supply of maternal sIgA onto the epithelial surface of the infant's gastrointestinal tract. This sIgA might prevent exposure of microorganisms to the infant's humoral immune system, and translocation of gut bacteria.[34,35] Thus, fewer naive B cells might be activated to differentiate into memory B cells (Supplemental Fig 1). Indeed, in suckling mice it has been suggested that maternal sIgA blocks mucosal B cell responses in the offspring.[36,37] Alternatively, the unchanged numbers of T-cell independent CD27-IgA+ B cells suggest that breastfeeding does not affect local IgA responses. It is therefore possible that maternal IgA helps to block translocation of intestinal bacteria, thereby only preventing systemic T-cell dependent B-cell memory formation. Other growth factors in breast milk that can reduce exposure of microbiota to B cells are epidermal growth factor, IGF-1, TGF-β, leptin and prolactin. These factors enhance maturation of the epithelial barrier,[33,38] and decrease uptake of foreign protein antigens. Factors such as lactoferrin, oligosaccharides and lipids may directly prevent attachment of the bacterial outer membrane to the mucosal surface.[33] Finally, the passive transfer of functional Ig-secreting plasma cells in breast milk may prevent bacterial or viral transmission.[39,40]

Total T-cell numbers nor CD4 T-cell numbers were associated with breastfeeding exposure or breastfeeding duration. However, infants who were breastfed until 3 months had higher CD8 T-cell numbers than infants who were never breastfed, and this change persisted when breastfeeding was prolonged until 6 months, suggesting an ongoing activation of CD8 T cells by exposure to breast milk, but no accumulation over time. These results extend previous observations in small studies (n<40) of increased frequencies of CD8 T cells or decreased CD4/CD8 ratios.[21,22,24] In addition, an increase in CD8 T cells from 8 to 10 months of age was observed before,[24] suggesting a longer-lasting effect of breastfeeding. However, the children received breastfeeding until 8 months of age.[24] In addition, the number of children in the breastfeeding group was small (n=35), and no detailed analysis of CD8 subsets was performed. Therefore, future studies will be needed to validate these findings.

Within total CD8 T cells, the central memory subset was most significantly increased in breastfed children. This expansion will be the result of stimulation of mature T cells.

Candidate immune stimulatory factors in breast milk are lactoferrin and exosomes. Lactoferrin has direct microbicidal properties, including binding of the bacterial cell wall and the concomitant release of lipopolysaccharide (LPS, endotoxin). Moreover, lactoferrin is known to enhance T-cell proliferation.[21,33] Exosomes are carrier vesicles of 50-100 nm that can bud from the membrane of eukaryotic cells.[41] Exosomes formed by B cells or dendritic cells contain MHC class I and class II, and have the potential to stimulate T cells.[42,43] Human breast milk has been found to contain immune modulatory exosomes that express MHC molecules, IL-2, IFNγ and TNFα. In line with our findings that breastfeeding did not affect memory CD4 T cells, these vesicles did not appear to stimulate CD4 T cells.[44] It is, however, likely that immunostimulatory compounds such as exosomes, do have CD8 T-cell stimulating capacities and contribute to the increase in central memory T cells in breastfed infants (S1 Fig.). In addition, it has been described that the infants take up live and functional maternal immune cells from the breast milk.[45-47] This might result in stimulation of CD8 T cells. Thus, it remains unclear whether the increase in central memory CD8 T cells can be attributed to immunostimulatory factors in breast milk, such as lactoferrin and exosomes, or reflects a role for maternal immune cells. Nevertheless, central memory T cells might mediate reactive immunity,[48] because they circulate between the spleen, blood and lymph nodes and proliferate extensively in response to a second encounter of an antigen.[13,14,15] Therefore, it could be hypothesized that an increase in central memory T cells is associated with increased reactive immunity.

Beneficial effects of breastfeeding on adaptive immune responses have previously been demonstrated in vaccination studies.[7,8,9] Breasted children showed increased interferon-γ production, and increased frequencies of CD8+ T cells after vaccination with measles, mumps and rubella.[9] Furthermore, breastfeeding had beneficial effects on virus-specific immune responses to poliovirus, diphtheria toxoid and tetanus toxoid,[49] whereas the responses to rotavirus are not clearly enhanced.[50,51] These studies included relatively few children, and confounding factors could not be taken into account. Thus, our results extend these previous observations from vaccination studies. Unfortunately, we were not able to study virus-specific memory cells and studies determining associations between central memory cell numbers and functional immunity are lacking. Thus, interpretations regarding cell-mediated immunity should be made with caution.

We did not observe associations between breastfeeding and naive T cells or effector memory T cells. In contrast, a previous study reported lower frequencies of naive CD4 T cells [[21]], but no differences in memory T-cell frequencies. Therefore, it was suggested that the adaptive immune system develops slower in breastfed infants.[21] However, the study reported on a relatively small number of children in the breastfeeding group (n=34), and lacked data on absolute B- and T-cell counts.[21]

Our study was conducted in a large population-based prospective birth cohort. Previous studies that addressed the effects of breastfeeding on adaptive immunity, had smaller sample size and/or lacked follow-up.[21,22,23,24] Most of these retrospective studies were

based on recall of infant feeding after several years, making recall bias of feeding habits an important concern.[52] Because of the prospective design of our study, detailed information on the duration of breastfeeding was collected at multiple time points shortly after breastfeeding was finished, thereby limiting potential recall bias.[52] Still, data on breastfeeding was collected retrospectively, and due to the use of questionnaires, misclassification may occur.[52] Because any misclassification would be independent of laboratory determined T and B cell numbers, it is unlikely that these affected the outcome. In addition, our study design provided information on a large number of potential confounders. We used an unbiased approach, investigating a broad panel of determinants on lymphocyte numbers and frequencies. Moreover, we performed measurements of lymphocytes at different ages, enabling us to study the effect of breastfeeding on adaptive immunity over a longer period of time. Finally, using detailed 6-color flow cytometry, we were able to discriminate multiple, functionally distinct B- and T-cell subsets.[11,12,53] Thus, we could evaluate the effects of breastfeeding on specific aspects of adaptive immunity.

At 14 months or age, we observed the same trends of decreasing B-cell numbers with longer breastfeeding duration, and higher CD8 T-cell numbers with breastfeeding exposure, as we did for 6 months, although effects were not significant. The sample size at 14 months was smaller than the sample size at 6 months. Hence, we cannot exclude that non-significant effects at 14 months are due to loss of statistical power, and some effects of breastfeeding remain at this age.

Our measurements were limited to peripheral blood. Thus, it is not possible to deduce whether B- and T-cell numbers in lymphoid tissue were affected. Theoretically, preferential homing to or away from tissue could result in changes in blood lymphocyte counts [54]. If breastfeeding affects this preferential homing, the effects will be lost once breastfeeding is discontinued. Nearly all children were no longer receiving breast milk and still showed lower memory B cell numbers and higher CD8 T-cell numbers. Thus, we conclude that the effects of breastfeeding on preferential tissue homing was limited.

In conclusion, this prospective population-based cohort study among a large number of healthy children showed that CD27+IgA+, CD27+IgM+ and CD27-IgG+ memory B-cell numbers decreased with a longer breastfeeding duration. CD8 T cells, and especially CD8 central memory T-cell numbers, were higher in breastfed children up to 6 months of age. This suggests that breastfeeding enhances T-cell maturation in the first 6 months of life. On top of the protective effects of maternal IgA in breast milk, this might contribute to the protective effect against infectious diseases in infancy.

REFERENCES

1. Lonnerdal B. Bioactive proteins in breast milk. J Paediatr Child Health. 2013;49 Suppl 1:1-7.
2. Duijts L, Ramadhani MK, Moll HA. Breastfeeding protects against infectious diseases during infancy in industrialized countries. A systematic review. Matern Child Nutr. 2009;5(3):199-210.
3. Walker A. Breast milk as the gold standard for protective nutrients. J Pediatr. 2010;156(2 Suppl):S3-7.
4. Zinkernagel RM. Maternal antibodies, childhood infections, and autoimmune diseases. N Engl J Med. 2001;345(18):1331-5.
5. Wilson AC, Forsyth JS, Greene SA, Irvine L, Hau C, Howie PW. Relation of infant diet to childhood health: seven year follow up of cohort of children in Dundee infant feeding study. BMJ. 1998;316(7124):21-5.
6. Silfverdal SA, Bodin L, Hugosson S, Garpenholt O, Werner B, Esbjorner E, et al. Protective effect of breastfeeding on invasive Haemophilus influenzae infection: a case-control study in Swedish preschool children. Int J Epidemiol. 1997;26(2):443-50.
7. Pabst HF, Godel J, Grace M, Cho H, Spady DW. Effect of breast-feeding on immune response to BCG vaccination. Lancet. 1989;1(8633):295-7
8. Pabst HF, Spady DW. Effect of breast-feeding on antibody response to conjugate vaccine. Lancet. 1990;336(8710):269-70.
9. Pabst HF, Spady DW, Pilarski LM, Carson MM, Beeler JA, Krezolek MP. Differential modulation of the immune response by breast- or formula-feeding of infants. Acta Paediatr. 1997;86(12):1291-7.
10. Hanson LA, Korotkova M, Lundin S, Haversen L, Silfverdal SA, Mattsby-Baltzer I, et al. The transfer of immunity from mother to child. Ann N Y Acad Sci. 2003;987:199-206.
11. Berkowska MA, Driessen GJ, Bikos V, Grosserichter-Wagener C, Stamatopoulos K, Cerutti A, et al. Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways. Blood. 2011.
12. Appay V, van Lier RA, Sallusto F, Roederer M. Phenotype and function of human T lymphocyte subsets: consensus and issues. Cytometry A. 2008;73(11):975-83.
13. Mueller SN, Gebhardt T, Carbone FR, Heath WR. Memory T cell subsets, migration patterns, and tissue residence. Annu Rev Immunol. 2013;31:137-61.
14. Mackay CR. Dual personality of memory T cells. Nature. 1999;401(6754):659-60.
15. Pepper M, Jenkins MK. Origins of CD4(+) effector and central memory T cells. Nat Immunol. 2011;12(6):467-71.
16. Comans-Bitter WM, de Groot R, van den Beemd R, Neijens HJ, Hop WC, Groeneveld K, et al. Immunophenotyping of blood lymphocytes in childhood. Reference values for lymphocyte subpopulations. J Pediatr. 1997;130(3):388-93.
17. Driessen GJ, Dalm VA, van Hagen PM, Grashoff HA, Hartwig NG, van Rossum AM, et al. Common variable immunodeficiency and idiopathic primary hypogammaglobulinemia: two different conditions within the same disease spectrum. Haematologica. 2013;98(10):1617-23.
18. Shearer WT, Rosenblatt HM, Gelman RS, Oyomopito R, Plaeger S, Stiehm ER, et al. Lymphocyte subsets in healthy children from birth through 18 years of age: the Pediatric AIDS Clinical Trials Group P1009 study. J Allergy Clin Immunol. 2003;112(5):973-80.
19. van Gent R, van Tilburg CM, Nibbelke EE, Otto SA, Gaiser JF, Janssens-Korpela PL, et al. Refined characterization and reference values of the pediatric T- and B-cell compartments. Clin Immunol. 2009;133(1):95-107.
20. Tosato F, Bucciol G, Pantano G, Putti MC, Sanzari MC, Basso G, et al. Lymphocytes subsets reference values in childhood. Cytometry A. 2014. 21.Andersson Y, Hammarstrom ML, Lonnerdal B, Graverholt G, Falt H, Hernell O. Formula feeding skews immune cell composition toward adaptive immunity compared to breastfeeding. J Immunol. 2009;183(7):4322-8.

22. Carver JD, Pimentel B, Wiener DA, Lowell NE, Barness LA. Infant feeding effects on flow cytometric analysis of blood. J Clin Lab Anal. 1991;5(1):54-6.
23. Hawkes JS, Neumann MA, Gibson RA. The effect of breast feeding on lymphocyte subpopulations in healthy term infants at 6 months of age. Pediatr Res. 1999;45(5 Pt 1):648-51.
24. Jeppesen DL, Hasselbalch H, Lisse IM, Ersboll AK, Engelmann MD. T-lymphocyte subsets, thymic size and breastfeeding in infancy. Pediatr Allergy Immunol. 2004;15(2):127-32.
25. Jaddoe VW, van Duijn CM, Franco OH, van der Heijden AJ, van Iizendoorn MH, de Jongste JC, et al. The Generation R Study: design and cohort update 2012. Eur J Epidemiol. 2012;27(9):739-56.
26. van Rossem L, Oenema A, Steegers EA, Moll HA, Jaddoe VW, Hofman A, et al. Are starting and continuing breastfeeding related to educational background? The generation R study. Pediatrics. 2009;123(6):e1017-27.
27. Kiefte-de Jong JC, Escher JC, Arends LR, Jaddoe VW, Hofman A, Raat H, et al. Infant nutritional factors and functional constipation in childhood: the Generation R study. Am J Gastroenterol. 2010;105(4):940-5.
28. Statistics N. Dutch Standard Classification of Education 2006. Voorburg/Heerlen, Statistics Netherlands 2008
29. Statistics Netherlands. Welfare in the Netherlands. Income, welfare and spending of households and individuals. Den Haag/Heerlen: Statistics Netherlands, 2012.
30. Mickey RM, Greenland S. The impact of confounder selection criteria on effect estimation. Am J Epidemiol. 1989;129(1):125-37.
31. Bender R, Lange S. Adjusting for multiple testing--when and how? J Clin Epidemiol. 2001;54(4):343-9.
32. Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology. 1990;1(1):43-6.
33. M'Rabet L, Vos AP, Boehm G, Garssen J. Breast-feeding and its role in early development of the immune system in infants: consequences for health later in life. J Nutr. 2008;138(9):1782S-90S.
34. Pabst O. New concepts in the generation and functions of IgA. Nat Rev Immunol. 2012;12(12):821-32.
35. Wold AE, Adlerberth I. Does breastfeeding affect the infant's immune responsiveness? Acta Paediatr. 1998;87(1):19-22.
36. Kramer DR, Cebra JJ. Role of maternal antibody in the induction of virus specific and bystander IgA responses in Peyer's patches of suckling mice. Int Immunol. 1995;7(6):911-8.
37. Kramer DR, Cebra JJ. Early appearance of "natural" mucosal IgA responses and germinal centers in suckling mice developing in the absence of maternal antibodies. J Immunol. 1995;154(5):2051-62.
38. Lawrence RM, Pane CA. Human breast milk: current concepts of immunology and infectious diseases. Curr Probl Pediatr Adolesc Health Care. 2007;37(1):7-36.
39. Hanson LA, Silfverdal SA, Korotkova M, Erling V, Strombeck L, Olcen P, et al. Immune system modulation by human milk. Adv Exp Med Biol. 2002;503:99-106.
40. Tuaillon E, Valea D, Becquart P, Al Tabaa Y, Meda N, Bollore K, et al. Human milk-derived B cells: a highly activated switched memory cell population primed to secrete antibodies. J Immunol. 2009;182(11):7155-62.
41. Thery C, Ostrowski M, Segura E. Membrane vesicles as conveyors of immune responses. Nat Rev Immunol. 2009;9(8):581-93.
42. Raposo G, Nijman HW, Stoorvogel W, Liejendekker R, Harding CV, Melief CJ, et al. B lymphocytes secrete antigen-presenting vesicles. J Exp Med. 1996;183(3):1161-72.
43. Zitvogel L, Regnault A, Lozier A, Wolfers J, Flament C, Tenza D, et al. Eradication of established murine tumors using a novel cell-free vaccine: dendritic cell-derived exosomes. Nat Med. 1998;4(5):594-600.
44. Admyre C, Bohle B, Johansson SM, Focke-Tejkl M, Valenta R, Scheynius A, et al. B cell-derived exosomes can present allergen peptides and activate allergen-specific T cells to proliferate and produce TH2-like cytokines. J Allergy Clin Immunol. 2007;120(6):1418-24.

45. Jain L, Vidyasagar D, Xanthou M, Ghai V, Shimada S, Blend M. In vivo distribution of human milk leucocytes after ingestion by newborn baboons. Arch Dis Child. 1989;64(7 Spec No):930-3.
46. Schlesinger JJ, Covelli HD. Evidence for transmission of lymphocyte responses to tuberculin by breast-feeding. Lancet. 1977;2(8037):529-32.
47. Mohr JA. Lumphocyte sensitisation passed to the child from the mother. Lancet. 1972;1(7752):688.
48. Sallusto F, Geginat J, Lanzavecchia A. Central memory and effector memory T cell subsets: function, generation, and maintenance. Annu Rev Immunol. 2004;22:745-63.
49. Hahn-Zoric M, Fulconis F, Minoli I, Moro G, Carlsson B, Bottiger M, et al. Antibody responses to parenteral and oral vaccines are impaired by conventional and low protein formulas as compared to breast-feeding. Acta Paediatr Scand. 1990;79(12):1137-42.
50. Groome MJ, Moon SS, Velasquez D, Jones S, Koen A, van Niekerk N, et al. Effect of breastfeeding on immunogenicity of oral live-attenuated human rotavirus vaccine: a randomized trial in HIV-uninfected infants in Soweto, South Africa. Bull World Health Organ. 2014;92(4):238-45.
51. Rennels MB. Influence of breast-feeding and oral poliovirus vaccine on the immunogenicity and efficacy of rotavirus vaccines. J Infect Dis. 1996;174 Suppl 1:S107-11.
52. Li R, Scanlon KS, Serdula MK. The validity and reliability of maternal recall of breastfeeding practice. Nutr Rev. 2005;63(4):103-10.
53. Freud AG, Caligiuri MA. Human natural killer cell development. Immunol Rev. 2006;214:56-72.
54. Bunders MJ, van der Loos CM, Klarenbeek PL, van Hamme JL, Boer K, Wilde JC, et al. Memory CD4(+) CCR5(+) T cells are abundantly present in the gut of newborn infants to facilitate mother-to-child transmission of HIV-1. Blood. 2012;120(22):4383-90.

Supplemental Table 1. Maternal and infant characteristics of the study population at 14, 25 months and 6 years.

	14 months (n=166)		25 months (n=112)		6 years (n=332)	
	No breastfed (n=16)	**Breastfed (n=150)**	**No breastfed (n=16)**	**Breastfed (n=96)**	**No breastfed (n=37)**	**Breastfed (n=295)**
Maternal characteristics						
Age (Mean ± SD; years)	31 (4.2)	32 (3.9)	31 (4.1)	32 (3.8)	31 (3.6)	32 (3.6)
Educational level (n; %)						
Lower	**7 (44%)**	**59 (40%)**	**12 (75%)**	**31 (33%)**	**22 (59%)**	**96 (33%)**
Higher	**9 (56%)**	**90 (60%)**	**4 (25%)**	**64 (67%)***	**15 (41%)**	**196 (67%)***
Net household income per month (n; %)						
< € 2400	0 (0%)	21 (15%)	1 (8%)	13 (15%)	4 (12%)	33 (13%)
≥ € 2400	13 (100%)	112 (84%)	12 (92%)	73 (85%)	30 (88%)	227 (87%)
Smoking continued during pregnancy (n; %)	2 (14%)	16 (15%)	1 (7%)	8 (11%)	6 (20%)	32 (15%)
Alcohol use continued during pregnancy (n; %)	6 (43%)	41 (38%)	4 (29%)	29 (39%)	11 (37%)	83 (37%)
Body Mass Index before pregnancy (Mean ± SD; kg/m^2)	25 (5)	24 (5)	25 (5)	23 (4)	24 (5)	24 (4)
Fever in third trimester of pregnancy (n; %)	1(6%)	9 (6%)	2 (13%)	5 (6%)	2 (6%)	19 (7%)
Maternal atopy (eczema, allergy HDM, hay-fever)(n; %)	3 (21%)	51 (36%)	6 (40%)	41 (46%)	9 (27%)	109 (40%)
Paternal atopy (eczema, allergy HDM, hay-fever) (n; %)	3 (23%)	41 (29%)	5 (39%)	24 (28%)	13 (39%)	77 (29%)
Family history of asthma / atopy (n; %)	5 (31%)	77 (52%)	9 (56%)	56 (59%)	20 (54%)	156 (54%)
Any reported autoimmune disease (diabetes mellitus, SLE, arthritis, MS, thyroid disorder, or celiac disease) (n; %)	0 (0%)	3 (2%)	0 (0%)	2 (2%)	2 (5%)	8 (3%)
Mode of delivery (n; %)						
Vaginal	9 (56%)	19 (64%)	7 (47%)	48 (55%)	20 (59%)	182 (65%)
Forceps or vacuum assisted	2 (13%)	30 (21%)	2 (13%)	21 (24%)	5 (15%)	59 (21%)
Caesarian section	5(31%)	21 (15%)	6 (40%)	19 (22%)	9 (26%)	40 (14%)
Premature rupture of membranes (n; %)	1 (6%)	2 (1%)	1 (6%)	2 (2%)	3 (8%)	4 (2%)
Infant characteristics (n=258)						
Males (n; %)	9 (56%)	77 (51%)	7 (44%)	47 (49%)	20 (54%)	144 (49%)
Gestational age (Mean ± SD; weeks)	39.2 (2.2)	40.1 (1.9)	39.7 (1.9)	39.9(1.5)	39.7 (1.7)	40.1 (1.6)
Preterm birth (<37 weeks) (n; %)	1 (6%	3 (2%)	1 (6%)	4 (4%)	2 (5%)	8 (3%)
Birth weight (Mean ± SD; grams)	3,455 (738)	3,505 (513)	3,516 (614)	3,542 (480)	3,543 (554)	3,531 (508)
Apgar score at 5 min <7 (n; %)	0 (0%)	3 (2%)	0 (0%)	0 (0%)	0 (0%)	3 (1%)
Birth season (n; %)						

Supplemental Table 1. Maternal and infant characteristics of the study population at 14, 25 months and 6 years. (continued)

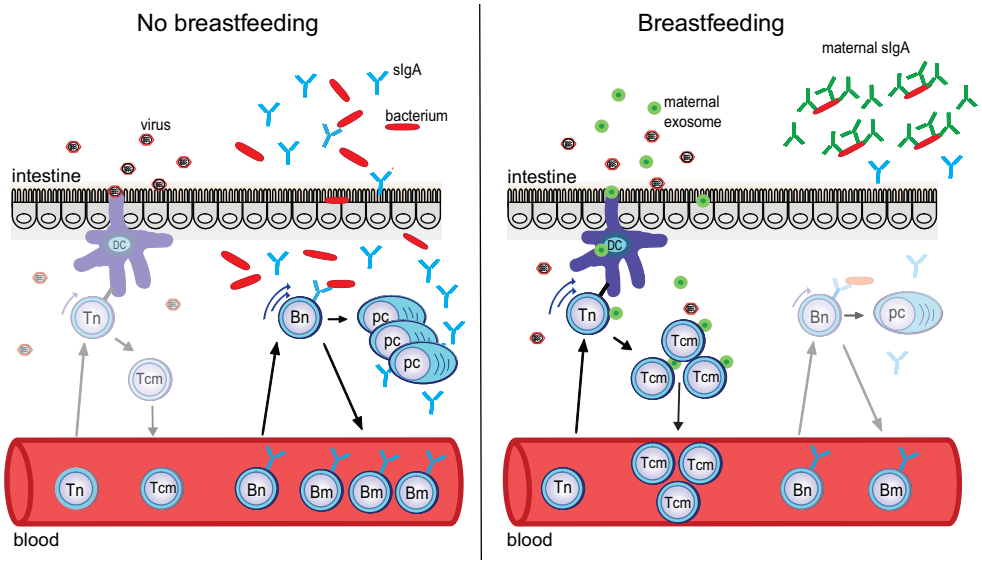
	14 months (n=166)		25 months (n=112)		6 years (n=332)	
	No breastfed (n=16)	Breastfed (n=150)	No breastfed (n=16)	Breastfed (n=96)	No breastfed (n=37)	Breastfed (n=295)
Winter (Dec-Jan-Feb)	2 (13%)	38 (25%)	3 (19%)	28 (29%)	6 (16%)	73 (25%)
Spring (Mar-Apr-May)	5 (31%)	41 (27%)	3 (19%)	26 (27%)	6 (16%)	93 (31%)
Summer (Jun-Jul-Aug)	3 (18%)	32 (21%)	4 (25%)	21 (22%)	11 (30%)	65 (22%)
Autumn (Sept-Oct-Nov)	6 (38%)	39 (26)	6 (37%)	21 (22%)	14 (38%)	64 (22%)
Siblings ≥1 (n; %)	2 (12%)	17 (11%)	1 (6%)	8 (8%)	2 (5%)	26 (9%)
Day-care >16 hours /week (n; %)	3 (33%)	51 (47%)	4 (25%)	33 (34%)	5 (33%)	99 (47%)
Fever in first 6 months (n; %)	7 (78%)	80 (63%)	3 (60%)	51 (62%)	13 (57%)	175 (67%)
Age at focus visit (Median ± range; months/years)	14.4 (13.2-16.3)	14.4 (13.1-17.4)	25.3 (23.4-27.5)	25.2 (23.3-29.8)	6.0 (5.8-6.6)	5.8 (5.1-7.2)

Values are means (SD), absolute numbers (percentages) or #medians (90% range). * Significantly different between breastfeeding and no breastfeeding groups.

Supplemental Table 3. Antibody panel used for 6-color flow cytometry.

Labeling	Conjugated monoclonal antibodies (clone)					
	FITC	**PE**	**PerCP-Cy5.5**	**PE-Cy7**	**APC**	**APC-Cy7**
1	CD3	CD16.56	CD45	CD4	CD19	CD8
	(SK7)	(B73.1C5.9)	(2D1)	(SK3)	(SJ25C1)	(SK1)
2	IgD	CD23	CD19	CD21	IgM	CD27
	(poly)	(EBVCS5)	(SJ25C1)	(B-ly-4)	(G20-127)	(M-T271)
3	IgA	IgG	CD19	CD40	IgM	CD27
	(poly)	(poly)	(SJ25C1)	(5C3)	(G20-127)	(M-T271)
4	CD28	CD197	CD3	CD8	CD45RO	CD27
	(CD28.2)	(3D13)	(SK7)	(SK1)	(UCHL-1)	(M-T271)

FITC= fluorescein isothiocyanate, PE=phycoerythrin, PerCPCy5.5= peridin chlorophyll protein, PE-Cy7= phycoerythrin-cyanin dye, APC=allophycocyanin and APC-Cy7= allophycocyanin-cyanin dye, poly = polyclonal antibody.



Supplemental Figure 1. Summarizing mechanism of how breastfeeding might affect adaptive memory. In absence of breast milk, the infant's B and T cells respond to microorganisms in the intestine and generate long-lived memory cells and IgA (blue) that circulate through the body (left). Breast milk contains immune modulating components (right). Of these, maternal sIgA (green) is able to catch microorganisms and prevent recognition of these by B-cells. This might inhibit B-cell responses and B-cell memory formation. Other immunostimulatory components, such as exosomes, might stimulate naive T cells and increase T-cell memory formation. Abbreviations: Bn, naïve B cell; Bm, memory B cell; DC, dendritic cell; pc,plasma cell; Tn, naive T cell; Tm, memory T cell.

Supplemental Table 2. Adjusted associations between breastfeeding and cell numbers at 25 months and 6 years

	Logtransformed cell numbers (*10e9/L)								
	25 months (n=112)					6 years (n=332)			
	B	NK	T	CD4	CD8		B	NK	T
DURATION						**DURATION**			
Breastfeeding (n=96) increase per month	-1 (-4,1)	0 (-3,3)	0 (-2,2)	1 (-2,3)	-2 (-4,1)	Breastfeeding (n=295) increase per month	-1(-2,1)	0(-2,1)	1(-1,2)
EXPOSURE						**EXPOSURE**			
Never (n= 16)	REF	REF	REF	REF	REF	Never (n=37)	REF	REF	REF
< 3 months (n=29)	14 (-10,38)	27 (-3, 58)	14 (-7,35)	11 (-10,32)	18 (-9,45)	< 3 months (n=112)	-6(-19,8)	11(-5,26)	1(-10,12)
≥ 3 < 6 months (n=30)	0 (-23,23)	7 (-23, 37)	-2 (-22,18)	-4 (-24,17)	1 (-25,27)	≥ 3 < 6 months (n=76)	-5(-19,10)	7(-10,24)	2(-10,14)
≥ 6 months (n=37)	11 (-14,35)	11 (-20,42)	12 (-9,33)	18 (-3,30)	-3 (-30, 24)	≥ 6 months (n=107)	-12(-26,3)	-1(-17,16)	3(-8,15)
Never (n=15)	REF	REF	REF	REF	REF	Never (n=52)	REF	REF	REF
≤ 6 months (n=50)	-17 (-35, 2)	3 (-22,27)	-7 (-23,10)	-11 (-28,6)	3 (-18,25)	≤ 6 months (n=150)	3(-8,13)	6(-6,18)	1(-8,10)
6-9 months (n=18)	-4 (-29,21)	1 (-33,34)	-2 (-24,21)	0 (-23,23)	-2 (-31,28)	6-9 months (n=50)	5(-10,20)	2(-15,19)	1(-11,13)
≥ 9 months (n=15)	3 (-26, 32)	-17 (-55, 21)	9 (-17,35)	10 (-17,36)	8 (-26,41)	≥ 9 months (n=34)	-3(-20,14	-6(-25,14)	1(-13,15)

Values are log transformed regression coefficients (95% confidence interval of cell numbers (*10^{e9}/L)) derived from linear regression models and reflect the increase or decrease (%) in log transformed cell numbers per months continuation of breastfeeding (duration), and the increase or decrease (%) in log transformed cell numbers of groups exposed to breastfeeding relative to the reference group (never breastfeeding). Missing in in categories 0-6-9 months (n=14 at 25 months, n=46 at 6 years). Breastfed groups include both partially and exclusively breastfed children. Adjusted for maternal education and maternal alcohol use during pregnancy. *P-value <0.05.

Chapter 4

Cytomegalovirus- and Epstein-Barr virus induced T-cell expansions in young children do not impair naive T-cell populations or vaccination responses:

The Generation R Study

Diana van den Heuvel, Michelle A.E. Jansen, Wim A. Dik, Halima Bouallouch-Charif, Dan Zhao, Kevin A.M. van Kester, Marja A.W. Smits-te Nijenhuis, Marion J. Kolijn-Couwenberg, Vincent W.V. Jaddoe, Ramon Arens, Jacques J.M. van Dongen, Henriette A. Moll, Menno C. van Zelm

J Infect Dis. 2016 Jan 15;213(2):233-42

ABSTRACT

Background

Cytomegalovirus (CMV) and Epstein Barr virus (EBV) induce effector memory T-cell expansions, which are variable and potentially depend on the age at primary exposure and co-infections. We evaluated the T-cell compartment and herpesvirus infections in 6-year-old children.

Methods

T-cell subsets and immunoglobulin G (IgG) seropositivity for CMV, EBV, herpes simplex virus-1 (HSV-1) and varicella zoster virus (VZV) were studied in 1079 6-year-old children. A random subgroup of 225 children was evaluated for CMV and EBV seropositivity before 2 years of age and for vaccination responses against measles and tetanus.

Results

CMV and EBV infections were associated with significant expansions of $CD27^-$ and $CD27^+$ effector memory T cells, respectively. These expansions were enhanced in CMV+EBV+ children and were independent of VZV or HSV-1 co-infection. Naive and central memory T-cell numbers were not affected, nor were anti-tetanus and anti-measles IgG levels. Children infected before 2 years of age showed smaller effector memory T-cell expansions than children infected between 2 and 6 years of age.

Conclusions

CMV- and EBV-related T-cell expansions do not impair naive T-cell numbers or maintenance of protective responses against non-related pathogens. Duration of infection was not directly related to larger expansions of effector memory T cells in children, suggesting that other mechanisms affect these expansions at later age.

INTRODUCTION

Cytomegalovirus (CMV) and Epstein Barr Virus (EBV) are ubiquitous in the human population and persist with presumed viral latency.[1,2] Infection occurs mainly in childhood, reaching approximately 50% seropositivity at around the age of 6 years and in 80-90% in adults within the Western world.[3,4] CMV and EBV are associated with changes in immunological memory. Infected adults display persistent expansions of virus-specific effector memory T cells in both the $CD8^+$ and the $CD4^+$ lineages.[5,6] CMV-specific $CD8^+$ T cells are predominantly $CCR7^-CD45RO^+CD27^{+/-}$ in young adults, and more clonal $CCR7^-CD45RA^+CD27^-$ populations are found in elderly.[5,7-14] EBV-specific T cells are mainly $CCR7^-CD45RO^+CD27^+$.[5,8,15,16] The majorities of these populations have an extended lifespan, but a poor response to T-cell mitogens.[7-9,16-18]

Co-infection with CMV and EBV can have both synergistic and antagonistic effects: CMV co-infection can restore defective vaccination responses in EBV-infected children, and EBV co-infection can boost CMV-induced NK-cell differentiation; together they affect the chance for developing allergic complications.[19-21] In contrast, infections with herpes simplex virus 1 (HSV-1) do not result in persistent effector T-cell expansions.[15] Less is known for co-infection with varicella zoster virus (VZV), because this already reaches 90-100% seropositivity in young adulthood in the Western world.

Accumulations of CMV- and EBV-induced effector memory T-cells have been suggested to overcrowd T cells with other specificities,[5,22] and to negatively affect immune responses to other infections and/or vaccinations.[9,17,23-25] Moreover, CMV and EBV persistence are associated with cardiovascular disease, infectious complications and with increased mortality rates, especially in immunosuppressed individuals or the very elderly.[23,26-31] Still, these effects are not consistently observed,[32-35] and despite high CMV and EBV seropositivity rates in elderly persons, clinical complications develop in only a minority. An explanation could be the variation in T-cell responses: virus-specific T-cell numbers vary from barely detectable to >30% of total memory T-cells in peripheral blood. This diversity might be affected by the infectious dose,[36] and long-term infection, and may thus especially develop in individuals who had been infected early in life. However, in contrast to infection after puberty, primary infections with CMV or EBV in early childhood are mostly asymptomatic and might even be protective against the development of celiac disease and allergies.[37,38]

To study whether the developing immune system in childhood provides more effective control of persistent viruses,[39] we studied 1079 6-year old children in the Generation R cohort. The combined analysis of CMV, EBV, HSV-1 and VZV infection in this large cohort allowed us to study virus-specific and modifying effects on the T-cell compartment and their relation with vaccination responses to tetanus and measles.

MATERIALS AND METHODS

Study subjects

This study was conducted in the context of the Generation R Study, a prospective population-based cohort study from fetal life until young adulthood.[40] We included 1079 6-year-old children (range 5.0-7.9 years) in whom CMV, EBV, HSV-1 and VZV IgG serology and detailed immunophenotyping of blood T cells was performed. From a random selection of 225 of these children, additional virus serology and immunophenotyping was performed on blood samples obtained in the second year of life (age range 13.1-29.9 months). Written informed consent was obtained from all parents of participating children. Ethical approval for the study was obtained from the Medical Ethical Committee of the Erasmus MC.

Serology

Blood plasma samples were subjected to enzyme-linked immunoassays for IgG antibodies against CMV, EBV capsid antigen, HSV-1 glycoprotein C1, VZV, tetanus toxoid or measles antigen (Euroimmun). Results were evaluated relative to a manufacturer-provided reference threshold sample. Seropositivity, and assumed virus persistence, was defined by a sample/threshold ratio above 0.6 (CMV), 0.8 (EBV capsid antigen), 1.6 (HSV-1) and 1.0 (VZV). Anti-measles and anti-tetanus IgG levels were analyzed in international units/mL (IU/mL) as a continuous variable and with plasma levels above 275 IU/mL (measles) or 0.5 IU/mL (tetanus) being defined as protective vaccination responses.

Immunophenotyping

Absolute numbers of $CD3^+$ T cells were obtained with a routine diagnostic lyse-no-wash protocol. Detailed analysis of T-cell subsets was performed with 6-color flowcytometry (Supplemental Table 1 and Supplemental Figure 1).[22,41]

Antigen-specific T-cells were detected in thawed post-Ficoll peripheral blood mononuclear cells of 15 CMV+ and 14 EBV+ children who carried the HLA-A*0201 allele, defined by single-nucleotide polymorphism tags rs2844821(G) and rs762324(C) in previously generated single-nucleotide polymorphism arrays.[40,42] Virus-specific $CD8^+$ T cells were detected with HLA-A*0201 tetramers loaded with CMV peptides NLVPMVATV (from pp65; allophycocyanin (APC)-labeled) or VLEETSVML (from IE-1; phycoerythrin (PE)-labeled) proteins and EBV peptides GLCTLVAML (from BMLF-1; APC-labeled) or YVLDHLIVV (BRLF-1; PE-labeled). Flow cytometric data were acquired on a FACSCalibur or LSRII (BD Biosciences) and analyzed using FACSDiva (BD Biosciences; version 6.2) and Infinicyte (Cytognos; version 1.7) analysis software.[43]

Statistical analyses

Differences in infection prevalence were assessed using Chi-square tests. Differences in lymphocyte counts (or frequencies) between uninfected controls and virus-infected groups

were assessed with Kruskal-Wallis tests, followed by post-hoc Dunn tests. Differences in longitudinal frequencies of lymphocyte populations were assessed using paired *t*-tests. Differences were considered statistically significant at $P<0.05$.

RESULTS

Herpesvirus seropositivity

Among the 1079 children included in the study, 36.2% were IgG-seropositive for CMV, 47.1% for EBV, 14.0% for HSV-1 and 92.1% for VZV. Because mainly CMV and EBV are reported to result in chronic T-cell expansions, we divided the 1079 children into 4 groups: no CMV or EBV infection (uninfected; n=399); CMV but no EBV infection (CMV infected; n=172); EBV but no CMV infection (EBV infection; n=289) and CMV and EBV coinfection (n=219) (Table 1).

Table 1. HSV-1 or VZV co-infection in correlation to CMV and EBV IgG seropositivity at 6 years of age.

	n	%	HSV n (% of group)	VZV n (% of group)	HSV and VZV n (% of group)
None	399	37.0	39 (9.8%)ref[a]	367 (92%)ref[a]	35 (8.8%)ref[a]
CMV only	172	15.9	18 (10.5%)	165 (95.9%)	18 (10.5%)
EBV only	289	26.8	46 (15.9%)*	260 (90%)	42 (14.5%)*
CMV and EBV	219	20.3	48 (21.9%)****	202 (92.2%)	40 (18.3%)***
Total	**1079**	100	151 (14%)	994 (92.1%)	135 (12.5%)

Abbreviations: CMV, cytomegalovirus; EBV Epstein-Barr virus, HSV-1, herpes simplex virus type 1, IgG immunoglobulin G, VZV varicella-zoster virus. [a] The significance of co-infection in the CMV infected, EBV infected and CMV+EBV+ coinfected group was tested relative to the CMV-EBV- uninfected controls, using Chi-square tests. * $P<0.05$. *** $P<0.001$.**** $P<0.0001$.

The rate of coinfection with VZV was >90% in all groups. HSV-1 infection was significantly more frequent in EBV-infected (p=0.02) or CMV+EBV+ coinfected (p<0.001) children (Table 1). Therefore, HSV-1 infection in particular needs to be considered for possible confounding effects in our analyses on EBV.

CMV– and EBV–specific effector memory T–cells

To analyze whether 6-year-old children carried virus-specific T-cell expansions, we phenotyped virus-specific $CD8^+$ T cells in 15 CMV+ and 14 EBV+ infected children using CD27 and CD45RA and HLA-A*0201 tetramers loaded with CMV-specific peptides of pp65 and IE-1 and EBV-specific peptides of BMLF-1 and BRLF-1 (Figure 1). Virus-specific $CD8^+$ T cells were detected in blood at frequencies of 0.01-0.52% (pp65-NLV) and 0.01-2.32% (IE-1-VLE) for CMV peptides and 0.02-0.64% (BMLF-1-GLC) and 0.02-2.6% (BRLF-1-YVL) for EBV peptides. The CMV- and EBV-specific $CD8^+$ T-cells were phenotypically diverse and predominantly consisted of $CD45RA^-CD27^{+/-}$ and $CD45RA^+CD27^-$ memory T-cells (Figure 1). Thus, 6-year old

children already display expansions of effector memory T cells directed against CMV and EBV antigens.

CMV and EBV associated memory T–cell expansions

To study the effects of CMV and EBV on the T-cell compartment, we immunophenotyped $CD8^+$ and $CD8^-$ ($CD4^+$) T cells in 18-20 randomly-selected children who were either uninfected, CMV ifected, EBV infected or CMV and EBV coinfected. T-cell data from each group of children were merged and subjected to automatic population separation with Infinicyte software,[43] based on the expression of CD3, CD8, CCR7, CD45RO, CD27 and CD28. Within both $CD8^+$ and $CD4^+$ T cells, four populations were distinguished in 2 dimensional plots of principle component 1 vs principal component 2 (Figure 2). CMV was associated with a relative increase of two $CD8^+$ populations: population 2 ($CCR7^-CD45RO^-$; 9.2% uninfected and 15.1% CMV infected) and population 3 ($CCR7^-CD45RO^+$; 2.8% uninfected and 5.9% CMV

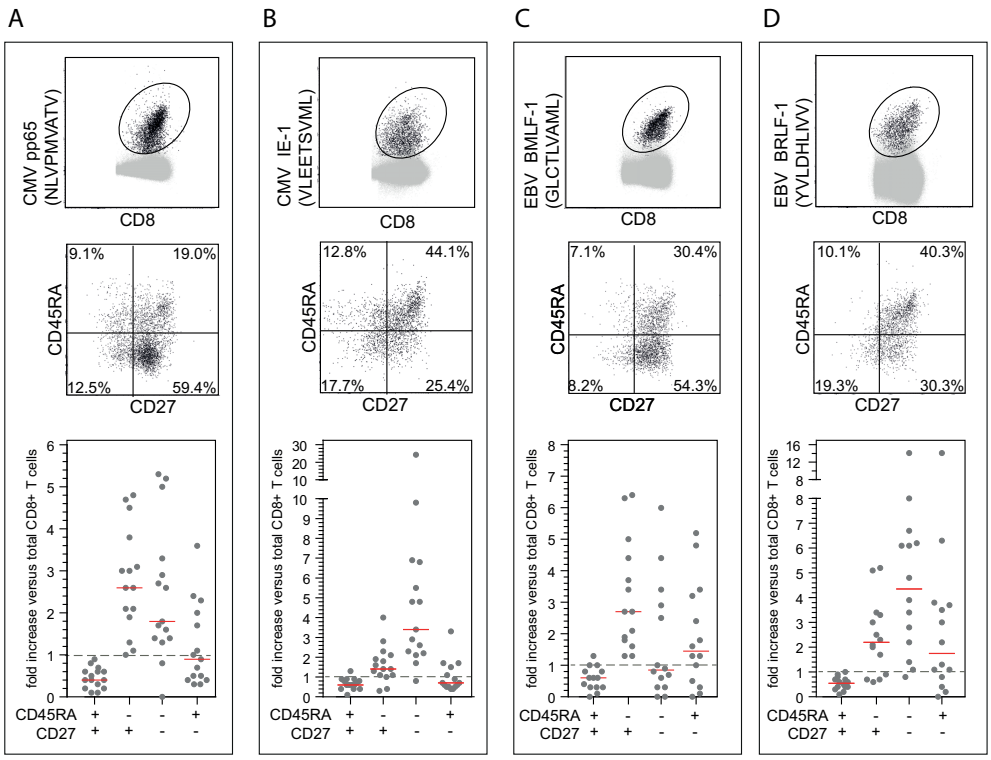
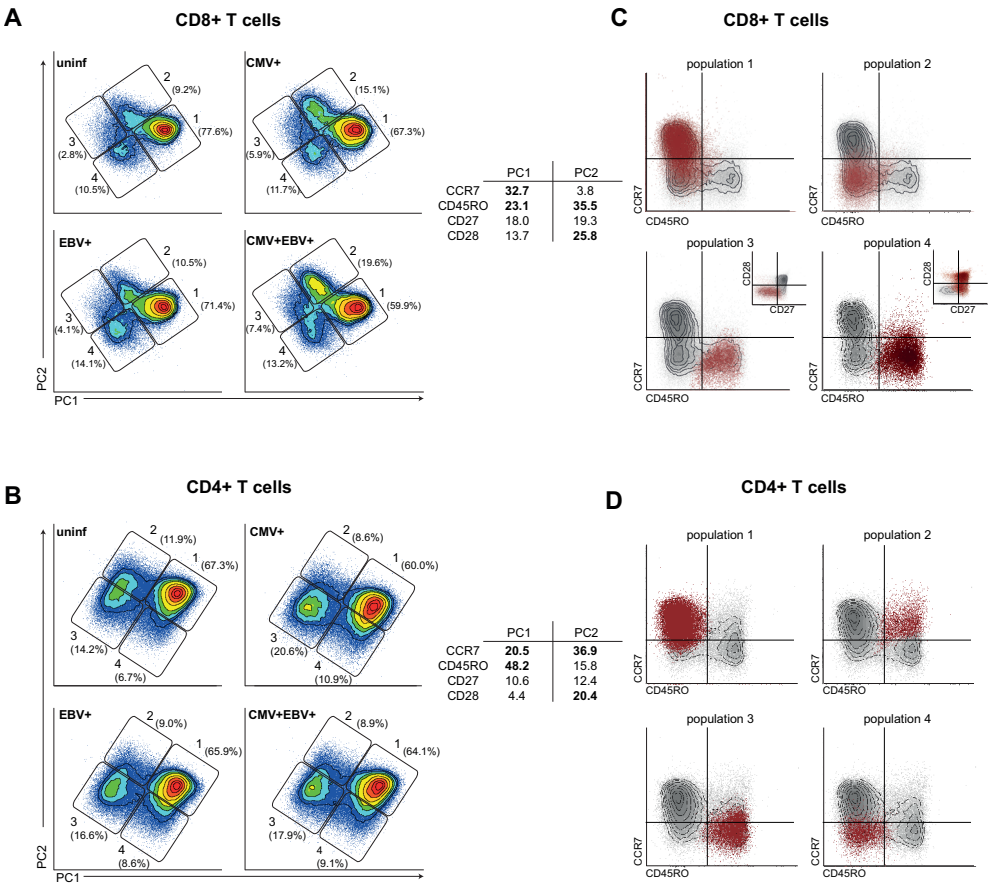


Figure 1. Frequencies and phenotypes of cytomegalovirus (CMV)- and Epstein-Barr virus (EBV) -specific $CD8^+$ T cells in 6-year-old children.

$CD8^+$ T cells recognizing CMV-specific (**A,B**) or EBV-specific (**C,D**) peptides were detected with multicolor flowcytometry and further analyzed for CD27 and CD45RA expression. Flowcytometry plots consist of merged data files of 14 representative CMV infected (A,B) or EBV infected (C,D) children who carried the HLA A*0201 allele. Bottom graphs depict for each individual the relative distributions of tetramer-positive cells as fold increase compared with the relative distribution of total $CD8^+$ T cells.

infected). EBV-infected children showed an increase in $CCR7^{-}CD45RO^{+}$ populations, both $CD27^{-}$ (population 3; 2.8% uninfected and 4.1% EBV infected) and $CD27^{+}$ (population 4; and 10.5% uninfected vs 14.1% EBV infected). CMV and EBV infections were also associated with $CD4^{+}$ T-cell memory expansions: $CCR7^{-}CD45RO^{+}$ (population 3; 14.2% uninfected, 20.6% CMV infected, and 16.6% EBV infected) and $CCR7^{-}CD45RO^{-}$ (population 4; 6.7% uninfected, 10.9% CMV infected, and 8.6% EBV infected) (Figure 2B/D). Thus, CMV and EBV infection in young children are associated with relative expansions of memory T cells.



	PC1	PC2
CCR7	**32.7**	3.8
CD45RO	**23.1**	**35.5**
CD27	18.0	19.3
CD28	13.7	**25.8**

	PC1	PC2
CCR7	**20.5**	**36.9**
CD45RO	**48.2**	15.8
CD27	10.6	12.4
CD28	4.4	**20.4**

Figure 2. Relative change in the T-cell phenotype in cytomegalovirus (CMV)- or Epstein-Barr virus (EBV)-infected children.

Unbiased flowcytometric analyses of $CD8^{+}$ T cells (**A**) and $CD8^{-}$ T cells (**B**) in uninfected, CMV infected, EBV infected and CMV+EBV+ coinfected children. Data from 18-20 children in each group were merged into a single file, and subjected to automatic population separation based on the expression of CD3, CD8, CCR7, CD45RO, CD27 and CD28. 2D projections of principle component (PC)1 versus PC2 revealed 4 populations. The relative contributions of CCR7, CD45RO, CD27 and CD28 to PC1 or PC2 are indicated. **C,D,** CD45RO versus CCR7 expression of the populations defined by automatic population separation. Populations 3 and 4 in C were both CCR7-CD45RO+ but differed in CD27 and CD28 expression (*small inset*).

Normal numbers of naive and central memory T cells in CMV– or EBV–infected children

To determine whether the relative memory T-cell expansions in CMV and/or EBV carriers also affected absolute T-cell numbers, we next evaluated the $CD4^+$ and $CD8^+$ T-cell lineages in 1079 6-year-old children. $CCR7^+CD45RO^-CD27^+CD28^+$ naive T cells were distinguished from $CCR7^+CD45RO^+CD27^+CD28^+$ central memory T cells (Tcm), $CCR7^-CD45RO^+$ effector memory (TemRO) and $CCR7^-CD45RO^-$ effector memory (TemRA) (Supplemental Figure 1A).[22,41] TemRO and TemRA $CD8^+$ T cells were significantly increased in children with CMV or EBV infection compared to uninfected controls (Figure 3B). For CMV, these specifically concerned a 4.4 fold increase of $CD27^-CD28^-$ late differentiated TemRA cells and 1.3-3.3 fold increase of $CD27^{-/+}CD28^-$ intermediate to late-differentiated TemRO cells compared with uninfected controls (Figure 3C-D). EBV-infected children showed a 1.2-1.3 fold increase in $CD27^{-/+}CD28^-$ intermediate to late-differentiated TemRA cells compared with uninfected controls and a 1.4-1.8 fold increase in all TemRO cell subsets (Figure 3C-D). Naive $CD8^+$ T cells and $CD8^+$ Tcm cells were present in normal numbers. Thus, the expansions of effector memory cells resulted in a significant 1.1 fold increase in the total number of $CD8^+$ T cells (Figure 3A). Combined, CMV and EBV infections resulted in a further increase in total $CD8^+$ T-cell numbers (Figure 3A), with seemingly additive effects of the two viruses in doubly infected children compared with singly infected children (Figure 3). Thus, in these

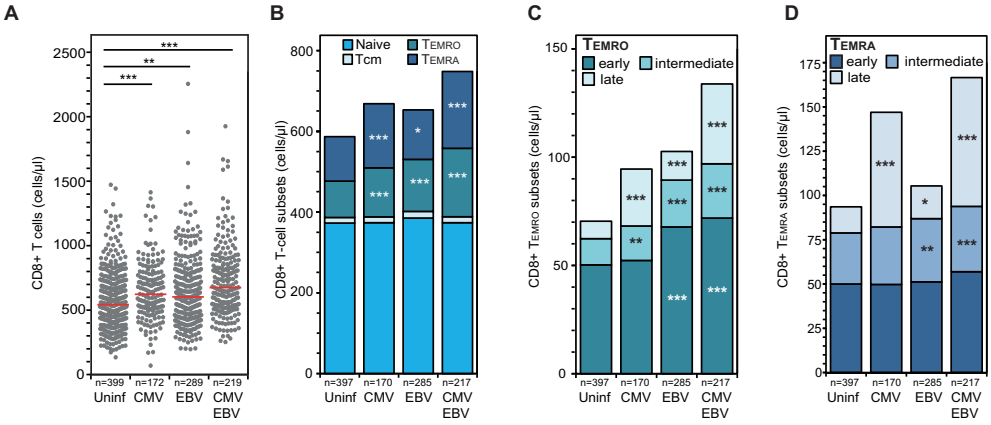


Figure 3. Absolute numbers of $CD8^+$ T-cell subsets in cytomegalovirus (CMV)- or Epstein-Barr virus (EBV)-infected children.

A, Absolute numbers of total $CD8^+$ T-cells in children uninfected with CMV or EBV, infected with only CMV, only EBV or both CMV and EBV. **B,** Similar findings as in A, for 4 main $CD8^+$ T-cell subsets: $CCR7^+CD45RO^-CD27^+CD28^+$ Naive, $CCR7^+CD45RO^+CD27^+CD28^+$ Central memory (Tcm), $CCR7^-CD45RO^+$ Effector memory (TemRO) and $CCR7^-CD45RO^-$ Effector memory (TemRA) cells [41]. **C,D,** Similar findings as in A, for $CD27^+CD28^+$ early, $CD27^+CD28^-$ intermediate and $CD27^-CD28^-$ late TemRO (C) and TemRA (D) populations [22]. Bars represent stacked median values per T-cell population. The number of individuals per category is indicated underneath each plot. Significance was tested first with a Kruskal-Wallis test per T-cell population relative to the uninfected controls, if results were significant (P<0.05) Dunn tests were performed to compare individual patient groups. Significance for the Dunn's test is indicated in the plots: * P<0.05; ** P<0.01; *** P <0.001; **** P <0.0001.

6-year-old children, infections with CMV and EBV were both associated with T-cell memory expansions. The phenotypes of the expanded populations differed for each of the viruses, did not affect each other in CMV+EBV+ double infection, and did not result in loss of naive $CD8^+$ T cells.

Total $CD4^+$ T-cell numbers were not affected by CMV and/or EBV infection (Supplemental Figure 1B). Still, EBV-infected children had 1.5 fold more $CD27^-CD28^-$ late-differentiated TemRO cells than uninfected controls. Furthermore, CMV-infected children showed a 1.3-2 fold increase in $CD27^-$ intermediate and late-differentiated TemRO cells, and a 1.9 fold increase in $CD27^-CD28^-$ late-differentiated TemRA cells compared with uninfected controls (Supplemental Figure 1B). Similar to $CD8^+$ T cells, CMV and EBV infection independently resulted in effector memory $CD4^+$ T-cell expansions, which did not affect naive T-cell numbers.

Effect of HSV-1 and VZV on T-cell subset numbers

Through combined analysis of children with single CMV, or single EBV, and double CMV/EBV infection, we could distinguish distinct effects of these viruses on the T-cell compartment in young children. Still, the observed effects could be influenced by infection with other viruses, such as HSV-1 or VZV, especially considering the increased HSV-1 positivity in EBV+ children (Table 1). Therefore, we separated the groups of uninfected controls, CMV

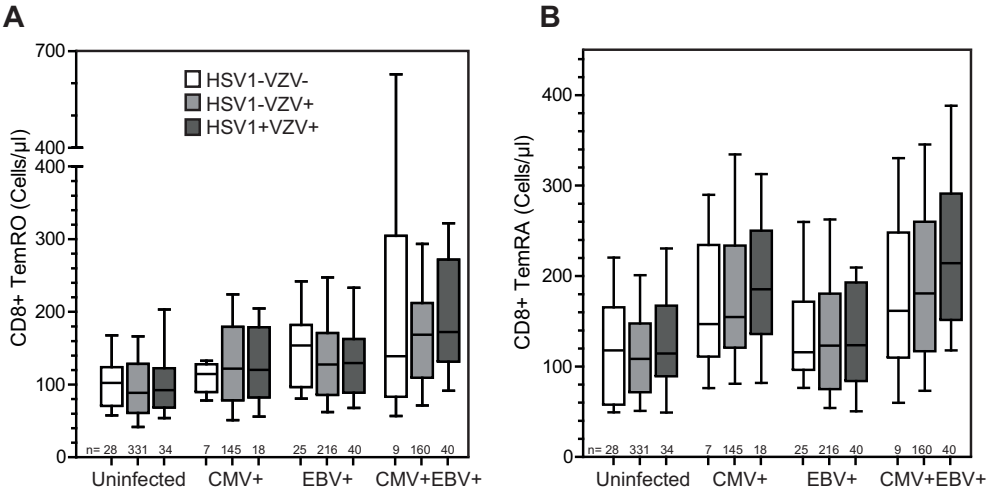


Figure 4. Effect of HSV-1 and VZV on blood $CD8^+$ TemRO and TemRA cell numbers.

Uninfected, CMV infected, EBV infected and CMV+EBV+ coinfected children were further subdivided in children without HSV-1 or VZV (***white***), co-infected with only VZV (***light gray***) or with both HSV-1 and VZV (***dark gray***). Absolute numbers of $CD8^+$ TemRO **(A)** and TemRA **(B)** cells in all population groups depicted as 10-90% box-whisker-plots; the number of individuals per category is indicated underneath each plot. Significance was tested with Kruskal-Wallis test per T-cell population, if results were significant (P<0.05) Dunn tests of individual patient groups were performed to compare HSV-1 and/or VZV subgroups within each category of uninfected, CMV infected, EBV infected or CMV+EBV+ coinfected children.

infected, EBV infected and CMV+EBV+ coinfected children further into 3groups: VZV and HSV-1 uninfected, VZV infected but HSV-1 uninfected, and VZV+ and HSV-1+ coinfected. Because VZV prevalence was >90% in our cohort, the effect of HSV-1, without VZV, could not be determined. Although the TemRO and TemRA populations differed significantly between all virus-infected groups (Kruskal Wallis; p<0.0001), these effects were only caused by CMV- or EBV-associated expansions. The presence of VZV and HSV-1 within the CMV infected or CMV+EBV+ coinfected group was associated only with a slight but not significant increase in TemRA cells (Figure 4). Thus, co-infection with VZV and HSV-1 did not significantly affect CMV- or EBV-associated effector memory T-cell expansion.

Normal vaccination responses in CMV– or EBV–infected children

The effector memory T-cell expansions in CMV- or EBV-infected children did not result in reduced naive and Tcm numbers (Figures 2 and 3). To determine whether immunological memory to other pathogens was normally present in these 6-year-old children, we tested their responses to previous vaccinations with tetanus at 2, 3, 4, 11 months and 4 years, and to measles at 14 months, according to the Dutch national vaccination protocol.[44] We defined vaccination responses in a randomly selected subgroup of 225 6-year-old children who showed similar seroprevalence for CMV and EBV (Supplemental Table 2), and similar effector memory T-cell expansions (data not shown) as the total cohort of 1079 children.

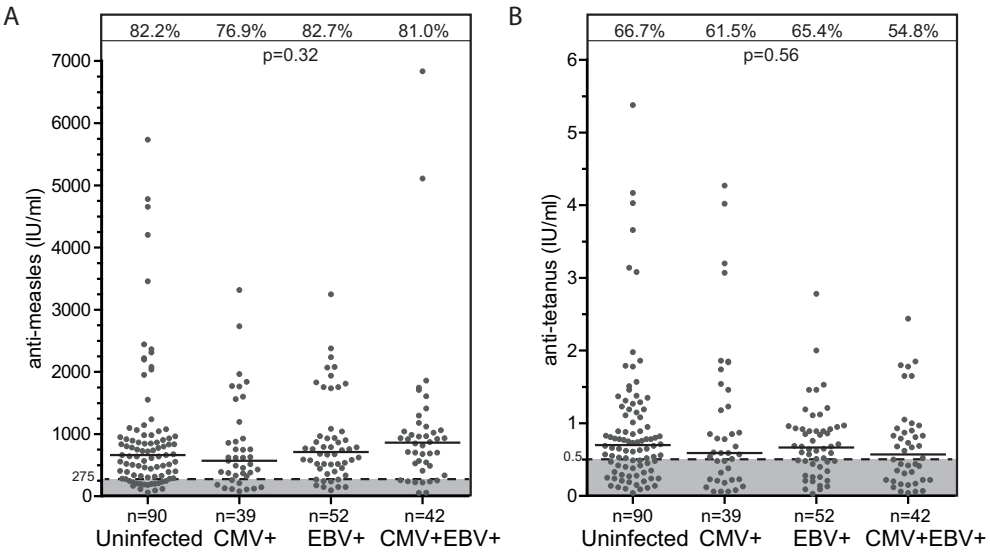


Figure 5. Vaccination responses in 6-year-old CMV or EBV infected children.

Anti-measles **(A)** and anti-tetanus **(B)** immunoglobulin G (IgG) levels in uninfected, CMV infected, EBV infected, or CMV+EBV+ coinfected children. Specific IgG levels did not differ significantly between the virus-positive groups and uninfected controls, as determined with the Kruskal-Wallis test. Percentages at the top represent the number of children with a protective anti-measles IgG response >275 international units/mL (IU/mL) **(A)** or anti-tetanus IgG response >0.5 IU/ml **(B)**. Percentages were not significantly different (Chi-square test).

Furthermore, the seroprevalence of coinfection with HSV-1 and VZV was not significantly different between the CMV+ or EBV+ infected groups in this selected cohort (Supplemental Table 3). In total, 81.2% of the children had IgG antibodies against measles, and 63.2% against tetanus. These frequencies did not differ significantly in children with CMV and/or EBV infection (Figure 5). Furthermore, median titers of anti-measles and anti-tetanus IgG did not differ between the four groups. Thus, EBV and CMV seropositivity did not impair immunity to measles and tetanus vaccination in 6-year-old children.

No association between T–cell expansions and duration of CMV or EBV infection

Our results indicate that the EBV- and CMV-specific T-cell expansions do not affect naive T-cell numbers or vaccination responses to tetanus and measles in 6-year old children. To determine whether infection early in childhood has different immunological consequences than infection later in childhood, we assessed CMV and EBV seropositivity around

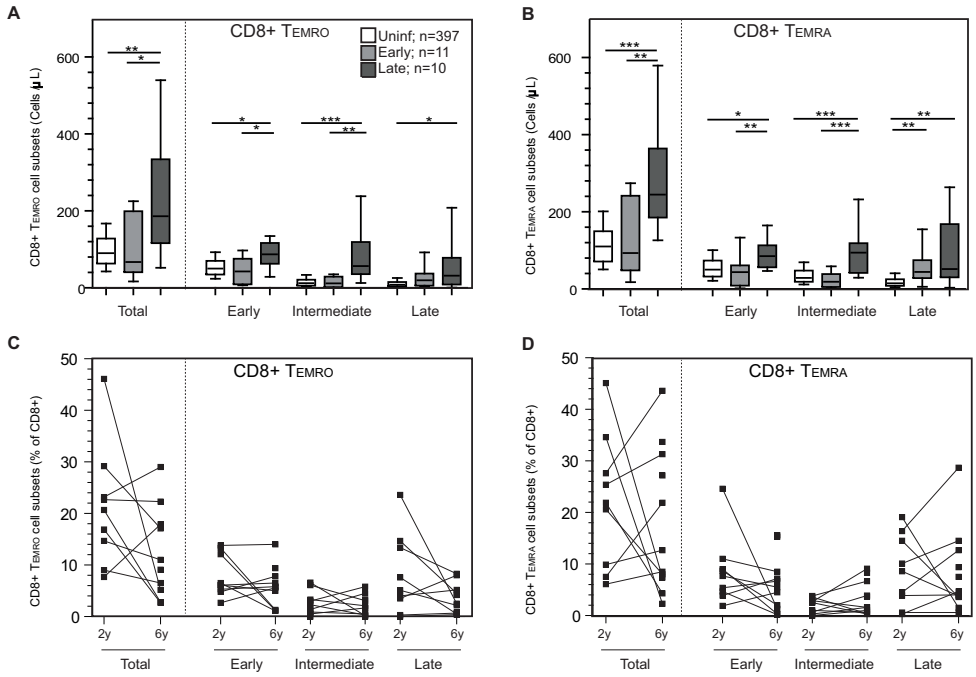


Figure 6. Limited effector memory T-cell expansions in children infected with EBV and CMV at age <2 years. Six-year old CMV+EBV+ coinfected children were divided into subgroups based on age at infection: <2 years (early; n=11) (***light gray***) or >2 years (late; n=10) (***dark gray***), and compared with CMV-EBV- uninfected controls (n=397) (***white***). **A,B,** Absolute number of CD8$^+$ TemRO (***A***) and TemRA (***B***) cell subsets, shown as 10-90% box-whisker plots. Significance was tested first with Kruskal-Wallis tests per T-cell population; if results were significant (***P***<0.05), Dunn tests of individual patient groups were performed. Significance for the Dunn's test is indicated in the plots: * ***P***<0.05; ** ***P*** <0.01; *** ***P*** <0.001; **** ***P*** <0.0001. **C,D)** Longitudinal follow-up of frequencies of CD8$^+$ TemRO (***C***) and TemRA (***D***) subsets at age 2 and 6 years in early-infected children. No significant differences between these ages were seen with paired T-test (***P***>0.12).

2 years of age in a subgroup of 225 children. We observed that of the CMV+EBV+ coinfected children, 80% already carried anti-CMV IgG before 2 years of age, and 46.8% anti-EBV IgG.

We subdivided the EBV+ and CMV+ infected children into early-infected (before 2 years) and late-infected subgroups (2-6 years), and compared their T-cell compartments at 6 years of age. Neither early nor late infection affected naive and Tcm cell numbers (not shown), but $CD4^+$ and $CD8^+$ effector memory T cells were significantly increased (Figure 6 and Supplemental Figure 1C). CMV+EBV+ coinfected children with early infection had significantly smaller $CD8^+$ TemRO and TemRA T-cell populations than late-infected children, with numbers in early-infected children similar to those in uninfected controls (Figure 6A). Moreover, retrospective longitudinal analysis of early-infected children indicated that $CD8^+$ effector memory T-cell numbers were stable between the ages of 2 and 6 years (Figure 6B). $CD4^+$ effector memory T-cell numbers did not differ significantly between early-infected and late-infected children. Still, some subsets in late-infected children were significantly higher than in uninfected controls, and these were not increased in early-infected children (Supplemental Figure 1C). Furthermore, $CD4^+$ effector memory T-cell numbers increased between the ages of 2 and 6 years in early-infected children (Supplemental Figure 1D). Together, these results suggest that $CD8^+$ effector memory T-cell numbers are not directly related to the time after infection. Rather, early infections were associated with smaller expansions of effector memory T cells at age 6 years.

DISCUSSION

We studied the effects of single or combined infection with CMV, EBV, HSV-1 and/or VZV on naive and memory T cells in a large cohort of >1000 6-year-old children. CMV and EBV infections each resulted in distinct effector memory T-cell expansions, which were additive in co-infection. HSV-1 and VZV infections did not significantly affect the T-cell compartment, but might slightly enhance CMV- and EBV-associated T-cell expansions. In contrast to adults, $CD8^+$ Tem (either TemRO or TemRA) cell expansions in CMV- or EBV-infected children did neither result in overcrowding of naive and Tcm compartments, nor in loss of vaccination responses to tetanus or measles. Notably, children infected <2 years showed fewer Tem cells at age 6 years than children infected between age 2 and 6 years. Thus, depending on the age at the time of infection, CMV and EBV infections seem to be controlled differently.

In our current study, we detected sizeable CMV- and EBV-specific $CD8^+$ memory T-cell expansions in 6-year-old children. Unfortunately, the real extent of virus-specific T-cell expansions is difficult to assess owing to the large number of viral epitopes and the diverse HLA backgrounds of the children. However, because 1) the phenotype of the antigen-specific T cells is largely similar to that of the expanded effector memory T-cell populations, both relatively and absolute, within the total $CD8^+$ T-cell pool in our cohort of children, and 2)

the antigen-specific phenotypes are in accordance with previous literature on CMV and EBV infections in adults,[45] we are convinced that phenotyping the total $CD8^+$ T-cell compartment generates an reliable representation of virus-associated changes in the immune compartment.

Our observation that CMV and EBV were associated with an increase of distinct $CD8^+$ Tem subsets is consistent with previous literature in adults and children.[5,7-12,15,46,47] Importantly, our large cohort allowed us to further analyze the possible antagonistic, synergistic or independent impact of CMV and EBV on the immune compartment. Importantly, in the group children coinfected with CMV and EBV, an additive effect was seen for the Tem cell expansions compared with that seen in singly infected patients. Furthermore, the size of the CMV-associated TemRA late cell population was not (negatively) affected by the size of the EBV-associated TemRO cell populations in individual patients (data not shown). Thus, the distinct effects of CMV and EBV on $CD8^+$ Tem cell expansions seem to be independent of each other.

In addition to $CD8^+$ Tem cell expansions, we found that $CD4^+$ Tem cell numbers were increased in both CMV and EBV infections, which is similar to observations in adults.[6] Although these expansions were smaller and did not significantly increase in total $CD4^+$ T-cell numbers, they were distinct between both viruses. Furthermore, the phenotypes of the $CD4^+$ Tem cell expansions were remarkably similar to those of their $CD8^+$ counterparts and were additive in CMV and EBV doubly infected children. Despite these seemingly minor expansions, $CD4^+$ T cells have an important role in controlling primary CMV infection.[48] The formation of sizeable Tem populations is apparently necessary to successfully suppress the virus. Importantly, these expansions did not result in overcrowding of more immature T-cell subsets, something that for CMV infection in elderly has been associated with poor $CD4^+$ memory T-cell responses to influenza proteins.[24]

Our large cohort allowed us to study the effects of HSV-1 and VZV in EBV and CMV uninfected children, as well as any modifying effects on CMV and/or EBV infections. Because virtually all adults are infected with VZV, little is known on the effects on T-cell memory[15] and we could for the first time conclude that VZV does not modify T-cell memory alone or in combination with CMV and/or EBV. Co-infection with HSV-1 was more frequent in EBV-infected and CMV+EBV+ coinfected children and might therefore be an important contributor to the immune modulation that is currently assumed to be EBV-associated. Still, HSV-1 did not affect Tem in EBV infected children, suggesting that these expansions are EBV specific and not due to HSV-1. The difference in T-cell modulation between CMV and EBV on the one hand and VZV and HSV-1 on the other is most likely due to tropism and anatomical localization of the viruses. In fact, it is well-possible that HSV-1 and VZV induce T-cell expansions, but these cells are thought to reside mostly in the human skin or locally around virus-infected cells.[49]

Despite the expansions of $CD8^+$ and $CD4^+$ Tem in children infected with CMV and/or EBV, these did not result in decreased naive or Tcm cell numbers. In fact, infected children had

significantly more total $CD8^+$ T cells. Especially in immunosuppressed individuals, but also in CMV-infected elderly persons, effector memory T-cell expansions with a subsequent loss of naive T cells and loss of vaccination responses have been described as hallmarks of CMV-associated immunosenescence.[9,17,24,25] However, fimdings in the literature are inconsistent,[32-35] and the loss of vaccination responsiveness in children is not consistently observed.[19] In our cohort, IgG titers to previous measles and tetanus vaccinations were similar in uninfected and infected children. Combined, our data suggest that 6-year-old CMV- and/or EBV-infected children do not lose their immune responsiveness owing to virus-related effects.

An important explanation for the discrepancy in the literature on the loss of naive T cells could be differences in analysis strategies. In fact, owing to the Tem cell expansions, the relative proportions of naive and Tcm were decreased in infected children in our study. Moreover, the expansion of total $CD8^+$ T-cell numbers cause a shift in the ratio of CD4 to CD8 T-cells. These relative shifts, however, are the results of data analysis, rather than real defects, because the absolute naive $CD8^+$ and total $CD4^+$ T-cell numbers were normal. This is why we have long advocated for consistent analysis of absolute cell numbers with age-matched controls.[50]

Although 90% of elderly persons are infected with CMV and EBV, only a minority of them shows loss of naive T-cell numbers, impaired vaccination responses and clinical complications. The reason for this is still unclear. Our findings in children at age 6 years indicate that early infection before age 2 years results in smaller Tem expansions than infection between 2 and 6 years. Because the size of the expanded virus-specific Tem pool has been directly correlated with the extent of CMV-induced endothelial damage, the children who were infected before 2 years might have a reduced risk of cardiovascular complications.[31] CMV-related clinical complications might occur only in individuals infected later in life, who show the strongest CMV-associated T-cell expansions. An important factor influencing the variability in the response might be the infectious dose.[36] It would therefore be interesting to correlate the age at infection and the subsequent size of the Tem-cell expansions with the infectious dose and the development of clinical complications in children and adults, in a separate cohort or in longitudinal follow-up of the uninfected children in our cohort.

In conclusion, we here provide evidence that in young children, CMV and EBV infections lead to sizeable Tem expansions but not to the associated immunosenescence. In fact, the Tem expansion result in increased total T-cell numbers, whereas naive and Tcm remain normally present, as do responses to previous vaccinations. Moreover, in 6-year old children infected before age 2 years, stable control of these persistent viruses was maintained with only limited Tem expansions. These new insights into the immunomodulatory effects of herpesviruses in young children are important in our understanding of herpesvirus-associated immunosenescence in the elderly.

REFERENCES

1. Sinclair J. Human cytomegalovirus: Latency and reactivation in the myeloid lineage. *J Clin Virol*. 2008; 41(3): 180-185.
2. Wada T, Toga A, Sakakibara Y, Toma T, Hasegawa M, Takehara K, et al. Clonal expansion of Epstein-Barr virus (EBV)-infected gammadelta T cells in patients with chronic active EBV disease and hydroa vacciniforme-like eruptions. *Int J Hematol*. 2012; 96(4): 443-449.
3. Bate SL, Dollard SC, Cannon MJ. Cytomegalovirus seroprevalence in the United States: the national health and nutrition examination surveys, 1988-2004. *Clin Infect Dis*. 2010; 50(11): 1439-1447.
4. Dowd JB, Palermo T, Brite J, McDade TW, Aiello A. Seroprevalence of Epstein-Barr virus infection in U.S. children ages 6-19, 2003-2010. *PLoS One*. 2013; 8(5): e64921.
5. Appay V, Dunbar PR, Callan M, Klenerman P, Gillespie GM, Papagno L, et al. Memory CD8+ T cells vary in differentiation phenotype in different persistent virus infections. *Nat Med*. 2002; 8(4): 379-385.
6. van Leeuwen EM, Remmerswaal EB, Vossen MT, Rowshani AT, Wertheim-van Dillen PM, van Lier RA, et al. Emergence of a CD4+CD28- granzyme B+, cytomegalovirus-specific T cell subset after recovery of primary cytomegalovirus infection. *J Immunol*. 2004; 173(3): 1834-1841.
7. Griffiths SJ, Riddell NE, Masters J, Libri V, Henson SM, Wertheimer A, et al. Age-associated increase of low-avidity cytomegalovirus-specific CD8+ T cells that re-express CD45RA. *J Immunol*; 190(11): 5363-5372.
8. Brunner S, Herndler-Brandstetter D, Weinberger B, Grubeck-Loebenstein B. Persistent viral infections and immune aging. *Ageing Res Rev*. 2011; 10(3): 362-369.
9. Weinberger B, Lazuardi L, Weiskirchner I, Keller M, Neuner C, Fischer KH, et al. Healthy aging and latent infection with CMV lead to distinct changes in CD8+ and CD4+ T-cell subsets in the elderly. *Hum Immunol*. 2007; 68(2): 86-90.
10. Almanzar G, Schwaiger S, Jenewein B, Keller M, Herndler-Brandstetter D, Wurzner R, et al. Long-term cytomegalovirus infection leads to significant changes in the composition of the CD8+ T-cell repertoire, which may be the basis for an imbalance in the cytokine production profile in elderly persons. *J Virol*. 2005; 79(6): 3675-3683.
11. Sylwester AW, Mitchell BL, Edgar JB, Taormina C, Pelte C, Ruchti F, et al. Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects. *J Exp Med*. 2005; 202(5): 673-685.
12. Chidrawar S, Khan N, Wei W, McLarnon A, Smith N, Nayak L, et al. Cytomegalovirus-seropositivity has a profound influence on the magnitude of major lymphoid subsets within healthy individuals. *Clin Exp Immunol*. 2009; 155(3): 423-432.
13. Henson SM, Riddell NE, Akbar AN. Properties of end-stage human T cells defined by CD45RA re-expression. *Curr Opin Immunol*. 2012; 24(4): 476-481.
14. Komatsu H, Inui A, Sogo T, Fujisawa T, Nagasaka H, Nonoyama S, et al. Large scale analysis of pediatric antiviral CD8+ T cell populations reveals sustained, functional and mature responses. *Immun Ageing*. 2006; 3(11.
15. Derhovanessian E, Maier AB, Hahnel K, Beck R, de Craen AJ, Slagboom EP, et al. Infection with cytomegalovirus but not herpes simplex virus induces the accumulation of late-differentiated CD4+ and CD8+ T-cells in humans. *J Gen Virol*. 2011; 92(Pt 12): 2746-2756.
16. Ouyang Q, Wagner WM, Walter S, Muller CA, Wikby A, Aubert G, et al. An age-related increase in the number of CD8+ T cells carrying receptors for an immunodominant Epstein-Barr virus (EBV) epitope is counteracted by a decreased frequency of their antigen-specific responsiveness. *Mech Ageing Dev*. 2003; 124(4): 477-485.

17. Khan N, Hislop A, Gudgeon N, Cobbold M, Khanna R, Nayak L, et al. Herpesvirus-specific CD8 T cell immunity in old age: cytomegalovirus impairs the response to a coresident EBV infection. *J Immunol*. 2004; 173(12): 7481-7489.
18. Khan N, Shariff N, Cobbold M, Bruton R, Ainsworth JA, Sinclair AJ, et al. Cytomegalovirus seropositivity drives the CD8 T cell repertoire toward greater clonality in healthy elderly individuals. *J Immunol*. 2002; 169(4): 1984-1992.
19. Holder B, Miles DJ, Kaye S, Crozier S, Mohammed NI, Duah NO, et al. Epstein-Barr virus but not cytomegalovirus is associated with reduced vaccine antibody responses in Gambian infants. *PLoS One*. 2010; 5(11): e14013.
20. Sidorchuk A, Wickman M, Pershagen G, Lagarde F, Linde A. Cytomegalovirus infection and development of allergic diseases in early childhood: interaction with EBV infection? *J Allergy Clin Immunol*. 2004; 114(6): 1434-1440.
21. Saghafian-Hedengren S, Sohlberg E, Theorell J, Carvalho-Queiroz C, Nagy N, Persson JO, et al. Epstein-Barr virus coinfection in children boosts cytomegalovirus-induced differentiation of natural killer cells. *J Virol*. 2013; 87(24): 13446-13455.
22. Appay V, van Lier RA, Sallusto F, Roederer M. Phenotype and function of human T lymphocyte subsets: consensus and issues. *Cytometry A*. 2008; 73(11): 975-983.
23. Savva GM, Pachnio A, Kaul B, Morgan K, Huppert FA, Brayne C, et al. Cytomegalovirus infection is associated with increased mortality in the older population. *Aging Cell*; 12(3): 381-387.
24. Derhovanessian E, Maier AB, Hahnel K, McElhaney JE, Slagboom EP, Pawelec G. Latent infection with cytomegalovirus is associated with poor memory CD4 responses to influenza A core proteins in the elderly. *J Immunol*. 2014; 193(7): 3624-3631.
25. Frasca D, Diaz A, Romero M, Landin AM, Blomberg BB. Cytomegalovirus (CMV) seropositivity decreases B cell responses to the influenza vaccine. *Vaccine*; 33(12): 1433-1439.
26. Jonasson L, Tompa A, Wikby A. Expansion of peripheral CD8+ T cells in patients with coronary artery disease: relation to cytomegalovirus infection. *J Intern Med*. 2003; 254(5): 472-478.
27. Sansoni P, Vescovini R, Fagnoni FF, Akbar A, Arens R, Chiu YL, et al. New advances in CMV and immunosenescence. *Exp Gerontol*. 2014; 55(54-62.
28. Pawelec G. Immunosenenescence: role of cytomegalovirus. *Exp Gerontol*. 2014; 54(1-5.
29. White DW, Suzanne Beard R, Barton ES. Immune modulation during latent herpesvirus infection. *Immunol Rev*. 2012; 245(1): 189-208.
30. van de Berg PJ, Yong SL, Remmerswaal EB, van Lier RA, ten Berge IJ. Cytomegalovirus-induced effector T cells cause endothelial cell damage. *Clin Vaccine Immunol*. 2012; 19(5): 772-779.
31. Bolovan-Fritts CA, Trout RN, Spector SA. High T-cell response to human cytomegalovirus induces chemokine-mediated endothelial cell damage. *Blood*. 2007; 110(6): 1857-1863.
32. den Elzen WP, Vossen AC, Cools HJ, Westendorp RG, Kroes AC, Gussekloo J. Cytomegalovirus infection and responsiveness to influenza vaccination in elderly residents of long-term care facilities. *Vaccine*. 2011; 29(29-30): 4869-4874.
33. Odumade OA, Knight JA, Schmeling DO, Masopust D, Balfour HH, Jr., Hogquist KA. Primary Epstein-Barr virus infection does not erode preexisting CD8(+) T cell memory in humans. *J Exp Med*. 2012; 209(3): 471-478.
34. Remmerswaal EB, Havenith SH, Idu MM, van Leeuwen EM, van Donselaar KA, Ten Brinke A, et al. Human virus-specific effector-type T cells accumulate in blood but not in lymph nodes. *Blood*. 2012; 119(7): 1702-1712.
35. van Leeuwen EM, Koning JJ, Remmerswaal EB, van Baarle D, van Lier RA, ten Berge IJ. Differential usage of cellular niches by cytomegalovirus versus EBV- and influenza virus-specific CD8+ T cells. *J Immunol*. 2006; 177(8): 4998-5005.

36. Redeker A, Welten SP, Arens R. Viral inoculum dose impacts memory T-cell inflation. *Eur J Immunol*; 44(4): 1046-1057.
37. Nilsson C, Larsson Sigfrinius AK, Montgomery SM, Sverremark-Ekstrom E, Linde A, Lilja G, et al. Epstein-Barr virus and cytomegalovirus are differentially associated with numbers of cytokine-producing cells and early atopy. *Clin Exp Allergy*. 2009; 39(4): 509-517.
38. Saghafian-Hedengren S, Sverremark-Ekstrom E, Linde A, Lilja G, Nilsson C. Early-life EBV infection protects against persistent IgE sensitization. *J Allergy Clin Immunol*. 2010; 125(2): 433-438.
39. Prendergast AJ, Klenerman P, Goulder PJ. The impact of differential antiviral immunity in children and adults. *Nat Rev Immunol*. 2012; 12(9): 636-648.
40. Kruithof CJ, Kooijman MN, van Duijn CM, Franco OH, de Jongste JC, Klaver CC, et al. The Generation R Study: Biobank update 2015. *Eur J Epidemiol*. 2014; 29(12): 911-927.
41. Sallusto F, Lenig D, Forster R, Lipp M, Lanzavecchia A. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature*. 1999; 401(6754): 708-712.
42. de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, et al. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet*. 2006; 38(10): 1166-1172.
43. Pedreira CE, Costa ES, Lecrevisse Q, van Dongen JJ, Orfao A, EuroFlow C. Overview of clinical flow cytometry data analysis: recent advances and future challenges. *Trends Biotechnol*. 2013; 31(7): 415-425.
44. F. Abbink, H.G.A.M. van de Avoort, W.A.M. Berbers, R.S. van Binnendijk, H.J. Boot, Y.T.H.P. van Duynhoven, et al. The National Immunisation Programme in the Netherlands_Developments in 2006 *Centre for Infectious Diseases Control, RIVM*. 2006.
45. Gamadia LE, van Leeuwen EM, Remmerswaal EB, Yong SL, Surachno S, Wertheim-van Dillen PM, et al. The size and phenotype of virus-specific T cell populations is determined by repetitive antigenic stimulation and environmental cytokines. *J Immunol*. 2004; 172(10): 6107-6114.
46. Huygens A, Dauby N, Vermijlen D, Marchant A. Immunity to cytomegalovirus in early life. *Front Immunol*. 2014; 5(552.
47. Kuijpers TW, Vossen MT, Gent MR, Davin JC, Roos MT, Wertheim-van Dillen PM, et al. Frequencies of circulating cytolytic, CD45RA+CD27-, CD8+ T lymphocytes depend on infection with CMV. *J Immunol*. 2003; 170(8): 4342-4348.
48. Gamadia LE, Remmerswaal EB, Weel JF, Bemelman F, van Lier RA, Ten Berge IJ. Primary immune responses to human CMV: a critical role for IFN-gamma-producing CD4+ T cells in protection against CMV disease. *Blood*. 2003; 101(7): 2686-2692.
49. Farber DL, Yudanin NA, Restifo NP. Human memory T cells: generation, compartmentalization and homeostasis. *Nat Rev Immunol*. 2014; 14(1): 24-35.
50. Comans-Bitter WM, de Groot R, van den Beemd R, Neijens HJ, Hop WC, Groeneveld K, et al. Immunophenotyping of blood lymphocytes in childhood. Reference values for lymphocyte subpopulations. *J Pediatr*. 1997; 130(3): 388-393.

Supplemental Table 1. Antibody details

Antibody	clone	manufacturer
CD45-PerCP	2D1	BD Biosciences
CD3-FITC	SK7	BD Biosciences
CD3-PerCP	SK7	BD Biosciences
CD4-PE-Cy7	SK3	BD Biosciences
CD8-APC-H7	SK1	BD Biosciences
CD8-PE-Cy7	SK1	BD Biosciences
CD45RO-APC	UCHL-1	BD Biosciences
CD28-FITC	CD28.2	BD Biosciences
CD27-APC-H7	M-T271	BD Biosciences
CD197(CCR7)-PE	3D13	e-Bioscience

Supplemental Table 2. CMV- and EBV- IgG seropositivities in the total and selected cohorts

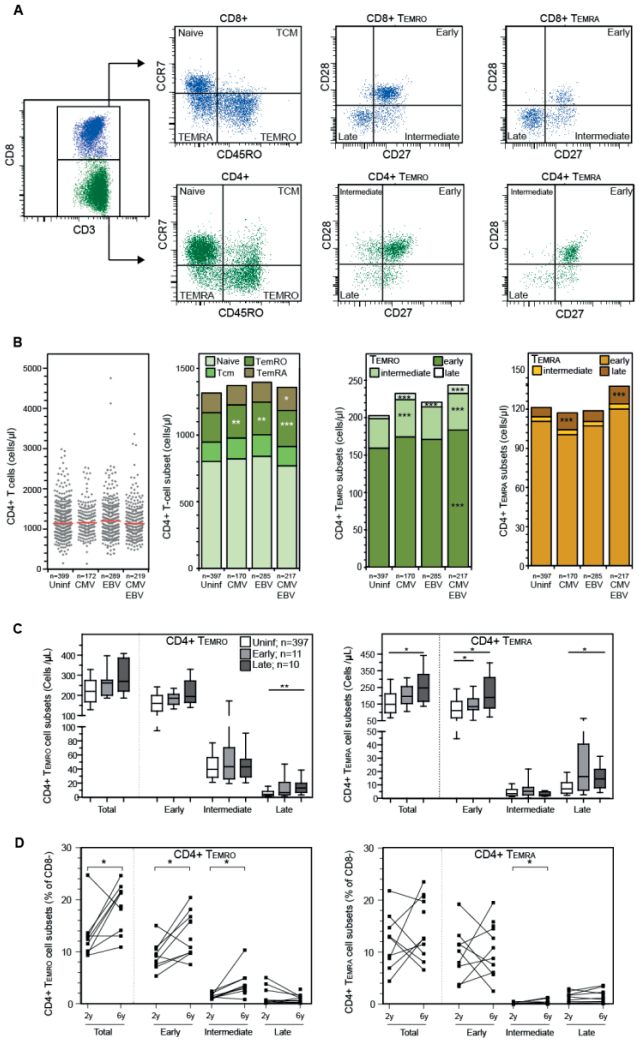
	Total cohort n(%)	Selected cohort n(%)
Uninfected	399 (37.0)	90 (40.0)
CMV+	172 (15.9)	41 (18.2)
EBV+	289 (26.8)	52 (23.1)
CMV+EBV+	219 (20.3)	42 (18.7)
total	1079	225

No significant differences were observed in IgG seropositivity between the two cohorts using the Chi-square test.

Supplemental Table 3. HSV-1 or VZV co-infection in relation to CMV- and EBV-carriership.

	n	%	HSV+ n(% of group)	VZV+ n(% of group)	HSV+VZV+ n(% of group)
Uninfected	90	40.0	9 (10%)ref[a]	83 (92.2%)ref[a]	8 (8.9%)ref[a]
CMV+	41	18.2	5 (12.2%)	41 (100%)	5 (12.2%)
EBV+	52	23.1	10 (19.2%)	48 (92.3%)	9 (17.3%)
CMV+EBV+	42	18.7	6 (14.3%)	39 (92.9%)	6 (14.3%)
total	225	100.0	30 (13.3%)	211 (93.8%)	28 (12.4%)

[a] No significant differences were observed of co-infection in the CMV infected, EBV infected or CMV+EBV+ coinfected groups relative to the CMV-EBV-
uninfected controls using the Chi-square test.



Supplemental Figure 1. CD4$^+$ T cell subset analyses

A, Gating strategy for delineating T-cell subsets. Total CD3$^+$ T cells were subdivided into CD8$^+$ (*blue*) and CD8$^-$ (CD4$^+$) (*green*) T cells. Both CD4$^+$ and CD8$^+$ T cells were further subdivided into CCR7$^+$CD45RO$^-$ Naive, CCR7$^+$CD45RO$^+$ central memory (TCM), CCR7$^-$CD45RO$^+$ effector memory (TEMRO) and CCR7$^-$CD45RO$^-$ effector memory (TEMRA) cells.[41] TEMRO and TEMRA subsets were further subdivided into CD27$^+$CD28$^+$ early, CD27$^+$CD28$^-$ (CD8$^+$) or CD27$^-$CD28$^+$ (CD4$^+$) intermediate and CD27$^-$CD28$^-$ late populations.[22] **B,** Absolute numbers of total CD4$^+$ T-cells and CD4$^+$ T-cell subsets in children uninfected with CMV or EBV (uninf), infected with only CMV (CMV), only EBV (EBV) or both CMV and EBV (CMV EBV). Bars depict stacked median values per T-cell population. The number of individuals per category is indicated underneath each plot. **C,** Absolute numbers of CD8$^+$ TEMRO (left) and TEMRA (right) cell subsets in 6-year-old children who were either CMV-EBV-, became CMV+EBV+ before the age of 2 years (early) or CMV+EBV+ between 2 and 6 years of age (late). Plots depict 10-90% box-whisker plots. **B,C,** Significance was tested first with a Kruskal-Wallis test per lymphocyte population, if results were significant (P<0.05) Dunn tests of individual patient groups were performed. Significance for Dunn tests is indicated in the plots: *P<0.05; ** P <0.01; *** P <0.001. **D,** Longitudinal follow-up of frequencies of CD8$^-$ (CD4$^+$) TEMRO (left) and TEMRA (right) subsets at the age of 2 and 6 years in 9 early infected children. Significance between 2 and 6 years was tested with paired T-tests: *P<0.05.

Chapter 5

No Interactive Effects of Sex and Persistent Cytomegalovirus on Immune Phenotypes in Young Children:

The Generation R Study

Michelle A.E. Jansen, Diana van den Heuvel , Vincent W.V. Jaddoe , Henriette A. Moll, Menno C. van Zelm MC

J. Infect Dis 2017 March 15;215(6):883-888

ABSTRACT

Persistent infections with cytomegalovirus (CMV) differentially affect the host immune phenotype in middle-aged males and females. As CMV already impacts on T-cell memory at a young age, we here studied whether these effects were modified by sex in 1,079 children with an average age of 6-years. Sex and CMV independently impacted on multiple B-cell and T-cell subsets. However, there was no significant effect of their interaction. Importantly, the effects of sex and CMV were in part explained by age and infection with other herpesviruses. Thus, immune aging is likely to be more complex, with involvement of hormonal changes with age, socioeconomic status, birth characteristics, and pathogen exposure

INTRODUCTION

Cytomegalovirus (CMV) is a β-herpesvirus that persists following primary infection, and is actively suppressed by the host's immune system.[1] Primary infections are usually mild, but complications can arise for the fetus in pregnant women or as a result of viral reactivation in immunocompromised individuals.[2] Furthermore, CMV infection is associated with impaired immune responses and increased mortality in elderly persons,[3] a process referred to as immunosenescence.[4] These effects were initially ascribed to CMV-associated expansions of effector memory T (Tem) cells within the CD4 and CD8 lineages[5,6] and the resulting relative decline in naive T (Tn) cells.[7] However, absolute naive T-cell numbers are normal in most CMV+ individuals,[8] CMV-specific effector Tem cells do not accumulate in lymph nodes,[9] and CMV-positive elderly can mount responses to influenza vaccination [10]. Moreover, CMV-driven T-cell expansions are already present in young children and are not associated with age of infection, decreased naive Tn-cell numbers or impaired protective immune responses.[11] Thus, the effects of CMV on early immunosenescence are still inconclusive and are likely impacted on by other determinants.

We recently identified differential effects of CMV carriage on the immune phenotype in pre-elderly males and females.[12] Middle-aged, CMV-infected males had lower numbers of Tn and central memory T (Tcm) cells within the CD4 and CD8 lineages, as well as lower follicular helper T (Tfh) cells, regulatory T cells, and memory B-cells than CMV-infected female subjects.[12] However, it remains unclear if the observed sex-specific differences are related to older age and whether these are impacted on by confounding factors.

Therefore, we here analyzed the combined effects of CMV infection and sex in our previously published cohort of 1,079 children with an average age of 6 years from the Generation R Study.[11] All effects on B- and T-cell subsets were adjusted for socioeconomic factors, ethnicity, breastfeeding and co-infection with Epstein-Barr virus (EBV) and herpes simplex virus type 1 (HSV-1).

METHODS

Study subjects

This study was conducted in the context of the Generation R Study, a prospective population-based cohort study from fetal life until young adulthood.[13] We included 1,079 children (mean = 6.1 y; range = 5.0-7.9 y) from whom CMV Immunoglobulin G (IgG) serology and detailed immunophenotyping of blood B- and T-cells was performed.[11] Data on sex, gestational age, birth weight, parity and mode of delivery were obtained from obstetric records from hospitals and mid-wife practices. Age-adjusted standard deviation scores (SDSs) for weight and body mass index (BMI) were obtained by using Dutch reference growth curves (Growth Analyzer 3.0; Dutch Growth Research Foundation, Rotterdam, Netherlands), and

data on sociodemographic and lifestyle factors from prenatal and postnatal questionnaires completed by both parents, as described before.[14] Written informed consent was obtained from all parents of participating children. Ethical approval for the study was obtained from the Medical Ethical Committee of the Erasmus MC.

Herpesvirus serology

Blood plasma samples were subjected to enzyme-linked immunoassays for IgG antibodies against CMV, EBV (viral capsid antigen [VCA]) and HSV-1 (EUROIMMUN, Lübeck, Germany).[11]

Immunophenotyping

Absolute numbers of $CD3^+$ T cells and $CD19^+$ B cells were obtained with a routine diagnostic lyse-no-wash protocol. Detailed analysis of B- and T-cell subsets was performed with 6-color flowcytometry.[15]

Statistical analyses

Leukocyte subset numbers were natural log-transformed to normalize the data. To test whether maternal and child characteristics were different between the 4 groups, chi-square tests, independent *t* tests and non-parametric Mann-Whitney U tests were applied. Independent associations of CMV seropositivity with each leukocyte subset number were examined using multivariable linear regression model, wherein the *P* value indicates significance for association with the natural log-transformed absolute numbers of immune cells. The ratio (antilog of β coefficient) indicates the strength of the association; the higher the value of β, the larger the difference in mean between either female vs male, CMV+ vs CMV-, male CMV+ vs male CMV-, or female CMV+ vs female CMV-. A positive value indicates a higher number of cells, and a negative value a lower number. Effect modification by sex was evaluated by adding the interaction term (ie, product term of the independent variable (CMV) and subgroup [sex]; CMV^{x}sex, as covariate in the model. The interaction term compares the logs of the 2 ratios, and reflects whether CMV-induced immune responses are differentially affected by sex. The age-adjusted model included only CMV status (0=seronegative, 1=seropositive), sex (0=male, 1=female), age (years) and the interaction term CMV^{x}sex. Because of previously identified significant effects,[11,14] the model was additionally adjusted for ethnic background (Western vs non-Western), EBV seropositivity, HSV-1 seropositivity, breastfeeding at 6 months (yes/no), daycare in first year of life (yes/no), maternal educational level (low/high), and antibiotic use in first year of life (yes/no). The results were reported as β, ratios (antilog of β) and 95% confidence intervals (CIs); $P \leq 0.05$ were considered statistically significant. Additional adjustment for multiple comparisons in the regression analyses (n=27 tests) was performed by the Bonferroni method; only values with $P < 0.0019$ (p=0.05/27) were considered significant. Statistical analyses were performed in SPSS 24.0 for Windows.

RESULTS AND DISCUSSION

Characteristics of study participants

A total of 1,079 children with mean age 6.1 years (range 5.0-7.9 years) participated in the study. Of these, 553 were boys (51.3%), and 391 (36.2%) were IgG+ for CMV (Supplemental Table 1).[11] Median age and mean BMI were similar between boys and girls, as well as between CMV+ and CMV- children. However, CMV+ children were more often of non-Western origin, were more frequently coinfected with EBV and HSV-1, were breastfed longer, went to daycare more often, more frequently needed antibiotics during their first year of life, and their mothers were lower educated than CMV- children (Supplemental Table 1).[14]

Effects of sex on circulating leukocyte subsets

Numbers of multiple B- and T-cell subsets were significantly different between girls and boys. The largest effect of sex was observed for CD27- memory B cells: girls had 1.16-fold (95% CI 1.07-1.26) more CD27- memory B cells than boys, followed by 1.12-fold (95% CI 1.02-1.24) more CD8 Tem cells expressing CD45RO+ (TemRO) cells. Additionally, girls had 1.11-fold (95% CI 1.04-1.19) more CD4 Tn cells, 1.09-fold (95% CI 1.02-1.17) more CD4 Tcm cells, 1.08-fold (95% CI 1.02-1.13) more CD4 T cells, and 1.06-fold (95% CI 1.00-1.11) more total T cells. The smallest effect of sex was observed for total lymphocyte numbers: girls had 1.05-fold (95% CI 1.00-1.11) more total lymphocyte numbers than boys (Table 2, fully adjusted model 2). Similar differences have been reported for pre-elderly persons, although middle-aged males lack expansions of memory T cells, and have additional decreases in CD8 Tn cells and all CD4 T-cell subsets [12].

Effects of Cytomegalovirus status on circulating leukocyte subsets

We previously reported expansions of Tem cells in the CD8 and CD4 lineages (all $P<0.001$) in CMV+ children using unadjusted comparisons [11]. Following adjustment, we here observed the largest effects of CMV for the CD8 Tem CD45RO- (TemRA) late and CD8 TemRO late cells: CMV+ children had 4.04-fold (95% CI 3.47-4.71) and 2.83-fold (95% CI 2.40-3.35) more CD8 TemRA late and CD8 TemRO late cell numbers, respectively, than CMV- children, followed by 1.84-fold (95% CI 1.58-2.14) more CD4 TemRA late cells, and 1.77-fold (95% CI 1.51-2.08) more CD4 TemRO late cells (Table 2, model 2). In addition, CD27- memory B cells were increased in CMV+ children (ratio of 1.14; 95% CI 1.04-1.23) (Table 2).

Effects of Cytomegalovirus status on circulating leukocyte subsets in boys and girls

To determine whether CMV infection was associated with differential effects on the immune phenotype between males and females, the study cohort was divided into 4 groups: CMV− boys, CMV+ boys, CMV− girls, and CMV+ girls (Table 1). In both girls and boys, CMV positivity was associated with significantly increased numbers of γδT cells, CD8 T lymphocytes, and CD4 and CD8 Tem cell populations, and a significantly lower CD4/CD8 ratio.

Table 1. Absolute numbers of leukocyte subsets in blood of boys and girls with and without Cytomegalovirus

n=1079	Male CMV- n=348	Male CMV+ n=205	Ratio [#] (95% CI)	Female CMV- n=340
Leukocyte subsets, cells/µl blood				
granulocyte	3545 (3372 – 3728)	3396 (3194 – 3611)	0.96 (0.88;1.04)	3769 (3567 – 3983)
monocyte	**648 (623 – 674)**	**601 (572 – 631)**	**0.93 (0.87;0.99)***	647 (621 – 675)
lymphocyte	2924 (2836 – 3016)	3033 (2907 – 3163)	1.04 (0.99;1.09)	3027 (2934 – 3123)
NK cell	**176 (168 - 185)**	**210 (196 - 225)**	**1.19 (1.10;1.29)*****	191 (183 - 200)
T cell	2037 (1971 – 2105)	2111 (2021 – 2204)	1.04 (0.98;1.09)	**2122 (2052 – 2195)**
γδ **T cell**	**182 (173 – 191)**	**199 (185 – 213)**	**1.09 (1.01;1.19)***	**171 (163 – 180)**
CD4 T cell	1353 (1309 – 1399)	1347 (1289 – 1409)	1.00 (0.94;1.05)	1449 (1400 – 1500)
CD8 T cell	**667 (640 – 694)**	**745 (708 – 784)**	**1.12 (1.05;1.19)*****	**658 (632 – 684)**
CD4/CD8 ratio	**2.03 (1.97 – 2.09)**	**1.81 (1.74 – 1.88)**	**0.89 (0.85;0.94)*****	**2.20 (2.14 – 2.27)**
CD4 T naive	766 (731 – 802)	716 (673 – 762)	0.94 (0.87;0.99)	851 (813 – 891)
CD4 Tcm	143 (137 – 150)	140 (132 – 149)	0.98 (0.91;1.06)	155 (148 – 162)
CD4 TemRO	**216 (207 – 225)**	**251 (237 – 265)**	**1.16 (1.08;1.25)*****	226 (217 – 235)
CD4 TemRO late	**5.65 (5.11 – 6.26)**	**9.83 (8.69 – 11.1)**	**1.74 (1.48;2.04)*****	**5.20 (4.69 – 5.76)**
CD4 TemRA	**147 (138 – 156)**	**164 (152 - 177)**	**1.12 (1.01;1.23)***	143 (134 – 153)
CD4 TemRA late	**7.91 (7.20 – 8.70)**	**14.57 (12.74-16.67)**	**1.84 (1.57;2.16)*****	**6.92 (6.28 – 7.62)**
CD8 T naive	378 (359-398)	357 (333-382)	0.94 (0.87;1.03)	373 (355-392.)
CD8 Tcm cells	14.5 (13.6-15.6)	14.0 (12.8-15.3)	0.97 (0.86; 1.08)	14.7 (13.7-15.7)
CD8 TemRO	**98.2 (92.3-104.5)**	**129.4 (118.8-141.0)**	**1.32 (1.19;1.46)*****	**105.0 (98.7-111.7)**
CD8 TemRO late	**9.70 (8.79-10.70)**	**28.73 (25.19-32.78)**	**2.96 (2.52;3.49)*****	**9.38 (8.47-10.38)**
CD8 TemRA	**114 (108-121)**	**168.8 (156.0-182.7)**	**1.48 (1.34;1.63)*****	**105.7 (99.0-112.9)**
CD8 TemRA late	**17.8 (16.3-19.6)**	**66.9 (59.4-75.3)**	**3.75 (3.23;4.35)*****	**14.47 (13.16-15.91)**
B cell	601 (577-626)	597 (564-633)	0.99 (0.93;1.07)	613 (588-639)
transitional	60.5 (56.9-64.3)	61.7 (57.6-66.2)	1.02 (0.93;1.12)	58.8 (55.2-62.7)
naive mature	390.6 (370.8-411.4)	388.0 (364.0-413.6)	0.99 (0.91;1.08)	400.4 (382.3-419.3)
natural effector	18.9 (17.6-20.3)	18.7 (17.0-20.4)	0.99 (0.88;1.11)	19.8 (18.5-21.1)
CD27- memory	**44.6 (42.2-47.1)**	**49.3 (45.7-53.1)**	**1.11 (1.01;1.21)***	50.3 (47.9-52.8)
CD27+ memory	23.5 (22.0-25.0)	25.1 (23.0-27.5)	1.07 (0.96;1.19)	25.2 (23.8-26.8)

The interaction term compares the logs of the 2 ratios, and reflects whether cytomegalovirus (CMV)-induced immune responses are differentially affected by sex. The *P* values indicate whether the interaction effect of sex is significant for model 1 and model 2. Model 1 includes the predicting variables sex, CMV and age. Model 2 additionally includes the variables: ethnicity, EBV and HSV-1 serostatus, breastfeeding at 6 months, daycare attendance, maternal educational level and antibiotic use in the first year of life.

Values represent geometric mean cells per microliter of blood [95% CI]. Differences in geometric means between the groups were calculated by unadjusted linear regression analysis.Bolding indicates significant differences ($P < 0.05$). Abbreviations: CI, confidence interval; Tcm, central memory T cells; TemRO, CD45RO+ effector memory T cells; TemRA, CD45RO- effector memory T cells.

[#]Ratio of geometric means (95% CI) comparing CMV+ and CMV- males; [¶] Ratio of geometric means (95% CI) comparing CMV+ and CMV- females. *p<0.05 **p<0.01 ***p<0.001.

Female CMV+ n=186	Ratio ¶ (95% CI)	β Interaction (95% CI) age adjusted	P value, model 1	P value, model 2
3663 (3410 - 3935)	0.97 (0.87;1.07)	1.02 (0.91;1.17)	0.67	0.53
611 (583 - 640)	0.94 (0.88;1.01)	1.03 (0.94;1.13)	0.51	0.15
3185 (3068 - 3306)	1.05 (1.00;1.11)	1.02 (0.94;1.10)	0.66	0.98
204 (191 - 218)	1.07 (0.99;1.16)			
2264 (2174 - 2358)	**1.07 (1.01;1.13)***	1.03 (0.95;1.11)	0.52	0.88
191 (180 - 204	**1.12 (1.03;1.21)****	1.01 (0.90;1.14)	0.81	0.86
1478 (1416 - 1543)	1.02 (0.96;1.08)	1.02 (0.94;1.11)	0.69	0.98
780 (745 - 818)	**1.19 (1.11;1.26)*****	1.05 (0.95;1.15)	0.32	0.63
1.89 (1.82 - 1.97)	**0.86 (0.45;2.27)*****	0.97 (0.90;1.04)	0.39	0.54
817 (768 - 870)	0.96 (0.89;1.04)	1.01 (0.90;1.13)	0.83	0.71
157 (147 - 167)	1.01(0.94;1.09)	1.02 (0.91;1.14)	0.72	0.92
257 (242 - 272)	1.14 (1.06;1.22)	0.97 (0.88;1.07)	0.54	0.43
10.17 (8.95 - 11.55)	**1.96 (1.66;2.31)*****	1.13 (0.89;1.43)	0.32	0.46
156 (143 – 171)	1.09 (0.98;1.22)	0.99 (0.85;1.15)	0.90	0.81
12.70 (11.39-14.17)	**1.84 (1.57;2.14)*****	1.03 (0.82;1.29)	0.80	0.64
368 (347-391)	0.99 (0.91;1.07)	1.02 (0.90;1.15)	0.75	0.88
15.1 (13.6-16.7)	1.03 (0.91;1.15)	1.06 (0.90;1.26)	0.49	0.90
144.2 (131.2-158.5)	**1.37 (1.23;1.53)*****	1.01 (0.87;1.18)	0.90	0.43
30.71 (26.79-35.20)	**3.27 (2.76;3.88)*****	1.06 (0.83;1.35)	0.63	0.54
172.8 (158.3-188.6)	**1.63 (1.46;1.82)*****	1.10 (0.95;1.28)	0.21	0.28
58.20 (50.81-66.67)	**4.02 (3.42;4.73)*****	1.08 (0.86;1.35)	0.51	0.71
622 (594-651)	1.01 (0.94;1.08)	1.02 (0.93;1.13)	0.64	0.82
59.9 (55.8-64.4)	1.02 (0.92;1.13)	0.98 (0.85;1.13)	0.80	0.66
405.4 (384.9-426.9)	1.01 (0.94;1.09)	1.02 (0.91;1.14)	0.73	0.93
19.1 (17.6-20.7)	0.96 (0.87;1.07)	0.99 (0.84;1.16)	0.87	0.71
54.4 (50.7-58.3)	1.08 (0.99; 1.17)	0.98 (0.87;1.12)	0.80	0.45
26.2 (24.0-28.6)	1.04 (0.94; 1.15)	0.97 (0.83;1.13)	0.68	0.35

Table 2. Linear regression analysis for the absolute cell numbers per leukocyte subset

Cell subset	Predicting variable	Ratio (95% CI) Age adjusted model[1]	p-value[1]	Ratio (95% CI) Fully adjusted model[2]	P value[2]
lymphocyte	**CMV status**	**1.05 (1.01; 1.09)**	**0.007**	**1.06 (1.01; 1.11)**	**0.02**
	sex	**1.04 (1.01; 1.08)**	**0.02**	**1.05 (1.00; 1.11)**	**0.03**
T cell	**CMV status**	**1.06 (1.02; 1.11)**	**0.003**	**1.07 (1.01; 1.12)**	**0.014**
	sex	**1.06 (1.01; 1.10)**	**0.006**	**1.06 (1.00; 1.11)**	**0.033**
CD4 T cell	CMV status	1.02 (0.97; 1.06)	0.47	1.02 (0.96; 1.07)	0.56
	sex	**1.08 (1.04; 1.12)**	**0.00**	**1.08 (1.02; 1.13)**	**0.004**
CD4 Tnaive	CMV status	**0.96 (0.91; 1.02)**	0.17	0.99 (0.92; 1.06)	0.67
	sex	**1.12 (1.06; 1.18)**	**0.000**	**1.11 (1.04; 1.19)**	**0.003**
CD4 Tcm	CMV status	1.01 (0.95; 1.06)	0.85	1.02 (0.96; 1.11)	0.34
	sex	**1.10 (1.04; 1.16)**	**0.00**	**1.09 (1.02; 1.17)**	**0.009**
CD4 TemRO	**CMV status**	**1.13 (1.08; 1.19)**	**0.00**	**1.09 (1.03; 1.17)**	**0.007**
	sex	1.04 (0.99; 1.09)	0.11	1.03 (0.96; 1.10)	0.39
CD4 TemRO late	**CMV status**	**1.82 (1.62; 2.05)**	**0.00**	**1.77 (1.51; 2.08)**	**0.00**
	sex	0.96 (0.86; 1.08)	0.49	0.98 (0.84; 1.15)	0.79
CD4 TemRA	CMV status	**1.10 (1.03; 1.19)**	**0.009**	1.07 (0.97; 1.18)	0.16
	sex	0.96 (0.89; 1.03)	0.23	0.96 (0.87; 1.05)	0.36
CD4 TemRA late	**CMV status**	**1.82 (1.63; 2.03)**	**0.00**	**1.84 (1.58; 2.14)**	**0.00**
	sex	**0.87 (0.78; 0.97)**	**0.01**	0.94 (0.81; 1.09)	0.38
CD8 T cell	**CMV status**	**1.16 (1.11; 1.22)**	**0.00**	**1.18 (1.11; 1.25)**	**0.00**
	sex	1.01 (0.97; 1.06)	0.62	1.02 (0.95; 1.08)	0.64
CD8 Tnaive	CMV status	0.98 (0.93; 1.04)	0.59	1.03 (0.95; 1.11)	0.50
	sex	1.00 (0.95; 1.06)	0.90	0.99 (0.92; 1.07)	0.84
CD8 Tcm	CMV status	1.01 (0.93; 1.10)	0.85	1.06 (0.95; 1.19)	0.30
	sex	1.03 (0.95; 1.12)	0.45	1.05 (0.93; 1.17)	0.44
CD8 TemRO	**CMV status**	**1.33 (1.23; 1.43)**	**0.00**	**1.28 (1.15; 1.41)**	**0.00**
	sex	**1.10 (1.02; 1.18)**	**0.01**	**1.12 (1.02; 1.24)**	**0.02**
CD8 TemRO late	**CMV status**	**3.05 (2.70; 3.44)**	**0.00**	**2.83 (2.40; 3.35)**	**0.00**
	sex	1.01 (0.90; 1.14)	0.82	1.05 (0.89; 1.23)	0.56
CD8 TemRA	**CMV status**	**1.55 (1.44; 1.67)**	**0.00**	**1.52 (1.37; 1.69)**	**0.00**
	sex	0.97 (0.90; 1.04)	0.34	0.98 (0.88; 1.09)	0.69
CD8 TemRA late	**CMV status**	**3.88 (3.47; 4.33)**	**0.00**	**4.04 (3.47; 4.71)**	**0.00**
	sex	**0.84 (0.76; 0.94)**	**0.001**	**0.84 (0.73; 0.98)**	**0.03**
B cell	CMV status	1.01 (0.96; 1.06)	0.71	1.02 (0.95; 1.09)	0.60
	sex	1.03 (0.98; 1.08)	0.19	**1.07 (1.00; 1.14)**	**0.04**
transitional	CMV status	1.02 (0.94; 1.09)	0.65	1.01 (0.92; 1.12)	0.80
	sex	0.98 (0.92; 1.05)	0.58	1.02 (0.92; 1.12)	0.75
naive mature	CMV status	1.01 (0.95; 1.07)	0.78	1.00 (0.93; 1.08)	0.92
	sex	1.04 (0.98; 1.10)	0.20	**1.08 (1.00; 1.15)**	**0.04**
natural effector	CMV status	0.97 (0.90; 1.05)	0.46	1.00 (0.90; 1.12)	0.94
	sex	1.04 (0.96; 1.12)	0.30	1.06 (0.95; 1.18)	0.31
CD27- memory	**CMV status**	**1.09 (1.03; 1.17)**	**0.005**	**1.14 (1.04; 1.23)**	**0.003**
	sex	**1.12 (1.05; 1.19)**	**0.00**	**1.16 (1.07; 1.26)**	**0.00**
CD27+ memory	CMV status	1.05 (0.98; 1.14)	0.18	1.11 (1.00; 1.24)	0.05
	sex	1.07 (1.00; 1.15)	0.07	1.10 (0.99; 1.23)	0.06

Linear regression was performed on natural log-transformed data by using the variables cytomegalovirus (CMV) status (0=seronegative, 1=seropositive), sex (0=male, 1= female), and Age (years). Because the CMV*sex interaction term was not significant in the crude, age-adjusted and fully adjusted model, the interaction-term was omitted. Bold text indicates significant differences within 1 subset. Italicized text indicates differences that remained significant after multiple testing correction (n=27 tests; p< 0.0019). The ratios (antilog of the β) compare the adjusted geometric means between sexes and between CMV+ and CMV- children. The ***P*** value indicates the association between the predicting variable and the absolute number of immune cells.[1] Abbreviations: CI, confidence interval; Tcm, central memory T cells; TemRO, CD45RO+ effector memory T cells; TemRA, CD45RO- effector memory T cells; Tn, naieve T cells. The age adjusted model includes the predicting variables sex, CMV, and age.[2] The fully adjusted model additionally includes the variables: ethnicity, EBV and HSV-1 serostatus, breastfeeding duration, daycare attendance, maternal educational level, and antibiotic use in the first year.

To study whether associations between CMV status and leukocyte subsets were differentially affected by sex, we performed linear regression analyses that included the following variables: sex, age, CMV status, and the interaction term CMV*sex (Table 2). In contrast with pre-elderly persons [12], sex did not significantly modify the associations between CMV and total CD4 T cells (β-interaction=1.02, *P*=0.69), CD4+ Tcm cells (β-interaction=1.02, *P*=0.72), CD4 TemRO (β-interaction=0.97, *P*=0.54), and CD8+ Tcm cells (β-interaction=1.06, *P*=0.49).

Confounding factors in the association between sex, Cytomegalovirus and leukocyte subsets

CMV seropositivity in children is associated with socioeconomic factors, crowding and ethnic backgrounds (Supplemental Table 1).[14] Because these differences between CMV+ and CMV- children can potentially mask the presence of effect modification by sex, multivariable linear regression analyses were performed with additional inclusion of the following potential confounding factors: ethnicity, breastfeeding, daycare attendance, education level of the mother, antibiotics usage, and EBV- and HSV-1 seropositivity (Table 1). Following correction for these confounders, the CMV*sex interaction remained nonsignificant for total CD4 (*P*=0.98), CD4 Tcm cells (*P*=0.92), CD4 TemRO cells (p=0.43) and CD8 Tcm cells (*P*=0.90); Table 1, model 2). Thus, the effects of CMV on the immune phenotypes were not sex-specific in both the unadjusted and adjusted models.

Furthermore, following correction for confounders, the initially observed positive effect estimates of CMV seropositivity reduced markedly. Hence, the effect of CMV can be partly attributed to the combination of variables included in the multivariable model, especially to EBV and HSV-1. Specifically, EBV carriership was an independent predictor for CD8+ TemRO and TemRO late cells,[11] with ratios of 1.32 (95% CI 1.19-1.46, *P*<0.001) and 1.49 (95% CI 1.26-1.75, *P*<0.001), respectively. In addition, EBV affected the association between CMV and B cell populations (esp. CD27- memory B cells), and HSV-1 affected associations between CMV and T cells. In contrast, the effect sizes of sex did not materially change after inclusion of the potential confounders in the fully adjusted model (Table 2). Notably, after inclusion of the confounding variables, girls were found to have 1.08-fold (*P*=0.04) more naive mature B cells than boys (without adjustment 1.04-fold, *P*=0.20). Thus, to enable

delineation of specific determinants such as sex and CMV (and their possibly interaction) in relation to immune cells, various confounders, such as other herpesviruses, but also breastfeeding and lifestyle characteristics should be taken into account. Arguably, some of the effects observed in pre-elderly persons could have been impacted upon by confounding factors such as coinfections. However, as nearly all pre-elderly persons in the study of Van der Heiden et al were EBV seropositive, the EBV serostatus is not likely to explain the observed sex-specific effects of CMV on CD8 Tcm cells.[12]

Technical considerations

In our study we were not able to analyze Tfh cells. However, because these are CD45RA-CCR7+, their definition mostly overlaps with CD4 Tcm cells. Therefore, it is unlikely that in young children these numbers are differentially affected by an interaction of sex and CMV.

A major strength of the present study is that we used a multivariable approach to assess the relative contributions of CMV seropositivity, age, sex and their possible interactions on immune subsets in a large prospective population-based cohort study of healthy children. In contrast with the previous study in pre-elderly persons,[12] we could adjust in our childhood cohort for a broad range of potential confounders, including socioeconomic status, ethnicity, breastfeeding, daycare, herpesvirus co-infections, and antibiotic use. These analyses limit the possibility of bias introduced by confounding. However, the observed associations might still be related to unmeasured variables, such as lifestyle, other viral infections and microbial burden. Hence, residual confounding cannot be fully excluded.

Conclusion

We here found that both CMV carriage and sex affect blood immune cell numbers. However, in contrast to pre-elderly persons, there is no combined effect of the two. The potential difference between these populations could be age-related and further affected by hormone balances. This would need further analysis in additional study cohorts, especially in teenagers and in young adults, with longitudinal follow-up data. This would ensure that not only CMV serostatus, but also other important characteristics of socioeconomic status, birth and pathogen exposure are taken into account and delineate the main determinants of immune aging.

REFERENCES

1. Sinclair J. Human cytomegalovirus: Latency and reactivation in the myeloid lineage. J Clin Virol 2008; 41:180-5.
2. Griffiths P, Baraniak I, Reeves M. The pathogenesis of human cytomegalovirus. J Pathol 2015; 235:288-97.
3. Savva GM, Pachnio A, Kaul B, et al. Cytomegalovirus infection is associated with increased mortality in the older population. Aging Cell 2013; 12:381-7.
4. Pawelec G. Immunosenenescence: role of cytomegalovirus. Exp Gerontol 2014; 54:1-5.
5. Appay V, Dunbar PR, Callan M, et al. Memory CD8+ T cells vary in differentiation phenotype in different persistent virus infections. Nat Med 2002; 8:379-85.
6. van Leeuwen EM, Remmerswaal EB, Vossen MT, et al. Emergence of a CD4+CD28- granzyme B+, cytomegalovirus-specific T cell subset after recovery of primary cytomegalovirus infection. J Immunol 2004; 173:1834-41.
7. Derhovanessian E, Maier AB, Hahnel K, McElhaney JE, Slagboom EP, Pawelec G. Latent infection with cytomegalovirus is associated with poor memory CD4 responses to influenza A core proteins in the elderly. J Immunol 2014; 193:3624-31.
8. Wertheimer AM, Bennett MS, Park B, et al. Aging and cytomegalovirus infection differentially and jointly affect distinct circulating T cell subsets in humans. J Immunol 2014; 192:2143-55.
9. Remmerswaal EB, Havenith SH, Idu MM, et al. Human virus-specific effector-type T cells accumulate in blood but not in lymph nodes. Blood 2012; 119:1702-12.
10. den Elzen WP, Vossen AC, Cools HJ, Westendorp RG, Kroes AC, Gussekloo J. Cytomegalovirus infection and responsiveness to influenza vaccination in elderly residents of long-term care facilities. Vaccine 2011; 29:4869-74.
11. van den Heuvel D, Jansen MA, Dik WA, et al. Cytomegalovirus- and Epstein-Barr Virus-Induced T-Cell Expansions in Young Children Do Not Impair Naive T-cell Populations or Vaccination Responses: The Generation R Study. J Infect Dis 2016; 213:233-42.
12. van der Heiden M, van Zelm MC, Bartol SJ, et al. Differential effects of Cytomegalovirus carriage on the immune phenotype of middle-aged males and females. Sci Rep 2016; 6:26892.
13. Kooijman MN, Kruithof CJ, van Duijn CM, et al. The Generation R Study: design and cohort update 2017. Eur J Epidemiol 2016; 31:1243-64.
14. Jansen MA, van den Heuvel D, Bouthoorn SH, et al. Determinants of Ethnic Differences in Cytomegalovirus, Epstein-Barr Virus, and Herpes Simplex Virus Type 1 Seroprevalence in Childhood. J Pediatr 2016; 170:126-34 e1-6.
15. van den Heuvel D, Jansen MA, Nasserinejad K, et al. Effects of nongenetic factors on immune cell dynamics in early childhood: The Generation R Study. J Allergy Clin Immunol 2016.

Supplemental Table 1. Participant characteristics.

	Male	Female	CMV-	CMV+	Male CMV-	Male CMV+	Female CMV-	Female CMV+	Total
No. (%)	553 (51%)	526 (49%)	688 (64%)	391 (36%)	348 (32%)	205 (19%)	340 (32%)	186 (17%)	1,079 (100%)
Median age (years; range)	6.0 (5.0-7.9)	6.0 (5.2-7.6)	6.0 (5.0-7.8)	6.0 (5.4-7.9)	6.0 (5.0-7.8)	6.1 (5.4-7.9)	6.0 (5.2-7.5)	6.1 (5.7-7.6)	6.0 (5.0-7.9)
Mean BMI (SDS; SD)	0.19 (0.8)	0.21 (0.9)	0.20 (0.8)	0.20 (0.8)	0.19 (0.8)	0.19 (0.8)	0.20 (0.8)	0.20 (0.9)	0.20 (0.8)
Western ethnicity (n; %)	475 (86%)	457 (87%)	619 (90%)	**313 (80%)^{***a}**	314 (90%)	**161 (79%)^{***b}**	305 (90%)	**152 (82%)^{**c}**	932 (86%)
EBV seropositivity (n; %)	272 (49%)	236 (45%)	289 (42%)	**219 (56%)^{***a}**	157 (45%)	**115 (56%)^{*b}**	132 (39%)	**104 (56%)^{***c}**	508 (47%)
HSV-1 seropositivity (n; %)	83 (15%)	68 (13%)	85 (12%)	**66 (17%)^{*a}**	44 (13%)	**39 (19%)^{*b}**	41 (12%)	27 (15%)	151 (14%)
Breastfeeding duration (n; %)									
Never	43 (11%)	37 (10%)	67 (14%)	**13 (5%)^{***a}**	36 (15%)	**7 (5%)^{**b}**	31 (13%)	**6 (5%)^{*c}**	80 (11%)
<3 months	127 (33%)	122 (33%)	171 (35%)	**78 (29%)**	85 (35%)	42 (30%)	86 (35%)	36 (28%)	249 (33%)
3-6 months	69 (18%)	87 (23%)	90 (18%)	**66 (25%)**	40 (17%)	29 (21%)	50 (20%)	37 (29%)	156 (21%)
>6 months	143 (37%)	128 (34%)	162 (33%)	**109 (41%)**	111 (30%)	61 (44%)	80 (32%)	48 (38%)	271 (36%)
Breastfeeding at 6 months (n; %)	142 (32%)	126 (29%)	160 (28%)	**108 (35%)^{*a}**	82 (34%)	**61 (39%)^{*b}**	79 (27%)	47 (31%)	268 (30%)
Daycare in first year of life (n; %)	267 (92%)	269 (89%)	331 (87%)	**223 (94%)^{**a}**	152 (89%)	115 (95%)	161 (85%)	**108 (94%)^{*c}**	536 (90%)
Siblings >1	236 (43%)	216 (42%)	297 (44%)	155 (40%)	154 (45%)	82 (40%)	143 (43%)	73 (40%)	452 (43%)
Low maternal educational level (n; %)*	175 (33%)	159 (31%)	189 (28%)	**145 (39%)^{***a}**	95 (28%)	**90 (41%)^{**b}**	94 (29%)	65 (37%)	334 (32%)
Net household income <2200 €/month	149 (31%)	132 (27%)	181 (28%)	100 (30%)	98 (31%)	51 (30%)	83 (26%)	49 (29%)	281 (29%)
Antibiotic use 1st year (n; %)	149 (38%)	133 (33%)	172 (33%)	110 (40%)	87 (33%)	**62 (47%)^{*b}**	85 (32%)	48 (34%)	282 (35%)
Helicobacter pylori infection (n; %)	36 (7%)	36 (7%)	42 (6%)	30 (8%)	23 (7%)	13 (6%)	19 (6%)	17 (9%)	72 (7%)

BMI, body mass index; SDS, standard deviation score. To test for differences in characteristics between the groups, Chi-square (for categorical variables), Independent Samples T-tests (for normally distributes variables), and Mann-Whitney U tests (for non-normally distributed continuous variables) were applied. Bold numbers indicate significant differences between the groups $^{*}p<0.05$ $^{**}p<0.01$ $^{***}p<0.001$. We had missing data on BMI for age (n=4, <1%), breastfeeding duration (n=323; 30%), breastfeeding at 6 months (n=195, 18%), daycare 484 (45%), siblings n=15; 1%), educational level (n=41; 4%), net household income (109; 10%), antibiotic use (279 (26%), helicobacter pylori infection (n=2; <1%).

a. CMV+ versus CMV-

b. Male CMV+ versus male CMV-

c. Female CMV+ versus female CMV-

No significant differences in characteristics were observed between males and females, and between male CMV+ and female CMV+.

Chapter 6

Abnormalities in CD57+ cytotoxic T cells and Vδ1+ TCRγδ cells in subclinical celiac disease in childhood are affected by Cytomegalovirus: The Generation R Study

Michelle A.E. Jansen, Diana van den Heuvel , Vincent W.V. Jaddoe , Menno C. van Zelm, Henriette A. Moll

Clin Immunol. 2017 Apr 26. pii: S1521-6616(17)30278-4.

ABSTRACT

Celiac disease (CeD) is a digestive and autoimmune disorder driven by an immune response to modified gluten peptides. Affected intestines show infiltrates of various T-cell and NK-cell subsets. It is currently unclear if individuals with subclinical CeD have systemic abnormalities in immune cells. We here studied whether subclinical CeD is associated with changes in blood CD57-expressing and Vδ1-expressing lymphocytes in children, and whether cytomegalovirus (CMV) infection modifies this association. Included were 1,068 children from the Generation R Study. Serum Immunoglobulin G (IgG) levels against CMV were measured by ELISA; Tissue transglutaminase type 2 antibody (TG2A) levels with fluorescence enzyme immunoassay (FEIA). Duodenal biopsies, additional Human Leukocyte Antigen (HLA) DQ 2.2, 2.5 and 8 and endomysial antibody (EMA) typing were performed in TG2A positive children. Subclinical CeD cases (n=12) had 1.8 fold (95% CI 1.06; 3.1) fewer Vδ1+ T cells which was predominantly observed in CMV seronegative children (p-interaction 0.02), and 2.7 fold (95% CI 1.25; 5.99) more CD57+ T cells than HLA DQ2/-DQ8 positive controls (n=339). Hence, children with subclinical CeD have alterations in specific blood T cell subsets that are linked to viral pathology. The observed interaction effect between subclinical CeD and CMV may contribute to the understanding of disease pathogenesis.

INTRODUCTION

Celiac disease is a chronic T-cell mediated systemic condition, driven by an immune response to modified gluten peptides. The interaction between gluten-specific CD4+ T cells and IL-15 in the intestinal epithelium activates CD8+ intraepithelial T lymphocytes (IELs), causing a cascade of inflammatory responses leading to small intestinal enteropathy.[1, 2] As modified gluten peptides can only be presented by HLA DQ2.2, HLA DQ2.5 or HLA DQ8 molecules, carriership of these alleles is a prerequisite for CeD.[2, 3] Still, 40% of the general population carries one of these alleles, and only a minority develops celiac disease.[3, 4] Pathogenesis is therefore considered multifactorial, including a role for environmental factors, such as infections.[2]

Infiltration of both αβ and γδ T cells into the small intestinal epithelium marks one of the earliest stages in celiac disease pathogenesis.[5, 6] Although the majority of intestinal γδ T cells express a Vδ1 chain, Vδ1+ T cells are relatively rare in blood.[5] Still, they possess powerful cytotoxic, immune suppressive, tumoricidal and regulatory properties.[7-10] The majority of previous studies reported expansions of total γδ T and Vδ1+T cells in the gut of celiac disease patients.[5, 9, 11, 12] Based on a concurrent decrease in blood Vδ1+ T cell numbers,[5, 13] this is potentially due to recruitment to the intestinal epithelium in response to stressed epithelium or to repair tissue damage.[9, 14, 15] In patients with latent celiac disease, intraepithelial γδ T cells can already be prominent several years before appearance of mucosal atrophy.[16] This early recruitment of Vδ1+ T cells from blood to the intestinal mucosa potentially occurs even prior to the onset of intestinal inflammation, making blood Vδ1+ T-cell numbers a promising marker for subclinical celiac disease. Not all studies reported changes in blood cell numbers.[5, 11] However, the study of Kutlu et al included a heterogeneous study population consisting of both children and adults, with varying degrees of mucosal atrophy and differences in gluten intake.[11] Decreased blood Vδ1+ T cells have been found in adults with celiac disease, but were not affected in children.[5] Nonetheless, evidence in children is scarce: Only one study reported on these associations in children.[5]

In contrast to expanded γδ T-cell populations, decreased numbers of NK and NKT cells have been observed within IELs of patients with celiac disease.[5, 12] NK and NKT cell numbers were affected by gluten intake and degree of enteropathy, and were also found to be decreased in blood.[5, 12] Still, these associations have been found in adults, but not in children,[5] suggesting that associations may be different between children and adults. Chronic stimulation of T cells, as observed in autoimmunity or chronic infection, can result in development of CD8+ T cells that are capable of cytokine production, yet incapable of cell proliferation.[17] CD57 is a marker of such proliferative inability or T cell exhaustion. CD57-expressing T- and NK-cell numbers are reported to be increased in cytomegalovirus (CMV) carriers, as well as in patients with autoimmune diseases, including type 1 diabetes mellitus, rheumatoid arthritis and dermatomyositis.[17-19] Moreover, it is thought to predict the course of autoimmunity.[18-20] However, it is not known whether blood CD57+ expressing

lymphocytes are abnormal in celiac disease. Because CD57 is a marker of large granular leukemia of T-cell origin, one of the complications of refractory celiac disease,[21] it could be hypothesized that in subclinical celiac disease, chronic inflammation already drives the expansion of these cells. Although CMV can drive expansions of CD57+ T cells, NK cells and NKT cells, as well as Vδ1+ T cells,[22, 23] it has been inversely related to celiac disease.[24, 25] Thus, it remains unclear if CMV infection confounds or modifies associations between celiac disease and these blood cells.

We studied blood lymphocytes in children with subclinical celiac disease with the aim to identify associations with blood Vδ1+ γδT-cell, CD57+ T-cell and CD57+ NK-cell numbers, and to assess whether these associations are modified by CMV seropositivity.

METHODS

Study design

This study was embedded within the Generation R study, a prospective population based cohort study from fetal life until young adulthood, described in detail previously.[26] Ethics approval was obtained from the Medical Ethical Committee of Erasmus MC, University Medical Center Rotterdam, The Netherlands. Written informed consent was obtained from all participants. Children were born between April 2002 and January 2006. At the age of 6 years, 6,690 children visited the research center.[26, 27] During this visit, blood samples were collected from 4,593 children. Data on Tissue Transglutaminase Type 2 Antibody (TG2A) levels and CMV seropositivity was available in 4,442 children. Of these, HLA-DQ2 or HLA-DQ8 carriership was measured in 2,915 children (42% carried one of the alleles), and detailed immune phenotyping was available in 1,068 children. Combined, from a total of 352 HLA-DQ2 or HLA-DQ8 positive children, detailed immunophenotyping of leukocytes, serum TG2A levels and CMV seropositivity could be determined [Supplemental Figure 1].

Tissue Transglutaminase Type 2 Antibodies, HLA risk alleles and Celiac Disease diagnosis

Tissue Transglutaminase type 2 Antibody (TG2A) serum levels (IgA) were measured at the median age of 6.0 years using a fluorescence enzyme immunoassay (Elia Celikey IgA, Phadia ImmunoCAP 250, Phadia AB, Uppsala, Sweden).[27] Sera with a TG2A level of 7 U/ml or higher were considered to be positive per manufacturer's instructions. HLA-DQ2 or HLA-DQ8 carriership in the total cohort was assessed based on DNA from cord blood cells, using a tag single nucleotide polymorphism (SNP) approach,[28] as described in detail previously.[29] Of 13 TG2A positive children who carried one of the genetic risk alleles and anti-endomysial antibodies (EMA), 12 children were diagnosed with celiac disease according to the ESPGHAN guideline.[30] All 12 subclinical celiac disease cases were identified in a

screening-based setting, and did not report more frequent or more severe gastrointestinal symptoms than controls (own unpublished results) [Supplemental Methods].

Immunophenotyping of CD57+ T cell and Vδ1+ T–cell subsets

Detailed 6-color flow cytometry was performed within 24 hours following sampling on fresh whole blood at the median age of 6.0 years.[31] Absolute cell numbers and relative frequencies of blood CD3+ T cells, and CD3-CD16/56+ NK cells were obtained with a lyse-no-wash- protocol. 6-color flow cytometry was performed on an LSRII (BD Biosciences) to distinguish CD57-expressing T (CD3+), NK (CD3-CD16/56+) and NKT (CD3+CD16/56+) cells, as well as total, Vδ1+ and Vδ2+ γδT cells [Supplemental Tables 1 and 2].

Cytomegalovirus serology

Venous blood samples, taken at the median age of 6.0 years, were analysed using enzyme immunoassays for IgG antibodies against CMV (EUROIMMUN®, Lübeck, Germany).[31] Results were evaluated semiquantitatively. Per manufacturer's instructions, a ratio of the test sample over a manufacturer-provided threshold sample ≥1.1 was defined as positive.

Covariates

Data on potential confounders, such as gender and mode of delivery were obtained from obstetric records from hospitals and mid-wife practices.[32] Data on sociodemographic and lifestyle factors were obtained by a combination of pre- and postnatal questionnaires completed by both parents. Prenatal questionnaires included information on ethnicity (Western/non-Western),[33] maternal smoking during pregnancy, maternal educational level (low/high),[34] and household income per month (<€2,200 and >€2,200).[35] Postnatal questionnaires included information on breastfeeding, day-care attendance, gastro-intestinal and respiratory tract infections and antibiotic use during the first year of life.

Statistical analysis

First, Chi-square tests and non-parametric Mann-Whitney U tests were applied to test whether maternal and child characteristics were different between children with and without subclinical celiac disease. Mann-Whitney U tests were applied to test for differences in median blood cell numbers between children with and without subclinical celiac disease. Because cell numbers were not normally distributed, these were subsequently natural log transformed. Multivariable linear regression models with adjustment for potential confounders were created with celiac disease diagnosis as independent variable, and natural log transformed blood cell numbers as dependent variable. Effect estimates (beta, β) and their 95% confidence intervals (CI) on the natural logarithm scale were back transformed (exponent β) for interpretation purposes. To assess whether the association between celiac disease and blood CD57+ T-cell numbers were modified by CMV seropositivity, statistical interaction was evaluated by adding the product term of the independent

variable and subgroup (independent variable * subgroup) as covariate in the model. If the interaction term was significant (p<0.10), we stratified our analyses. Second, to examine whether CMV seropositivity was associated with natural log transformed blood CD57+ T cell and Vδ1+ T- cell numbers, multivariable linear regression models were created. The results were reported in this paper as beta's (β's) and 95% confidence intervals (CIs). A p-value ≤0.05 was considered statistically significant. Statistical analyses were performed in SPSS 20.0 for Windows (SPSS Inc., Chicago, IL).

RESULTS

Subject characteristics

Maternal and child characteristics are shown in [Table 1]. Of n=352 children, 13 had increased serum TG2A levels, of whom 12 children were classified as having subclinical celiac disease [Supplemental Figure 1, Table 1]. Children with subclinical celiac disease were slightly younger (5.9 vs 6.0 years) than HLA DQ2/DQ8 positive controls (n=339), but no significant differences were observed in socioeconomic position, child's gender, body mass index, ethnicity, gastrointestinal or respiratory tract infections in the first year of life [Table 1].

CeD is associated with increased CD57+ T–cell and decreased Vδ1+ T–cell numbers

Children with subclinical celiac disease had higher blood CD57+ T-cell numbers (median 110.9 cells/μL; range 20.3-365.8) than controls (median 53.1 cells/μL; range 1.8-613.5). This was due to higher CD8+CD57+ T cell numbers (77.3 cells/μL; range 14.3-297.7 versus 35.9 cells/μL range 1.3-540.7), and higher CD4+CD57+ T cell numbers (28.7 cells/μL; range 6.2-94.6 versus 15.2 cells/μL range 1.4-145.2). No significant differences were observed for CD57+ NKT cell nor CD57+ NK-cell numbers [Figures 1 and 2A]. Similar differences were found when frequencies of CD57-expressing cells within their parent populations were compared between subclinical celiac disease cases and controls [Supplemental Table 3]. Absolute numbers of total, Vδ1+ and Vδ2+ γδT cells were not significantly different between subclinical celiac disease cases and controls [Figure 1B]. Still, as a result of a slightly lower Vδ1+ cells and slightly higher Vδ2+ cells, Vδ1+T-cell frequencies were significantly lower (median 12.2%; range 9.6-30.5) in subclinical celiac disease cases than in controls (median 19.2%; range 3.4-60.9) [Supplemental Table 3].

Adjustment for potential confounding factors led to an increase in the differences for CD57+ T-cell numbers: subclinical celiac disease cases had 2.4-fold (95% CI 1.36; 4.31) more CD57+ T cells, 2.6-fold more (95% CI 1.35; 4.86) CD57+CD8+ T cells, and 2.0-fold more (95% CI 1.14; 3.68) CD57+ CD4+ T cells than controls (p=0.003, 0.004, 0.017 resp.) [Figure 2]. The adjustment did not improve fold changes nor p-values for total, Vδ1+ and Vδ2+ γδT cell numbers. The results for the observed associations persisted after including HLA-DQ2/-DQ8 negative controls [Supplemental Figure 2].

Table 1. Characteristics of study population

HLA DQ2/DQ8 carriers only (n=351) Characteristics	No Celiac Disease n=339 (96.6%)	Subclinical Celiac Disease n=12 (3.4%)
Maternal characteristics		
High educational level (n;%)	228 (69.1%)	6 (50%)
Household income per month (>2200 euro) (n;%)	232 (74.6%)	8 (66.7%)
Smoking during pregnancy (n;%)	85 (27.6%)	4 (40%)
Caesarian section (n;%)	34 (10.8%)	3 (27.3%)
Multiparity (n;%)	143 (42.3%)	3 (25%)
Child characteristics		
Female (n;%)	165 (48.7%)	8 (66.7%)
Western ethnicity (n;%)	308 (90.9%)	11 (91.7%)
Age (years; median; range)	**6.04 (5.5-7.9)**	**5.88 (5.8-6.2)**
SDS BMI for age	0.17 (-2.3 ; 3.0)	-0.29 (-2.0 - 1.7)
Gluten free/low diet (n;%)	1 (0.3%)	0 (0%)
Breastfeeding ever (n;%)	265 (89.2%)	12 (100%)
Day care attendance first year (n;%)	183 (91.5%)	7 (78%)
Gastro intestinal tract infections first year (n;%)	82 (47.7%)	5 (55.6%)
Respiratory tract infections first year (n;%)	135 (50.9%)	7 (70%)
Use of antibiotics / penicillin first year (n;%)	84 (32.1%)	4 (40%)
Herpesvirus seropositivity at 6 years (n;%)		
CMV	107 (32%)	3 (25%)
EBV	148 (44%)	5 (42%)
HSV-1	56 (17%)	1 (8%)

Values are medians (range). Bold values indicate significant differences (p<0.05). Differences in maternal and child characteristics between groups with and without subclinical celiac disease were calculated with the non-parametric Mann-Whitney U test for non-normally distributed continuous variables), and the chi-square test for categorical variables.

Cytomegalovirus modifies the association between celiac disease, CD57+ T cells, and Vδ1+ T cells.

We further studied associations between CMV infection, CD57+ T cells, and Vδ1+ T cells [Figure 3]. CMV seropositivity was associated with a 3.52 fold increase (95% CI 2.64; 4.71) in CD57+ T-cell numbers (p<0.0001), and a 1.3 fold increase (95% CI 1.01; 1.68) in Vδ1+ T-cell numbers (p=0.039) in the HLA DQ2/DQ8 positive children. Associations were similar between children with and without subclinical celiac disease (p-interaction>0.14), and did not change after inclusion of the HLA DQ2/DQ8 negative controls [data not shown].

CMV seropositivity did not significantly modify the association between CeD diagnosis and CD57+ T cells (p-interaction=0.17). Still, associations between CD57+ T-cell numbers and subclinical celiac disease were stronger and merely present in CMV seronegative children. Children with subclinical celiac disease who were CMV seronegative had 2.54 fold (95% CI 1.17; 5.53) more CD57+ T-cell numbers than children who were CMV seronegative without celiac disease, mainly because they had more blood CD8+CD57+ T cells [Figure 4].

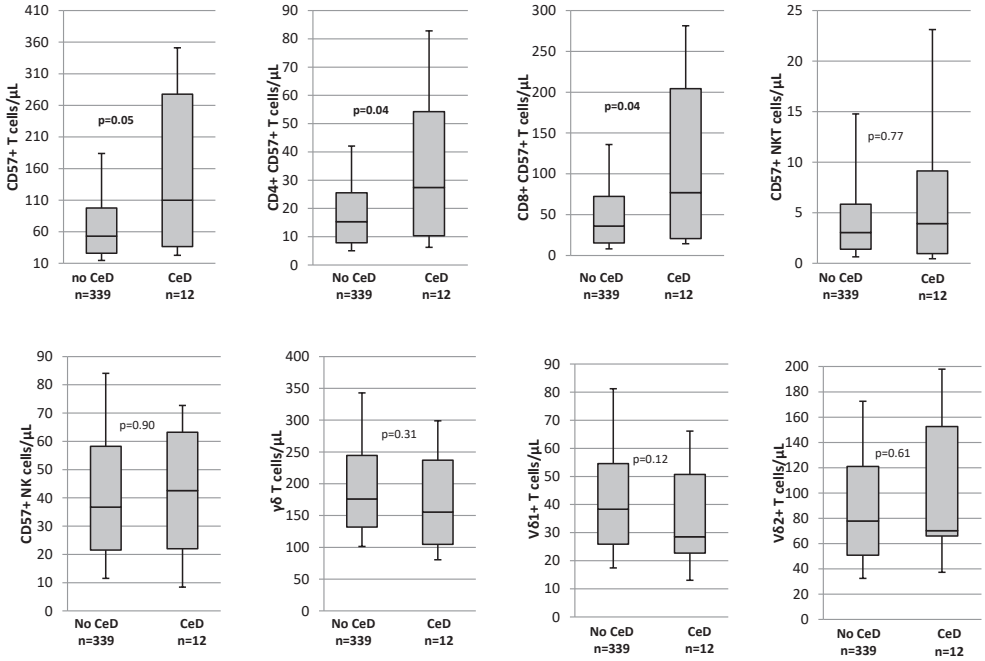


Figure 1. Difference in blood median lymphocyte cell numbers (cell/µL) between 6-year old children with and without subclinical celiac disease. A, Flowcytometric gating strategy. B, Children with subclinical CD (n=12), HLA DQ2 or DQ8 positive controls (n=339). Mann-Whitney U tests were performed to test for differences between continuous non-normally distributed variables.

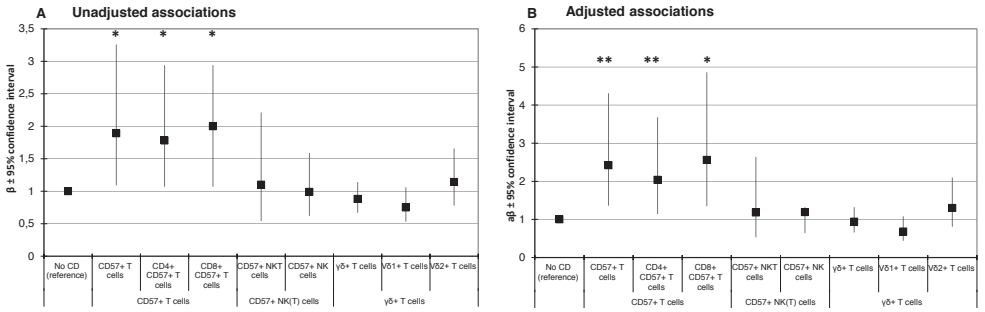


Figure 2. Relative differences in blood lymphocyte numbers between children with and without subclinical celiac disease. A, Mean lymphocyte numbers (cell/µL) in blood of children with subclinical CD (n=12), standardized to the same cell numbers in HLA DQ2/DQ8 positive controls (n=339) (unadjusted). B, Mean lymphocyte numbers (cell/µL) in blood of children with subclinical CD, standardized to the same cell numbers in HLA DQ2/DQ8 positive controls (adjusted for child's age, SDS BMI for age, gender, CMV, EBV and HSV-1 seropositivity, ethnicity (western/non-western), socioeconomic position (maternal educational level), respiratory and gastrointestinal tract infections, daycare attendance (yes/no), and breastfeeding (ever/never). The estimated back-transformed (a)β and 95% CI (exp Ln β ± exp Ln (95% CI) are derived from univariable (A) and multivariable (B) linear regression models, and indicate the relative difference in mean cell numbers between children with and without subclinical CD. *p<0.05, **p<0.01.

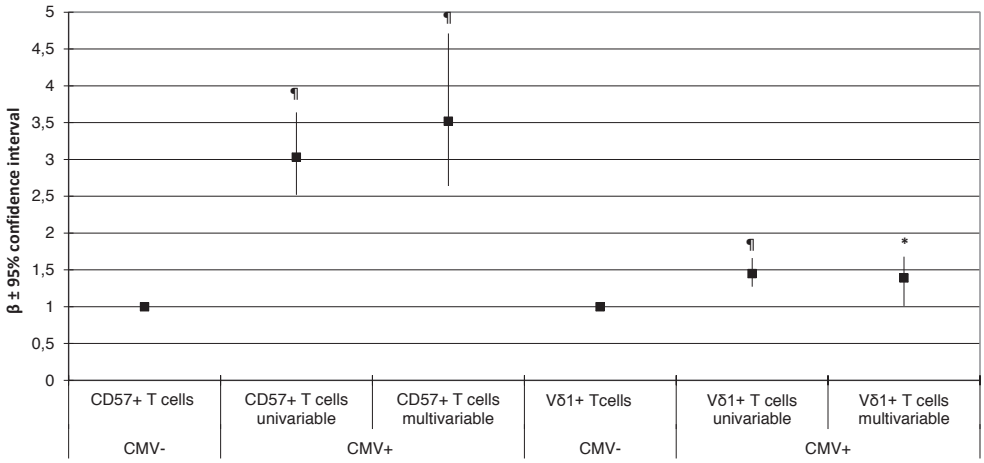


Figure 3. Associations between CMV seropositivity and blood CD57+ T cells, and Vδ1+ TCRγδ cells. Mean lymphocyte numbers (cell/μL) in peripheral blood of CMV seropositive children (n=110), standardized to the same cell numbers in CMV seronegative controls (n=241) within a HLA DQ2/DQ8 positive population. The estimated back-transformed (a)β and 95% CI (exp Ln β ± exp Ln (95% CI) are derived from univariable and multivariable linear regression models, and indicate the relative difference in mean cell numbers between children with and without CMV *<0.01, ¶p<0.0001**.** Multivariable models are adjusted for child's age, SDS BMI for age, gender, CMV, EBV and HSV-1 seropositivity, ethnicity (western/non-western), socioeconomic position (maternal educational level), respiratory and gastrointestinal tract infections, daycare attendance (yes/no), and breastfeeding (ever/never).

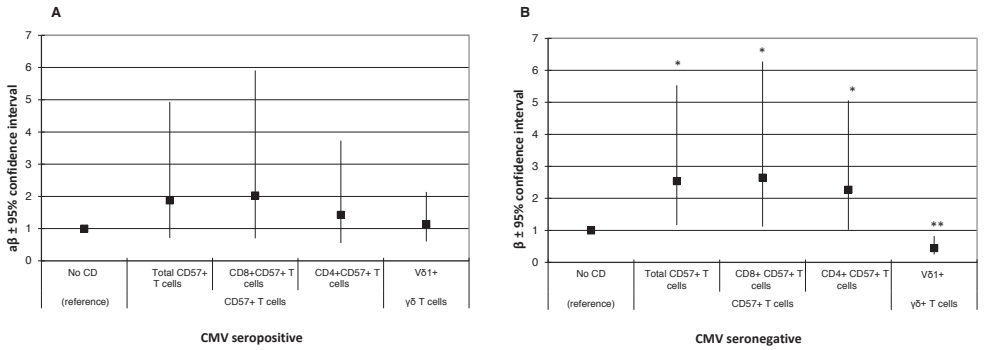


Figure 4. Associations between CD, blood CD57+ T cells, and Vδ1+ TCRγδ cells, stratified on CMV seropositivity.

A, Mean lymphocyte numbers (cell/μL) in peripheral blood of children with subclinical celiac disease, standardized to HLA DQ2/DQ8 positive controls in CMV seropositive children (n=110). B, Mean lymphocyte numbers (cell/μL) in peripheral blood of children with subclinical celiac disease, standardized to HLA DQ2/DQ8 positive controls in CMV seronegative children (n=241). The estimated back-transformed aβ and 95% CI (exp Ln β ± exp Ln (95% CI) are derived from multivariable linear regression models, and indicate the relative difference in mean cell numbers between children with and without subclinical celiac disease. *p<0.05, **p<0.01. Models are adjusted for child's age, SDS BMI for age, gender, CMV, EBV and HSV-1 seropositivity, ethnicity (western/non-western), socioeconomic position (maternal educational level), respiratory and gastrointestinal tract infections, daycare attendance (yes/no), and breastfeeding (ever/never).

The inverse association between Vδ1+ T cells and subclinical celiac disease was only observed in CMV seronegative children (p-interaction=0.02). Children with subclinical celiac disease who were CMV seronegative had 2.22 fold (95% CI 1.22; 4.0) lower Vδ1+T-cell numbers than children who were CMV seronegative without celiac disease [Figure 4]. The results for the observed stratified associations persisted after including HLA-DQ2/-DQ8 negative controls [Supplemental Figure 3]. Thus, children with subclinical celiac disease had more blood CD57+ expressing T cells, and fewer blood Vδ1+ expressing T cells than controls, particularly in CMV seronegative children.

EBV and HSV-1 seropositivity did not significantly modify the associations between CeD diagnosis and Vδ1**+** cells (p-interaction 0.19, and 0.75 resp.), nor between CeD diagnosis and CD57+ T cells (p-interaction 0.95 and 0.62 resp.) in the HLA DQ2/DQ8 positive group. Additionally, no statistical interactions were observed after inclusion of the HLA DQ2/DQ8 negative children.

DISCUSSION

We showed in a population-based cohort study that 6-year-old children with subclinical celiac disease had higher numbers of CD57+ CD8+ T-cells, and lower numbers of Vδ1+ T cells in blood than healthy controls of the same age who carried the genetic risk type for celiac disease. The inverse association between blood Vδ1+ T cells and subclinical celiac disease was most prominent in CMV seronegative children.

To the best of our knowledge, this is the first study examining blood immune cell numbers in children with subclinical celiac disease. Most studies have been performed in adult patients with symptomatic celiac disease,[5, 13] whereas studies including patients with subclinical disease are scarce.[36] In addition, our observation that CMV infection modifies the association between celiac disease and blood Vδ1+ T-cell numbers is novel, and extends the few previous studies that showed similar abnormalities in innate-like lymphocyte populations in celiac disease, specifically Vδ1+ T cells.[5, 13] However, most studies merely studied total γδ T cells,[11, 36] or total NK and T cells,[5, 12] but lacked detailed measurements of Vδ1+ or CD57+ expression. Moreover, only one study addressed these associations in children with celiac disease;[5] the majority of studies have been performed in adults. Besides, results on total NK and T cell-numbers have been contradictive.[5, 12] Conflicting results might be due to differences in age, stage of disease, variability in the quality of duodenal biopsies, genetic background, geographical region, and environmental factors, such as viral infections. The latter two factors are known to affect the composition of circulating T cells,[22, 36] but have not been included in previous studies. Moreover, we show that differences in blood cell numbers between cases and controls persisted after including HLA-DQ2/-DQ8 negative individuals from the control group. Thus, carriership of the HLA risk alleles does not affect the observed differences in cell numbers between

subclinical celiac disease cases and controls. Our observations are in line with a previous study that suggested that the altered distribution of γδ T-cell subsets in celiac patients is a consequence of an ongoing immunological process, rather than a phenomenon directly caused by the genetic risk allele.[13] Our observations do not provide direct insights into causality between abnormal cell numbers and celiac disease. Still, the cell types that were studied could be directly involved in disease pathogenesis.

First, we observed a positive association between celiac disease and blood CD57+ T-cell numbers, suggesting that celiac disease is associated with proliferative inability or T cell exhaustion.[17] Increased numbers of CD57+ T cells are reported in patients with frequent or chronic infections, such as CMV, or in patients with chronic stimulation of T cells.[17] Hence, the observed increased CD57 expression in subclinical celiac disease patients is likely to be a result of chronic inflammation, rather than a specific marker for celiac disease. Despite previous reports of increased CD57+ T-cell numbers in CMV carriers, our associations were still present, and even more pronounced, in the seronegative children. Hence, both CMV and celiac disease were independently associated with increased CD57+ T-cell numbers. Alternatively, this increase might reflect chronic immune activation induced by other chronic viruses, or chronic stimulation of T cells induced by celiac disease itself. Pathways involved in the proliferation of T cells, include signaling via IL-15 or IL-21.[37, 38] These pleiotropic cytokines are also greatly upregulated in the gut of celiac disease patients.[39] Moreover, IL-21 levels are increased in children with potential celiac disease, although to a lesser extent than in active disease.[40] Still, it is likely that on top of IL-15 and IL-21, other factors are involved, because CD57+ T cells are a heterogeneous mixture of cytotoxic and immunosuppressive subsets.[19]

Second, we identified an inverse association between blood Vδ1+ T cells and the presence of subclinical celiac disease. This is in line with previous observations in celiac disease patients of increased Vδ1+ T-cell numbers in gut epithelium,[5, 9, 11, 12] with a decrease in blood. As previous studies reported blood Vδ1+ T-cell numbers to be associated with the degree of small intestinal enteropathy,[12] and to depend on the amount of gluten exposure,[41] this association may be causal. The first stage of celiac disease is reflected by an increased number of IELs in the villi,[6] which are predominantly of the Vδ1 subtype.[5] These cells have a distinct and limited T cell receptor diversity, and play an important role in intestinal homeostasis and oral mucosal tolerance.[42-44] Although relatively little is known about the biological function of blood Vδ1+ T cells,[7] they have been shown to secret inflammatory, as well as immunosuppressive and regulatory cytokines,[8] including IFN-γ, TNF-α, IL-10 and IL-17, suggesting a role in autoimmune and allergic diseases.[42] In addition, Vδ1+ T cells may have a pivotal role in cancer surveillance based on their IL-2 and IL-15 mediated anti-tumor cytolytic function.[45] Our observed decrease in blood Vδ1+ T-cell numbers in celiac disease directly contrasts findings from inflammatory bowel disorders, where an increase in blood Vδ1+ T cells is noted.[5] This might directly reflect a difference in disease pathology. Still, the

functional significance of these changes in relation to subclinical celiac disease remains unclear.

Third, our results are in line with previous studies suggesting that CMV infection induces expansion of peripheral Vδ1+ T cells.[22] Combined with the suggested regulatory properties of Vδ1+ T cells, and the inverse association between celiac disease and CMV,[24] it is tempting to speculate that decreased blood Vδ1+ T-cell numbers reflect celiac disease activity, whereas early CMV infection might prevent celiac disease by increasing blood Vδ1+ T cells. On the other hand, it could be speculated that decreased peripheral Vδ1+ T cells in celiac disease patients reflect the absence of CMV infection, rather than a celiac disease specific pathway.

Methodological considerations

To our knowledge, this is the first population-based cohort study that examined associations between subclinical celiac disease, CD57+ T cells and Vδ1+ T cells in combination with CMV seropositivity in children. A major strength of the present study is that we used a multivariable approach to assess the relative contributions of both CMV seropositivity and subclinical celiac disease on immune subsets. Second, we examined children with an early presentation of celiac disease confirmed by biopsy, without clinical signs and symptoms, instead of symptomatic children. Furthermore, we performed our analyses in a large sample, and information on a broad range of potential confounders was available, including socioeconomic status, ethnicity, breastfeeding, daycare, gastrointestinal and respiratory tract infections and antibiotic use. However, some limitations should be taken into account. First, despite a large number of healthy controls, the number of subclinical CeD cases was limited to only 12 children. Therefore, the results of our analyses should be interpreted with caution. Second, causality remains uncertain, since detailed immune phenotyping and antibodies against CMV were measured at the same time as TG2A levels. Therefore, reverse causation cannot be excluded. Third, although we adjusted for a broad range of potential confounders, the observed associations might still be related to unmeasured variables, such as lifestyle, other viral infections and microbial burden. Moreover, CMV carriership may be associated with other T-cell subsets that we did not include in our analyses, thereby possibly influencing the observed associations between celiac disease and CD57+ T cells. Hence, residual confounding cannot be fully excluded. Last, we did not have data on effector functions of studied lymphocytes, such as the expression of effector markers (e.g. FOXP3), or cytokines such as IL-10 and TGF-β, precluding conclusions on effector or regulatory mechanisms of CD57+ T cells and Vδ1 cells.

Conclusion

This population-based cohort study showed that alterations in Vδ1 and CD57 expressing T-cell numbers were already present in children with subclinical CeD, prior to the occurrence of overt clinical symptoms. Associations were dependent on cytomegalovirus seropositiv-

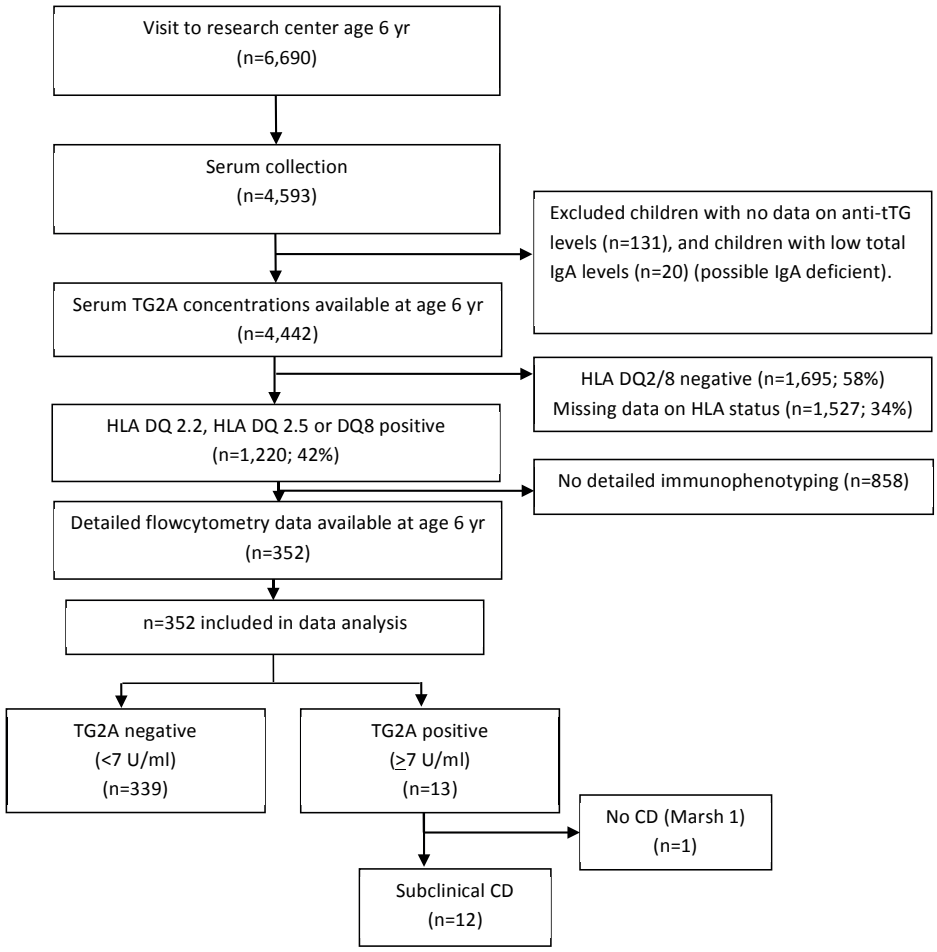
ity. Hence, alterations in blood T cells in subclinical CeD are linked to viral pathology, which may contribute to the understanding of CeD pathogenesis.

REFERENCES

1. N. Korneychuk, E. Ramiro-Puig, J. Ettersperger, J. Schulthess, N. Montcuquet, H. Kiyono, B. Meresse, N. Cerf-Bensussan, Interleukin 15 and CD4(+) T cells cooperate to promote small intestinal enteropathy in response to dietary antigen, Gastroenterology, 146 (2014) 1017-1027.
2. J. van Bergen, C.J. Mulder, M.L. Mearin, F. Koning, Local communication among mucosal immune cells in patients with celiac disease, Gastroenterology, 148 (2015) 1187-1194.
3. R. Dieli-Crimi, M.C. Cenit, C. Nunez, The genetics of celiac disease: A comprehensive review of clinical implications, J Autoimmun, 64 (2015) 26-41.
4. J. Romanos, A. Rosen, V. Kumar, G. Trynka, L. Franke, A. Szperl, J. Gutierrez-Achury, C.C. van Diemen, R. Kanninga, S.A. Jankipersadsing, A. Steck, G. Eisenbarth, D.A. van Heel, B. Cukrowska, V. Bruno, M.C. Mazzilli, C. Nunez, J.R. Bilbao, M.L. Mearin, D. Barisani, M. Rewers, J.M. Norris, A. Ivarsson, H.M. Boezen, E. Liu, C. Wijmenga, Improving coeliac disease risk prediction by testing non-HLA variants additional to HLA variants, Gut, 63 (2014) 415-422.
5. M.R. Dunne, L. Elliott, S. Hussey, N. Mahmud, J. Kelly, D.G. Doherty, C.F. Feighery, Persistent changes in circulating and intestinal gammadelta T cell subsets, invariant natural killer T cells and mucosal-associated invariant T cells in children and adults with coeliac disease, PLoS One, 8 (2013) e76008.
6. A. Ferguson, D. Murray, Quantitation of intraepithelial lymphocytes in human jejunum, Gut, 12 (1971) 988-994.
7. F. Hua, N. Kang, Y.A. Gao, L.X. Cui, D.N. Ba, W. He, Potential regulatory role of in vitro-expanded Vdelta1 T cells from human peripheral blood, Immunol Res, 56 (2013) 172-180.
8. G.M. Siegers, L.S. Lamb, Jr., Cytotoxic and regulatory properties of circulating Vdelta1+ gammadelta T cells: a new player on the cell therapy field?, Mol Ther, 22 (2014) 1416-1422.
9. G. Bhagat, A.J. Naiyer, J.G. Shah, J. Harper, B. Jabri, T.C. Wang, P.H. Green, J.S. Manavalan, Small intestinal CD8+TCRgammadelta+NKG2A+ intraepithelial lymphocytes have attributes of regulatory cells in patients with celiac disease, J Clin Invest, 118 (2008) 281-293.
10. W.H. Verbeek, B.M. von Blomberg, P.E. Scholten, D.J. Kuik, C.J. Mulder, M.W. Schreurs, The presence of small intestinal intraepithelial gamma/delta T-lymphocytes is inversely correlated with lymphoma development in refractory celiac disease, Am J Gastroenterol, 103 (2008) 3152-3158.
11. T. Kutlu, N. Brousse, C. Rambaud, F. Le Deist, J. Schmitz, N. Cerf-Bensussan, Numbers of T cell receptor (TCR) alpha beta+ but not of TcR gamma delta+ intraepithelial lymphocytes correlate with the grade of villous atrophy in coeliac patients on a long term normal diet, Gut, 34 (1993) 208-214.
12. S. Calleja, S. Vivas, M. Santiuste, L. Arias, M. Hernando, E. Nistal, J. Casqueiro, J.G. Ruiz de Morales, Dynamics of non-conventional intraepithelial lymphocytes-NK, NKT, and gammadelta T-in celiac disease: relationship with age, diet, and histopathology, Dig Dis Sci, 56 (2011) 2042-2049.
13. T.O. Kerttula, K. Holm, J. Partanen, A. Polvi, M. Maki, Circulating T lymphocyte subsets in coeliac disease (CoD) patients and healthy family members, Clin Exp Immunol, 111 (1998) 536-540.
14. S. Itohara, A.G. Farr, J.J. Lafaille, M. Bonneville, Y. Takagaki, W. Haas, S. Tonegawa, Homing of a gamma delta thymocyte subset with homogeneous T-cell receptors to mucosal epithelia, Nature, 343 (1990) 754-757.
15. A. Hayday, R. Tigelaar, Immunoregulation in the tissues by gammadelta T cells, Nat Rev Immunol, 3 (2003) 233-242.
16. M. Maki, K. Holm, P. Collin, E. Savilahti, Increase in gamma/delta T cell receptor bearing lymphocytes in normal small bowel mucosa in latent coeliac disease, Gut, 32 (1991) 1412-1414.
17. J.M. Brenchley, N.J. Karandikar, M.R. Betts, D.R. Ambrozak, B.J. Hill, L.E. Crotty, J.P. Casazza, J. Kuruppu, S.A. Migueles, M. Connors, M. Roederer, D.C. Douek, R.A. Koup, Expression of CD57 defines replicative senescence and antigen-induced apoptotic death of CD8+ T cells, Blood, 101 (2003) 2711-2720.

18. C.M. Nielsen, M.J. White, M.R. Goodier, E.M. Riley, Functional Significance of CD57 Expression on Human NK Cells and Relevance to Disease, Front Immunol, 4 (2013) 422.
19. M. Strioga, V. Pasukoniene, D. Characiejus, CD8+ CD28- and CD8+ CD57+ T cells and their role in health and disease, Immunology, 134 (2011) 17-32.
20. E.F. McKinney, J.C. Lee, D.R. Jayne, P.A. Lyons, K.G. Smith, T-cell exhaustion, co-stimulation and clinical outcome in autoimmunity and infection, Nature, 523 (2015) 612-616.
21. G. Malamut, B. Meresse, V. Verkarre, S. Kaltenbach, N. Montcuquet, J.P. Duong Van Huyen, C. Callens, J. Lenglet, G. Rahmi, E. Samaha, B. Ranque, E. Macintyre, I. Radford-Weiss, O. Hermine, N. Cerf-Bensussan, C. Cellier, Large granular lymphocytic leukemia: a treatable form of refractory celiac disease, Gastroenterology, 143 (2012) 1470-1472 e1472.
22. A. Roux, G. Mourin, M. Larsen, S. Fastenackels, A. Urrutia, G. Gorochov, B. Autran, C. Donner, D. Sidi, J. Sibony-Prat, A. Marchant, M. Stern, D. Sauce, V. Appay, Differential impact of age and cytomegalovirus infection on the gammadelta T cell compartment, J Immunol, 191 (2013) 1300-1306.
23. E. Sohlberg, S. Saghafian-Hedengren, E. Rasul, G. Marchini, C. Nilsson, E. Klein, N. Nagy, E. Sverremark-Ekstrom, Cytomegalovirus-seropositive children show inhibition of in vitro EBV infection that is associated with CD8+CD57+ T cell enrichment and IFN-gamma, J Immunol, 191 (2013) 5669-5676.
24. L. Plot, H. Amital, O. Barzilai, M. Ram, N. Bizzaro, Y. Shoenfeld, Infections may have a protective role in the etiopathogenesis of celiac disease, Ann N Y Acad Sci, 1173 (2009) 670-674.
25. M.A. Jansen, D. van den Heuvel, K.V. van der Zwet, V.W. Jaddoe, A. Hofman, J.C. Escher, P.L. Fraaij, H. Hooijkaas, M.C. van Zelm, H.A. Moll, Herpesvirus Infections and Transglutaminase type 2 Antibody Positivity in Childhood: The Generation R Study, J Pediatr Gastroenterol Nutr, (2016).
26. C.J. Kruithof, M.N. Kooijman, C.M. van Duijn, O.H. Franco, J.C. de Jongste, C.C. Klaver, J.P. Mackenbach, H.A. Moll, H. Raat, E.H. Rings, F. Rivadeneira, E.A. Steegers, H. Tiemeier, A.G. Uitterlinden, F.C. Verhulst, E.B. Wolvius, A. Hofman, V.W. Jaddoe, The Generation R Study: Biobank update 2015, Eur J Epidemiol, 29 (2014) 911-927.
27. M.A. Jansen, J.C. Kiefte-de Jong, R. Gaillard, J.C. Escher, A. Hofman, V.W. Jaddoe, H. Hooijkaas, H.A. Moll, Growth trajectories and bone mineral density in anti-tissue transglutaminase antibody-positive children: the Generation R Study, Clin Gastroenterol Hepatol, 13 (2015) 913-920 e915.
28. A.J. Monsuur, P.I. de Bakker, A. Zhernakova, D. Pinto, W. Verduijn, J. Romanos, R. Auricchio, A. Lopez, D.A. van Heel, J.B. Crusius, C. Wijmenga, Effective detection of human leukocyte antigen risk alleles in celiac disease using tag single nucleotide polymorphisms, PLoS One, 3 (2008) e2270.
29. J.C. Kiefte-de Jong, V.W. Jaddoe, A.G. Uitterlinden, E.A. Steegers, S.P. Willemsen, A. Hofman, H. Hooijkaas, H.A. Moll, Levels of antibodies against tissue transglutaminase during pregnancy are associated with reduced fetal weight and birth weight, Gastroenterology, 144 (2013) 726-735 e722.
30. S. Husby, S. Koletzko, I.R. Korponay-Szabo, M.L. Mearin, A. Phillips, R. Shamir, R. Troncone, K. Giersiepen, D. Branski, C. Catassi, M. Lelgeman, M. Maki, C. Ribes-Koninckx, A. Ventura, K.P. Zimmer, European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease, J Pediatr Gastroenterol Nutr, 54 (2012) 136-160.
31. D. van den Heuvel, M.A. Jansen, W.A. Dik, H. Bouallouch-Charif, D. Zhao, K.A. van Kester, M.A. Smits-Te Nijenhuis, M.J. Kolijn-Couwenberg, V.W. Jaddoe, R. Arens, J.J. van Dongen, H.A. Moll, M.C. van Zelm, Cytomegalovirus- and Epstein-Barr Virus-Induced T-Cell Expansions in Young Children Do Not Impair Naive T-cell Populations or Vaccination Responses: The Generation R Study, J Infect Dis, 213 (2016) 233-242.
32. V.W. Jaddoe, C.M. van Duijn, O.H. Franco, A.J. van der Heijden, M.H. van Iizendoorn, J.C. de Jongste, A. van der Lugt, J.P. Mackenbach, H.A. Moll, H. Raat, F. Rivadeneira, E.A. Steegers, H. Tiemeier, A.G. Uitterlinden, F.C. Verhulst, A. Hofman, The Generation R Study: design and cohort update 2012, Eur J Epidemiol, 27 (2012) 739-756.

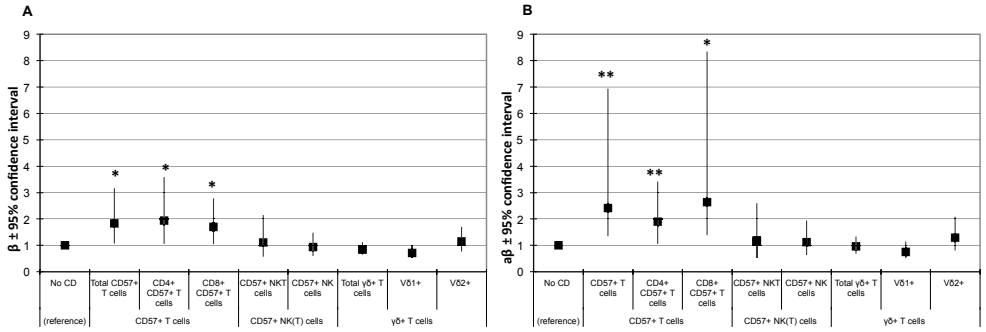
33. O. Swertz, P. Duimelaar, J. Thijssen, Migrants in the Netherlands 2004. Voorburg/Heerlen: Statistics Netherlands, 2004.
34. Statistics Netherlands, Dutch Standard Classification of Education 2003. Voorburg/Heerlen: Statistics Netherlands, 2004.
35. Statistics Netherlands, Welfare in the Netherlands. Income, welfare and spending of households and individuals., Den Haag/Heerlen: Statistics Netherlands, 2012.
36. H.H. Nataj Arab, M. Masjedi, F. Alsahebfosoul, M. Mokhtari, N. Jamali, M.H. Emami, A. Saffaei, Phenotypical characterization of the peripheral blood T cells in patients with celiac disease: does it differentiate suspicious celiac disease cases?, Gastroenterol Hepatol Bed Bench, 8 (2015) 132-139.
37. W.J. Leonard, R. Spolski, Interleukin-21: a modulator of lymphoid proliferation, apoptosis and differentiation, Nat Rev Immunol, 5 (2005) 688-698.
38. X.C. Li, G. Demirci, S. Ferrari-Lacraz, C. Groves, A. Coyle, T.R. Malek, T.B. Strom, IL-15 and IL-2: a matter of life and death for T cells in vivo, Nat Med, 7 (2001) 114-118.
39. R.W. DePaolo, V. Abadie, F. Tang, H. Fehlner-Peach, J.A. Hall, W. Wang, E.V. Marietta, D.D. Kasarda, T.A. Waldmann, J.A. Murray, C. Semrad, S.S. Kupfer, Y. Belkaid, S. Guandalini, B. Jabri, Co-adjuvant effects of retinoic acid and IL-15 induce inflammatory immunity to dietary antigens, Nature, 471 (2011) 220-224.
40. M. Borrelli, C. Gianfrani, G. Lania, R. Aitoro, K. Ferrara, M. Nanayakkara, D. Ponticelli, D. Zanzi, V. Discepolo, S. Vitale, M.V. Barone, R. Troncone, R. Auricchio, M. Maglio, In the Intestinal Mucosa of Children With Potential Celiac Disease IL-21 and IL-17A are Less Expressed than in the Active Disease, Am J Gastroenterol, 111 (2016) 134-144.
41. A. Han, E.W. Newell, J. Glanville, N. Fernandez-Becker, C. Khosla, Y.H. Chien, M.M. Davis, Dietary gluten triggers concomitant activation of CD4+ and CD8+ alphabeta T cells and gammadelta T cells in celiac disease, Proc Natl Acad Sci U S A, 110 (2013) 13073-13078.
42. R. Zheng, Q. Yang, The role of the gamma delta T cell in allergic diseases, J Immunol Res, 2014 (2014) 963484.
43. N.R. Locke, S. Stankovic, D.P. Funda, L.C. Harrison, TCR gamma delta intraepithelial lymphocytes are required for self-tolerance, J Immunol, 176 (2006) 6553-6559.
44. A.S. Ismail, K.M. Severson, S. Vaishnava, C.L. Behrendt, X. Yu, J.L. Benjamin, K.A. Ruhn, B. Hou, A.L. DeFranco, F. Yarovinsky, L.V. Hooper, Gammadelta intraepithelial lymphocytes are essential mediators of host-microbial homeostasis at the intestinal mucosal surface, Proc Natl Acad Sci U S A, 108 (2011) 8743-8748.
45. D.V. Correia, M. Fogli, K. Hudspeth, M.G. da Silva, D. Mavilio, B. Silva-Santos, Differentiation of human peripheral blood Vdelta1+ T cells expressing the natural cytotoxicity receptor NKp30 for recognition of lymphoid leukemia cells, Blood, 118 (2011) 992-1001.



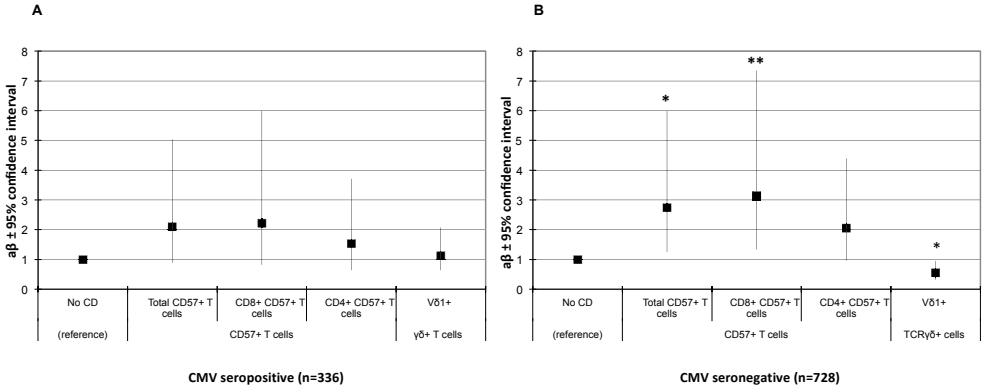
Supplemental Figure 1. Flowchart of participants included in the study
The population for analyses included 12 children with subclinical celiac disease, and n=339 HLA DQ2/DQ8 positive controls.

SUPPLEMENTAL METHODS

Of 13 TG2A positive children who carried one of the genetic risk alleles and anti-endomysial antibodies (EMA), 12 children were diagnosed with celiac disease according to the ESPGHAN guideline. Of these, nine had intestinal biopsy confirmed enteropathy (Marsh 2 n=1; Marsh 3a n=7, Marsh 3b n=1), and in one child biopsy was omitted because of high levels of TG2A and EMA. Two children did not provide consent for intestinal biopsies. Still, these children were considered celiac disease cases, as their serum TG2A levels were high at two time points, and they carried the genetic HLA risk allele (HLADQ 2.2 and HLA DQ8). One child had Marsh grade 1 and was therefore considered as having potential celiac disease (excluded from analysis).



Supplemental Figure 2. Relative differences in blood lymphocyte numbers between children with subclinical celiac disease and HLA DQ2/DQ8 positive and negative controls (n=1,065). Figure Legend: A, Mean lymphocyte numbers (cell/μL) in blood of children with subclinical CeD (n=12), standardized to the same cell numbers in to HLA DQ2/DQ8 positive and negative controls (n=1,053) (unadjusted). B, Mean lymphocyte numbers (cell/μL) in blood of children with subclinical CeD, standardized to the same cell numbers in to HLA DQ2/DQ8 positive and negative controls (adjusted for child's age, SDS BMI for age, gender, CMV, EBV and HSV-1 seropositivity, ethnicity (western/non-western), socioeconomic position (maternal educational level), respiratory and gastrointestinal tract infections, daycare attendance (yes/no), and breastfeeding (ever/never). The estimated back-transformed (a)β and 95% CI (exp Ln β ± exp Ln (95% CI) are derived from univariable (A) and multivariable (B) linear regression models, and indicate the relative difference in mean cell numbers between children with and without subclinical CD. *p<0.05, **p<0.01.



Supplemental Figure 3. Associations between celiac disease, blood CD57+ expressing T cells, and Vδ1+ expressing T cells in HLA DQ2/DQ8 positive and negative children, stratified on CMV seropositivity. Figure legend: A, Mean lymphocyte numbers (cell/μL) in peripheral blood of children with subclinical celiac disease, standardized to HLA DQ2/DQ8 positive and negative controls in CMV seropositive children (n=728). B, Mean lymphocyte numbers (cell/μL) in peripheral blood of children with subclinical celiac disease, standardized to HLA DQ2/DQ8 positive and negative controls in CMV seronegative children (n=336). The estimated back-transformed aβ and 95% CI (exp Ln β ± exp Ln (95% CI) are derived from multivariable linear regression models, and indicate the relative difference in mean cell counts between children with and without subclinical celiac disease. *p<0.05, **p<0.01.

Supplemental Table 1. Definition of lymphocyte subsets

Cell type	Phenotype definition
CD57+ T cells	SSC^{low}CD45+CD3+CD56-CD57+
CD8+CD57+ T cells	SSC^{low}CD45+CD3+CD8+CD56-CD57+
CD4+CD57+ T cells	SSC^{low}CD45+CD3+CD8-CD56-CD57+
CD57+ NKT cells	SSC^{low}CD45+CD3+CD56+CD57+
CD57+ NK cells	SSC^{low}CD45+CD3-CD56+CD57+
TCR γδ+ T cells (% within CD3+)	SSC^{low}CD3+TCRαβ-TCRγδ+
Vδ1+ T cells (% within γδ+ T cells)	SSC^{low}CD3+TCRαβ- Vδ1+ Vδ2-
Vδ2+ T cells (% within γδ+ T cells)	SSC^{low}CD3+TCRαβ- Vδ1- Vδ2+

Supplemental Table 2. Antibody details

		Fluorochrome					
Tube		**FITC**	**PE**	**PerCP-Cy5.5**	**PE-Cy7**	**APC**	**APC-Cy7**
1	**Antibody**	**CD3**	**CD16+CD56**	**CD45**	**CD4**	**-**	**CD8**
	Clone	SK7	B73.1+C5.9	2D1	SK3	-	SK1
	Manufacturer	BD Biosciences	BD Biosciences + Dako	BD Biosciences	BD Biosciences	-	BD Biosciences
2	**Antibody**	**TCRαβ**	**TCRγδ**	**CD3**	**CD4**	**CD8**	**-**
	Clone	WT31	11F2	SK7	SK3	SK1	-
	Manufacturer	BD Biosciences	BD Biosciences	BD Biosciences	BD Biosciences	BD Biosciences	-
3	**Antibody**	**Vδ2**	**Vδ1 ***	**CD3**	**CD4**	**TCRαβ**	**CD8**
	Clone	B6.1	R9.12	SK7	SK3	IP26	SK1
	Manufacturer	BD Biosciences	Beckman Coulter	BD Biosciences	BD Biosciences	eBiosciences	BD Biosciences
4	**Antibody**	**CD57**	**CD38**	**CD3**	**CD56**	**HLA-DR**	**CD8**
	Clone	HNK-1	HB7	SK7	B159	L243	SK1
	Manufacturer	BD Biosciences	BD Biosciences	BD Biosciences	BD Biosciences	BD Biosciences	BD Biosciences

*, unconjugated antibody, detected with Goat anti-Mouse IgG PE (polyclonal; Invitrogen)

Supplemental Table 3. Blood T and NK cell subset frequencies in HLA DQ2/DQ8 positive children with and without celiac disease

HLA DQ2/DQ8 carriers only (n=351)	Cell frequencies within their parent populations (%) (median, range)		
Cell type	No Celiac Disease n=339 (96.6%)	Subclinical Celiac Disease n=12 (3.4%)	P-value
CD57+ T cells (% of CD3+)	2.8 (0.1-46.6)	5.8 (1.6-13.5)	0.059
CD8+CD57+ (% of CD8+)	**5.8 (0.2-53.5)**	**11.5 (3.6-34.6)**	**0.05**
CD4+CD57+ (% of CD4+)	**1.1 (0.1-13.0)**	**2.2 (0.5-4.8)**	**0.04**
CD57+ NK cells (% of CD3-CD56+)	25.1 (1.5-72.6)	23.3 (6.5-49.1)	0.96
CD57+ NKT cells (% of CD3+CD56+)	9.9 (0.8-103.1)	13.2 (2.5-32.7)	0.44
γδ+ T cells (% of CD3+)	8.8 (2.2-21.5)	7.0 (3.3-16.0)	0.20
Vδ1 (% of γδ+ T cells)	**19.2 (3.4-60.9)**	**12.2 (9.6-30.5)**	**0.018**
Vδ2 (% of γδ+ T cells)	42.4 (8.7-82.7)	44.1 (30.8-64.0)	0.26

Values are medians (range), p value reflects differences in median cell frequencies within their parent populations (%) between children with and without subclinical celiac disease in the HLA DQ2/DQ8 positive population, as calculated with the non-parametric Mann-Whitney U test.

Part II

Determinants of Celiac Disease Autoimmunity

Chapter 7

Infant feeding and anti-tissue transglutaminase antibody concentrations in

The Generation R Study

Michelle A.E. Jansen, Ilse I.M. Tromp, Jessica C. Kiefte-de Jong, Vincent W.V. Jaddoe , Albert Hofman, Johanna C. Escher, Herbert Hooijkaas, Henriette A. Moll

Am J Clin Nutr 2014; Oct;100(4):1095-101.

ABSTRACT

Background

Celiac disease (CeD) has emerged as a common, but largely undiagnosed health problem. Numerous studies examined the influence of infant nutrition on the development of diagnosed CeD. However, results are still inconsistent. In addition, the effect of infant feeding practices on the development of potential forms of CeD might be different.

Objective

To examine whether the timing of gluten introduction and breastfeeding duration are associated with Celiac Disease Autoimmunity (CDA) in children at the age of 6 y.

Design

This study was embedded in the Generation R study, a population-based prospective cohort study. Participants included 1679 Dutch children positive for (HLA) DQ2/DQ8. Data on the timing of gluten introduction (<6 mo compared to ≥6 mo) and duration of breastfeeding (<6 months vs. ≥6 months) were obtained by questionnaire. Serum samples were analyzed for anti-tissue transglutaminase (tTG) levels at age 6 y. Anti-tTG levels were categorized into negative (<7 U/ml) and positive (≥7 U/ml) values. Positive anti-tTG levels were further categorized based on the ≥10 times upper limit of normal (ULN) values of the test kit (≥7-70 U/ml and ≥70 U/ml). Multivariable logistic regression analyses were performed.

Results

Positive anti-tTG concentrations were found in 43 children, 26 of whomn had concentrations ≥10 times the ULN (≥70 IU/ml). The introduction of gluten from the age of 6 mo onward and breastfeeding for ≥6 mo were not significantly associated with positive anti-tTG levels. In addition, the timing of gluten introduction and duration of breastfeeding were not significantly associated with positive anti-tTG concentrations below or above the 10 times ULN.

Conclusions

Delayed introduction of gluten beyond the age of 6 mo does not increase the risk of CDA. Also, breastfeeding for ≥6 mo does not decrease the risk of CDA in children at 6 y of age.

INTRODUCTION

Celiac disease (CeD) has emerged as a common, but largely undiagnosed health problem.[1, 2] Untreated CeD is associated with excess morbidity in children and adults.[3] CeD is characterized by an adaptive T-cell mediated response against gluten, classically resulting in chronic inflammation of the small intestinal mucosa and gastrointestinal complaints.[4, 5] However, the clinical presentation of CeD has changed over the past few decades to include milder, non-classic forms.[6, 7] Therefore, diagnosed CeD only represents the more visible tip of the iceberg, ie children having clinical symptoms, whereas asymptomatic or atypical cases are often missed.[6]

Both human leucocyte antigen (HLA) DQ2/DQ8 carrier status and gluten exposure are prerequisites to develop CeD. Breastfeeding and timing of gluten exposure might influence the risk of CeD. Underlying mechanisms remain uncertain, but might involve a complex interplay between innate and adaptive immune responses, gut colonization, intestinal membrane permeability, genetic predisposition and environmental factors such as infections and infant feeding habits.[8-10]

Numerous studies examined the influence of infant nutrition, including breastfeeding and the timing of gluten introduction, on the development of diagnosed CeD. On the basis of current literature, the European Society for Pediatric Gastroenterology, Hepatology , and Nutrition (ESPGHAN) Committee on Nutrition recommends to avoid both early (< 4 mo) and late (> 7 mo) introduction of gluten, and to introduce gluten gradually while the child is still being breastfed.[11] However, results are still inconsistent in this respect.[7, 9, 12-19] The lack of consistency between studies may be due to the majority of these studies being retrospective (ie comparing history of infant feeding practices in CeD cases compared with controls)[13, 17, 18, 20] or ecological (ie, linking incidence rates of diagnosed CeD to national infant feeding practices).[1] These latter studies were mainly based on clinical CeD. However, because subclinical, silent or latent forms of CeD might exist,[4] it could be speculated that these forms have different etiologies but these are not fully understood. Therefore, the effect of infant feeding practices on the development of subclinical, silent or latent forms of CeD might be different.[17, 21] For example, a Swedish study found a decreased prevalence of symptomatic CeD after new infant feeding recommendations were introduced but did not find any difference in prevalence of undiagnosed (screening detected) CeD.[21] However, a later replication within the same cohort study did not support these findings.[19] The aim of the current study, was to examine whether breastfeeding duration and the timing of gluten introduction was associated with celiac disease autoimmunity (CDA) in a population-based prospective cohort study of 6-y-old children.

SUBJECTS AND METHODS

Participants and study design

This study was embedded in the Generation R study, a prospective population based cohort study from fetal life onward, and has been described in detail previously.[22] A total of 8305 mothers with a delivery date from April 2002 through January 2006 provided consent for school age follow-up. Ethical approval for the study was obtained from the Medical Ethical Committee of the Erasmus MC, University Medical Center Rotterdam.

Duration of breastfeeding

Information regarding breastfeeding was obtained by a combination of delivery reports and postnatal questionnaires at the ages of 2, 6 and 12 mo.[22] Mothers were asked by questionnaire whether they had ever breastfed their child and, if yes, at what age (in mo) they had stopped breastfeeding. Breastfeeding duration was categorized as reported by Stordal et al(14) but later dichotomized as <6 mo and ≥6 mo because only 1 infant still received breast milk at the age of 13 mo. Questionnaire response rates were 82%, 73% and 72% at age 2, 6 and 12 mo, respectively.

Introduction of gluten in the first year of life

At the child's age of 6 and 12 mo parents were asked by questionnaire the age of first-time introduction of gluten in their infant's diet. In addition, parents were asked to complete a short food-frequency questionnaire consisting of food products frequently consumed according to a Dutch food consumption survey in infants.[23] The timing of gluten introduction was cross checked with the consumption of bread and biscuits and type of porridge (based on wheat or oats instead of rice) at the age of 6 and 12 mo as described previously.[24] The timing of introduction of gluten was categorized as <6 mo and ≥6 mo.

Anti–Tissue Transglutaminase concentrations

Anti-tissue transglutaminase (anti-tTG) concentrations were assessed in venous blood serum samples by using a fluorescence enzyme immunoassay (ELiA Celikey IgA, PhadiaImmunocap 250, Phadia AB, Upsala Sweden) at the Department of Immunology, Erasmus MC-University Medical Center Rotterdam, the Netherlands. The intra and inter-assay CVs were below 10% and 15%, respectively. Of 8305 children participating at the age of 6 y, serum anti-tTG was available in 53% of children. We excluded 20 children in whom IgA concentrations were low, possibly indicating IgA deficiency. None of the participants were aware of tTG-IgA determination. Median anti-tTG concentration of the study population was 0.25 U/ml, varying from 0.10 U/ml to 125 U/ml. Anti-tTG concentrations were categorized into negative and positive, using the cutoff for clinical practice of 7 U/ml [Figure 1]. Positive anti-tTG concentrations were further categorized into 2 categories based on the ≥10 times the upper limit of normal (ULN) values of the test kit (cutoffs for positive anti-tTG: ≥7-70 U/ml and ≥70 U/ml) [Figure 1].[25]

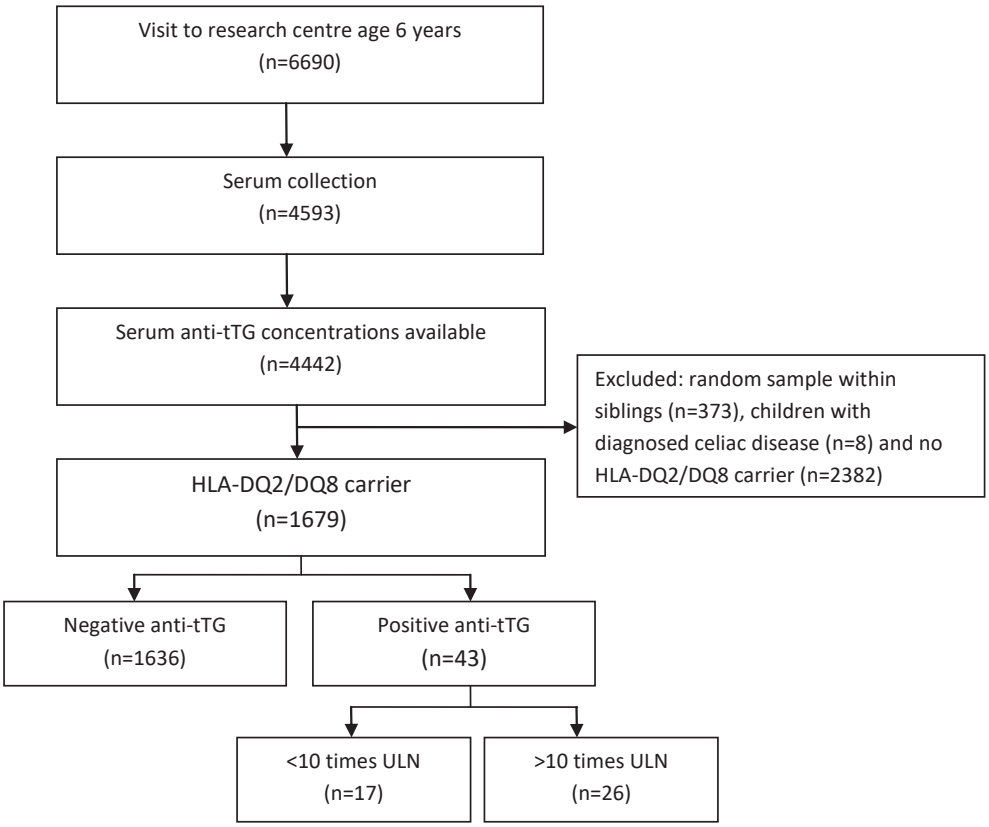


Figure 1. Flowchart of the participants within the Generation R Study

Detection of Human Leukocyte Antigen Risk Alleles

To capture whether the children carried the HLA-DQ risk type DQ2 or DQ8, a tag nucleotide polymorphisms (single nucleotide polymorphism) approach was used as has been described in detail previously.[26, 27] Children were genotyped for these single nucleotide polymorphisms for HLA-DQ2 (rs2187668, rs2395182, rs4713586 and rs7775228) and DQ8 (rs7454108)[26] by using genome-wide Illumina 610 Quad Array. Genotype and allele frequencies were in Hardy-Weinberg equilibrium (rs2187668, *p=0.71*; rs2395182, *p=0.85*; rs4713586, *p*=0.57; rs7775228, *p*=0.92; rs7454108, *p*=0.76).

Covariates

Data on potential confounders as fetal sex, gestational age, birth weight and caesarian section (no differentiation was made between planned and unplanned) were obtained from obstetric records assessed in mid-wife practices and hospitals.[22] Maternal anti-tTG concentrations were measured during pregnancy. Data on other sociodemographic and lifestyle factors were obtained by a combination of pre-and postnatal questionnaires completed by both parents. This included information on ethnicity (Western or non-Western),[28]

parity, maternal smoking during pregnancy, maternal educational level (low or high),[29] and household income per month (≤ € 2200 and > € 2200).[30] Postnatal questionnaires included information on vitamin D supplementation, gastrointestinal and respiratory tract infections and day-care attendance in the first year.

Population for analyses

Of all children with anti-tTG data available (n=4442), we excluded from the analysis those with diagnosed celiac disease (n=8) at the age of 6 years, and children who were HLA-DQ2/DQ8 negative (n=2382) from the analysis. To prevent clustering, only one child per family within the Generation R cohort was included by random selection (n=373). In total, 1679 children were available for statistical analyses [Figure 1].

Statistical analysis

First, independent Student's t-test and Chi-Square tests were performed to test for differences in characteristics between groups of anti-tTG concentrations. Second, logistic regression analysis was performed with anti-tTG concentrations at the age of 6 y as dependent variable. Breastfeeding and the timing of gluten introduction were analyzed separately as independent variables and adjusted for major confounders. Due to small numbers in the positive anti-tTG group, the selection of potential confounders in the multivariate model was restricted to those with an alteration of ≥10% in ORs.[31] To assess whether the association between breastfeeding or the timing of gluten introduction and anti-tTG concentrations differed by ethnicity and children with and without gastrointestinal infections during infancy, statistical interactions were evaluated by adding the product term of independent variable and subgroup (independent variable × subgroup) as covariate to the univariate model. To reduce attrition bias, multiple imputation of the exposures (i.e. breastfeeding and timing of gluten introduction) and covariates was performed (n=5 imputations). The multiple imputation procedure was based on the correlation between each variable with missing values with other subject characteristics.[32] The pooled results of the 5 imputed datasets are reported in this paper as ORs and 95% confidence intervals (95% CIs). A p-value <0.05 was considered as statistically significant. The Statistical analyses were performed in SPSS 20.0 for Windows (SPSS Inc).

RESULTS

Maternal and child characteristics of the study population are shown in table 1. In all 1679 HLA- DQ2/DQ8 positive children, 2.6% had positive anti-tTG concentrations and 97.4% had negative anti-tTG concentrations [Table 1 and figure 1]. Of children with positive anti-tTG concentrations (n=43), 60% of children (n=26) had values ≥10 times the ULN (≥70 IU/ml) [Table 1 and figure 1]. Breastfeeding for ≥6 mo was reported in 47% of children with positive and negative anti-tTG levels [Figure 2]. The introduction of gluten from the age of

6 mo onward was reported in 56% and 64% of children with positive and negative anti-tTG concentrations, respectively [Figure 2].

Table 1: Maternal and child characteristics according to anti-tTG concentrations.

N=1679	Negative anti-tTG		Positive anti-tTG			
	n= 1636 (97.4%)		7-70 IU/ml		>70 IU/ml	
			n=17 (1%)		n=26 (1.6%)	
	N	(%)	*N*	(%)	*N*	(%)
Maternal characteristics						
Educational level						
Low	369	23	6	35	2	8
Mid	854	52	7	41	15	58
High	413	25	4	24	9	34
Household income per month						
≤ 2200 euro	732	45	9	53	5	19
> 2200 euro	904	55	8	47	21	81*
Smoking during pregnancy						
Never	1203	74	13	76	20	77
Smoked during pregnancy	433	26	4	24	6	23
Caesarean section	213	13	2	12	2	8
Parity	711	43	7	41	10	38
Maternal anti-tTG during pregnancy [1]						
Negative	1242	99	12	100	20	100
Positive	6	1	0	0	0	0
Child characteristics						
Male	843	52	4	24 *	8	31 *
Ethnicity						
Western	1107	68	11	65	25	96 *
Non-Western	529	32	6	35	1	4
Birth weight ***z***-score m*ean (SD)*	-0.04	0.99	0.16	1.13	-0.28	1.14
Breastfeeding						
0 to <3 mo	461	28	3	18	9	35
3 to <6 mo	402	25	3	18	8	30
≥ 6 mo	773	47	11	64	9	35
Introduction of gluten						
< 6 mo	586	36	5	29	14	54 *
≥ 6 mo	1050	64	12	71	12	46
Gastro intestinal tract infections first year	1072	66	11	65	18	69
Respiratory tract infections first year	893	55	10	59	18	69
Vitamin D supplementation age 6-12 months	874	53	10	59	12	46
Day care attendance first year	1099	67	12	71	24	92 *

*Significantly different from negative anti-tTG levels

[1] Was not multiple imputed

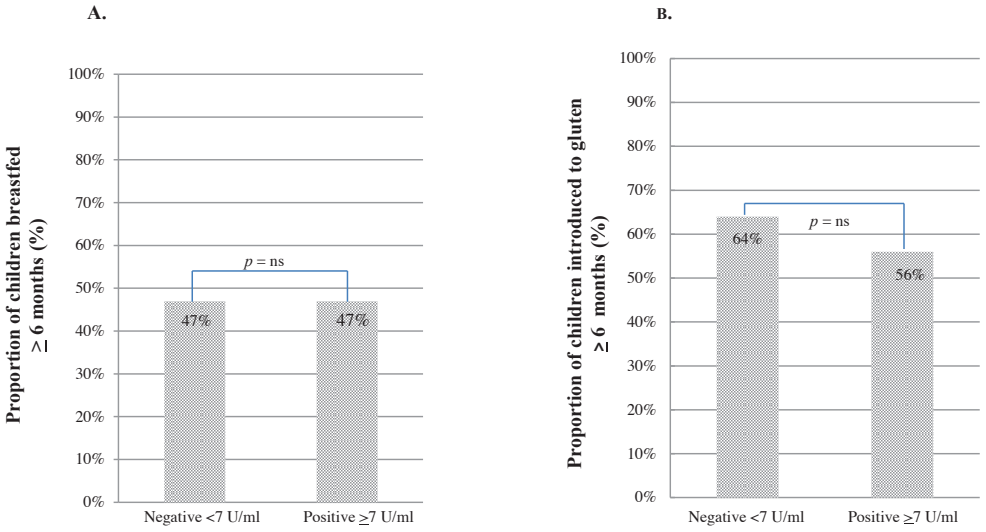


Figure 2. Duration of breastfeeding and timing of gluten introduction according to anti-tTG concentration.
A, Proportion of children breastfed ≥6 mo according to anti-tTG concentration (<7 U/ml n=1636, or ≥7U/ml n=43).
B, Proportion of children introduced to gluten from the age of 6 mo onward according to anti-tTG concentration (<7 U/ml n=1636, or ≥7 U/ml n=43).

Breastfeeding and anti–tTG levels

Relative to breastfeeding for ≤6 mo, breastfeeding for ≥6 mo was not significantly associated with positive anti-tTG concentrations [Table 2]. In addition, breastfeeding for ≥6 mo was not significantly associated with positive anti-tTG concentrations below or above 10 times the ULN (≥70 IU/ml) [Table 3]. In comparison with breastfeeding for ≤3 mo, breastfeeding between 3 and 6 mo and for ≥6 mo was not significantly associated with anti-tTG concentrations [data not shown]. No significant interaction was found between breastfeeding and ethnicity or gastrointestinal infections during infancy (p-interaction>0.50).

Table 2. Association between breastfeeding duration and positive anti-tTG concentrations.

N=1679 Breastfeeding	N (%)	Crude model OR, 95% CI	Multivariate model aOR, 95% CI[a]
< 6 mo	885 (53%)	*Reference*	*Reference*
≥ 6 mo	794 (47%)	1.07 (0.52,2.22)	1.20 (0.56,2.59)

OR: Odds ratio; 95% CI: 95% confidence interval.
[a] Adjusted for smoking during pregnancy, caesarean section, parity, day-care attendance first year, respiratory tract infections first year, vitamin D supplementation age 6-12 mo, household income per month, maternal educational level. Additional adjustment for sex, ethnicity, birth weight ***z***-score, timing of gluten introduction, gastrointestinal tract infections in the first year and maternal anti-tTG concentration did not provide an alteration of ≥ 10% in ORs.

Timing of gluten introduction and anti–tTG levels

In comparison with the introduction of gluten before the age of 6 mo, the introduction of gluten from the age of 6 mo onward was not significantly associated with positive anti-

tTG concentrations [Table 4]. In addition, the introduction of gluten from the age of 6 mo onward was not significantly associated with positive anti-tTG concentrations below or above 10 times the ULN (≥70 IU/ml) [Table 5]. No significant interaction was found between the timing of gluten introduction and ethnicity or gastrointestinal infections during infancy (p-interaction >0.69).

Table 3. Association between breastfeeding duration and positive anti-tTG concentrations below and above 10 times the ULN.

N=1679 Breastfeeding	N (%)	7-70 IU/ml N=17		≥70 IU/ml N=26	
		Crude model OR, 95% CI	Multivariate model aOR, 95% CI [a]	Crude model OR, 95% CI	Multivariate model aOR, 95% CI [a]
< 6 mo	885 (53%)	*Reference*	*Reference*	*Reference*	*Reference*
≥ 6 mo	794 (47%)	2.57 (0.56,11.75)	2.52 (0.59,10.84)	0.64 (0.24,1.73)	0.78 (0.27,2.27)

OR: Odds ratio; 95% CI: 95% confidence interval.

[a] Adjusted for smoking during pregnancy, caesarean section, parity, day-care attendance first year, respiratory tract infections first year, vitamin D supplementation age 6-12 mo, household income per month, maternal educational level.

Additional adjustment for sex, ethnicity, birth weight ***z***-score, timing of gluten introduction, gastrointestinal tract infections in the first year and maternal anti-tTG concentration did not provide an alteration of ≥ 10% in ORs.

Table 4. Association between the introduction of gluten and positive anti-tTG concentrations.

N=1679 Gluten introduction	N (%)	Crude model OR, 95% CI	Multivariate model aOR, 95% CI [a]
< 6 mo	605 (36%)	*Reference*	*Reference*
≥ 6 mo	1074 (64%)	0.68 (0.34,1.35)	0.64 (0.31,1.31)

OR: Odds ratio; 95% CI: 95% confidence interval.

[a] Adjusted for ethnicity and household income per month.

Adjustment for sex, birth weight ***z***-score, smoking during pregnancy, breastfeeding, parity, caesarean section, day-care attendance, vitamin D supplementation, gastrointestinal and respiratory tract infections in the first year, maternal educational level and maternal anti-tTG concentration did not provide an alteration of ≥ 10% in ORs.

Table 5. Association between the introduction of gluten and positive anti-tTG levels below and above 10 times the ULN.

N=1679 Gluten introduction	N (%)	7-70 IU/ml N=17		≥70 IU/ml N=26	
		Crude model OR, 95% CI	Multivariate model aOR, 95% CI [a]	Crude model OR, 95% CI	Multivariate model aOR, 95% CI [a]
< 6 mo	605 (36%)	*Reference*	*Reference*	*Reference*	*Reference*
≥ 6 mo	1074 (64%)	1.27 (0.37,4.37)	1.30 (0.38,4.41)	0.47 (0.19,1.13)	0.42 (0.17,1.02)

OR: Odds ratio; 95% CI: 95% confidence interval.

[a] Adjusted for ethnicity and household income per month.

Additional adjustment for sex, birth weight ***z***-score, smoking during pregnancy, breastfeeding, parity, caesarean section, day-care attendance, vitamin D supplementation, gastrointestinal and respiratory tract infections in the first year, maternal educational level and maternal anti-tTG concentration did not provide an alteration of ≥ 10% in ORs.

DISCUSSION

In this population-based prospective birth cohort, we did not find an association with breastfeeding duration and the timing of gluten introduction and positive anti-tTG concentrations in children with CDA at the age of 6 y.

The ESPGHAN Committee on Nutrition recommends that both early and late introduction of gluten should be avoided. In addition, it is recommended to introduce gluten in small amounts into the diet when the infant is still being breastfed.[11] However, evidence for avoiding very early and late introduction of gluten for the prevention of CeD is inconsistent. We did not find a relationship between timing of gluten introduction and the risk of CeD in childhood, which is in accordance with other studies.[13, 15, 17] In contrast, findings of 2 other prospective studies support a role for the timing of gluten introduction.[9, 14] Norris et al[9] found an increased risk of biopsy-confirmed CeD in CDA positive children for both early (≤3 mo) and late (≥7 mo) introduction of gluten.[9] Strikingly, before restricting their analyses to biopsy-confirmed CeD, Norris et al did not find late introduction to increase the risk of CDA. This suggests that timing of gluten introduction may be differentially associated with biopsy-confirmed and non biopsy-confirmed CeD. However, diagnosed CeD is subjected to bias, depending on the awareness of the pediatricians to make the diagnosis. In addition, Norris et al focused on high risk children defined as having a first degree relative with type 1 diabetes or having HLA genotypes associated with CeD and type 1 diabetes. Therefore, the results cannot be directly extrapolated to our study. Very recently, the Norwegian Mother and Child Cohort Study[14] found late (≥7 mo), but not early (≤4 mo) introduction, to be borderline associated with an increased risk for CeD (adjusted OR: 1.27 95% CI 1.01, 1.65). We could not confirm these results in our study, nor did we find the effect estimates in the same direction (adjusted OR: 0.64, 95% CI 0.31, 1.3). However, the Norwegian Mother and Child Cohort Study included children with clinical CeD instead of CDA which may explain the different results.

The results of our study on breastfeeding and the development of CDA are in line with the findings of other prospective studies.[9, 15, 33] Stordal et al [14] also did not find breastfeeding for ≥6 mo (6-12 mo) to be associated with CeD, but found breastfeeding beyond 12 mo to be associated with an increased risk of CeD. The majority of children in our cohort did not breastfeed after the age of 12 mo; therefore we were unable to replicate this analysis.[14] Several studies did find breastfeeding to reduce the risk of CeD,[12, 17] especially breastfeeding at the time of ([13, 34, 35]) and beyond[12, 13, 19, 35] the introduction of gluten, as well as with gradual introduction of gluten while breastfeeding.[12, 13, 19] However, most of these studies were based on retrospective collected data which may lead to recall bias (ie, differential reporting on infant feeding practices in those with and without CeD diagnosis).[13, 17, 18, 20, 35] In addition, it remains unclear whether the protective effect of breastfeeding is persistent or only delays the onset of symptoms, and therefore CeD diagnosis.[12, 20, 34, 35] For example, previous studies found longer breastfeeding[35] and breastfeeding exclusivity[18] only to delay

the onset of CeD in infancy. This may explain why we did not find any association between breastfeeding duration and CDA. Although we did not find breastfeeding duration and the timing of gluten introduction to be associated with CDA, we do not exclude the role of infant feeding in the development of CeD. Breastfeeding and the timing of gluten introduction might not be protective for the development of subclinical, silent or latent forms of CeD, but could protect against developing symptomatic CeD. Also, gluten introduction while breastfeeding and the amount of gluten, could be more important than the duration of breastfeeding and the timing of gluten introduction, as suggested by Ivarsson et al.[13]

An important strength of this study is first, the assessment of serum anti-tTG concentrations and HLA testing in the general pediatric population, which are good methods to detect clinical silent CeD.[6] Previous studies selected children with diagnosed CeD,[17] or with symptomatic CeD leading to CeD diagnosis,[13, 15] or focused on high-risk children.[9] Second, we subdivided positive anti-tTG concentrations (≥7 U/ml) into 2 groups based on 10 times the ULN level (≥70 IU/ml), because positive anti-tTG concentrations ≥10 times the ULN show high diagnostic accuracies.[25] Of all anti-tTG positive children in our study population, > 60% had anti-tTG concentrations ≥10 times the ULN, suggesting that these children likely have CeD. Third, our study design provided information on a broad range of potential confounders, including ethnicity, socioeconomic status, smoking during pregnancy, and infections.

Some limitations should be taken into account in the interpretation of the results. Information on breastfeeding duration and the timing of gluten introduction was obtained by parental self-report. Nevertheless, only if misclassification of infant feeding practices were related to CeD diagnosis would it have influenced our results, which is unlikely because questionnaires on feeding practices were completed prior to anti-tTG measurement at 6 y of age. In addition, participants in our study were unaware of tTG-IgA determination, so a response bias is therefore highly unlikely. Also, it may be questioned whether our study had sufficient power to detect small differences in gluten introduction practices between children with and without CDA development. Our study had a power of 80% (at an α of 0.05) to detect at least a 22% difference in gluten ≥6 mo to be statistically significant between children with and without CDA. A recent study[19] (n=13,279) found a difference of 2.5% in gluten >6 mo to be statistically significant, whereas Norris et al[9] found a difference of 12% to be significant (n=1,560). Although the sensitivity and specificity of anti-tTG are high,[25, 36] only when anti-tTG concentrations are ≥10 times the ULN in combination with symptoms and positive anti-endomysial antibody can (clinical) CeD be diagnosed without duodenal biopsy.[25] According to the ESPGHAN, clinically diagnosed CeD concerns children who visit the pediatrician because of symptoms, whereas subclinical CeD is found by screening and must be verified by biopsy specimens. However, because our study did not have biopsy specimens, we examined the development of CDA. Therefore, final conclusions concerning (subclinical) CeD diagnosis should be made with caution. Furthermore, we adjusted for potential confounders in our analysis, but residual confounding cannot be fully excluded.

Another limitation of this study is that our study cannot examine the effect of gluten introduction before the age of 4 mo, as well as the amount of gluten introduced in relation to anti-tTG concentrations. Although a previous study in Dutch infants showed that the majority of children receive gluten between 3 and 6 mo of age and not before that time,[37] a window of opportunity in which gluten introduction might prevent CDA cannot be fully ruled out. In relation to breastfeeding, this study cannot examine the effect of breastfeeding at the time of and beyond gluten introduction. Hence, our study precludes conclusions on the effect of very early introduction of gluten, gradual introduction of gluten, and the introduction while breastfeeding.

In conclusion, the results suggest that a delayed introduction of gluten beyond the age of 6 mo does not increase the risk of CDA. In addition, breastfeeding for ≥6 mo does not decrease the risk of CDA in children at 6 y of age. Our study precludes conclusions on the timing of gluten introduction while breastfeeding.

REFERENCES

1. Ivarsson A, Persson LA, Nystrom L, Ascher H, Cavell B, Danielsson L, Dannaeus A, Lindberg T, Lindquist B, Stenhammar L, et al. Epidemic of coeliac disease in Swedish children. Acta Paediatr 2000;89(2):165-71.
2. Maki M, Collin P. Coeliac disease. Lancet 1997;349(9067):1755-9.
3. Ludvigsson JF, Rubio-Tapia A, van Dyke CT, Melton LJ, 3rd, Zinsmeister AR, Lahr BD, Murray JA. Increasing incidence of celiac disease in a North American population. Am J Gastroenterol 2013;108(5):818-24.
4. Sapone A, Bai JC, Ciacci C, Dolinsek J, Green PH, Hadjivassiliou M, Kaukinen K, Rostami K, Sanders DS, Schumann M, et al. Spectrum of gluten-related disorders: consensus on new nomenclature and classification. BMC Med 2012;10:13.
5. Guandalini S, Assiri A. Celiac Disease: A Review. JAMA Pediatr 2014; 168(3):272-8.
6. Maki M, Mustalahti K, Kokkonen J, Kulmala P, Haapalahti M, Karttunen T, Ilonen J, Laurila K, Dahlbom I, Hansson T, et al. Prevalence of Celiac disease among children in Finland. N Engl J Med 2003;348(25):2517-24.
7. Ludvigsson JF, Leffler DA, Bai JC, Biagi F, Fasano A, Green PH, Hadjivassiliou M, Kaukinen K, Kelly CP, Leonard JN, et al. The Oslo definitions for coeliac disease and related terms. Gut 2013;62(1):43-52.
8. Prescott SL, Smith P, Tang M, Palmer DJ, Sinn J, Huntley SJ, Cormack B, Heine RG, Gibson RA, Makrides M. The importance of early complementary feeding in the development of oral tolerance: concerns and controversies. Pediatr Allergy Immunol 2008;19(5):375-80.
9. Norris JM, Barriga K, Hoffenberg EJ, Taki I, Miao D, Haas JE, Emery LM, Sokol RJ, Erlich HA, Eisenbarth GS, et al. Risk of celiac disease autoimmunity and timing of gluten introduction in the diet of infants at increased risk of disease. JAMA 2005;293(19):2343-51.
10. Szajewska H, Chmielewska A, Piescik-Lech M, Ivarsson A, Kolacek S, Koletzko S, Mearin ML, Shamir R, Auricchio R, Troncone R, et al. Systematic review: early infant feeding and the prevention of coeliac disease. Aliment Pharmacol Ther 2012;36(7):607-18.
11. Agostoni C, Decsi T, Fewtrell M, Goulet O, Kolacek S, Koletzko B, Michaelsen KF, Moreno L, Puntis J, Rigo J, et al. Complementary feeding: a commentary by the ESPGHAN Committee on Nutrition. J Pediatr Gastroenterol Nutr 2008;46(1):99-110.
12. Hornell A, Lagstrom H, Lande B, Thorsdottir I. Breastfeeding, introduction of other foods and effects on health: a systematic literature review for the 5th Nordic Nutrition Recommendations. Food Nutr Res 2013;Apr 12:57.
13. Ivarsson A, Hernell O, Stenlund H, Persson LA. Breast-feeding protects against celiac disease. Am J Clin Nutr 2002;75(5):914-21.
14. Stordal K, White RA, Eggesbo M. Early Feeding and Risk of Celiac Disease in a Prospective Birth Cohort. Pediatrics 2013;132(5):e1209-9
15. Welander A, Tjernberg AR, Montgomery SM, Ludvigsson J, Ludvigsson JF. Infectious disease and risk of later celiac disease in childhood. Pediatrics 2010;125(3):e530-6.
16. Shamir R. Can feeding practices during infancy change the risk for celiac disease? Isr Med Assoc J 2012;14(1):50-2.
17. Peters U, Schneeweiss S, Trautwein EA, Erbersdobler HF. A case-control study of the effect of infant feeding on celiac disease. Ann Nutr Metab 2001;45(4):135-42.
18. D'Amico MA, Holmes J, Stavropoulos SN, Frederick M, Levy J, DeFelice AR, Kazlow PG, Green PH. Presentation of pediatric celiac disease in the United States: prominent effect of breastfeeding. Clin Pediatr (Phila) 2005;44(3):249-58.
19. Ivarsson A, Myleus A, Norstrom F, van der Pals M, Rosen A, Hogberg L, Danielsson L, Halvarsson B, Hammarroth S, Hernell O, et al. Prevalence of childhood celiac disease and changes in infant feeding. Pediatrics 2013;131(3):e687-94.

20. Ludvigsson JF, Fasano A. Timing of introduction of gluten and celiac disease risk. Ann Nutr Metab 2012;60 Suppl 2:22-9.
21. Carlsson A, Agardh D, Borulf S, Grodzinsky E, Axelsson I, Ivarsson SA. Prevalence of celiac disease: before and after a national change in feeding recommendations. Scand J Gastroenterol 2006;41(5):553-8.
22. Jaddoe VW, van Duijn CM, Franco OH, van der Heijden AJ, van Iizendoorn MH, de Jongste JC, van der Lugt A, Mackenbach JP, Moll HA, Raat H, et al. The Generation R Study: design and cohort update 2012. Eur J Epidemiol 2012;27(9):739-56.
23. Hulshof K.F.A.M BBC. Results of the study on nutrient intake in young toddlers 2002.Zeist, The Netherlands TNO Nutrition, 2002.
24. Kiefte-de Jong JC, Escher JC, Arends LR, Jaddoe VW, Hofman A, Raat H, Moll HA. Infant nutritional factors and functional constipation in childhood: the Generation R study. Am J Gastroenterol 2010;105(4):940-5.
25. Husby S, Koletzko S, Korponay-Szabo IR, Mearin ML, Phillips A, Shamir R, Troncone R, Giersiepen K, Branski D, Catassi C, et al. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr 2012;54(1):136-60.
26. Monsuur AJ, de Bakker PI, Zhernakova A, Pinto D, Verduijn W, Romanos J, Auricchio R, Lopez A, van Heel DA, Crusius JB, et al. Effective detection of human leukocyte antigen risk alleles in celiac disease using tag single nucleotide polymorphisms. PLoS One 2008;3(5):e2270.
27. Kiefte-de Jong JC, Jaddoe VW, Uitterlinden AG, Steegers EA, Willemsen SP, Hofman A, Hooijkaas H, Moll HA. Levels of antibodies against tissue transglutaminase during pregnancy are associated with reduced fetal weight and birth weight. Gastroenterology 2013;144(4):726-35 e2.
28. Swertz O, Duimelaar P, Thijssen J. Migrants in the Netherlands 2004. Voorburg/Heerlen: Statistics Netherlands, 2004.
29. Statistics Netherlands. Dutch Standard Classification of Education 2003. Voorburg/Heerlen, Netherlands: Statistics Netherlands, 2004.
30. Statistics Netherlands. Welfare in the Netherlands: income, welfare and spending of households and individuals. The Hague/Heerlen, Netherlands: Statistics Netherlands, 2012.
31. Greenland S, Mickey RM. Re: "The impact of confounder selection criteria on effect estimation. Am J Epidemiol 1989;130(5):1066.
32. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393.
33. Roberts SE, Williams JG, Meddings D, Davidson R, Goldacre MJ. Perinatal risk factors and coeliac disease in children and young adults: a record linkage study. Aliment Pharmacol Ther 2009;29(2):222-31.
34. Akobeng AK, Ramanan AV, Buchan I, Heller RF. Effect of breast feeding on risk of coeliac disease: a systematic review and meta-analysis of observational studies. Arch Dis Child 2006;91(1):39-43.
35. Radlovic NP, Mladenovic MM, Lekovic ZM, Stojsic ZM, Radlovic VN. Influence of early feeding practices on celiac disease in infants. Croat Med J 2010;51(5):417-22.
36. Burgin-Wolff A, Dahlbom I, Hadziselimovic F, Petersson CJ. Antibodies against human tissue transglutaminase and endomysium in diagnosing and monitoring coeliac disease. Scand J Gastroenterol 2002;37(6):685-91.
37. Hopman EG, Kiefte-de Jong JC, le Cessie S, Moll HA, Witteman JC, Bleeker SE, Mearin ML. Food questionnaire for assessment of infant gluten consumption. Clin Nutr 2007;26(2):264-71.

Chapter 8

Herpesvirus infections and Transglutaminase type 2 Antibody positivity in Childhood: The Generation R Study

Michelle A.E. Jansen, Diana van den Heuvel , Kirsten V.M. van der Zwet, Vincent W.V. Jaddoe, Albert Hofman, Johanna C. Escher, Pieter LA Fraaij, Herbert Hooijkaas, Menno C. van Zelm MC, Henriette A. Moll

J Pediatr Gastroenterol Nutr. 2016 Oct;63(4):423-30

ABSTRACT

Objectives

Persistent viral infections have been implicated in the etiology of autoimmune diseases in adulthood, but it is not known whether herpesviruses are associated with the development of celiac disease autoimmunity (CDA) in childhood. We assessed whether herpesvirus infections are associated with transglutaminase type 2 antibody (TG2A) concentrations in children at 6 years of age.

Methods

This study was embedded within a population-based prospective cohort study. Serum immunoglobulin G levels against Epstein-Barr virus, cytomegalovirus (CMV), and herpes simplex virus type 1 were measured by enzyme-linked immunosorbent assay, and TG2A concentrations with fluorescence enzyme immunoassay in 4420 children at 6 years of age. Children were categorized based on TG2A concentrations into negative (<7 U/ml), positive (≥7-70 U/ml) and strongly positive (≥70 U/ml), that is 10 times upper limit normal (ULN).

Results

Fifty-nine children (1.3%) were TG2A positive, and of these 31 (53%) had concentrations 70 U/ml or more. Children with TG2A concentrations 70 U/ml or more were less often infected with CMV (adjusted odds ratio (aOR) 0.38; 95% CI 0.14, 0.98; p=0.04) and with any of the three viruses (aOR 0.38; 95% CI 0.18, 0.78; p<0.01) than children with TG2A negative concentrations. In addition, children with TG2A concentrations 70 U/ml or more were less often infected with 2 or more viruses than children with TG2A negative concentrations (aOR 0.15; 95 % CI 0.03, 0.65; p=0.01).

Conclusions

Both CMV single infection and combined CMV, Epstein-Barr virus and/or herpes simplex virus type 1 infections are inversely associated with strongly TG2A positivity. This may indicate a protective effect of herpesvirus infections in the pathogenesis of celiac disease autoimmunity.

INTRODUCTION

Species of the human herpesvirus family have been implicated in the pathogenesis of autoimmune diseases in adulthood,[1] but their role in childhood autoimmunity remains poorly studied.[2] Celiac disease (CeD) is one of the most prevalent, immune-mediated systemic disorders in childhood, and is characterised by villous atrophy causing enteropathy and a wide variety of clinical manifestations.[3] Carriership of either the HLA-DQ2 or HLA-DQ8 haplotype is a prerequisite for CeD.[3-5] Furthermore, the presence of IgA antibodies against the enzyme transglutaminase type 2 (TG2A) is a specific disease marker. CeD diagnosis is based on biopsy specimens, but these are omitted in HLA DQ2-DQ8 positive children having TG2A concentrations >10 times upper limit normal (ULN; 70 U/ml), in accordance with both clinical complaints and positive anti-endomysial antibody levels.[6, 7] The disease may present at any age. Still, the majority of patients initially lack typical clinical symptoms,[8] resulting in delayed diagnosis or even under diagnosis. Screening for TG2A positivity is, thereforem considered an effective method to detect potential and subclinical CeD.[9, 10]

Similar to many autoimmune diseases, the prevalence of CeD has been increasing rapidly with up to a 4-fold increase in the last 50-years. This is considered a true rise, rather than a result of improved detection.[11-14] Given its rapidity, this cannot be merely attributed to genetic factors. Environmental factors early in life, such as the amount of infections, might be involved.[4] Herpesviruses, including Epstein-Barr virus (EBV), Cytomegalovirus (CMV) and Herpes Simplex virus type 1 (HSV-1) are ubiquitous and highly contagious, but the age at primary infection is delayed in Western countries,[15-17] which together with the increase in CeD prevalence may suggest that infections early in life might have a protective role in the etiology of CeD. In addition, seroprevalences of herpesviruses depend on ethnic /geographic background and socioeconomic position,[18, 19] which may further explain some of the ethnic/geographic variation in CeD incidence.[20] In the present study, we used herpesviruses as a proxy for the amount of infections in early life. To our knowledge, associations between herpesvirus infections and TG2A concentrations in childhood have never been studied in a population based cohort. The objective of this study, therefore, was to examine whether EBV, CMV and/or HSV-1 infections were inversely associated with the presence of TG2A positivity as marker for celiac disease autoimmunity in children at 6 years of age.

METHODS

Study design

This present study was embedded within the Generation R study, a prospective population based cohort study from fetal life until young adulthood. The background, design and aims have been described in detail previously.[21] Ethics approval for the study design, the

research aims and all the specific measurements were obtained from the Medical Ethical Committee of Erasmus MC, University Medical Centre Rotterdam, The Netherlands. Written informed consent was obtained from all participants. A total of 9778 mothers with a delivery data between April 2002 and January 2006 were enrolled in the study. At the age of 6 years, n=8305 children (85% of the original cohort) still participated in the study, of whom 81% (n=6690) visited the research centre.[21] During this visit, blood samples were collected from 4593 children. From a total of 4436 children both serological viral tests and TG2A concentrations could be determined at the median age of 6.0 years. We excluded children with a questionnaire reported low or gluten free diet (GFD) (n=10), because TG2A concentrations drop as a results of the GFD, and we excluded children with a questionnaire-reported CeD diagnosis, but negative TG2A concentrations (n=6). Of 9,778 children participating in the original cohort, the resulting population for analysis consisted of 4,420 children (45%). [**Supplementary Figure 1**].

Herpes virus serology

Venous blood plasma samples were analysed using enzyme immunoassays for IgG antibodies against EBV (native mixture of several viral capsid antigens), CMV (purified native antigens from strain "AD169"), and HSV-1 (glycoprotein C1; all from EUROIMMUN, Lübeck, Germany).[22] Results were evaluated semiquantitatively relative to a manufacturer-provided reference threshold sample. Per manufacturer's instructions, an optical density in the patient sample >10% above the provided threshold sample was defined as positive. The inter-assay and intraassay coefficients of variation (CV) as provided by the manufacturer were <10%, based on 20 measurements performed on 3 sera for the intraassay CV, and 4 measurements for the interassay CV. In our study, we used 53 kits, and the interassay CV of the positive controls that were provided with the kits were 6% (CMV), 5% (EBV) and 11% (HSV-1).

Detection of transglutaminase type 2 antibodies

Serum IgA antibody concentrations against transglutaminase type 2 were measured using a fluorescence enzyme immunoassay (Elia Celikey IgA, Phadia ImmunoCAP 250, Phadia AB, Uppsala, Sweden). The intra- and interassay CV was <10% and 15% respectively. Sera with a TG2A concentration of 7 U/ml or higher were considered to be positive per manufacturer's instructions. In addition, TG2A positive concentrations were further categorized into 2 categories based on the 10 times or more upper limit normal (ULN) of the test kit (≥7-70 U/ml, and ≥70 U/ml) [Supplemental Figure 1].[6, 23]

Detection of HLA Risk Alleles

A genome-wide association scan (Illumia 610K) of child DNA was taken from n=5908 cord blood samples. Missing cord blood samples were mainly due to logistical constraints during delivery.[21] To identify the children carrying HLA-DQ2 or –DQ8 alleles, a tag single

nucleotide polymorphism approach was used as described in detail previously.[24, 25] Children were genotyped for single nucleotide polymorphisms linked to HLA-DQ2 (rs2187668, rs2395182, rs4713586 and rs7775228) and DQ8 (rs7454108) by using genome-wide Illumina 610 Quad Array (San Diego, CA, USA). Genotype and allele frequencies were in Hardy Weinberg Equilibrium (rs218677, p=0.88; rs2395182, p=0.78; rs4713586, p=0.96; rs7775228, p=0.86; rs7454108, p=0.90).

Covariates

Data on potential confounders, such as foetal sex, gestational age, birth weight, birth season, and mode of delivery were obtained from obstetric records from hospitals and mid-wife practices. Data on sociodemographic and lifestyle factors were obtained by a combination of pre- and postnatal questionnaires completed by both parents. Prenatal questionnaires included information on ethnicity,[26] parity, maternal smoking during pregnancy, maternal educational level[27] and family net household income per month.[28] Child's ethnic background was dichotomized into 'Western' and 'non-Western' ethnicity according to the Dutch agency for statistics. In addition, information on Caucasian versus non-Caucasian background was derived from GWAS data from cord blood.[21] The cut-off of 2200 euro/month was used based on the average income per household in the study area. Information on the highest level of achieved maternal education was subsequently dichotomized into 'low' and 'high' (university or PhD degree). Postnatal questionnaires included information on breastfeeding, timing of gluten introduction, vitamin D supplementation, day-care attendance, gastrointestinal and respiratory tract infections, and antibiotic use during the first 6 years of life. Information on the duration and exclusiveness of breastfeeding was asked by parental reported questionnaires at the ages of 2, 6 and 12 months and crosschecked. Because breastfeeding duration was stronger related to CMV seropositivity than breastfeeding exclusiveness, we included breastfeeding duration in the multivariable model. Information on gastrointestinal, upper and lower respiratory tract infections, and antibiotic use was assessed by parental reported questionnaires which were administered yearly from birth until 6 years of age.

Statistical analysis

First, independent student's t tests and Chi-square tests were used to test whether maternal and child characteristics were different between TG2A positive and negative groups. Second, logistic regression analyses was performed with TG2A positivity at the age of 6 years as dependent variable. Viral seropositivity (IgG against either EBV, CMV or HSV-1) was analyzed separately as independent dichotomous variable. The variable 'frequency of viral infection' was calculated by taking the sum of EBV, CMV and/or HSV-1 IgG seropositivity, resulting in 3 categories: 0, 1 or ≥2 seropositive results. Multivariable models were created, with stepwise adjustment for potential confounders. Because of small numbers in the TG2A positive group (n=59), the selection of potential confounders was restricted to

those who attained a 10% or more change in odds ratio (OR).[29] To assess whether the association between viral antibody levels and TG2A concentrations was different by ethnicity and children with and without prior gastrointestinal infections during infancy, statistical interaction was evaluated by adding the product term of the independent variable and subgroup (independent variable ^{x} subgroup) as covariate in the model. If the test was significant ($p<0.05$), we also stratified the analysis by the variables. To reduce attrition bias, multiple imputation of the covariates was performed (n=10 imputations). The multiple imputation procedure was based on the correlation of the variable with missing values with other subject characteristics.[30] Regression coefficients were pooled by taking the average of the coefficients of the 10 imputed datasets. The pooled standard error was then calculated using Rubin rule. The pooled results of the 10 imputed datasets were reported in this article as ORs and 95% confidence intervals (CIs). A p-value ≤0.05 was considered as statistically significant. Statistical analyses were performed in SPSS 20.0 for Windows (SPSS Inc., Chicago, IL).

RESULTS

Subject characteristics

Maternal and child characteristics are shown in [Table 1]. Within the total group of 4420 children, 59 children (1.3%) were TG2A positive, and of these 31 (53%) had concentrations >10 ULN (≥70 U/ml) [Table 1, Supplementary Figure 1]. Among these latter 31, 92% carried the HLA-DQ2 or –DQ8 alleles [Table 1]. TG2A positivity was significantly related to higher socio-economic status (net household income per month), female sex and Western ethnicity [Table 1]. No significant differences were observed in maternal educational level, maternal smoking during pregnancy, mode of delivery, parity, birth weight, birth season, feeding habits including breastfeeding, and timing of gluten introduction,[23] day-care attendance in the first year, gastrointestinal and respiratory tract infections in the first year and antibiotic use during the first 6 years.

Seroprevalences of herpesvirus infections

Within the total group of 4420 children with a median age of 6.0 years, 1,672 (38%) were CMV infected, 2,257 (51%) were EBV infected, and 871 (20%) were HSV-1 infected. [Table 2]. Twenty-nine percent were seronegative for all 3 viruses, 39% was seropositive for 1 virus and 32% was seropositive for at least 2 viruses [Table 3]. Among children infected with CMV, infection with EBV and HSV-1 was more prevalent (respectively 59% and 25%) than among children who were not infected with CMV (respectively. 46% and 17%). [Supplementary Table 3].

Table 1. Maternal and child characteristics of study population. Imputed dataset

N=4420	TG2A negative <7 IU/ml		TG2A positive ≥7 U/ml			
			7-70 IU/ml		≥70 IU/ml	
Characteristics	n=4361 (98.7%)		n=28 (0.6%)		n=31 (0.7%)	
Maternal characteristics						
Age (years; mean, sd)	30.7	5.1	30.7	5.4	**32.8***	**3.4**
High educational level (n;%)	3276	75%	23	82%	19	61%
Household income per month (>2200 euro) (n;%)	2349	54%	13	46%	**25**	**81%****
Smoking during pregnancy (n;%)	1127	26%	8	29%	8	26%
Caesarean section (n;%)	572	13%	4	9%	2	6%
Multiparity (n;%)	1971	45%	16	57%	12	39%
Child characteristics						
Age (years; mean, sd)	6.2	0.6	6.3	0.6	6.1	0.3
Male (n;%)	2258	52%	**8**	**29%***	12	39%
Western ethnicity (n;%)	2901	67%	18	64%	**29**	**94%****
HLA DQ2 DQ8 carriership[a] (n;%)	1164	41%	**15**	**88%*****	**23**	**92%*****
Missing data (n;%)	*1512*	*35%*	*11*	*39%*	*6*	*19%*
Birth weight ***z***-score ***Mean (SD)***	-0.07	1.00	0.18	1.19	-0.21	1.15
Birth season (autumn/winter) (n;%)	2221	51%	10	36%	14	45%
Breastfeeding (n;%)						
< 6 mo	2485	57%	14	50%	20	65%
≥ 6 mo	1876	43%	14	50%	11	35%
Introduction of gluten (n;%)						
< 6 mo	2413	55%	18	64%	14	45%
≥ 6 mo	1,948	45%	10	36%	17	55%
Day care attendance first year (n;%)	2925	67%	17	61%	27	87%
Gastro intestinal tract infections first year (n;%)	1960	45%	10	36%	17	55%
Respiratory tract infections first year (n;%)	2443	56%	13	46%	21	68%
Use of antibiotics / penicillin (n;%)						
First year	2141	49%	12	43%	11	35%
Ever	3371	77%	21	75%	23	74%

*****Significantly different from TG2A negative group. *P<0.05, ** p<0.01, *** P<0.001.

[a] Was not multiple imputed. Of 4420 children, data on HLA DQ2 or DQ8 carriership was present in n=2891 (65%) children. Of these, n=1202 (42%) were HLA DQ2/DQ8 positive, and 1689 (58%) children were HLA DQ2/DQ8 negative. Data on HLA DQ2/ DQ8 risk type was missing in n=1529 (35%)children.

TG2A strongly positive children less often experienced herpesvirus infections

Children with TG2A concentrations 70 U/ml or more were less often infected with CMV than TG2A negative children (adjusted odds ratio [aOR] 0.38; 95% CI 0.14, 0.98; *p*=0.04) [Table 2, Figure 1]. The same tendencies were found for EBV and HSV-1 [Table 2], although these associations did not remain significant after adjustment for ethnicity, income and

Table 2. Association between herpesvirus infection and TG2A concentrations

	Total number of children	**TG2A negative <7 U/ml**	**TG2A positive ≥7 U/ml**			
	n= 4420	n= 4361 (98.7%)	**7-70 U/ml** n= 28 (0.6%)		**≥70 U/ml** n= 31 (0.7%)	
	N (%)	*N* (%)	*N* (%)	[a] OR (95% CI)	*N* (%)	[a] OR (95% CI)
CMV	1672 (38%)	1657 (38%)	10 (36%)	0.87 (0.39;1.92)	5 (16%)	**0.38 (0.14;0.98)***
No CMV	2748 (62%)	2704 (62%)	18 (64%)	*Reference*	26 (84%)	*Reference*
EBV	2257 (51%)	2234 (51%)	13 (46%)	0.79 (0.37;1.68)	10 (32%)	0.57 (0.27;1.21)
No EBV	2163 (49%)	2127 (49%)	15 (54%)	*Reference*	21 (68%)	*Reference*
HSV-1	871(20%)	865 (20%)	5 (18%)	0.82 (0.1;2.21)	1 (3%)	0.18 (0.02;1.29)
No HSV-1	3549 (80%)	3496 (80%)	23 (82%)	*Reference*	30 (97%)	*Reference*
Any (EBV, CMV and/or HSV-1)	3117 (71%)	3085 (71%)	19 (68%)	0.82 (0.36;1.86)	13 (42%)	**0.38 (0.18;0.78)****
None	1303 (29%)	1276 (29%)	9 (32%)	*Reference*	18 (58%)	*Reference*

aOR: adjusted Odds ratio; 95% CI: 95% confidence interval. ORs are from multiple imputed data and derived from logistic regression models.[a] Adjusted for ethnicity, net household income per month, breast-feeding and day-care attendance. Additional adjustment for maternal educational level, maternal smoking during pregnancy, parity, sex, birth weight ***z***-score, caesarean section, season of birth, timing of gluten introduction, vitamin D supplementation age 6-12 months, gastrointestinal and respiratory tract infections in the first year (yes/no), and antibiotic use did not provide an alteration of ≥ 10% in OR. Unadjusted associations are shown in Supplemental Table S2. *****Significantly different from TG2A negative group. *P<0.05, ** p<0.01

breast-feeding (aOR EBV 0.57; 95% CI 0.27, 1.21; *p*=0.14) (aOR HSV-1 0.18; 95% CI 0.02, 1.29; *p*=0.10) [Table 2]. Unadjusted associations are shown in Supplementary Table 2. In addition, children with TG2A concentrations 70 U/ml or more were less often infected with any of the 3 viruses (aOR 0.38; 95% CI 0.18, 0.78; *p*<0.01). No significant differences in viral infection rates were observed between TG2A positive children with concentrations <70 U/ml and TG2A negative children [Table 2].

Within the total group of 4,420 children, CMV positive children were more often infected with other viruses than CMV negative children [Supplementary Table 3b]. Still, the inverse association between CMV infection and TG2A positivity was independent of EBV and HSV-1 infection. [Figure 1].

Associations between CMV infection and TG2A positivity were not significantly different between children with and without HLA-DQ2 or HLA-DQ8 (p>0.93). After stratifying the analysis on HLA-DQ2/DQ8 carriership, the same tendencies were observed for associations between herpesvirus infections and TG2A positivity in HLA-DQ2/DQ8 positive children. [Supplementary Table 4] No significant statistical interactions were found between combinations of viral infections, ethnicity (western/non-western, Caucasian/non-Caucasian), or gastrointestinal tract infection during the first year of life (p>0.18; 0.66; 0.50; and 0.33 respectively). In line with this, associations between herpesvirus infections and TG2A positivity were not materially different between Caucasians and non-Caucasians [Supplementary

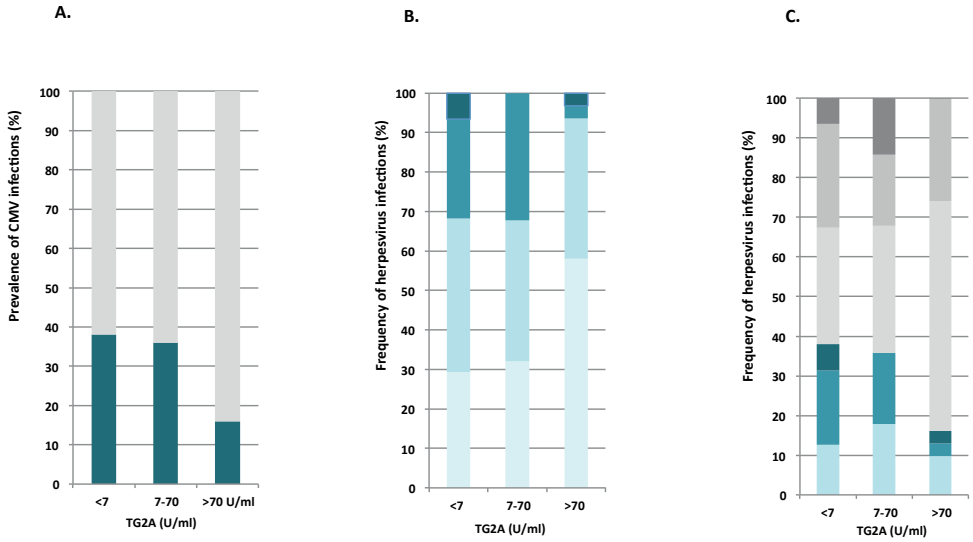


Figure 1. Association between herpesvirus infections and TG2A concentrations

A, Association between CMV infection and TG2A concentrations. Dark blue represents CMV seropositives, Light grey represents CMV seronegatives.

B, Association between multiple herpesvirus infections and TG2A concentrations. White represents 0 infections; Light blue 1 infection; Middle blue 2 infections; and dark blue 3 infections.

C, Association between multiple herpesvirus infections and TG2A concentrations according to CMV infection. Light blue represents 1 infection (CMV); Middle blue 2 infections (CMV+EBV+ or CMV+HSV-1+); Dark blue 3 infections (CMV+EBV+HSV-1+); Light grey no infections; Middle grey 1 infection (EBV+ or HSV-1+); and dark grey 2 infections (EBV+ and HSV-1+).

Table 3. Association between herpesvirus infection rate and TG2A concentrations

	Total number of children	**TG2A negative <7 U/ml**	**TG2A positive >7 U/ml**			
	n= 4420	n=4361 (98.7%)	**7-70 U/ml** n= 28 (0.6%)		**>70 U/ml** n= 31 (0.7%)	
	N (%)	*N* (%)	*N* (%)	[a] OR (95% CI)	*N* (%)	[a] OR (95% CI)
Infection frequency						
0 infections	1303 (29%)	1276 (29%)	9 (32%)	*Reference*	18 (58%)	*Reference*
1 infection	1723 (39%)	1702 (39%)	10 (36%)	0.80 (0.32; 1.99)	11 (36%)	0.52 (0.24;1.10)
≥ 2 infections	1394 (32%)	1383 (32%)	9 (32%)	0.85 (0.32; 2.25)	2 (6%)	**0.15 (0.03;0.65)***

aOR: adjusted Odds ratio; 95% CI: 95% confidence interval. ORs are from multiple imputed data and derived from logistic regression models. [a] Adjusted for ethnicity, net household income per month, breastfeeding and day-care attendance. Additional adjustment for maternal educational level, maternal smoking during pregnancy, parity, sex, birth weight ***z***-score, caesarean section, season of birth, timing of gluten introduction, vitamin D supplementation age 6-12 months, gastrointestinal and respiratory tract infections in the first year (yes/no), and antibiotic use did not provide an alteration of ≥ 10% in OR. Unadjusted associations are shown in Supplemental Table S2. ***Significantly different from TG2A negative group. *P<0.05

Table 5]. However, non-Caucasians experienced more often herpesvirus infections than Caucasians [Supplementary Table 6].

TG2A strongly positive children were less often infected with multiple viruses

Children with TG2A concentrations 70 U/ml or more were less often infected with 2 or more viruses than TG2A negative children (aOR 0.15; 95 % CI 0.03, 0.65; *p*=0.01) [Table 3, Figure 1]. However, associations between viral infections and TG2A positivity <70 U/ml were not significant.

DISCUSSION

This population based cohort study showed an inverse association between CMV single infection and strongly TG2A positivity. In addition, multiple (≥2) infections (EBV, CMV or HSV-1) were inversely related to the presence of TG2A. This could indicate an accumulated protective effect of multiple herpesvirus infections on the development of celiac disease autoimmunity in childhood.

Our results confirm and extent a case-control study that examined the association between CMV infection and TG2A positivity previously.[31] The complex interplay between infections and autoimmunity has only been partly understood.[32] More important, it should be considered that the observed associations in our study might not be causal, as herpes viridae can be seen as a proxy for the amount of infections early in life. CMV and EBV infection are, however, known to increase effector memory CD8+ T-cell numbers,[33-35] which influence the composition and diversity of the adaptive immune repertoire.[34, 36] As a consequence, the development of autoreactive antibodies, such as TG2A, may be influenced. Strikingly, some matched (nested) case-control studies,[37, 38] suggest that serum TG2A could be produced during a recent or ongoing EBV or CMV infection, irrespective of celiac disease diagnosis.[38] In these studies, data on TG2A was, however, collected retrospectively in patients presenting with clinical evidence of infection, such as glandular fever and hepatitis,[37] or patients were recruited at the emergency department,[38] which makes comparison with our study population difficult. Provocative effects of EBV and CMV have also been suggested in several other autoimmune diseases.[1, 37] These associations, however, may be multidirectional and multifaceted.[39] Hence, a common general pathophysiological mechanism to explain our results in relation to the development of TG2A may not exist.

Our results do not support the protective effects suggested for EBV.[31] A potential explanation could lie in the study groups based on adult populations. Furthermore, TG2A concentrations were not used as a screening tool for CeD, but clinically diagnosed celiac disease patients were studied.[31] In another study, viral seroprevalences and TG2A concentrations were studied in adults with diabetes mellitus type 1.[40] In these patients, lower levels of antibodies against infection (CMV and EBV) and higher tTG-IgG antibodies were

observed than in the healthy control group. Therefore, they speculate that infections may have a protective role in development of type 1 diabetes. Their results were, however, not adjusted for major confounders such as ethnicity and socioeconomic status. The results of our univariable analysis indeed suggested an inverse association between EBV and CeD, but after adjusting for ethnicity, socioeconomic status, breast-feeding, and day care, the results for EBV did not remain significant. Thus, the earlier observed associations for EBV may have been indirect effects of CMV or multiple herpes virus infections. Although it could be hypothesized that CMV-specific and EBV-specific immune responses act synergistic or antagonistic in the pathogenesis of CeD because of the different T-cell responses they elicit,[41, 42] effects on autoimmunity may also rely on the accumulating effects of infectious agents, irrespective of type of infection, also known as the 'multiple hit concept'.[43] Intriguingly, our results indicate a cumulative effect of CMV, EBV and HSV-1 in relation to TG2A concentrations, irrespective of type of viral infection. The odds of TG2A positive concentrations of at least 70 U/ml were much stronger for multiple (≥2) herpesvirus infections (aOR 0.15), than for the virus-specific effects of CMV (aOR 0.38), suggesting that the effect on TG2A may be especially attributed to multiple herpesvirus infections, rather than to virus-specific infections. Although accumulated viral and microbial exposure have been suggested to influence CeD risk ([44]), this was not consistently repeated.[45] In fact, provocative effects, rather than protective, have been observed for infections early in life. For example, neonatal infections ([46]) and repeated infectious episodes early in life[44] were associated with increased CeD risk. Neonatal infections were, however, based on a retrospective analysis of the birth register, and infectious episodes early in life were based on parental reports. Hence, these analysis could not discriminate between various kinds of infections or pathogens. Potential infectious agents could be enterovirus,[37] adenovirus,[47] rotavirus[48],and campylobacter,[49] which affect gut microbiota or intestinal membrane permeability, rather than influencing adaptive immune responses.

Methodological considerations

To our knowledge, this is the first population-based cohort study that examined associations between herpesviruses and TG2A concentrations in children. A major strength of the present study is the large sample, representing a typical Western multicultural community and information on a broad range of potential confounders, including socioeconomic status, ethnicity, breast- bfeeding, daycare and antibiotic use. Second, we examined screening-identified patients instead of clinically identified CeD patients. Considering the fact that the majority of CeD patients remain undiagnosed due to often silent or atypical forms of the disease,[3, 11] a screening-based method will give the opportunity to study a more diverse and representative study population. In addition, primary infection with herpesviruses is often asymptomatic in childhood. Herpesviruses remain persistent after primary infection, even though virus particles can be undetectable. Therefore, IgG antibodies against these viruses provide a good correlate of viral infections in a healthy paediatric population.

Some limitations should, however, be taken into account. Firstly, causality remains uncertain, since antibodies against EBV-CA, CMV and HSV-1 were measured at the same time as TG2A. Therefore, reverse causation cannot be excluded: children with a more 'celiac disease orientated immune profile' may be less prone to these infections. Primary infection is, however, related to crowding,[50] rather than it is related to immune (dys)function. Hence, reverse causation is not likely. Unfortunately, we could not study the effect of timing of herpesvirus seroconversion on TG2A positivity, because we did not have data on IgM virus-specific antibodies, nor data on virus-specific antibodies at earlier ages. In addition, we only determined TG2A concentrations ate the age of 6 years, therefore it remains uncertain at what age children became TG2A positive. Secondly, although we adjusted for a broad range of confounders, the protective effect of viral infections may still be related to unmeasured variables. Hence, residual confounding cannot be fully excluded. Thirdly, we had 35% missing data on HLA-DQ2/DQ8 risk type because of missing cord blood samples. The presence of HLA-DQ2/DQ8 was, however, defined from cord blood samples, thus availability did not depend on CeD diagnosis. Therefore, missing data on HLA-DQ2/DQ8 are not likely to have influenced the results. Because previous studies have shown associations between TG2A positivity irrespective of CeD diagnosis,[38] and because the specificity of genetic risk typing is not 100%, we included the HLA-DQ2/DQ8 negative children. From a clinical perspective, it may be, however, argued that children with HLA-DQ2/DQ8 negative are not likely to develop CeD, and therefore should not be included in the analyses. Although the HLA-DQ2/DQ8*CMV interaction was nonsignificant, we stratified our analysis on the presence of HLA-DQ2/DQ8 for clinical interpretation purposes. The same trends between herpesvirus infections and strongly TG2A positivity were, however, observed in the HLA-DQ2/DQ8 positive group. Fourthly, we performed our studies in a large cohort, but because of the small number of patients stratified by high TG2A positivity, the limited statistical power may have affected the size and significance of the effect estimates. Furthermore, correcting for multiple confounding factors was not always possible. These factors may partly explain the nonsignificant results in the stratified analysis for the HLA-DQ2/DQ8 positive children. Lastly, although the sensitivity and specificity of TG2A concentrations are high, we studied Celiac Disease Autoimmunity (CDA) and not CeD. CeD can merely be diagnosed without intestinal biopsy on condition that TG2A concentrations are increased 10 ULN or more in accordance with both clinical symptoms and elevated endomysial antibody concentrations.[6] Therefore, conclusions concerning final CeD diagnosis should be made with caution. Nevertheless, even if children with strongly TG2A positivity lack criteria for final CeD diagnosis, the associations between herpesvirus infections and TG2A positivity remain valid.

Conclusion

In conclusion, this population based cohort study showed an inverse association between CMV infection and strongly TG2A positivity. In addition, infection with 2 or more herpes-

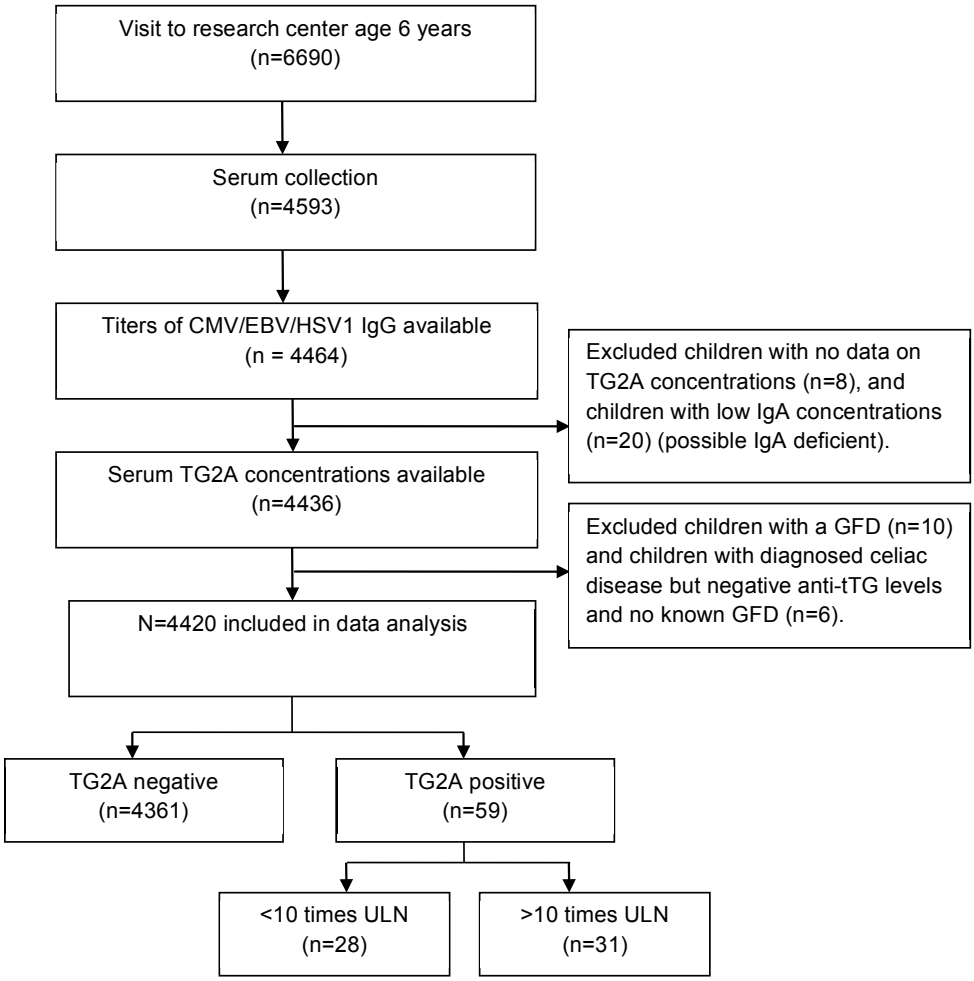
viruses (EBV, CMV or HSV-1) was inversely related to TG2A positivity. These 2 effects were independent of each other and indicate an accumulated protective effect of persistent viral infections on the development of celiac disease autoantibodies in childhood.

REFERENCES

1. Barzilai O, Sherer Y, Ram M, et al. Epstein-Barr virus and cytomegalovirus in autoimmune diseases: are they truly notorious? A preliminary report. Ann N Y Acad Sci 2007;1108:567-77.
2. Thomas D, Karachaliou F, Kallergi K, et al. Herpes virus antibodies seroprevalence in children with autoimmune thyroid disease. Endocrine 2008;33(2):171-5.
3. Guandalini S, Assiri A Celiac disease: a review. JAMA Pediatr 2014;168(3):272-8.
4. Liu E, Lee HS, Aronsson CA, et al. Risk of pediatric celiac disease according to HLA haplotype and country. N Engl J Med 2014;371(1):42-9.
5. Auricchio R, Tosco A, Piccolo E, et al. Potential celiac children: 9-year follow-up on a gluten-containing diet. Am J Gastroenterol 2014;109(6):913-21.
6. Husby S, Koletzko S, Korponay-Szabo IR, et al. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr 2012;54(1):136-60.
7. Tonutti E, Bizzaro N Diagnosis and classification of celiac disease and gluten sensitivity. Autoimmun Rev 2014;13(4-5):472-6.
8. Hopper AD, Hadjivassiliou M, Butt S, et al. Adult coeliac disease. BMJ 2007;335(7619):558-62.
9. Hill ID What are the sensitivity and specificity of serologic tests for celiac disease? Do sensitivity and specificity vary in different populations? Gastroenterology 2005;128(4 Suppl 1):S25-32.
10. Tommasini A, Not T, Kiren V, et al. Mass screening for coeliac disease using antihuman transglutaminase antibody assay. Arch Dis Child 2004;89(6):512-5.
11. Rubio-Tapia A, Kyle RA, Kaplan EL, et al. Increased prevalence and mortality in undiagnosed celiac disease. Gastroenterology 2009;137(1):88-93.
12. Catassi C, Kryszak D, Bhatti B, et al. Natural history of celiac disease autoimmunity in a USA cohort followed since 1974. Ann Med 2010;42(7):530-8.
13. Lohi S, Mustalahti K, Kaukinen K, et al. Increasing prevalence of coeliac disease over time. Aliment Pharmacol Ther 2007;26(9):1217-25.
14. White LE, Merrick VM, Bannerman E, et al. The rising incidence of celiac disease in Scotland. Pediatrics 2013;132(4):e924-31.
15. Takeuchi K, Tanaka-Taya K, Kazuyama Y, et al. Prevalence of Epstein-Barr virus in Japan: trends and future prediction. Pathol Int 2006;56(3):112-6.
16. Nilsson C, Larsson Sigfrinius AK, Montgomery SM, et al. Epstein-Barr virus and cytomegalovirus are differentially associated with numbers of cytokine-producing cells and early atopy. Clin Exp Allergy 2009;39(4):509-17.
17. Aarnisalo J, Ilonen J, Vainionpaa R, et al. Development of antibodies against cytomegalovirus, varicella-zoster virus and herpes simplex virus in Finland during the first eight years of life: a prospective study. Scand J Infect Dis 2003;35(10):750-3.
18. Pembrey L, Raynor P, Griffiths P, et al. Seroprevalence of cytomegalovirus, Epstein Barr virus and varicella zoster virus among pregnant women in Bradford: a cohort study. PLoS One 2013;8(11):e81881.
19. Dowd JB, Aiello AE, Alley DE Socioeconomic disparities in the seroprevalence of cytomegalovirus infection in the US population: NHANES III. Epidemiol Infect 2009;137(1):58-65.
20. West J, Fleming KM, Tata LJ, et al. Incidence and prevalence of celiac disease and dermatitis herpetiformis in the UK over two decades: population-based study. Am J Gastroenterol 2014;109(5):757-68.
21. Kruithof CJ, Kooijman MN, van Duijn CM, et al. The Generation R Study: Biobank update 2015. Eur J Epidemiol 2014;29(12):911-27.

22. van den Heuvel D, Jansen MA, Dik WA, et al. Cytomegalovirus- and Epstein-Barr Virus-Induced T-Cell Expansions in Young Children Do Not Impair Naive T-cell Populations or Vaccination Responses: The Generation R Study. J Infect Dis 2015.
23. Jansen MA, Tromp, II, Kiefte-de Jong JC, et al. Infant feeding and anti-tissue transglutaminase antibody concentrations in the Generation R Study. Am J Clin Nutr 2014;100(4):1095-101.
24. Monsuur AJ, de Bakker PI, Zhernakova A, et al. Effective detection of human leukocyte antigen risk alleles in celiac disease using tag single nucleotide polymorphisms. PLoS One 2008;3(5):e2270.
25. Kiefte-de Jong JC, Jaddoe VW, Uitterlinden AG, et al. Levels of antibodies against tissue transglutaminase during pregnancy are associated with reduced fetal weight and birth weight. Gastroenterology 2013;144(4):726-35 e2.
26. Swertz O, Duimelaar P, Thijssen J Migrants in the Netherlands 2004. Voorburg/Heerlen: Statistics Netherlands, 2004.
27. Statistics Netherlands Dutch Standard Classification of Education 2003. Voorburg/Heerlen: Statistics Netherlands, 2004.
28. Statistics Netherlands Welfare in the Netherlands. Income, welfare and spending of households and individuals. Den Haag/Heerlen: Statistics Netherlands, 2012.
29. Greenland S, Mickey RM Re: "The impact of confounder selection criteria on effect estimation. Am J Epidemiol 1989;130(5):1066.
30. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338(b2393.
31. Plot L, Amital H, Barzilai O, et al. Infections may have a protective role in the etiopathogenesis of celiac disease. Ann N Y Acad Sci 2009;1173(670-4.
32. Kivity S, Agmon-Levin N, Blank M, et al. Infections and autoimmunity--friends or foes? Trends Immunol 2009;30(8):409-14.
33. Appay V, van Lier RA, Sallusto F, et al. Phenotype and function of human T lymphocyte subsets: consensus and issues. Cytometry A 2008;73(11):975-83.
34. Almanzar G, Schwaiger S, Jenewein B, et al. Long-term cytomegalovirus infection leads to significant changes in the composition of the CD8+ T-cell repertoire, which may be the basis for an imbalance in the cytokine production profile in elderly persons. J Virol 2005;79(6):3675-83.
35. Brunner S, Herndler-Brandstetter D, Weinberger B, et al. Persistent viral infections and immune aging. Ageing Res Rev 2011;10(3):362-9.
36. Khan N, Shariff N, Cobbold M, et al. Cytomegalovirus seropositivity drives the CD8 T cell repertoire toward greater clonality in healthy elderly individuals. J Immunol 2002;169(4):1984-92.
37. Sarmiento L, Galvan JA, Cabrera-Rode E, et al. Type 1 diabetes associated and tissue transglutaminase autoantibodies in patients without type 1 diabetes and coeliac disease with confirmed viral infections. J Med Virol 2012;84(7):1049-53.
38. Ferrara F, Quaglia S, Caputo I, et al. Anti-transglutaminase antibodies in non-coeliac children suffering from infectious diseases. Clin Exp Immunol 2010;159(2):217-23.
39. Sfriso P, Ghirardello A, Botsios C, et al. Infections and autoimmunity: the multifaceted relationship. J Leukoc Biol 2010;87(3):385-95.
40. Krause I, Anaya JM, Fraser A, et al. Anti-infectious antibodies and autoimmune-associated autoantibodies in patients with type I diabetes mellitus and their close family members. Ann N Y Acad Sci 2009;1173(633-9.
41. Wada T, Muraoka M, Yokoyama T, et al. Cytokine profiles in children with primary Epstein-Barr virus infection. Pediatr Blood Cancer 2013;60(7):E46-8.
42. Kallas EG, Reynolds K, Andrews J, et al. Cytomegalovirus-specific IFNgamma and IL-4 are produced by antigen expanded human blood lymphocytes from seropositive volunteers. Immunol Lett 1998;64(2-3):63-9.

43. Janson C, Asbjornsdottir H, Birgisdottir A, et al. The effect of infectious burden on the prevalence of atopy and respiratory allergies in Iceland, Estonia, and Sweden. J Allergy Clin Immunol 2007;120(3):673-9.
44. Myleus A, Hernell O, Gothefors L, et al. Early infections are associated with increased risk for celiac disease: an incident case-referent study. BMC Pediatr 2012;12(194.
45. Welander A, Tjernberg AR, Montgomery SM, et al. Infectious disease and risk of later celiac disease in childhood. Pediatrics 2010;125(3):e530-6.
46. Sandberg-Bennich S, Dahlquist G, Kallen B Coeliac disease is associated with intrauterine growth and neonatal infections. Acta Paediatr 2002;91(1):30-3.
47. Kagnoff MF, Paterson YJ, Kumar PJ, et al. Evidence for the role of a human intestinal adenovirus in the pathogenesis of coeliac disease. Gut 1987;28(8):995-1001.
48. Stene LC, Honeyman MC, Hoffenberg EJ, et al. Rotavirus infection frequency and risk of celiac disease autoimmunity in early childhood: a longitudinal study. Am J Gastroenterol 2006;101(10):2333-40.
49. Plot L, Amital H Infectious associations of Celiac disease. Autoimmun Rev 2009;8(4):316-9.
50. Crowcroft NS, Vyse A, Brown DW, et al. Epidemiology of Epstein-Barr virus infection in pre-adolescent children: application of a new salivary method in Edinburgh, Scotland. J Epidemiol Community Health 1998;52(2):101-4.



Supplemental Figure 1. Flowchart of the participants within the Generation R Study

Supplemental Table 1. Maternal and child characteristics according to TG2A concentrations. Observed Dataset

N=4420	TG2A negative <7 U/ml		TG2A positive ≥7 U/ml			
			7-70 U/ml		≥70 U/ml	
Characteristics	n= 4361 (98.7%)		n= 28 (0.6%)		n= 31 (0.7%)	
Maternal characteristics						
Age (years; mean, sd)	30.7	5.1	30.7	5.4	**32.8**	3.4*
High educational level (n;%)	2931	74%	18	78%	17	61%
Household income per month (>2200 euro) (n;%)	1914	58%	11	58%	**21**	**81%***
Smoking during pregnancy (n;%)	946	25%	6	30%	6	24%
Caesarean section (n;%)	478	13%	3	12%	2	7%
Multiparity (n;%)	1890	45%	15	56%	12	39%
Child characteristics						
Male (n;%)	2258	52%	8	29%*	12	39%
Western ethnicity (n;%)	2825	65%	17	61%	**29**	**94%*** **
HLA DQ2 DQ8 carriership (n;%) [a]	1164	41%	**15**	**88%*** **	**23**	**92%*** **
Missing data (n;%)	*1512*	*35%*	*11*	*39%*	*6*	*19%*
Birth season (autumn/winter) (n;%)	2221	51%	10	36%	14	45%
Breastfeeding (n;%)						
< 6 months	1829	64%	8	50%	15	65%
≥ 6 months	1023	36%	8	50%	8	35%
Age (years; mean, sd)	6.2	0.6	6.3	0.6	6.1	0.3
Introduction of gluten (n;%)						
< 6 months	1260	43%	5	31%	15	58%
≥ 6 months	1644	57%	11	69%	11	42%
Day care attendance first year (n;%)	1665	88%	8	80%	21	96%
Gastro intestinal tract infections (n;%)						
First year	981	59%	3	33%	13	72%
Year 4-5						
Respiratory tract infections first year (n;%)						
First year	1642	60%	8	50%	18	72%
Year 4-5	840	24%	3	18%	6	22%
Vitamin D supplementation age 6-12 months (n;%)	1226	45%	9	56%	13	52%
Use of antibiotics / penicillin (n;%)						
First year	1203	(52%)	6	43%	7	33%
Ever	2461	61%	13	54%	19	63%

* Significantly different from TG2A negative group. *P<0.05, ** p<0.01, *** P<0.001.

[a] Of 4420 children, data on HLA DQ2/DQ8 carriership was present in n=2891 (65%) children. Of these, n=1202 (42%) were HLA DQ2/DQ8 positive, and 1689 (58%) children were HLA DQ2/DQ8 negative. Data on HLA DQ2/ DQ8 was missing in n=1529 (35%)children.

Supplemental Table 2. Association between herpesvirus infection and TG2A concentrations

	Total number of children	**TG2A negative <7 U/ml**	**TG2A positive ≥7 U/ml**			
	n=4420	n=4361 (98.7%)	**7-70 U/ml** n= 28 (0.6%)		**≥70 U/ml** n= 31 (0.7%)	
	N (%)	N (%)	N (%)	OR (95% CI)	N (%)	OR (95% CI)
CMV	1672 (38%)	1657 (38%)	10 (36%)	0.91 (0.42;1.96)	5 (16%)	**0.31 (0.12;0.82)***
No CMV	2748 (62%)	2704 (62%)	18 (64%)	*Reference*	26 (84%)	*Reference*
EBV	2257 (51%)	2234 (51%)	13 (46%)	0.82 (0.39;1.74)	10 (32%)	**0.45 (0.21;0.96)***
No EBV	2163 (49%)	2127 (49%)	15 (54%)	*Reference*	21 (68%)	*Reference*
HSV-1	871(20%)	865 (20%)	5 (18%)	0.88 (0.33;2.32)	1 (3%)	**0.14 (0.02;0.98)***
No HSV-1	3549 (80%)	3496 (80%)	23 (82%)	*Reference*	30 (97%)	*Reference*
Any (EBV, CMV and/or HSV-1)	3117 (71%)	3085 (71%)	19 (68%)	0.87 (0.39;1.94)	13 (42%)	**0.30 (0.15;0.61)****
None	1303 (29%)	1276 (29%)	9 (32%)	*Reference*	18 (58%)	*Reference*
Infection frequency						
0 infections	1303 (29%)	1276 (29%)	9 (32%)	*Reference*	18 (58%)	*Reference*
1 infection	1723 (39%)	1702 (39%)	10 (36%)	0.83 (0.34;2.06)	11 (36%)	**0.46 (0.21;0.97)***
≥ 2 infections	1394 (32%)	1383 (32%)	9 (32%)	0.92 (0.37;2.33)	2 (6%)	**0.10 (0.02;0.44)****

* Significantly different from TG2A negative group. *P<0.05, ** p<0.01, *** P<0.001. OR: Odds ratio; 95% CI: 95% confidence interval. ORs are from multiple imputed data and derived from logistic regression models. ORs are not adjusted for potential confounders. Adjusted ORs are shown in Table 2.

Supplemental Table 3. Frequency of co-infections in population for analysis

a. Prevalence of infections among groups

	N=4,420				
Viral Infections	EBV + *N*=2,257	CMV + *N*=1,672	HSV-1 + *N*=871	Any (EBV or CMV or HSV1) *N*=3,117	2 or more (EBV, CMV, and/or HSV-1) *N*=1,394
EBV + n (%)	-	987 (59%)	574 (66%)	2257 (72%)	1272 (91%)
CMV + n (%)	987 (44%)	-	411 (47%)	1672 (54%)	1109 (80%)
HSV-1 + n (%)	574 (25%)	411 (25%)	-	871 (28%)	696 (50%)

b. Prevalence of infections within CMV+ and CMV- groups

	N=4,420						
Viral Infections	CMV+ *N*=1,672	CMV – *N*=2,748	p-value	**Frequency viral co-Infections**	CMV+ *N*=1,672	CMV – *N*=2,748	p-value
				0 n (%)	0	1303 (47%)	<0.001
EBV + n (%)	987 (59%)	1270 (46%)	<0.001	1 n (%)	563 (34%)	1160 (42%)	<0.001
HSV-1 + n (%)	411 (25%)	460 (17%)	<0.001	2 n (%)	820 (49%)	285 (10%)	<0.001
EBV+ HSV-1+ n (%)	289 (17%)	285 (10%)	<0.001	3 n (%)	289 (17%)	0	<0.001

Supplemental Table 4. Association of herpesvirus infections and TG2A concentrations in HLA DQ2/DQ8 positive children.

	Total number of HLA DQ2/ DQ8 positive [1] children	HLA DQ2/ DQ8 positive TG2A negative <7 U/ml	HLA DQ2/ DQ8 positive TG2A positive ≥7 U/ml			
	n=1,202	n= 1,164	7-70 U/ml n=15 (1.2%)		≥70 U/ml n=23 (1.9%)	
			n (%)	OR (95% CI)	n (%)	OR (95% CI)
CMV	453 (38%)	443 (38%)	5 (33%)	0.81 (0.28, 2.40)	5 (22%)	0.45 (0.17, 1.23)
No CMV	749 (62%)	721 (62%)	10 (66%)	*Reference*	18 (78%)	*Reference*
EBV	603 (50%)	586 (50%)	7 (47%)	0.86 (0.31, 2.40)	10 (43%)	0.76 (0.33, 1.74)
No EBV	599 (50%)	578 (50%)	8 (53%)	*Reference*	13 (57%)	*Reference*
HSV-1	224 (19%)	221 (19%)	2 (13%)	0.66 (0.15, 2.93)	1 (4%)	0.19 (0.03, 1.45)
No HSV-1	978 (81%)	943 (81%)	13 (87%)	*Reference*	22 (96%)	*Reference*
Any (EBV, CMV and/or HSV-1)	839 (70%)	816 (70%)	10 (66%)	0.85 (0.29, 2.51)	13 (57%)	0.55 (0.24, 1.28)
None	363 (30%)	348 (30%)	5 (33%)	*Reference*	10 (43%)	*Reference*
Infection frequency						
0 infections	363 (30%)	348 (30%)	5 (33%)	*Reference*	10 (43%)	*Reference*
1 infection	473 (39%)	456 (39%)	6 (40%)	0.92 (0.28, 3.03)	11 (48%)	0.84 (0.35, 1.99)
≥ 2 infections	366 (30%)	360 (31%)	4 (27%)	0.77 (0.21, 2.90)	2 (9%)	**0.19 (0.04, 0.89)***

* Significantly different from TG2A negative group. OR: Odds ratio; 95% CI: 95% confidence interval. ORs are from multiple imputed data and derived from logistic regression models. ORs are not adjusted due to limited number of cases in each group.

[1] Of 4420 children, data on HLA DQ2 or DQ8 carriership was present in n=2891 (65%) children. Of these, n=1202 (42%) were HLA DQ2/DQ8 positive, and 1689 (58%) children did not carry the HLA DQ2/DQ8 risk type. Data on HLA DQ2/ DQ8 risk type was missing in n=1529 (35%)children.

Supplemental Table 5. Association of herpesvirus infections and TG2A positivity in Caucasian and non-Caucasian children.

n=2,887	TG2A positive (n=42) ≥7 U/ml	
	Caucasian n=28 (1.8%)	**Non-Caucasian** n=14 (1.0%)
	OR (95% CI)	OR (95% CI)
CMV	0.79 (0.33; 1.86)	0.43 (0.13; 1.36)
No CMV	*Reference*	*Reference*
EBV	0.91 (0.42; 1.96)	0.68 (0.24; 1.96)
No EBV	*Reference*	*Reference*
HSV-1	0.51 (0.12; 2.17)	0.25 (0.03; 1.91)
No HSV-1	*Reference*	*Reference*
Any (EBV, CMV and/or HSV-1)	0.82 (0.38; 1.74)	0.35 (0.12; 1.02)
None	*Reference*	*Reference*
Infection frequency		
0 infections	*Reference*	*Reference*
1 infection	0.99 (0.45; 2.18)	0.38 (0.11; 1.35)
≥ 2 infections	0.46 (0.13; 1.66)	0.33 (0.09; 1.19)

* Significantly different from TG2A negative group. OR: Odds ratio; 95% CI: 95% confidence interval. ORs are from multiple imputed data and derived from logistic regression models. ORs are not adjusted due to limited number of cases in each group.[1] Of 4420 children, data on herpesvirus serology and ethnic background (GWAS Caucasians versus non-Caucasians) was present in n=2,887 (65%) children.

Supplemental Table 6. Characteristics in Caucasians and non-Caucasians

n=2,887	Caucasian	Non-Caucasian
	n=1,540 (53%)	n=1,347 (47%)
Age (median; range)	**6.0 (4.9 – 9.0)*** **	6.1 (5.0 – 9.0)
Gender (male)	784 (50.9%)	682 (50.6%)
HLA DQ2 or DQ8 positive	**679 (44.1%)*** **	520 (38.6%)
TG2A positivity (≥7 U/ml)	28 (1.8%)	14 (1.0%)
Strong positivity (≥70 U/ml)	18 (1.2%)	7 (0.5%)
CMV	**457 (29.7%)*** **	650 (48.3%)
No CMV	1083 (70.3%)	697 (51.7%)
EBV	**639 (41.5%)*** **	799 (59.3%)
No EBV	901 (58.5%)	548 (40.7%)
HSV-1	**200 (13%)*** **	316 (23.5%)
No HSV-1	1340 (87%)	1031 (75.5%)
Any (EBV, CMV and/or HSV-1)	**954 (61.9%)*** **	1062 (78.8%)
None	586 (38.1%)	285 (21.2%)
Infection frequency		
0 infections	**586 (38.1%)*** **	285 (21.2%)
1 infection	**642 (41.7%)*** **	498 (37.0%)
≥ 2 infections	**312 (20.2%)*** **	564 (41.8%)

* Significantly different from Non-Caucasian group, based on chi-square test for categorical variables, and Mann-Whitney U test for non-normally distributed variables

Chapter 9

Ethnic differences in coeliac disease autoimmunity in Childhood:

The Generation R Study

Michelle A.E. Jansen, Sytske A. Beth, Diana van den Heuvel, Jessica C. Kiefte-de Jong, Hein Raat, Vincent W.V. Jaddoe, Menno C. van Zelm, Henriette A. Moll

Arch Dis Child. 2017 Jun; 102(6):529-534

ABSTRACT

Objective

The aim was to identify whether ethnic differences in celiac disease autoimmunity (CDA) in children at 6 years of age exist, and when present, to evaluate how these differences may be explained by sociodemographic and environmental factors.

Design

This study was embedded within a multi-ethnic population-based prospective cohort study.

Setting and patients

4442 six-year old children born between 2002 and 2006 were included. Information on ethnicity, environmental and lifestyle characteristics was assessed by questionnaires. Ethnicity was categorised into Western (Dutch, European, Indonesian, American, Oceanian) and non-Western (Turkish, Moroccan, Cape Verdean, Antillean, Surinamese). Serum transglutaminase type 2 antibody (TG2A) levels were measured with fluorescence enzyme immunoassay. Serum IgG levels against cytomegalovirus (CMV) were measured by ELISA.

Main outcome measures

TG2A positivity was defined as TG2A $\geq$7 U/ml, strong TG2A positivity as TG2A $\geq$10 upper limit normal (70 U/ml).

Results

Of 4442 children, 60 (1.4%) children were TG2A positive, of whom 31 were strong positive. 66% of children were Western, 33% non-Western. Western ethnicity, high socioeconomic position and daycare attendance were positively associated with strong TG2A positivity (odds ratio (OR) 6.85 (1.62-28.8) p<0.01, OR 3.70 (1.40-9.82) p<0.01, OR 3.90 (1.38-11.0) p=0.01 resp.), whereas CMV seropositivity was inversely related to strong TG2A positivity (OR 0.32 (0.12-0.84) p=0.02). Together, these factors explained up to 47% (-67 to -17; p=0.02) of the ethnic differences in TG2A positivity between Western and Non-Western children.

Conclusions

Ethnic differences in children with CDA are present in childhood. Socioeconomic position, daycare attendance and CMV seropositivity partly explained these differences, which may serve as targets for prevention strategies for CDA.

INTRODUCTION

Celiac Disease (CeD) is a common autoimmune disease with an increasing prevalence of 1% in the general population.[1] The rising prevalence of CeD cannot be merely attributed to increased diagnosis among physicians, nor changes in genetic factors.[2-6] Hence, socio-economic and lifestyle factors should contribute. In addition, CeD has been reported to be more common in Western countries among whites than in individuals from other ethnicities.[2-9] Intriguingly, a recent study showed that the prevalence of CeD autoimmunity (CDA) was four to eight times higher among non-Hispanic whites than among other ethnicities.[10] This suggests that ethnic differences are present both in CeD and in CDA. Although these individuals lack symptoms suggestive for CeD, antibodies against transglutaminase type 2 (TG2A) are present, thereby possibly indicating the presence of subclinical or potential CeD.[11] Furthermore, incidence and prevalence of CeD vary across different geographical areas of the world.[8] Apart from differences in human leucocyte antigen (HLA)-DQ2/DQ8 carriership among different ethnic groups,[12, 13] variations in environmental factors, such as gluten consumption and smoking behaviour, might explain some of the geographical variations in prevalence of CeD.[10, 14] Other studies suggested socioeconomic variation as an explanatory factor, as children with a higher socioeconomic position (SEP) may face different infectious agents during early childhood.[15-20] Consequently, a different programming of their immune system occurs.[21] Furthermore, these children might seek healthcare more often.[15, 20] However, it still remains unclear to what extent ethnic differences in childhood CDA are related to environmental factors. Therefore, we here aimed to identify whether ethnic differences in CDA in children aged 6 years exist, and when present, to evaluate how these differences may be explained by sociodemographic and environmental factors.

PATIENTS AND METHODS

Study design

This study was part of the Generation R Study, a prospective multi-ethnic population-based cohort study from early pregnancy onward. All children were born between April 2002 and January 2006 in Rotterdam. The study area covers more than one-half of the city's inhabitants, reflecting Rotterdam's multi-ethnic population.[22] The study was approved by the Medical Ethics Committee of the Erasmus Medical Center. Written informed consent was obtained from all of the participating parents.[23]

At the median age of 6 years, 6690 children visited our research centre, of whom 4593 provided serum. After exclusion of children with no data on TG2A levels (n=131) and children with serum IgA levels below the detection limit (n=20), a population of 4442 children remained eligible for analysis [Supplemental Figure 1].

Ethnic background

Ethnicity of the child was determined by country of birth of the parents.[24-26] The child was of non-native Dutch origin if one of the parents was born in another country than the Netherlands.[24] If both parents were born abroad, the mother's country of birth prevailed. To be able to distinguish between all different ethnicities, detailed information on parental ethnic background based on country of birth was assessed by questionnaires at enrolment. Briefly, because of recent migration history, we were able to differentiate between Western and non-Western ethnicity. Western ethnicity included Dutch, European, American Western (including North American), Asian Western (including Indonesian and Japanese) and Oceanian. Non-Western ethnicity included Turkish, Moroccan, Surinamese, Antillean, Cape Verdean, African, Asian non-Western (all Asian counties except Indonesia and Japan) and American non-Western (including South American and Central American).[25, 27-29] [Supplemental Table 1].

Detection of Transglutaminase Type 2 Antibody Levels

Serum IgA levels against TG2A were measured using a fluorescence enzyme immunoassay (Elia Celikey IgA, Phadia ImmunoCAP 250, Phadia AB, Uppsala, Sweden), at the child's median age of 6.0 years. TG2A levels of ≥7 U/ml were considered to be positive as per the manufacturer's instructions. In addition, TG2A-positive levels were further subdivided into two categories based on the ≥10 times upper limit normal (ULN) of the test kit (≥7-70 U/ml, and ≥70 U/ml, ie, TG2A strong positives).[30-32]

Potential Mediating variables

SEP, parity (as indicator for siblings/family size), daycare and herpesvirus infections were suggested to be potential explanatory variables in the pathway between ethnic background and TG2A positivity.[2-4, 7, 8, 20] Information on household income,[33] and educational level,[34] both indicators of SEP, was obtained from prenatal questionnaires. Herpesvirus seropositivity (IgG against Cytomegalovirus (CMV, Epstein-Barr virus (EBV), and herpes simplex virus type 1 (HSV-1)) was determined from plasma samples collected at 6 years of age.[21] Data on gestational age, birth weight, parity and mode of delivery were obtained from obstetric records. Prenatal questionnaires included information on maternal lifestyle, including smoking and alcohol use during pregnancy. Data on breastfeeding was obtained by a combination of delivery reports and postnatal questionnaires at 2, 6 and 12 months. Information on the timing of gluten introduction was assessed by questionnaires at the age of 6 and 12 months.[32] Postnatal questionnaires included information on daycare attendance, respiratory and gastrointestinal tract infections.

Potential confounding variables

Child's age and gender (data obtained from obstetric records) were treated as potential confounders in the association between ethnicity and TG2A positivity, because they cannot be considered indisputable explanatory variables.[35, 36]

Statistical analysis

First, χ^2 tests, independent student's T-tests and Mann-Whitney U tests were used to test whether characteristics were different between Western and Non-Western children. Logistic regression analyses were performed with TG2A positivity as dependent variables (y/n). To study to which extent associations between ethnicity and TG2A (strong) positivity were explained by factors associated with SEP, crowding, breastfeeding behaviour and maternal lifestyle, we used Baron and Kenney's causal step approach.[35] Only those factors that were unequally distributed across the ethnic groups, and were significantly associated with TG2A (strong) positivity (independent of ethnicity, gender and age), were added separately to the model.[35] To assess their mediating effects, the corresponding percentages of attenuation of effect estimates were calculated by comparing differences of model 1 with the adjusted ones (100 x (OR model 1 – OR model 2 with explanatory factor)/ (OR model 1)). Finally, a full model containing ethnicity and all the explanatory factors assessed the joint effects of explanatory factors. To reduce attrition bias, multiple imputation of the determinant and covariates was performed (n=5 imputations). Regression coefficients were pooled by taking the average of the coefficients of the five imputed datasets. The pooled results of the five imputed datasets were reported in this paper as Odds Ratios (ORs) and 95% confidence intervals (CI's). A 95% CI was calculated around the percentage attenuation using a bootstrap method with 1000 resamplings per imputed dataset in the statistical program R. All other statistical analyses were performed in SPSS 21.0 for Windows (IBM Corp., Armonk, NY, USA).

RESULTS

Of 4442 children, n=2854 (66%) were Western, and n=1472 (33%) non-Western. Western children were slightly younger *($p<0.001$)*, had a slightly lower body mass index (*$p<0.001$*), their mothers were higher educated (*$p<0.001$*), and their family household income was higher (*$p<0.001$*) than non-Western children. They also had a slightly higher birth weight, were less often breastfed (*$p<0.001$*), and more frequently introduced to gluten before 6 months of age (*$p<0.001$*) than non-Western children. In addition, Western children had less often gastrointestinal and respiratory tract infections during the first year of life (*$p<0.001$*), and Western children were less often infected with CMV, EBV and HSV-1 than non-Western children *($p<0.001$)*,[Table 1].

Table 1. Associations between ethnic background and potential confounders and mediators

	Ethnic background			p-value[a]
	Total	Western	Non-Western	
	(n=4442)	(n= 2854)	(n=1472)	
Potential confounders[b]				
Age (years; mean; SD)	6.2 (0.6)	6.1 (0.47)	6.3 (0.64)	**<0.001**
Female (n; %)	2,155 (48.5)	1,409 (49.3)	687 (46.7)	0.09
Potential mediators				
Maternal educational level (n; %)				
Low	863 (21.4)	366 (13.3)	497 (38.3)	**<0.001**
Mid-low	1,230 (30.5)	741 (27.1)	489 (37.7)	
Mid-high	886 (22.0)	697 (25.5)	188 (14.5)	
High	1,056 (26.2)	933 (34.1)	123 (9.5)	
Net household income (n; %)				
≤ 2200 euro	1,386 (41.4)	633 (27.2)	732 (73.7)	**<0.001**
> 2200 euro	1,959 (58.6)	1,696 (72.8)	260 (26.3)	
Maternal smoking (n; %)	960 (24.9)	634 (24.7)	315 (24.9)	0.90
Maternal alcohol use (n; %)	1,745 (52.1)	1,138 (52.3)	559 (51.4)	0.63
Birth weight (grams; mean; SD)	3413 (569)	3460 (575)	3330 (540)	**<0.001**
Gestational age (weeks; median; range)	40.0 (26.3-43.6)	40.1 (26.9-43.3)	39.9 (27.1-43.5)	**<0.001**
Caesarean section (n; %)	485 (12.7)	328 (13.5)	141 (10.9)	0.03
Daycare (n; % yes)	1,543 (60.3)	1,317 (68.1)	225 (36.8)	**<0.001**
Multiparity (n; %)	1,927 (45.0)	1,100 (40.0)	772 (53.8)	**<0.001**
Breastfeeding (n; %)				
yes	3,232 (92.5)	2,210 (91.4)	968 (95.3)	**<0.001**
no	263 (7.5)	209 (8.6)	48 (4.7)	
Gluten introduction >6 months (n; %)	1,285 (43.4)	846 (38.7)	432 (57.3)	**<0.001**
Gastrointestinal tract infection (n; %)	1,001 (59.0)	731 (55.4)	265 (71.6)	**<0.001**
Respiratory tract infection (n; %)	1,679 (59.8)	1,185 (56.8)	481 (68.2)	**<0.001**
CMV seropositivity (n; %)	1,681 (37.9)	872 (30.6)	751 (51.0)	**<0.001**
EBV seropositivity (n; %)	2265 (51.1)	1,258 (44.2)	923 (62.7)	**<0.001**
HSV-1 seropositivity (n; %)	873 (19.7)	405 (14.2)	425 (28.9)	**<0.001**
SDS BMI child (kg/m^2)	0.28 (0.91)	0.17 (0.82)	0.48 (1.03)	**<0.001**

Abbreviations: BMI, Body Mass index (kg/m^2); CMV, cytomegalovirus; EBV Epstein-Bar virus; HSV-1 herpes simplex virus type 1; SDS, standard deviation score; BMI was adjusted for age. Values are percentages, means (SD) or medians (range), based on observed dataset. [a] p-values are calculated with the Chi-square test for categorical variables, Mann-Whitney U test for non- normally distributed variables and Student's T test for continuous normal distributed variables, and reflects differences in characteristics between Western and Non-western groups. [b] Data was missing for maternal educational level (9.2%), net household income (24.7%), maternal smoking (13.3%), maternal alcohol use (24.6%), birth weight (<0.01%), gestational age (<1%), daycare (42.4%), parity (3.6%), breastfeeding (21.3%), timing of gluten introduction (33.3%), GI 61.8%, RTI (36.8%), CMV (<0.01%), EBV (0.1%), HSV-1 (<0.01%) (22.3%).

Ethnic background and TG2A positivity

Western ethnicity was positively associated with TG2A positivity (adjusted odds ratio (aOR) 1.94; 95% CI 1.02-3.70). The association was mainly driven by the children who were TG2A strong positive; (i.e. TG2A levels ≥70 U/ml; aOR Western ethnicity 6.85; 95% CI 1.62- 28.8),[Table 3, model 1]. CMV seropositivity was related to both ethnicity and TG2A positivity,[Table 1 and 2]. Therefore, CMV seropositivity was selected as explanatory variable in the association between ethnicity and strong TG2A positivity (35). CMV seropositivity tended to explain 9% (95% CI -18 to 0.5, *P=0.059*) of the differences in TG2A positivity between Western and Non-Western children, although the attenuation in OR was borderline significant,[Table 3]. However, variations in SEP, daycare and CMV prevalence together explained up to 29% (95% CI -47 to -5, *p=0.018*) of the ethnic differences in TG2A positivity. EBV and HSV-1 added another 4% to the explanation of ethnic differences in TG2A positivity. In total, 33% (95% CI -50 to -13, *p=0.004*) of differences in TG2A positivity between Western and non-Western could be explained by differences in SEP, daycare, and herpesvirus seropositivity [Table 3].

Ethnic background and strong TG2A positivity

High SEP (high maternal educational level and high net household income), frequent daycare attendance and CMV seropositivity were related to ethnic background and strong TG2A positivity [Table 1 and Table 2]. Therefore, these variables were selected as explanatory variables in the association between ethnicity and strong TG2A positivity.[35] SEP tended to be the most important contributor to the association between ethnicity and TG2A strong positivity [Table 3]. For example, net household income tended to explain 31% (95% CI -58 to -3, *p=0.07*) of the differences in TG2A strong positivity between Western and Non-Western children, followed by maternal educational level, which tended to explain 22% (95% CI -37 to 5, *p=0.13*) [Table 3]. Daycare attendance tended to explain 26% (95% CI -36 to 0.6, *p=0.056*) of the association between ethnicity and strong TG2A positivity. However, the alterations in effect estimates induced by the single explanatory variables did not reach statistical significance. In contrast, CMV seropositivity significantly explained 14% (95% CI -25 to -0.4; *p=0.04*) of the ethnic differences in strong TG2A positivity. Overall, SEP, daycare and CMV seropositivity together explained up to 47% (95% CI -67 to -17, *p=0.02*) of the ethnic difference in TG2A strong positivity [Table 3]. EBV and HSV-1 tended to explain another 4% to 51% (95% CI -68 to 21) of ethnic differences in strong TG2A positivity, although the alteration of the OR in the fully adjusted model no longer remained significant (*p=0.08*) [Table 3].

Table 2. Associations between potential mediators and TG2A positivity (n=4442)

	TG2A positivity			
	TG2A pos. (n=60; ≥7 U/ml)	p-value [a]	TG2A Strong positivity (n=31; ≥70 U/ml)	p-value [b]
Potential mediators				
Maternal educational level (%)				
Low	Ref		Ref	
High	1.63 (0.95; 2.81)	0.08	**3.10 (1.22; 7.93)**	**0.02**
Net household income (%)				
≤ 2200 euro	Ref		Ref	
> 2200 euro	1.85 (0.98; 3.49)	0.056	**3.70 (1.40; 9.82)**	**0.009**
Maternal smoking (%)	1.03 (0.57; 1.88)	0.92	1.02 (0.42; 2.51)	0.96
Maternal alcohol use (%)	1.26 (0.69, 2.30)	0.45	1.09 (0.52; 2.31)	0.82
Birth weight (grams)	1.00 (1.00;1.00)	0.56	1.00 (1.00;1.00)	0.77
Gestational age (weeks)	1.03 (0.89;1.18)	0.76	1.06 (0.86; 1.31)	0.61
Caesarean Section (%)	0.75 (0.29; 1.95)	0.56	0.58 (0.12; 2.72)	0.48
Daycare (% yes)	1.70 (0.98; 2.94)	0.06	**3.90 (1.38; 11.0)**	**0.011**
Multiparity (%)	1.08 (0.64; 1.84)		0.80 (0.39; 1.66)	0.55
Breastfeeding (%)				
yes	Ref		Ref	
no	1.09 (0.46: 2.57)	0.85	2.05 (0.33; 12.65)	0.43
Timing of gluten introduction (%)				
<6 months	Ref		Ref	
>6 months	0.86 (0.36; 2.02)	0.70	0.58 (0.25; 1.35)	0.21
Gastrointestinal Tract infection (%)	1.07 (0.46; 2.50)	0.86	1.55 (0.73; 3.13)	0.26
Respiratory Tract infection (%)	1.17 (0.66; 2.09)	0.59	1.69 (0.72; 3.98)	0.22
CMV seropositivity (%)	**0.53 (0.30; 0.96)**	**0.04**	**0.32 (0.12; 0.84)**	**0.02**
EBV seropositivity (%)	0.65 (0.38; 1.09)	0.10	0.49 (0.23; 1.04)	0.06
HSV-1 seropositivity (%)	0.46 (0.20; 1.07)	0.07	0.15 (0.02; 1.09)	0.06

Abbreviations: OR, Odds Ratio; CI, confidence interval. Values are derived from multivariable logistic regression models (based on imputed dataset) and represent effect estimates (ORs; 95% confidence intervals), adjusted for child's gender and age.

[a]p-value reflects differences in characteristics between TG2A positive children (≥7 U/ml) and TG2A negative controls.

[b]p-value reflects differences in characteristics between TG2A strong positive children (≥70 U/ml) and TG2A negative controls.

DISCUSSION

This large multi-ethnic prospective study confirms differences in CDA between children from Western and non-Western ethnicity. Furthermore, SEP, daycare attendance and CMV seropositivity partly explained ethnic differences in TG2A positivity.

Table 3. Attenuation of the OR of TG2A positivity for Western Ethnicities relative to Non-Western ethnicities after individual adjustment for the explanatory variables

	TG2A positivity (n=60; ≥7 U/ml)			Strong TG2A positivity (n=31; ≥70 U/ml)		p[b]
	Western ethnicity OR[a] (95% CI)	Attenuation[b] (95% CI)	p[b]	Western ethnicity OR[a] (95% CI)	Attenuation[b] (95% CI)	
Model 1 (includes child sex and current age)	**1.94 (1.02; 3.70)***	-		**6.85 (1.62; 28.8)****	-	
Socioeconomic status						
Model 1+ education	1.75 (0.89; 3.44)	-10% (-25 to 6)	0.18	**5.36 (1.24; 23.11)***	-22% (-37 to 5)	0.13
Model 1+ household income	1.60 (0.77; 3.30)	-17% (-37 to 8)	0.17	**4.72 (1.06; 21.05)***	-31% (-58 to 3)	0.07
Model 1 + daycare	1.73 (0.88; 3.40)	-11% (-23; 3)	0.14	**5.08 (1.18; 21.89)***	-26% (-36 to 0.6)	0.056
Model 1 + CMV carriership	1.76 (0.92; 3.39)	-9% (-18; 0.5)	0.059	**5.89 (1.39; 24.89)***	**-14% (-25; -0.4)***	0.04
Model 1 + EBV carriership	1.83 (0.95; 3.51)	-6% (-15; 3)	0.18	**6.26 (1.48; 26.48)***	-9% (-20; 4)	0.17
Model 1 + HSV-1 carriership	1.80 (0.94; 3.45)	**-8% (-14; -0.1)***	0.047	**6.00 (1.42; 25.29)***	-12% (-19; 107)	0.10
Fully adjusted model without EBV and HSV-1	1.37 (0.66; 2.86)	**-29% (-47; -5)***	0.018	3.47 (0.77; 15.5)	**-47% (-67 to -17)***	0.02
Fully adjusted model with EBV and HSV-1	1.28 (0.62; 2.68)	**-33% (-50;-13)****	**0.004**	3.17 (0.71; 14.15)	-51% (-68; 21)	0.08

Abbreviations: OR, Odds Ratio; CI, confidence interval.

[a] The OR's represent the attenuated effect estimate and their 95% CIs relative to model 1 after individual adjustment for explanatory factors.

[b] Attenuations represents the attenuations of effect estimates for Western ethnicities relative to model 1 (includes confounders) after adjustment for explanatory variables (100x ($OR_{model\ 1} - OR_{model\ 1\ with\ explanatory\ factor}$) / ($OR_{model\ 1}$)). Non-Western ethnicity is the reference group. * p<0.05, ** p<0.01, *** p<0.001. The pooled OR's and 95% CI's are derived from imputed datasets analyzed in SPSS, the pooled attenuation of effect estimates and 95% CI's of the attenuation are derived from pooled datasets analyzed in R.

Our results confirm previous studies by the observation that CDA is more common in Western than in non-Western populations.[2, 3, 7, 9, 10] It needs to be considered though that we lacked data on intestinal biopsies, thereby impeding conclusions on final CeD diagnosis. Still, given the high sensitivity and specificity of TG2A, even in unselected screened populations,[37] these children are at increased risk to develop CeD.

Interestingly, recent studies show an increase in CeD prevalence in non-Western, urban areas of the de world.[38, 39] Possible explanations for this finding are more awareness of the disease, a wider use of serological screening tests and the worldwide shift to Western, gluten-rich dietary patterns. Though, the inter-country differences in CeD prevalence cannot be convincingly explained by different predictability of HLA DQ2/DQ8 haplotypes and dietary patterns only. Other factors, such as SEP, environmental exposures (e.g. infections) and/or the presence of protective genes in non-whites,[7] may also contribute to the ethnic and geographic variety in CeD prevalence.[8, 13-18] Still, studies investigating the relative contribution of each factor by multivariate models in an undiagnosed population are scarce.[9, 15-20] One study compared two Western populations with the same genetic ancestry, living in comparable geographical areas and equal exposures to gluten-containing cereals, but with markedly different lifestyle and socioeconomic circumstances. In line with our results, they found a significantly lower prevalence of CeD in screened children with a lower SEP.[15] Differences in lifestyle and SEP, including the frequency of intestinal infections, antibiotic use and a variety of dietary factors other than gluten, may lead to differences in the gut microbial flora, thereby affecting the programming of the immune system.[15] This so called hygiene or 'old friends hypothesis' [40] has been linked to the emergence of atopic- and autoimmune disease. Individuals living in better hygienic conditions, that is, in urban environments of rich developed countries, face fewer infections during childhood.[41] Consequently, immune regulatory mechanisms become less efficient, leading to intolerance to common harmless allergens.[40, 41] Other environmental factors, such as delayed exposure to viruses or vitamin D deficiency may add to this explanation.[40, 42] Disturbances in immunoregulatory processes might act indirectly via changes in the microbiota of the gut.[40, 43] Still, it remains debated whether CeD patients have a different microbiome.[44] Another explanation for the association between SEP and CeD may be that children/parents with a higher SEP are more likely to seek healthcare. [20] We and others [15] confirmed this socioeconomic gradient in a screened and asymptomatic group of children in the healthy population. However, children in our study were unaware of TG2A assessment, and no differences were observed in gastrointestinal complaints.[31] Moreover, none of the TG2A positive children had visited a general practitioner for their complaints. Hence, the role of seeking healthcare may be less important. On the basis of our results, though, daycare attendance and herpesvirus infections may contribute. Children who attend daycare are likely to face various pathogens early in life. Several studies reported conflicting results on associations between infectious diseases in early life and CeD risk.[45, 46] Although some have suggested a causative linkage between autoimmune disease and CMV infection, overall

evidence does not convincingly support this.[45, 46] Still, CMV might play a role in either the onset, progression or amelioration of autoimmune diseases.[38] In addition, other herpesviruses might protect against the development of CeD.[46] In line with this, we found that EBV and HSV-1 added both to the explanation of ethnic differences in CDA. Previously, we observed inverse associations between combined CMV, EBV and/or HSV-1 infections and TG2A positivity, which may indicate a cumulative effect of herpesviruses, irrespective of type of viral infection.[47] CMV-infected children were more often infected with other herpesviruses. Hence, CMV infection itself may not partly explain the ethnic variety in CDA prevalence, but it may reflect a diverse viral and microbial exposure in early life, which, together with genetic and other environmental factors, might influence the development of CDA.

Methodological considerations

A major strength of this study is our large screened population-based multi-ethnic cohort. Mothers and children were unaware of TG2A screening at the moment of blood sampling, thereby minimizing the risk of selection bias. To the best of our knowledge, explaining ethnic differences in CDA by using mediation analysis is novel. Since environmental factors, besides genetic susceptibility, contribute to the pathogenesis of CeD, it might be of value that we were able to assess their relative contributions. Despite the large amount of immigrants in Rotterdam, the number of cases, though, was too small to divide groups per ethnicity. Therefore, we divided our study population into Western and non-Western background. Literature is scarce on ethnic differences in CeD other than Western compared to non-Western. This might well be explained by the low prevalence of CeD in non-Western groups,[7, 8, 10] although recent findings have shown that the CeD prevalence is increasing in non-Western populations as in Asia and Africa.[38, 39] SEP, daycare and CMV seropositivity together explained 47% of the ethnic differences in CDA. The strong correlation between these factors is likely to be responsible for this finding. Only CMV seropositivity explained a significant 14% on its own. Although reverse causation cannot be totally excluded, the majority (80%) of CMV positive children in our study was infected before the age of 2 years.[21] Therefore, we consider reverse causation less likely. Last, we cannot completely exclude the possibility of residual confounding. That is, other environmental factors, such as gluten intake or infectious disease in early childhood, might play a role. Hence, the possible interactions between genetics, gluten and other environmental determinants should be further investigated.

CONCLUSION

In this multi-ethnic population-based cohort, we found higher seroprevalence of TG2A positivity in Western children relative to non-Western children at 6 years of age. High so-

cioeconomic position, daycare attendance and CMV infection partly explained the ethnic differences in TG2A positivity. These factors may serve as targets for primary and secondary prevention strategies for CDA.

What's known about this subject:

- Incidence and prevalence of Celiac Disease (CeD) and CeD Autoimmunity vary across different ethnicities and geographical areas of the world.
- It is less clear to what extent these ethnic differences can be explained by socioeconomic position and other lifestyle-related factors.

What this study adds:

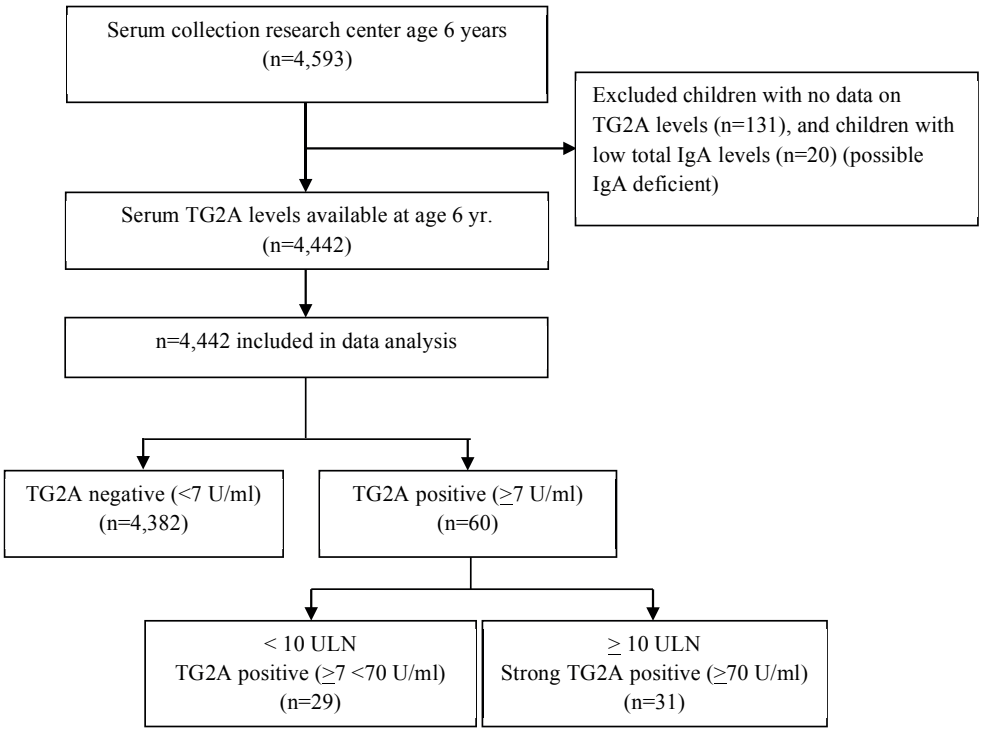
- Nearly 50% of the ethnic differences in transglutaminase type 2 antibody (TG2A) positivity between Western and non-Western children can be explained by socioeconomic position and lifestyle-related factors.
- High socioeconomic position, and the absence of cytomegalovirus (CMV) infection in Western children are associated with increased risk of CeD Autoimmunity.
- Variation in socioeconomic position was considered to be the most important explaining factor, followed by cytomegalovirus (CMV) infection.

REFERENCES

1. Guandalini S, Assiri A. Celiac disease: a review. *JAMA Pediatr*. 2014;168:272-8
2. Lohi S, Mustalahti K, Kaukinen K, *et al*. Increasing prevalence of coeliac disease over time. *Aliment Pharmacol Ther*. 2007;26:1217-25
3. Riddle MS, Murray JA, Porter CK. The Incidence and Risk of Celiac Disease in a Healthy US Adult Population. *American Journal of Gastroenterology*. 2012;107:1248-55
4. Rubio-Tapia A, Kyle RA, Kaplan EL, *et al*. Increased Prevalence and Mortality in Undiagnosed Celiac Disease. *Gastroenterology*. 2009;137:88-93
5. West J, Fleming KM, Tata LJ, Card TR, Crooks CJ. Incidence and prevalence of celiac disease and dermatitis herpetiformis in the UK over two decades: population-based study. *Am J Gastroenterol*. 2014;109:757-68
6. White LE, Merrick VM, Bannerman E, *et al*. The rising incidence of celiac disease in Scotland. *Pediatrics*. 2013;132:e924-31
7. Choung RS, Ditah IC, Nadeau AM, *et al*. Trends and racial/ethnic disparities in gluten-sensitive problems in the United States: findings from the national health and nutrition examination surveys from 1988 to 2012. *American Journal of Gastroenterology*. 2015;110:455-61
8. Kang JY, Kang AHY, Green A, Gwee KA, Ho KY. Systematic review: worldwide variation in the frequency of coeliac disease and changes over time. *Alimentary Pharmacology & Therapeutics*. 2013;38:226-45
9. Burger JP, Roovers EA, Drenth JP, Meijer JW, Wahab PJ. Rising incidence of celiac disease in the Netherlands; an analysis of temporal trends from 1995 to 2010. *Scand J Gastroenterol*. 2014;49:933-41
10. Mardini HE, Westgate P, Grigorian AY. Racial Differences in the Prevalence of Celiac Disease in the US Population: National Health and Nutrition Examination Survey (NHANES) 2009-2012. *Dig Dis Sci*. 2015;60:1738-42
11. Ludvigsson JF, Leffler DA, Bai JC, *et al*. The Oslo definitions for coeliac disease and related terms. *Gut*. 2013;62:43-52
12. Monsuur AJ, de Bakker PI, Zhernakova A, *et al*. Effective detection of human leukocyte antigen risk alleles in celiac disease using tag single nucleotide polymorphisms. *PLoS One*. 2008;3:e2270
13. Liu E, Lee HS, Aronsson CA, *et al*. Risk of pediatric celiac disease according to HLA haplotype and country. *N Engl J Med*. 2014;371:42-9
14. van Zanten SV. Could population changes in smoking habits help explain the change in incidence and prevalence of celiac disease? *American Journal of Gastroenterology*. 2014;1953 109:1953
15. Kondrashova A, Mustalahti K, Kaukinen K, *et al*. Lower economic status and inferior hygienic environment may protect against celiac disease. *Ann Med*. 2008;40:223-31
16. Ludvigsson JF, Group AS. Socio-economic characteristics in children with coeliac disease. *Acta Paediatr*. 2005;94:107-13
17. Olen O, Bihagen E, Rasmussen F, Ludvigsson JF. Socioeconomic position and education in patients with coeliac disease. *Dig Liver Dis*. 2012;44:471-6
18. Whyte LA, Kotecha S, Watkins WJ, Jenkins HR. Coeliac disease is more common in children with high socio-economic status. *Acta Paediatr*. 2014;103:289-94
19. Wingren CJ, Bjorck S, Lynch KF, Ohlsson H, Agardh D, Merlo J. Coeliac disease in children: a social epidemiological study in Sweden. *Acta Paediatr*. 2012;101:185-91
20. Zingone F, West J, Crooks CJ, *et al*. Socioeconomic variation in the incidence of childhood coeliac disease in the UK. *Arch Dis Child*. 2015;100:466-73
21. van den Heuvel D, Jansen MA, Dik WA, *et al*. Cytomegalovirus- and Epstein-Barr Virus-Induced T-Cell Expansions in Young Children Do Not Impair Naive T-cell Populations or Vaccination Responses: The Generation R Study. *J Infect Dis*. 2016;213:233-42

22. Kruithof CJ, Kooijman MN, van Duijn CM, *et al.* The Generation R Study: Biobank update 2015. *Eur J Epidemiol.* 2014;29:911-27
23. Jaddoe VW, van Duijn CM, Franco OH, *et al.* The Generation R Study: design and cohort update 2012. *Eur J Epidemiol.* 2012;27:739-56
24. Swertz O, Duimelaar P, Thijssen J. Migrants in the Netherlands 2004. *Statistics Netherlands*, 2016 https://www.cbs.nl/nr/rdonlyres/3c60b3e9-09e0-491f-87f2-99b8e54936a1/0/2004a3pub.pdf.
25. Year report Integration 2012. *Statistics Netherlands*, 2016 https://www.cbs.nl/nr/rdonlyres/cf430ff1-4623-4d60-b34b-cbe33749006a/0/2012b72pub.pdf.
26. https://www.volksgezondheidenzorg.info/onderwerp/bevolking/regionaal-internationaal/etniciteit. 2016; Online.
27. Wijtzes AI, Jansen W, Jaddoe VW, *et al.* Ethnic background and television viewing time among 4-year-old preschool children: the generation R study. *J Dev Behav Pediatr.* 2013;34:63-71
28. Jansen MA, van den Heuvel D, Bouthoorn SH, *et al.* Determinants of Ethnic Differences in Cytomegalovirus, Epstein-Barr Virus, and Herpes Simplex Virus Type 1 Seroprevalence in Childhood. *J Pediatr.* 2015; 170:126-34.e1-6
29. Troe EJ, Raat H, Jaddoe VW, *et al.* Smoking during pregnancy in ethnic populations: the Generation R study. *Nicotine Tob Res.* 2008;10:1373-84
30. Husby S, Koletzko S, Korponay-Szabo IR, *et al.* European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. *J Pediatr Gastroenterol Nutr.* 2012;54:136-60
31. Jansen MA, Kiefte-de Jong JC, Gaillard R, *et al.* Growth Trajectories and Bone Mineral Density in Anti-Tissue Transglutaminase Antibody-positive Children: The Generation R Study. *Clin Gastroenterol Hepatol.* 2014; 13:913-20,e5
32. Jansen MA, Tromp, II, Kiefte-de Jong JC, *et al.* Infant feeding and anti-tissue transglutaminase antibody concentrations in the Generation R Study. *Am J Clin Nutr.* 2014;100:1095-101
33. Statistics Netherlands Welfare in the Netherlands. Income, Welfare and spending of households and individuals. *Den Haag/Heerlen: Statistics Netherlands.* 2012.
34. Statistics Netherlands. Dutch Standard Classification of Education 2003 *Voorburg/Heerlen: Statistics Netherlands* 2004.
35. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol.* 1986;51:1173-82
36. McNamee R. Confounding and confounders. *Occup Environ Med.* 2003;60:227-34; quiz 164, 234
37. Mustalahti K, Catassi C, Reunanen A, *et al.* The prevalence of celiac disease in Europe: results of a centralized, international mass screening project. *Ann Med.* 2010;42:587-95
38. Halenius A, Hengel H. Human cytomegalovirus and autoimmune disease. *Biomed Res Int.* 2014;2014:472978
39. Catassi C, Gatti S, Lionetti E. World perspective and celiac disease epidemiology. *Dig Dis.* 2015;33:141-6
40. Singh P, Arora S, Singh A, Strand TA, Makharia GK. Prevalence of celiac disease in Asia: A systematic review and meta-analysis. *J Gastroenterol Hepatol.* 2016;31:1095-101
41. Rook GA. Hygiene hypothesis and autoimmune diseases. *Clin Rev Allergy Immunol.* 2012;42:5-15
42. Brown EM, Arrieta MC, Finlay BB. A fresh look at the hygiene hypothesis: how intestinal microbial exposure drives immune effector responses in atopic disease. *Semin Immunol.* 2013;25:378-87
43. Round JL, Mazmanian SK. The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol.* 2009;9:313-23
44. Sanz Y, De Pama G, Laparra M. Unraveling the ties between celiac disease and intestinal microbiota. *Int Rev Immunol.* 2011;30:207-18

45. de Meij TG, Budding AE, Grasman ME, Kneepkens CM, Savelkoul PH, Mearin ML. Composition and diversity of the duodenal mucosa-associated microbiome in children with untreated coeliac disease. *Scand J Gastroenterol*. 2013;48:530-6
46. Plot L, Amital H. Infectious associations of Celiac disease. *Autoimmun Rev*. 2009;8:316-9
47. Plot L, Amital H, Barzilai O, Ram M, Bizzaro N, Shoenfeld Y. Infections may have a protective role in the etiopathogenesis of celiac disease. *Ann N Y Acad Sci*. 2009;1173:670-4
48. Jansen MA, van den Heuvel D, van der Zwet KV*, et al*. Herpesvirus Infections and Transglutaminase type 2 Antibody Positivity in Childhood: The Generation R Study. *J Pediatr Gastroenterol Nutr*. 2016;63:423-30



Supplemental Figure 1. Flowchart of participants included in the study

Supplemental Table 1. TG2A levels per ethnic group

	Population for analysis (n=4,442)		
	TG2A neg. (≤7U/ml) n=4382	**TG2A pos. (7-70 U/ml) n=60**	**TG2A ULN (>70U/ml) n=31**
Western ethnicity			
Indonesian	25	0	0
Dutch	2450	42	26
European	326	5	3
American, western	25	0	0
Asia, western	6	0	0
Oceanie	8	0	0
Non-Western ethnicity			
Turkish	307	0	0
Moroccan	247	8	1
Surinamese	314	1	0
Dutch Antilles	141	0	0
Cape Verdian	125	2	0
African	100	0	0
Asia, non-western	110	0	0
American, non-western	83	1	1
Missing	*0*	*1*	*0*
Total	**4382**	**60**	**31**

Chapter 10

Determinants of Ethnic differences in Cytomegalovirus, Epstein-Barr Virus, and Herpes Simplex Virus Type 1 Seroprevalence in Childhood

The Generation R Study

Michelle A.E. Jansen, Diana van den Heuvel, Selma H. Bouthoorn, Vincent W.V. Jaddoe, Herbert Hooijkaas, Hein Raat, Pieter L.A. Fraaij, Menno C. van Zelm, Henriette A. Moll

J Pediatr. 2016 Mar;170:126-34.e1-6.

ABSTRACT

Objective

To identify whether there are ethnic differences in cytomegalovirus (CMV), Epstein-Barr virus (EBV) and herpes simplex virus type 1 (HSV-1) seroprevalence rates in children at 6 years of age, and when present, to evaluate how these differences can be explained by sociodemographic and environmental factors.

Study design

This study was embedded within a multi-ethnic population-based prospective cohort study. Serum IgG levels against CMV, EBV and HSV-1 were measured by enzyme-linked immunosorbent assay in 4464 children (median age 6.0 years). Information on demographics and characteristics were assessed by questionnaires. Herpesvirus seroprevalences between Surinamese-Creole, Surinamese-Hindustani, Turkish, Moroccan, Cape Verdean, Antillean and Native Dutch children were compared.

Results

Non-western ethnicity was an independent risk factor for CMV (aOR, 2.16; 95% CI 1.81-2.57), EBV (1.76; 1.48-2.09) and HSV-1 seropositivity (1.52; 1.39-1.66). Among the ethnic groups, CMV seroprevalences ranged between 29 and 65%, EBV between 43 and 69% and HSV-1 between 13 and 39%. Low family net household income, low maternal educational level, crowding and lifestyle factors explained up to 48% of the ethnic differences in HSV-1 seroprevalences, and up to 39% of the ethnic differences in EBV seroprevalences. These factors did not explain ethnic differences in CMV seroprevalences.

Conclusion

Socioeconomic position and factors related to lifestyle only explain a part of the large ethnic differences in EBV and HSV-1 seroprevalences, whereas they do not explain ethnic differences in CMV seroprevalences in childhood.

INTRODUCTION

Cytomegalovirus (CMV), Epstein-Barr virus (EBV), and Herpes simplex virus 1 (HSV-1) are ubiquitous herpesviruses, and often are acquired in childhood. They viruses usually are transmitted by contact with infected saliva (all 3 viruses), breastfeeding, placental transfer (CMV), or through contact with infectious skin lesions or secretions (HSV-1).[1, 2] The viruses establish a lifelong latency after primary infection, and periodically can reactivate with shedding of the virus. Most primary infections are mild or asymptomatic in childhood, but they can cause serious complications in fetuses, immunocompromised individuals, or in elderly.[3-9] The incidence and severity of herpesvirus associated diseases vary geographically, and many studies link age at primary infection to the pathogenesis of herpesvirus associated diseases.[10-14]

CMV seroprevalences vary between 40 and 60% in children aged 4-12 years, [8, 15, 16] with exclusion of African countries where almost all children seroconvert by 3 years of age. Approximately 50% of children in the US and Europe are seropositive for EBV in the first years of life, but seroprevalences vary from 20-80% depending on age, race/ethnicity, geographic location and socioeconomic development of the country.[9, 12, 17, 18] HSV-1 is widespread during the first few years of life in developing countries, whereas in some European countries primary infection occurs at an older age.[19-22] Besides geographical differences, ethnic differences also have been described in population subgroups.[23] Ethnic and geographical differences in CMV, EBV and HSV-1 seroprevalences might be explained by differences in behavior that facilitate transmission, such as breastfeeding, daycare, family size and socioeconomic position. [17, 22, 24, 25] Because socioeconomic position and ethnicity are strongly related, it remains unclear to what extent socioeconomic factors explain ethnic differences in herpesvirus seroprevalences.[26] In addition, knowledge on other lifestyle related determinants is mainly based on prevalence studies from the late 1990s,[27-30] but social conditions and lifestyle have changed over time.[31, 32] Furthermore, the relative importance of each risk factor needs to be assessed,[26] and many studies lack the assessment of multiple risk factors. Previously, it has been suggested that early life exposures, in particular breastfeeding, are more important for CMV infection than adulthood exposures.[33] On the other hand, the ALADDIN birth cohort study found no strong influence of these early life risk factors.[34] Understanding transmission dynamics of herpesvirus infections within populations, and understanding how socioeconomic and ethnic groups are differentially exposed to these infections, can improve existing preventive policies and interventions, eventually including vaccination programs.[35]

The first aim of our study is to assess whether there are ethnic differences in acquisition of CMV, EBV and HSV-1 in children by 6 years of age. The second aim is to evaluate how possible ethnic differences in acquisition of CMV, EBV and HSV-1 can be explained by sociodemographic and environmental factors.

METHODS

Study design

This study was embedded within the Generation R Study, a prospective population-based cohort study from fetal life onwards.[36, 37] All children were born between 2002 and 2006 in Rotterdam, the second largest city in The Netherlands. The study area covers more than one-half of the city's inhabitants, reflecting Rotterdam's multi-ethnic population.[36] The largest ethnic groups are of Dutch, Surinamese, Turkish, Moroccan, Dutch Antilles and Cape Verdean origin.[36] In total, 8305 children participated in the postnatal phase of the study, of whom 81% (n=6690) visited the research centre (median age 6.0 years).[37] During this visit, blood samples were collected from 4593 children. Serologic data on EBV, CMV, HSV-1 status were provided on 4464 children [Supplemental Figure 1]. The study was approved by the Medical Ethical Committee of Erasmus MC, University Medical Centre Rotterdam. Parents of the children gave written informed consent.

Herpesvirus serology

Venous blood plasma samples were analysed using enzyme immunoassays for IgG antibodies against CMV (purified native antigens strain "AD169"), EBV capsid antigen (EBV-CA, native mixture of several viral capsid antigens), HSV-1 native glycoprotein C1, (all from EUROIMMUN®, Lübeck, Germany).[38] The assays are specific for the particular viruses and discriminate between HSV-1 and HSV-2. Results were evaluated semiquantitatively relative to a manufacturer-provided reference threshold sample. Per manufacturer's instructions, an optical density in the patient sample >10% above the provided threshold sample was defined as positive.

Ethnic background

Ethnicity of the child was determined by country of birth of the parents. The child was of nonnative Dutch origin if one of the parents was born in another country than The Netherlands.[39] If both parents were born abroad, the country of birth of the mother was taken. To study whether ethnic background was related to herpesvirus infections, we constructed a dichotomous variable 'Western/non-western' ethnicity. Next, to study to what extent socioeconomic and lifestyle factors explained associations between ethnicity and herpesvirus infections, and to study the effect of cultural background of the mother (most often primary caregivers),[40] a distinction was made between mothers of Dutch, Turkish, Moroccan, Cape Verdean, Antillean, Surinamese Creole and Surinamese Hindustani ethnic background. Maternal ethnic background was based on country of birth of mother's parents and was assessed by questionnaires at enrollment. Mothers were considered non-native Dutch if one of her parents was born abroad.[41] If both parents were born abroad, country of birth of the mother's mother prevailed. Generational status of non-native Dutch mothers was based on their own country of birth. Foreign-born mothers were classified as 'first generation' and mothers born in the Netherlands were classified as 'second generation'. [40]

Other determinants

Other potential determinants of herpesvirus infections were selected based on existing literature. [2, 12, 24, 32, 33, 42-44] These are: socioeconomic position, family size/parity, maternal lifestyle, breastfeeding, daycare and sex. Data on sex, gestational age, birth weight, parity and mode of delivery were obtained from obstetric records from hospitals and mid-wife practices. Data on sociodemographic and lifestyle factors were obtained by a combination of prenatal and postnatal questionnaires completed by both parents. Prenatal questionnaires included information on maternal lifestyle and maternal educational level. Information on family net household income was determined by questionnaires obtained at 6 years of age. Information on breastfeeding and daycare was obtained by postnatal questionnaires. All questionnaires were available in 3 languages (Dutch, English and Turkish) and further support for verbal translation of questionnaires was available in Arabic, Portuguese and French.

Explanatory variables

Most effects of ethnicity on herpes virus infections are probably indirect ones, acting through more proximal determinants of CMV, EBV, and HSV-1. We considered the following factors to be potential explanatory variables: socioeconomic position, breastfeeding, parity as indicator for siblings/family size, and day-care attendance.[2, 12, 24, 32, 33, 43, 45] Information on household income[46] and educational level,[47] both indicators of socioeconomic position, were obtained using prenatal questionnaires. The Dutch Standard Classification of Education was used to categorize 4 subsequent levels of education: 1=high, 2= mid-high, 3=mid-low, and 4=low. [47] For additional analyses, income was subdivided into three categories: 1. Below modal <2,000 euro per month (<US $2511), 2. Between 2000 and 3200 euro per month (US $2511- $4018), 3. >3200 euro per month (>US $4018).[48, 49] Child's age and sex were treated as potential confounders in the association between ethnicity and herpesvirus seroprevalences.

Statistical analysis

First, independent Student *t* test, and χ^2 tests were used to test whether characteristics were different between children who were seropositive for CMV, EBV and HSV-1, and those who were not. To investigate which determinants were associated with CMV, EBV and HSV-1 seropositivity (yes/no), logistic regression analyses were performed with herpesvirus seropositivity at the age of 6 years as dependent variables (CMV yes/no, EBV yes/no, and HSV-1 yes/no). To study to which extent associations between ethnicity and CMV, EBV, and HSV-1 seropositivity were explained by factors associated with socioeconomic position, crowding, breastfeeding behaviour and maternal lifestyle, we used Baron and Kenny's causal step approach.[50] Only those factors that were unequally distributed across the ethnicity groups (determinant) [Supplemental Table 1], and were significantly ($p<0.05$) associated with CMV, EBV and HSV-1 seropositivity (outcomes) (independent of ethnicity, sex and age) [Supplemental Table 2], were added separately to the model.[50] To assess their mediating effects, the corresponding percentages of attenuation of effect estimates were calculated by comparing differences of

model 1 with the adjusted ones (100 x (OR model 1 – OR model 2 with explanatory factor)/ (OR model 1)). Finally, a full model containing ethnicity and all the explanatory factors assessed the joint effects of explanatory factors. Interaction terms between ethnic background and the explanatory variables were tested for significance. If the test was significant ($p<0.05$), we also stratified the analysis by the variables. To reduce attrition bias, multiple imputation of the covariables was performed (n=5 imputations).[51] Regression coefficients were pooled by taking the average of the coefficients of the 5 imputed datasets. The pooled results of the 5 imputed datasets were reported in this paper as Odds Ratios (ORs) and 95% confidence intervals (CI's). A 95% CI was calculated around the percentage attenuation using a bootstrap method with 1000 resamplings per imputed dataset in the statistical program R.[52] All other statistical analyses were performed in SPSS 20.0 for Windows (SPSS Inc, Chicago, Illinois).

RESULTS

Maternal and child characteristics related to CMV, EBV and HSV-1 are shown in [Table 3]. Within the total group of 4464 children (median age 6.0 years), n=1692 (38%) were infected with CMV, n=2279 (51%) were infected with EBV, and n=877 (20%) were infected with HSV-1. Thirty percent (n=1319) were seronegative for CMV, EBV and HSV-1, 39% (n=1735) were infected with one of these viruses, and 32% (n=1410) were infected with two or more viruses. Of children with available data on herpesvirus serology, a total of 2386 children were Native Dutch, 181 were Cape Verdean, 247 Moroccan, 109 Dutch-Antillean, 131 Suriname-Creole, and 136 children were of Surinamese-Hindustani origin [Supplemental Figure 1].

Determinants of CMV, EBV and HSV–1 seropositivity

CMV seropositive children were more likely of non-Western ethnicity (aOR 2.16; 95% CI 1.81-2.57), to be girls (aOR 1.18, 95% CI 1.02-1.36), and to be breastfed for 6 months or longer (aOR 1.96; 95% CI 1.25-3.07) than CMV seronegative children. Multiparity was inversely associated with CMV seropositivity (aOR 0.76; 95% CI 0.65-0.90). In contrast, EBV and HSV-1 seropositive children more often had 2 or more siblings (aOR $EBV_{parity \geq 2}$ 1.35; 95% CI 1.05-1.74, aOR $HSV\text{-}1_{parity \geq 2}$ 1.31; 95% CI 1.14-1.50), than EBV and HSV-1 seronegative children. Similar to CMV, non-Western ethnicity was an independent risk factor for EBV (aOR 1.76; 95% CI 1.48-2.09) and HSV-1 (aOR 1.52; 95% CI 1.39-1.66). Daycare and socioeconomic position were only associated with HSV-1 seropositivity (aOR $_{daycare}$ 0.89; 95% CI 0.81-0.97). Mothers of HSV-1 seropositive children were lower educated (aOR 1.55; 95% CI 1.35-1.79) and their family household income was lower (aOR 1.11; 95% CI 1.01-1.23), than mothers of HSV-1 seronegative children [Table 4; online]. Additional analyses between the seroprevalence of herpesvirus infections and net family household income in three groups strengthened the inverse associations between socioeconomic position and herpesvirus seroprevalences [Table 5].

Table 3. Maternal and infant characteristics of the study population (n=4,464)

	CMV		EBV		HSV-1	
	CMV - (n=2772)	CMV + (n=1692)	EBV - (n=2185)	EBV + (n=2279)	HSV-1 - (n=3592)	HSV-1 + (n=877)
Maternal characteristics						
Age (Mean ± SD; years)	31.1 (5.0)	30.1 (5.3)***	31.4 (4.9)	30.1 (5.3)***	31.0 (5.0)	29.5 (5.4)***
Educational level (n; %)						
Low	495 (19%)	373 (25%)	333 (16%)	535 (27%)	613 (19%)	255 (35%)
Mid low	794 (31%)	440 (29)	604 (30%)	630 (31%)	996 (30%)	238 (32%)
Mid high	582 (23%)	312 (21%)	504 (25%)	390 (19%)	768 (23%)	126 (17%)
Higher	689 (27%)	368 (25%)	595 (29%)	462 (23%)	936 (28%)	121 (16%)
Net household income per month (n; %)						
< € 2200	802 (28%)	591 (48%)***	594 (35%)	799 (48%)***	1061 (38%)	332 (57%)***
≥ € 2200	1320 (63%)	650 (52%)	1102 (65%)	864 (52%)	1718 (62%)	252 (43%)
Smoking during pregnancy (n; %)	349 (16%)	187 (14%)	235 (14%)	301 (17%)**	419 (15%)	117 (18%)
Alcohol use during pregnancy (n; %)	918 (44%)	519 (39%)**	782 (45%)	695 (39%)***	1262 (44%)	215 (32%)***
BMI before pregnancy (Mean ± SD; kg/m^2)	23.7 (4.3)	23.5 (4.0)	23.3 (4.0)	23.9 (4.3)***	23.4 (1.6)	24.3 (4.9)**
Fever in 3rd trimester of pregnancy (n; %)	28 (7%)	11 (6%)	22 (7%)	17 (7%)	33 (7%)	6 (8%)
Maternal atopy (n; %)	845 (35%)	453 (32%)*	670 (35%)	628 (33%)	1069 (34%)	229 (32%)
Family history of asthma atopy (n; %)	1263 (49%)	692 (46%)*	1018 (50%)	937 (46%)*	1611 (49%)	344 (45%)
Parity						
0	495 (19%)	373 (25%)*	1198 (55%)	1172 (54%)***	1947 (57%)	423 (51%)*
1	794 (31%)	440 (30%)	661 (32%)	677 (31%)	1072 (31%)	266 (32%)
2	582 (23%)	312 (21%)	192 (9%)	257 (12%)	345 (10%)	104 (12%)
≥3	689 (27%)	368 (25%)	36 (2%)	76 (4%)	95 (2%)	51 (6%)
Mode of delivery (n; %)						
Vaginal	1746 (74%)	1066 (72%)	1342 (71%)	1468 (75%)	2240 (72%)	572 (76%)
Forceps or vacuum assisted	305 (13%)	228 (15%)	292 (16%)	240 12%)	449 (15%)	84 (11%)
Caesarian section	315 (13%)	190 (13%)	249 (13%)	255 (13%)	408 (13%)	97 (13%)
Premature rupture of membranes (n; %)	100 (4%)	81 (5%)	96 (5%)	85 (4%)	148 (4%)	33 (4%)
Infant characteristics						
Males (n; %)	1469 (53%)	836 (49%)*	1121 (51%)	1183 (52%)	1855 (52%)	450 (51%)
Western ethnicity (n;%)¶	1991 (73%)	877 (54%)***	1598 (74%)	1266 (58%)***	2426 (70%)	407 (49%)***
Gestational age (Mean ± SD; weeks)	39.8 (1.8)	39.8 (1.8)	39.7 (1.9)	39.8 (1.8)	39.8 (1.8)	39.7 (1.8)*
Preterm birth (<37 weeks) (n; %)	181 (7%)	119 (7%)	159 (7%)	140 (6%)	246 (7%)	54 (6%)
Birth weight (Mean ± SD; grams)	3427 (576)	3391 (555)*	3,427 (584)	3,401 (553)	3,424 (523)	3,369(549)*

Table 3. Maternal and infant characteristics of the study population (n=4,464) (continued)

	CMV		EBV		HSV-1	
	CMV - (n=2772)	CMV + (n=1692)	EBV - (n=2185)	EBV + (n=2279)	HSV-1 - (n=3592)	HSV-1 + (n=877)
Day-care in first year (n; %)	1069 (86%)	646 (92%)***	940 (88%)	773 (88%)	1493 (89%)	222 (84%)*
Breastfeeding (n; %)						
Never	211 (11%)	53 (5%)	140 (9%)	124 (9%)	215 (9%)	49 (10%)
< 3 months	650 (35%)	341 (33%)	518 (34%)	472 (34%)	812 (34%)	179 (36%)
3-6 months	395 (21%)	224 (22%)	305 (20%)	313 (23%)	524 (22%)	95 (19%)
>6 months	626 (33%)	421 (40%)	565 (37%)	479 (35%)	870 (36%)	177 (35%)
Infections first year (n; %)						
RTI	1054 (58%)	631 (63%)**	877 (58%)	808 (61%)	1404 (60%)	281 (61%)
LRTI	241 (14%)	131 (14%)	190 (13%)	182 (15%)	314 (14%)	58 (14%)
URTI	975 (51%)	593 (56%)**	810 (52%)	758 (54%)	1296 (52%)	272 (55%)
GI	623 (56%)	380 (64%)**	530 (57%)	473 (61%)	833 (58%)	170 (61%)
Atopy first year (n; %)¶¶						
Eczema	486 (26%)	305 (30%)*	440 (29%)	351 (26%)	678 (28%)	113 (24%)
Wheezing (≥ 1 episode)	680 (45%)	385 (48%)	545 (44%)	520 (48%)*	884 (45%)	181 (48%)
Asthma diagnosis at age 6 (n; %)	168 (8%)	83 (7%)	133 (7%)	118 (7%)	209 (7%)	42 (7%)
Age (Median ± range; years)	6.2 (5.0-9.0)	6.2 (4.9-9.1)***	6.1 (5.0-9.0)	6.2 (4.9-9.1)***	6.2 (4.9-9.0)	6.2(5.0-9.1)***
BMI for age (mean; SD; SDS)	0.25 (0.90)	0.32 (0.92)**	0.21 (0.87)	0.34 (0.94)***	0.25 (0.89)	0.41 (0.97)***

BMI, Body Mass Index; RTI, respiratory tract infections; LRTI lower respiratory tract infections; URTI upper respiratory tract infections; GI gastrointestinal infections.

Values are means (SD), absolute numbers (percentages) or #medians (range). Significantly different from not infected:

*P<0.05.

**P <0.01.

*** P <0.001.

Table is based on observed dataset.

¶ Western ethnicity includes children from Dutch, European, American Western, and Asian Western origin. Non-Western ethnicity included children from Turkish,

Moroccan, Cape Verdean, Antillean, Surinamese-Creole and Surinamese-Hindustani origin.

¶¶ Atopy includes eczema, house dust mite allergy and hay-fever. Any reported autoimmune disease includes diabetes mellitus, systemic lupus erythematosus, arthritis, multiple sclerosis, thyroid disorder or celiac disease.

Ethnic background and CMV seropositivity

Children of Cape Verdean, Moroccan, Turkish, Surinamese-Creole and Surinamese Hindustani mothers were more often seropositive for CMV, than children of native Dutch mothers. CMV seroprevalences in the non-native Dutch groups ranged between 44 and 65%, compared with 29% in the native Dutch group [Figure 2; online].

To explain the observed ethnic differences in CMV seroprevalences, the following planatory variables were selected: maternal educational level, parity, daycare and breastfeed-

Table 5. Odds Ratios for Infections and Coinfections According to Socioeconomic Position

	CMV OR (95% CI)	*P* Value	EBV OR (95% CI)	*P* Value	HSV-1 OR (95% CI)	*P* Value	CMV and HSV-1 OR (95% CI)	*P* Value	EBV and HSV-1 OR (95% CI)	*P* Value	EBV and CMV Odds Ratio (95% CI)	*P* Value
Income (€/month)												
<2000	**1.70 (1.45; 2.01)**	<0.0001	**1.92 (1.63; 2.27)**	<0.0001	**2.56 (2.09; 3.12)**	<0.0001	**3.43 (2.60; 4.54)**	<0.0001	**3.10 (2.44; 3.93)**	<0.0001	**2.19 (1.81; 2.64)**	<0.0001
2000-3200	0.93 (0.79; 1.10)	0.41	**1.32 (1.13; 1.55)**	0.001	**1.52 (1.23; 1.88)**	<0.0001	**1.66 (1.21; 2.28)**	0.002	**1.58 (1.21; 2.06)**	0.001	1.14 (0.93; 1.40)	0.22
>3200	1.00 (ref)		1.00 (ref)		1.00 (ref)		1.00 (ref)		1.00 (ref)		1.00 (ref)	
Education												
Low	**1.36 (1.08; 1.72)**	0.011	**1.92 (1.59; 2.33)**	<0.0001	**2.63 (2.11; 3.27)**	<0.0001	**3.21 (2.37; 4.35)**	<0.0001	**2.93 (2.21; 3.87)**	<0.0001	**2.01 (1.50; 2.68)**	<0.0001
Mid Low	0.97 (0.80; 1.19)	0.79	**1.23 (1.04; 1.46)**	0.02	**1.47 (1.17; 1.86)**	0.001	**1.40 (1.02; 1.94)**	0.04	**1.62 (1.22; 2.15)**	0.001	1.21 (0.95; 1.54)	0.13
Mid High	0.95 (0.79; 1.14)	0.58	0.94 (0.77; 1.14)	0.52	1.07 (0.77; 1.94)	0.67	1.02 (0.69; 1.51)	0.93	1.05 (0.71; 1.57)	0.79	1.12 (0.88; 1.42)	0.37
High	1.00 (ref)		1.00 (ref)		1.00 (ref)		1.00 (ref)		1.00 (ref)		1.00 (ref)	

Abbreviations: CI, confidence interval; CMV, cytomegalovirus, EBV, Epstein-Barr virus; HSV-1, herpes simplex virus type 1; OR Odds Ratio.
Data on family net household income: <2000 €/month, n=870; between 2000-3200 €/month, n=915; and >3200 €/month, n=1827.
Data on maternal educational level: low n=1064; mid low n=1258; mid high n=940; and high n=1202
Bold values represent significant differences.

Table 7. Attenuation of the OR of Epstein-Barr virus seropositivity for different ethnicities after individual adjustment for the explanatory variables

	Surinamese-Creoles[a] OR (95% CI)	Attenuation a (95% CI)	Surinamese-Hindustani [b] OR (95% CI)	Attenuation b (95% CI)	Turkish [c] OR (95% CI)	Attenuation c (95% CI)
EBV						
Model 1 (includes child sex and current age)	1.77 (1.24, 2.54)		2.85 (1.96, 4.13)		2.69 (2.11, 3.44)	
Socioeconomic status						
Model 1+ education	1.66 (1.15, 2.40)	-15% (-43 to 0.2)	2.71 (1.85, 3.95)	-7% (-17 to 3)	2.37 (1.84, 3.08)	**-19% (-31 to -7)****
Model 1+ income	1.63 (1.13, 2.36)	**-19% (-50 to -5)***	2.60 (1.80, 3.86)	**-13% (-24 to -4)***	2.36 (1.86, 3.16)	**-19% (-33 to -5)***
Crowding						
Model 1 + parity	1.80 (1.25, 2.59)	+4% (-9 to 24)	2.86 (1.96, 4.19)	1% (-10 to 16)	2.58 (2.02, 3.31)	-6% (-13 to 0.7)
Fully adjusted model	1.61 (1.12, 2.37)	-21% (-56 to 3)	2.59 (1.77, 3.87)	-14% (-30 to 5)	2.14 (1.67, 2.88)	**-33% (-48 to -17)*****

Abbreviations: CI, confidence interval; OR, Odds Ratio. Bold values represent significant differences.

[a] The ORs represent the attenuated effect estimate and their 95% CIs relative to model 1 after individual adjustment for explanatory factors.

[b] Attenuations a to f represent the attenuations of effect estimates for the different ethnicities relative to model 1 (includes confounders) after adjustment for explanatory variables (100x ($OR_{model\ 1}$ - $OR_{model\ 1\ with\ explanatory\ factor}$) / ($OR_{model\ 1}$)). Dutch is the reference group.

ing duration [Supplemental Table 1 and 2].[50] However, after including all the explanatory factors in the model, differences in CMV seroprevalences compared with Dutch children increased with 55% (95% CI 24-124) for Surinamese-Hindustani children, 36% (95% CI 19-60) for Moroccan children, 34% (95% CI 14-66) for Surinamese-Creole children, 33% (95% CI 13-59) for Turkish children, and 23% (95% CI 5-47) for Cape Verdean children [Supplemental Table 6].

Ethnic background and EBV seropositivity

Children in non-native Dutch groups were more often EBV seropositive than native Dutch children. EBV seroprevalences in the non-native Dutch minority groups ranged between 59 and 69%, compared with 43% in the native Dutch group [Supplemental Figure 2]. Maternal educational level, net household income and parity were selected to explain ethnic differences in EBV seroprevalences [Supplemental Table 1 and 2; online].[50] Socioeconomic position was the most important contributor to the association between ethnicity and EBV seroprevalence. For example, net household income explained 22% (95% CI -41 to -6) of the differences in EBV seroprevalences between Cape Verdean and native Dutch children, followed by maternal educational level, which explained 18% (95% CI -31 to -7) [Table

Moroccan [d] OR (95% CI)	Attenuation d (95% CI)	Cape Verdean [e] OR (95% CI)	Attenuation e (95% CI)	Antillean [f] OR (95% CI)	Attenuation f (95% CI)
2.40 (1.82, 3.17)		2.18 (1.59, 2.99)		1.98 (1.33, 2.96)	
2.19 (1.63, 2.91)	**-15% (-26 to -6)****	1.97 (1.41, 2.70)	**-18% (-31 to -7)****	1.86 (1.24, 2.80)	**-12% (-32 to -0.6)***
2.11 (1.60, 2.91)	**-21% (-36 to -6)***	1.92 (1.41, 2.74)	**-22% (-41 to -6)***	1.77 (1.19, 2.73)	**-22% (-50 to -6)***
2.21 (1.66, 2.94)	**-14% (-25 to -4)***	2.03 (1.47, 2.81)	-12% (-28 tot 1)	1.97 (1.32, 2.96)	-1% (-12 to 14)
1.89 (1.42, 2.61)	**-36% (-55 to -19)*****	1.72 (1.23, 2.44)	**-39% (-66 to -19)*****	1.74 (1.17, 2.69)	**-25% (-60 to -3)***

* *P*<0.05.
** *P*<0.01.
*** *P*<0.001.

7]. Overall, 39% (95% CI -66 to -19) of the difference in EBV seroprevalences prevalence between Cape Verdean and Dutch children was explained by differences in socioeconomic position and family size. These factors explained 36% (95% CI -55 to -19) of the differences in EBV seroprevalences between Moroccan and Dutch children, 33% (95% CI -48 to -17) of the differences in EBV seroprevalences between Turkish and Dutch children, and 25% (95% CI -60 to -3) of the differences in EBV seroprevalences between Antillean and Dutch children [Table 7].

Ethnic background and HSV–1 seropositivity

All ethnic minority groups, except for Surinamese-Creole, were more often HSV-1 seropositive than native Dutch children. HSV-1 seroprevalences in the non-native Dutch groups ranged between 22 and 39%, compared with 13% in the native Dutch group [Supplemental Figure 2]. To explain ethnic differences in HSV-1 seroprevalences, maternal educational level, net household income and breastfeeding duration were selected as explanatory variables [Supplemental Table 1 and 2].[50] Maternal educational level and net household income were the most important factors to explain ethnic differences in HSV-1 seroprevalence. For example, 27% (95% CI -38 to -16) of this differences in HSV-1 seroprevalences

Table 8. Attenuation of the OR of Herpes simplex virus 1 seroprevalence for different ethnicities after individual adjustment for the explanatory variables

	Surinamese-Creoles[a] OR (95% CI)	Attenuation a (95% CI)	Surinamese-Hindustani [b] OR (95% CI)	Attenuation b (95% CI)	Turkish [c] OR (95% CI)
CMV					
Model 1 (includes child sex and current age)	ns	NA	2.10 (1.40, 3.16)		4.08 (3.18, 5.24)
Socioeconomic status					
Model 1+ education			1.82 (1.19, 2.73)	**-26% (-61 to -10)****	3.24 (2.49, 4.26)
Model 1+ income			1.92 (1.24, 2.87)	**-17% (-41 to -0.1)***	3.58 (2.66, 4.68)
Crowding					
Model 1 + breastfeeding			2.04 (1.40, 3.18)	**-6% (-14 to -1)***	4.29 (3.29, 5.50)
Fully adjusted model			1.78 (1.17, 2.75)	**-30% (-68 to -10)****	3.33 (2.43, 4.43)

Abbreviations: CI, confidence interval; ns, nonsignificant; NA, not applicable; OR, Odds Ratio. Bold values represent significant differences.

[a] The ORs represent the attenuated effect estimate and their 95% CIs relative to model 1 after individual adjustment for explanatory factors.

[b] Attenuations b to f represent the attenuations of effect estimates for the different ethnicities relative to model 1 (includes confounders) after adjustment for explanatory variables ($100x (OR_{model\ 1} - OR_{model\ 1\ with\ explanatory\ factor}) / (OR_{model\ 1})$). Dutch is the reference group.

between Turkish and Dutch children was explained by maternal educational level [Table 8]. Socioeconomic position and breastfeeding together explained 24% (95% CI -40 to -7) of the difference in HSV-1 seroprevalence between Turkish and Dutch children; they explained 48% (95% CI -150 to -18) of the differences in HSV-1 seroprevalences between Cape Verdean and Dutch children; 30% (95% CI -68 to -10) of the differences in HSV-1 prevalence between Surinamese-Hindustani and Dutch children and 22% (95% CI -36 to -5) of the differences in HSV-1 seroprevalence between Moroccan and Dutch children. Maternal educational level was the most important explanatory factor in the association between ethnicity and HSV-1 seroprevalence [Table 8].

Ethnic background and multiple herpesvirus infections

We observed that children who were seropositive for a single virus, were more likely to be infected with multiple herpesviruses. Specifically, EBV-infected children were 1.69 (95% CI 1.50-1.91) times more likely to be seropositive for CMV and 2.17 (95% CI 1.86-2.53) times more likely to be seropositive for HSV-1 than children with EBV seronegativity. Furthermore, the odds ratio for both CMV and HSV-1 seropositivity was 2.95 (95% CI 2.02-3.11) times greater in children with seropositivity than in children with EBV seronegativity (data not

Attenuation c (95% CI)	Moroccan [d] OR (95% CI)	Attenuation d (95% CI)	Cape Verdean [e] OR (95% CI)	Attenuation e (95% CI)	Antillean [f] OR (95% CI)	Attenuation f (95% CI)
	3.73 (2.80, 4.98)		1.75 (1.20, 2.54)		1.70 (1.06, 2.73)	
-27% (-38 to -16)* **	3.15 (2.28, 4.17)	**-21% (-30 to -12)*** **	1.45 (0.95, 2.07)	**-39% (-119 to -19)** **	1.47 (0.91, 2.36)	**-33% (-160 to -7)***
-16% (-30 to -1)*	3.29 (2.35, 4.42)	**-16% (-30 to -0.5)***	1.54 (1.01, 2.24)	-28% (-92 to 1)	1.52 (0.91, 2.43)	-26% (-115 to 11)
+7% (2 to 14) **	3.80 (2.89, 5.16)	+2% (0 to 6)	1.67 (1.16, 2.47)	**-10% (-32 to -2)***	1.65 (1.02, 2.65)	-8% (-43 to 2)
-24% (-40 to -7) **	3.14 (2.21, 4.26)	**-22% (-36 to -5)***	1.39 (0.90, 2.02)	**-48% (-150 to -18)***	1.42 (0.85, 2.30)	-40% (-213 to 9)

No significant difference in HSV-1 prevalence was observed for the Surinamese-Creole group compared with the Dutch group, therefore attenuations in effect estimates for the Surinamese-Creole group are not presented.
*P<0.05.
**P <0.01.
*** P <0.001.

shown). Finally, CMV, EBV and HSV-1 seropositive children tended to be more frequently born to 'first' generation mothers than seronegative children, although differences were not significant [Supplemental Figure 3]. Mothers of seronegative children were more often higher educated, and had a higher net household income, than mothers of children who were seropositive for 1 or 2 herpesviruses. Seronegative children were more often of Western origin and had fewer siblings than children who had 1 or 2 herpesvirus infections [Supplemental Table 9].

DISCUSSION

In this large and multi-ethnic population-based cohort, we found large ethnic differences in seroprevalences of EBV (range 43-69%), CMV (range 29-65%) and HSV-1 (range 13-39%) among 6 year old children, and the highest seroprevalence in non-Dutch children. Socioeconomic factors and factors related to crowding early in life partly explained ethnic differences in EBV and HSV-1 seroprevalence in childhood, but did not explain ethnic differences in CMV seroprevalence.

Our study design provided a unique opportunity to study determinants of these infections, the variation in seroprevalence by ethnic group, and to which extent socioeconomic and lifestyle factors explained these differences. Our results confirm and extend previous studies by observing that socioeconomic position partly explains ethnic variation in EBV and HSV-1 seroprevalences, but not in CMV seroprevalences.[12, 35, 45] Although these 3 viruses are all members of the herpesvirus family, and thus share significant biological properties, transmission dynamics between these viruses seem to be different. Our results suggest that ethnic differences in CMV seroprevalences in childhood are mainly determined by factors that facilitate vertical transmission, which underscores that breastfeeding is the most important source for CMV infection in childhood.[53-55] In contrast, ethnic differences in EBV and HSV-1 seroprevalences are partly determined by factors that facilitate horizontal transmission, such as family size and daycare, which are strongly related to socioeconomic position. In fact, we were able to explain up to 39% of the ethnic differences in EBV seroprevalences, of which socioeconomic position was the most important contributor, which is in line with previous studies.[42, 44] However, it has been suggested that ethnic differences could not be explained by crowding. Still, it could be speculated that socioeconomic position reflects more proximal determinants (e.g. factors that facilitate horizontal transmission, such as family size, bed sharing and childcare arrangements). Indeed, the results from our multivariable model confirm that parity is important in estimating the risk of EBV infection.[18] In fact, we observed a dose-response relationship between parity and the risk of EBV and HSV-1 seropositivity.

In contrast with the results on EBV and HSV-1, ethnic differences in CMV seroprevalences could not be explained by socioeconomic position, nor by other 'explanatory variables' such as breastfeeding, daycare and parity. In fact, these factors strengthened rather than weakened the association between ethnicity and CMV. We expected breastfeeding to partly mediate this association, because breastfeeding can transmit CMV, and breastfeeding behavior is influenced by ethnic background.[56] Vertical transmission from mother to child is probably more important in explaining ethnic differences in CMV seroprevalences than factors that facilitate horizontal transmission. Given the economic burden of CMV associated diseases, it has been suggested that vaccination in the general population would be the most straightforward way to control for all CMV associated diseases.[3] Some opinions favor inclusion of 12-year old children, whereas others suggest vaccination of toddlers.[33,57] However, thus far, safe and effective vaccines have not been developed. Nevertheless, on the basis of our results, ethnic and socioeconomic differences should be taken into account when vaccination is contemplated.

HSV-1 prevalence was highest in Turkish and Moroccan children according to a previous Dutch study in adolescents.[45] Our results confirm that differences in HSV-1 seroprevalences between European countries could be partly explained by socioeconomic factors.[22] It has been shown that lower socioeconomic position indeed increases the risk for HSV-1.[45] All factors together, including socioeconomic position, breastfeeding and crowding, explained

up to 48% of ethnic differences in HSV-1 seroprevalence, of which socioeconomic position was the most important contributor, suggesting an important role for factors that facilitate horizontal transmission.

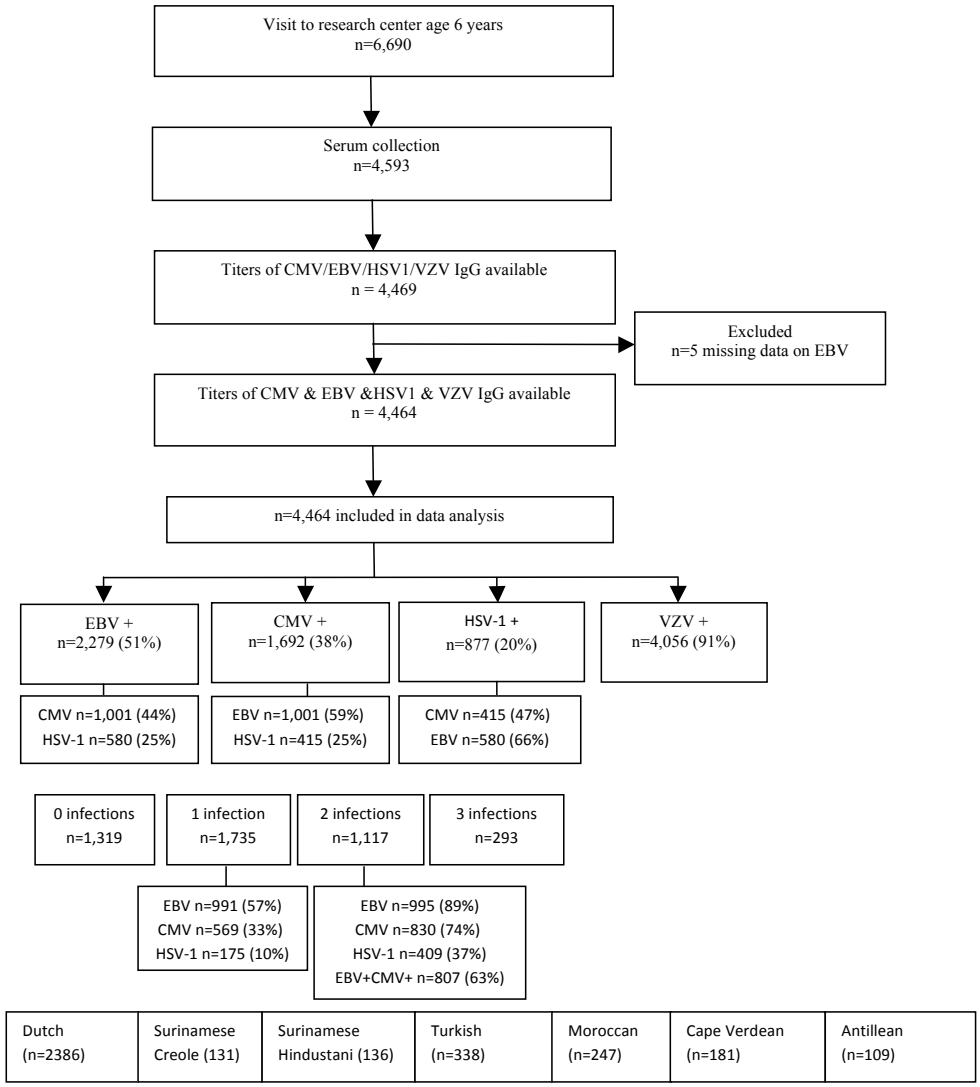
An important strength of this study is that we had a large multi-ethnic cohort study population drawn from the general population of Rotterdam. Because immigration is common in Western countries, our findings may be more broadly applicable. In addition, our study design enabled us to assess multiple risk factors, and to study the relative importance of each using multivariate models. Moreover, our use of a mediator approach to explain ethnic differences in herpesvirus infections is novel. Besides socioeconomic position, numerous other factors vary between countries, and adjustment for these factors might decrease ethnic/racial or geographical variation.[12] Because our study was performed within a population-based study of a region within a single country, we were able to adjust for all of these variables. A limitation of this study is that we did not have data on maternal CMV, EBV, HSV-1 seroprevalences in order to analyse mother-to-child transmission. A second limitation was lack of data on EBV, CMV and HSV-1 seroprevalences in fathers and siblings, precluding examination of their roles in horizontal transmission. When immunization against these viruses is considered, ethnic and socioeconomic differences should be taken into account.

REFERENCES

1. Crough T, Khanna R. Immunobiology of human cytomegalovirus: from bench to bedside. Clin Microbiol Rev. 2009;22:76-98, Table of Contents.
2. Boppana SB, Fowler KB. Persistence in the population: epidemiology and transmisson. 2007.
3. Griffiths PD. Burden of disease associated with human cytomegalovirus and prospects for elimination by universal immunisation. Lancet Infect Dis. 2012;12:790-8.
4. Griffiths P, Baraniak I, Reeves M. The pathogenesis of human cytomegalovirus. J Pathol. 2014.
5. Simanek AM, Dowd JB, Pawelec G, Melzer D, Dutta A, Aiello AE. Seropositivity to cytomegalovirus, inflammation, all-cause and cardiovascular disease-related mortality in the United States. PLoS One. 2011;6:e16103.
6. Dreyfus DH. Herpesviruses and the microbiome. J Allergy Clin Immunol. 2013;132:1278-86.
7. Longnecker R, Neipel F. Introduction to the human gamma-herpesviruses. 2007.
8. Sidorchuk A, Wickman M, Pershagen G, Lagarde F, Linde A. Cytomegalovirus infection and development of allergic diseases in early childhood: interaction with EBV infection? J Allergy Clin Immunol. 2004;114:1434-40.
9. Hjalgrim H, Friborg J, Melbye M. The epidemiology of EBV and its association with malignant disease. 2007.
10. Ogwang MD, Bhatia K, Biggar RJ, Mbulaiteye SM. Incidence and geographic distribution of endemic Burkitt lymphoma in northern Uganda revisited. Int J Cancer. 2008;123:2658-63.
11. Stefan DC, Lutchman R. Burkitt lymphoma: epidemiological features and survival in a South African centre. Infect Agent Cancer. 2014;9:19.
12. Flavell KJ, Biddulph JP, Powell JE, Parkes SE, Redfern D, Weinreb M, et al. South Asian ethnicity and material deprivation increase the risk of Epstein-Barr virus infection in childhood Hodgkin's disease. Br J Cancer. 2001;85:350-6.
13. Bagni R, Whitby D. Age of infection and risk of virally associated cancers: new clues to an old puzzle. J Infect Dis. 2012;205:873-4.
14. Hsu JL, Glaser SL. Epstein-barr virus-associated malignancies: epidemiologic patterns and etiologic implications. Crit Rev Oncol Hematol. 2000;34:27-53.
15. Grillner L, Strangert K. A prospective molecular epidemiological study of cytomegalovirus infections in two day care centers in Sweden: no evidence for horizontal transmission within the centers. J Infect Dis. 1988;157:1080-3.
16. Aarnisalo J, Ilonen J, Vainionpaa R, Volanen I, Kaitosaari T, Simell O. Development of antibodies against cytomegalovirus, varicella-zoster virus and herpes simplex virus in Finland during the first eight years of life: a prospective study. Scand J Infect Dis. 2003;35:750-3.
17. Balfour HH, Jr., Sifakis F, Sliman JA, Knight JA, Schmeling DO, Thomas W. Age-specific prevalence of Epstein-Barr virus infection among individuals aged 6-19 years in the United States and factors affecting its acquisition. J Infect Dis. 2013;208:1286-93.
18. Dowd JB, Palermo T, Brite J, McDade TW, Aiello A. Seroprevalence of Epstein-Barr virus infection in U.S. children ages 6-19, 2003-2010. PLoS One. 2013;8:e64921.
19. Mertz GJ, Rosenthal SL, Stanberry LR. Is herpes simplex virus type 1 (HSV-1) now more common than HSV-2 in first episodes of genital herpes? Sex Transm Dis. 2003;30:801-2.
20. Smith JS, Robinson NJ. Age-specific prevalence of infection with herpes simplex virus types 2 and 1: a global review. J Infect Dis. 2002;186 Suppl 1:S3-28.
21. Whitley RJ, Kimberlin DW, Roizman B. Herpes simplex viruses. Clin Infect Dis. 1998;26:541-53; quiz 54-5.
22. Pebody RG, Andrews N, Brown D, Gopal R, De Melker H, Francois G, et al. The seroepidemiology of herpes simplex virus type 1 and 2 in Europe. Sex Transm Infect. 2004;80:185-91.

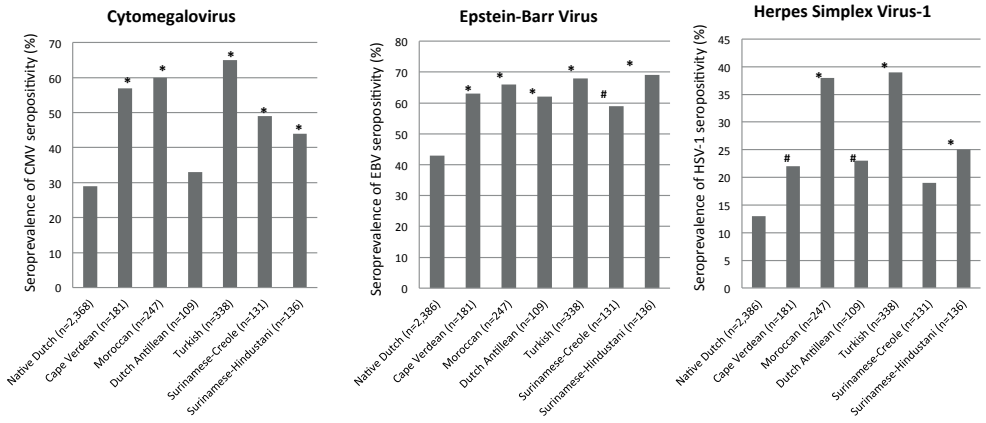
23. Wald A, Corey L. Persistence in the population: epidemiology, transmission. 2007.
24. Joseph SA, Beliveau C, Muecke CJ, Rahme E, Soto JC, Flowerdew G, et al. Cytomegalovirus as an occupational risk in daycare educators. Paediatr Child Health. 2006;11:401-7.
25. Staras SA, Flanders WD, Dollard SC, Pass RF, McGowan JE, Jr., Cannon MJ. Cytomegalovirus seroprevalence and childhood sources of infection: A population-based study among pre-adolescents in the United States. J Clin Virol. 2008;43:266-71.
26. Cannon MJ, Schmid DS, Hyde TB. Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. Rev Med Virol. 2010;20:202-13.
27. Schuster V, Kreth HW. Epstein-Barr virus infection and associated diseases in children. I. Pathogenesis, epidemiology and clinical aspects. Eur J Pediatr. 1992;151:718-25.
28. Kangro HO, Osman HK, Lau YL, Heath RB, Yeung CY, Ng MH. Seroprevalence of antibodies to human herpesviruses in England and Hong Kong. J Med Virol. 1994;43:91-6.
29. Leogrande G, Jirillo E. Studies on the epidemiology of child infections in the Bari area (south Italy). VII. Epidemiology of Epstein-Barr virus infections. Eur J Epidemiol. 1993;9:368-72.
30. Wang PS, Evans AS. Prevalence of antibodies to Epstein-Barr virus and cytomegalovirus in sera from a group of children in the People's Republic of China. J Infect Dis. 1986;153:150-2.
31. Svahn A, Berggren J, Parke A, Storsaeter J, Thorstensson R, Linde A. Changes in seroprevalence to four herpesviruses over 30 years in Swedish children aged 9-12 years. J Clin Virol. 2006;37:118-23.
32. Stagno S, Cloud GA. Working parents: the impact of day care and breast-feeding on cytomegalovirus infections in offspring. Proc Natl Acad Sci U S A. 1994;91:2384-9.
33. Pembrey L, Raynor P, Griffiths P, Chaytor S, Wright J, Hall AJ. Seroprevalence of cytomegalovirus, Epstein Barr virus and varicella zoster virus among pregnant women in Bradford: a cohort study. PLoS One. 2013;8:e81881.
34. Hesla HM, Gutzeit C, Stenius F, Scheynius A, Dahl H, Linde A, et al. Herpesvirus infections and allergic sensitization in children of families with anthroposophic and non-anthroposophic lifestyle - the ALADDIN birth cohort. Pediatr Allergy Immunol. 2013;24:61-5.
35. Dowd JB, Aiello AE, Alley DE. Socioeconomic disparities in the seroprevalence of cytomegalovirus infection in the US population: NHANES III. Epidemiol Infect. 2009;137:58-65.
36. Jaddoe VW, van Duijn CM, Franco OH, van der Heijden AJ, van Iizendoorn MH, de Jongste JC, et al. The Generation R Study: design and cohort update 2012. Eur J Epidemiol. 2012;27:739-56.
37. Kruithof CJ, Kooijman MN, van Duijn CM, Franco OH, de Jongste JC, Klaver CC, et al. The Generation R Study: Biobank update 2015. Eur J Epidemiol. 2014;29:911-27.
38. van den Heuvel D, Jansen MA, Dik WA, Bouallouch-Charif H, Zhao D, van Kester KA, et al. Cytomegalovirus- and Epstein-Barr Virus-Induced T-Cell Expansions in Young Children Do Not Impair Naive T-cell Populations or Vaccination Responses: The Generation R Study. J Infect Dis. 2015.
39. Swertz O, Duimelaar P, Thijssen J. Migrants in the Netherlands 2004. Voorburg/Heerlen: Statistics Netherlands, 2004.
40. Wijtzes AI, Jansen W, Jaddoe VW, Moll HA, Tiemeier H, Verhulst FC, et al. Ethnic background and television viewing time among 4-year-old preschool children: the generation R study. J Dev Behav Pediatr. 2013;34:63-71.
41. Statistics Netherlands. Jaarrapport Integratie 2010; Den Haag/Heerlen, the Netherlands: Statistics Netherlands. 2010.
42. Evans AS. Viral Infections of Humans: Epidemiology and Control1976. p. 209-33.
43. Tookey PA, Ades AE, Peckham CS. Cytomegalovirus prevalence in pregnant women: the influence of parity. Arch Dis Child. 1992;67:779-83.

44. Condon LM, Cederberg LE, Rabinovitch MD, Liebo RV, Go JC, Delaney AS, et al. Age-specific prevalence of Epstein-Barr virus infection among Minnesota children: effects of race/ethnicity and family environment. Clin Infect Dis. 2014;59:501-8.
45. Kramer MA, Uitenbroek DG, Ujcic-Voortman JK, Pfrommer C, Spaargaren J, Coutinho RA, et al. Ethnic differences in HSV1 and HSV2 seroprevalence in Amsterdam, the Netherlands. Euro Surveill. 2008;13.
46. Statistics Netherlands. Welfare in the Netherlands. Income, welfare and spending of households and individuals. Den Haag/Heerlen: Statistics Netherlands, 2012.
47. Statistics Netherlands. Dutch Standard Classification of Education 2003. Voorburg/Heerlen: Statistics Netherlands, 2004.
48. Wijtzes AI, Jansen W, Bouthoorn SH, Pot N, Hofman A, Jaddoe VW, et al. Social inequalities in young children's sports participation and outdoor play. Int J Behav Nutr Phys Act. 2014;11:155.
49. Netherlands Bureau for Economic Policy Analysis. http://cpb.nl. Accessed November 2014.
50. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol. 1986;51:1173-82.
51. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.
52. Venables WN SD. An Introduction to R. Notes on R: A Programming Environment for Data Analysis and Graphics Version 2.12.1 (2012-12-16). http://www.r-project.org. Accessed November 2014.
53. Forsgren M. Cytomegalovirus in breast milk: reassessment of pasteurization and freeze-thawing. Pediatr Res. 2004;56:526-8.
54. Hayes K, Danks DM, Gibas H, Jack I. Cytomegalovirus in human milk. N Engl J Med. 1972;287:177-8.
55. Ahlfors K, Ivarsson SA, Harris S. Report on a long-term study of maternal and congenital cytomegalovirus infection in Sweden. Review of prospective studies available in the literature. Scand J Infect Dis. 1999;31:443-57.
56. Kelly YJ, Watt RG, Nazroo JY. Racial/ethnic differences in breastfeeding initiation and continuation in the United kingdom and comparison with findings in the United States. Pediatrics. 2006;118:e1428-35.
57. Plotkin SA. Vaccination against cytomegalovirus, the changeling demon. Pediatr Infect Dis J. 1999;18:313-25; quiz 26.

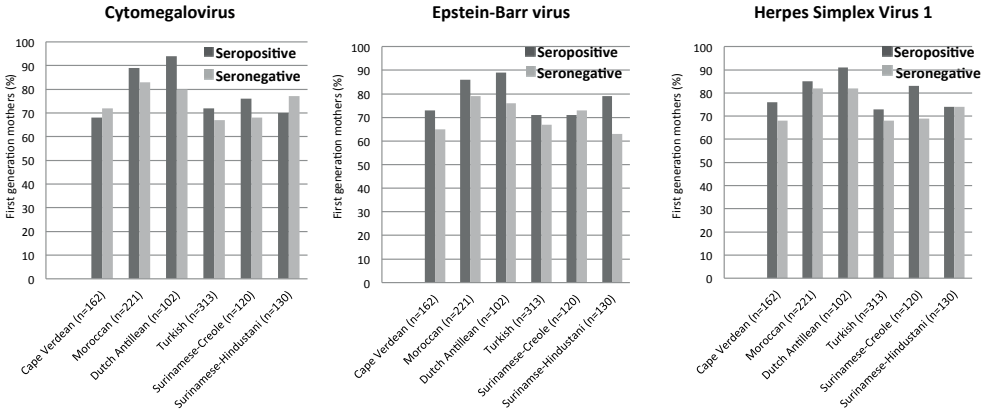


Supplemental Figure 1 Flowchart of the participants within the Generation R Study

A total of 4,464 children provided serological data. Of these, 2,279 children (51%) were EBV positive, 1,692 (38%) CMV positive, 877 (20%) HSV-1 positive, and 4,056 (91%) VZV positive.



Supplemental Figure 2 Seroprevalence of herpesvirus infections in childhood according to maternal ethnic background. Values are seroprevalences of CMV, EBV, and HSV-1 infections (%) per ethnicity category. CMV, EBV, and HSV-1 seroprevalences of Cape Verdean, Moroccan, Dutch Antillean, Turkish, Surinamese-Creole, and Surinamese-Hindustani children are compared to Dutch children. Footnotes: ¶ p-value <0.05, #P-value <0.01, *P-value <0.001 based on χ^2 test.



Supplemental Figure 3 Percentages of first generation mothers according to CMV, EBV or HSV-1 seropositivity in childhood.

Values are percentages of 'first' generation mothers according to CMV, EBV or HSV-1 seroprevalences in childhood. Generational status of non-native Dutch mothers was based on their country of birth. Foreign-born mothers were classified as 'first generation' and mothers born in the Netherlands were classified as 'second generation'.

Supplemental Table 1. Associations between ethnic background and potential confounders and mediators (n=3,528)

Ethnic background	Dutch	Surinamese-creoles	Surinamese-hindustani	Turkish	Moroccan	Cape Verdean	Antillean	*P*-value[a]
	(n=2386)	(n=131)	(n=136)	(n=338)	(n=247)	(n=181)	(n=109)	
Potential confounders[b]								
Age (years)	6.1 (0.4)	6.4 (0.7)	6.3 (0.6)	6.3 (0.6)	6.4 (0.7)	6.3 (0.7)	6.6 (0.9)	<0.001
Parity (%)								
Nulliparous	60.3	54.3	53.8	43.7	37.9	47.7	51.9	<0.001
Potential mediators								
Maternal educational level (%)								
Low	12.7	32.8	31.5	53.3	45.1	50.7	36.0	<0.001
Mid-low	25.7	50.4	49.6	33.0	37.9	39.5	49.0	
Mid-high	26.7	14.3	15.0	10.0	14.6	8.6	9.0	
High	34.9	2.5	3.9	3.7	2.4	1.3	6.0	
Net household income (%)								
≤ 2200 euro	25.1	70.5	70.2	87.1	88.2	91.6	83.3	<0.001
> 2200 euro	74.9	29.5	29.8	12.9	11.8	8.4	16.7	
Daycare (% yes)	89.5	83.3	70.4	57.1	65.4	92.1	82.6	<0.001
Breastfeeding duration (%)								
never	10.1	8.3	4.2	1.2	3.7	11.7	13.0	<0.001
<3 months	31.8	41.7	56.9	35.5	41.1	51.9	52.2	
3-6 months	22.7	18.1	20.8	18.7	18.7	18.2	8.7	
>6 months	35.4	31.9	18.1	44.6	36.4	18.2	26.1	
Maternal smoking (% continued)	14.0	27.2	15.2	32.3	4.9	17.8	20.4	<0.001
Maternal alcohol use (% continued)	53.4	40.4	18.6	6.8	3.0	28.5	36.3	<0.001
BMI child (kg/m^2)	15.9 (1.4)	16.9 (2.7)	15.8 (2.1)	17.2 (2.3)	16.9 (2.0)	17.0 (2.3)	17.2 (2.5)	<0.001
Maternal BMI (kg/m^2)	23.2 (3.9)	24.9 (4.9)	23.8 (4.5)	24.9 (4.8)	25.2 (4.3)	23.7 (3.6)	25.5 (5.6)	<0.001

Values are percentages or means (SD).

[a] P-values are calculated with the Chi-square test for categorical variables and ANOVA for continuous variables.

[b] Data was missing for maternal educational level (9.2%), net household income (24.8%), daycare (56.3%), breastfeeding duration (34.7%), maternal smoking (13.3%), maternal alcohol use (21.0%), BMI child (0.2%) and maternal BMI (26.1%)

Supplemental Table 2. Associations between potential mediators and herpesvirus seropositivity

	Cytomegalovirus		Epstein-Barr virus		Herpes Simplex virus 1	
	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value
Potential mediators						
Parity	0.89 (0.82,0.98)	0.012	1.08 (0.99,1.17)	0.096	1.07 (0.96,1.18)	0.22
Maternal educational level						
Low	0.67 (0.53,0.85)	0.001	1.34 (1.07,1.69)	0.011	1.88 (1.41,2.51)	<0.001
Mid-low	0.66 (0.54,0.82)	<0.001	0.99 (0.81,1.21)	0.93	1.37 (1.03,1.83)	0.03
Mid-high	0.89 (0.72,1.11)	0.31	0.92 (0.75,1.13)	0.43	1.14 (0.80,1.63)	0.46
High	Ref		Ref		Ref	
Net household income						
≤ 2200 euro	0.87 (0.72, 1.06)	0.16	1.20 (1.00, 1.44)	0.057	1.29 (1.02,1.64)	0.036
> 2200 euro	Ref		Ref		Ref	
Daycare						
No	Ref		Ref		Ref	
Yes	1.38 (1.09, 1.76)	0.009	0.97 (0.76, 1.23)	0.77	0.86 (0.68, 1.09)	0.20
Breastfeeding duration						
never	Ref		Ref		Ref	
<3 months	1.34 (0.94,1.90)	0.10	0.89 (0.66,1.20)	0.44	0.74 (0.51,1.08)	0.12
3-6 months	1.48 (1.07,2.06)	0.019	1.01 (0.72,1.42)	0.95	0.67 (0.43,1.03)	0.067
>6 months	1.63 (1.13,2.36)	0.012	0.83 (0.60,1.14)	0.23	0.67 (0.44,1.02)	0.061

OR, odds ratio; CI, confidence interval.

Table is based on imputed dataset (potential confounders and mediators are multiple imputed, determinant observed).

Values are derived from logistic regression models and represent effect estimates (95% confidence intervals), adjusted for sex, age and ethnic background.

Supplemental Table 4. Determinants of herpesvirus seropositivity in children at 6 years of age

n=4,464	Cytomegalovirus	Epstein-Barr Virus	Herpes simplex virus 1
Determinants	(aOR 95% CI)	(aOR 95% CI)	(aOR 95% CI)
Gender			
Male	1.0	1.0	1.0
Female	**1.18 (1.02; 1.36)***	0.99 (0.87; 1.15)	1.05 (0.97; 1.13)
Ethnicity			
Western	1.0	1.0	1.0
Non western	**2.16 (1.81; 2.57)*****	**1.76 (1.48; 2.09)*****	**1.52 (1.39; 1.66)***
Income			
Low (<2200)	1.18 (0.98; 1.42)	1.09 (0.91; 1.30)	**1.11 (1.01; 1.23)***
High (≥2200)	1.0	1.0	1.0
Maternal educational level			
Low	0.95 (0.73; 1.23)	1.09 (0.85; 1.41)	**1.55 (1.35; 1.79)*****
Mid-low	0.82 (0.66; 1.01)	0.93 (0.76; 1.15)	**1.19 (1.05; 1.35)****
Mid-high	0.96 (0.77; 1.19)	0.94 (0.77; 1.15)	**1.16 (1.03; 1.31)***
High	1.0	1.0	1.0
Daycare	1.20 (1.00; 1.43)	1.06 (0.86; 1.31)	**0.89 (0.81; 0.97)***
Breastfeeding duration			
Never	1.0	1.0	1.0
<3 months	**1.53 (1.04; 2.23)***	0.98 (0.73; 1.33)	1.02 (0.89; 1.16)
3-6 months	**1.77 (1.21; 2.60)****	1.20 (0.85; 1.69)	1.08 (0.93; 1.26)
>6 months	**1.96 (1.25; 3.07)****	1.08 (0.78; 1.85)	1.06 (0.92; 1.22)
Parity			
0	1.0	1.0	1.0
1	**0.76 (0.65; 0.90)****	1.11 (0.95; 1.30)	**1.18 (1.08; 1.29)*****
2	0.90 (0.69; 1.18)	**1.35 (1.05; 1.74)***	**1.31 (1.14; 1.50)*****
≥3	0.85 (0.54; 1.33)	**1.58 (1.01; 2.48)***	**1.60 (1.29; 1.99)*****

Abbreviations: aOR, adjusted Odds Ratio; CI, confidence interval. Values reflect aORs (95% CI). Table is based on multiple imputed dataset. *P<0.05.

** P<0.01.

*** P<0.001. Models were adjusted for age, sex, socioeconomic position (ethnicity, family net household income, maternal educational level), crowding (daycare, breastfeeding, parity) and maternal lifestyle (maternal age, smoking and alcohol use during pregnancy, and maternal BMI).

Supplemental Table 6. Attenuation of the OR of Cytomegalovirus seropositivity for different ethnicities after individual adjustment for the explanatory variables

	Surinamese-Creoles[a] OR (95% CI)	Attenuation a (95% CI)	Surinamese-Hindustani [b] OR (95% CI)	Attenuation b (95% CI)	Turkish [c] OR (95% CI)
CMV					
Model 1 (includes child sex and current age)	2.29 (1.61, 3.28)		1.92 (1.35, 2.72)		4.41 (3.47, 5.62)
Socioeconomic status					
Model 1+ education	2.72 (1.87, 3.88)	**+33% (17 to 60)*** **	2.26 (1.56, 3.19)	**+37% (19 to 80)*** **	5.36 (4.01, 6.76)
Crowding					
Model 1 + parity	2.29 (1.60, 3.28)	0% (-10 to 10)	1.89 (1.33, 2.71)	-3% (-20 to 14)	4.49 (3.51, 5.74)
Model 1 + daycare	2.37 (1.62, 3.32)	**+6% (1 to 14)****	2.14 (1.49, 3.07)	**+25% (8 to 58)****	5.08 (3.84, 6.42)
Model 1 + breastfeeding	2.35 (1.63, 3.33)	**+5% (0.4 to 12)***	1.99 (1.37, 2.79)	**+8% (2 to 19)****	4.24 (3.33, 5.43)
Fully adjusted model	2.73 (1.81, 3.83)	**+34% (14 to 66)*** **	2.42 (1.62, 3.42)	**+55% (24 to 124)*** **	5.53 (3.98, 6.97)

Abbreviations: OR, Odds Ratio; CI, confidence interval.

[a] The ORs represent the attenuated effect estimate and their 95% CIs relative to model 1 after individual adjustment for explanatory factors.

[b] Attenuations a to e represent the attenuations of effect estimates for the different ethnicities relative to model 1 (includes confounders) after adjustment for explanatory variables (100x ($OR_{model\ 1} - OR_{model\ 1\ with\ explanatory\ factor}$) / ($OR_{model\ 1}$)). Dutch is the reference group. No significant difference in CMV prevalence was observed for the Antillean group compared with the Dutch group, therefore attenuations in effect estimates for the Antillean group are not presented.

* p<0.05.

** p<0.01.

*** p<0.001.

Attenuation c (95% CI)	Moroccan [d] OR (95% CI)	Attenuation d (95% CI)	Cape Verdean [e] OR (95% CI)	Attenuation e (95% CI)	Antillean [f] OR (95% CI)	Attenuation f (95% CI)
	3.64 (2.77, 4.79)		3.17 (2.33,4.32)		n.s.	NA
+28% (14 to 44)* **	4.23 (3.17, 5.61)	**+22% (11 to 36)*** **	3.70 (2.69, 5.12)	**+24% (12 to 40)*** **		
+2% (-4 to 9)	3.94 (2.97, 5.24)	**+11% (0.5 to 24)***	3.14 (2.29, 4.32)	-1% (-12 to 11)		
+20% (7 to 36) **	3.90 (3.00, 5.40)	**+10% (3 to 19)*** **	3.15 (2.28, 4.24)	-1% (-3 to 2)		
-5% (-8 to -2) **	3.62 (2.70, 4.69)	-1% (-4 to1)	3.33 (2.45, 4.58)	**+8% (2 to 15)** **		
+33% (13 to 59) **	4.60 (3.36, 6.31)	**+36% (19 to 60)*** **	3.68 (2.6-, 5.07)	**+23% (5 to 47)** **		

Supplemental Table 9. Maternal and infant characteristics of the study population

(n=4464)	Epstein Barr virus, cytomegalovirus or herpes simplex virus type1		
	0 infections (n=1319)	**1 Infection (n=1735)**	**≥2 infections (n=1410)**
Maternal characteristics			
Age (Mean ± SD; years)	**31.7 (4.7)**	**30.8 (5.0)*** **	**29.7 (5.4)*** **
Educational level (n; %)			
Low	**180 (14.4%)**	**322 (20.3%)*** **	**366 (30.1%)*** **
Mid-Low	376 (30.0%)	476 (30.0%)	382 (31.5%)
Mid-High	329 (26.3%)	336 (21.2%)	229 (18.9%)
High	368 (29.3%)	452 (28.5%)	237 (19.5%)
Net household income per month (n; %)			
< € 2000	177 (16%)	325 (23%)	**368 (35%)*** **
€ 2000-3200	291 (22%)	370 (26%)	**254 (24%)**
> € 3200	668 (59%)	716 (51%)	**443 (42%)**
Smoking continued during pregnancy (n; %)	138 (13%)	**232 (17%)****	166 (15%)
Alcohol use continued during pregnancy (n; %)	487 (46%)	595 (43.4)	**395 (35.7%)*** **
BMI before pregnancy (Mean ± SD; kg/m^2)	23.4 (3.9)	**23.6 (1.7)**	**23.9 (4.3)****
Maternal atopy (n; %)	396 (33.8%)	548 (36.9%)	354 (30.2%)
Family history of asthma / atopy (n; %)	605 (48.7%)	809 (51.1%)	**541 (43.7%)***
Parity			
0	**717 (57%)**	**917 (55%)***	**736 (54%)*** **
1	**404 (32%)**	**533 (32%)**	**401 (29%)**
2	**118 (9%)**	**162 (10%)**	**169 (12%)**
≥3	**23 (2%)**	**61 (3%)**	**62 (5%)**
Mode of delivery (n; %)			
Vaginal	825 (73.4%)	1071 (72%)	914 (74.6%)
Forceps or vacuum assisted	166 (14.8%)	203 (13.6%)	163 (13.3%)
Caesarian section	133 (11.8%)	214 (14.4%)	148 (12.1%)
Premature rupture of membranes (n; %)	51 (4.1%)	72 (4.4%)	58 (4.3%)
Infant characteristics (n=258)			
Males (n; %)	683 (51.8%)	922 (53.1%)	699 (49.6%)
Western ethnicity (n;%)	1047 (80%)	**1162 (69%)*** **	**655 (49%)*** **
Gestational age (Mean ± SD; weeks)	39.8 (1.9)	39.8 (1.7)	39.8 (1.8)
Preterm birth (<37 weeks) (n; %)	76 (6.0%)	104 (6.0%)	79 (5.6%)
Birth weight (Mean ± SD; grams)	3442 (599)	3413 (557)	**3386 (552)****
Low birth weight (<2500 gr)	75 (5.7%)	98 (5.7%)	67 (4.8%)
Day-care in first year (n; %)	567 (85%)	**699 (89%)***	447 (88%)
Breastfeeding (n; %)			
Never	**98 (10.2%)**	114 (10.0%)	**52 (6.3%)***
< 3 months	**340 (35.5%)**	364 (32.0%)	**286 (34.8%)**

Supplemental Table 9. Maternal and infant characteristics of the study population (continued)

(n=4464)	Epstein Barr virus, cytomegalovirus or herpes simplex virus type1		
	0 infections (n=1319)	**1 Infection (n=1735)**	**≥2 infections (n=1410)**
3-6 months	**195 (20.4%)**	237 (20.8%)	**186 (22.7%)**
>6 months	**324 (33.9%)**	423 (37.2%)	**297 (36.2%)**
Infections first year (n; %)			
RTI	**538 (57.2%)**	652 (59.4%)	**495 (63.3%)***
LRTI	125 (13.8%)	140 (13.5%)	107 (15.0%)
URTI	**490 (49.8%)**	607 (52.6%)	**471 (56.1%)****
GI	**329 (55.7%)**	378 (58.0%)	**296 (64.1%)****
Atopy first year			
Eczema	264 (27.4%)	312 (28.2%)	215 (27.2%)
Wheezing (≥ 1 epsiode)	**338 (42.8%)**	412 (45.5%)	**315 (49.5%)***
Nasopharyngeal bacterial carriership 6 years (n; %)			
No pathogen	606 (47.3%)	794 (46.8%)	606 (44.2%)
S Aureus	338 (26.4%)	459 (27.0%)	408 (29.8%)
S. Pneumoniae	255 (19.9%)	296 (17.4%)	246 (17.9%)
H. Influenza	125 (9.8%)	197 (11.6%)	152 (11.1%)
M. Catarrhalis	117 (9.1%)	156 (9.2%)	156 (11.4%)
Antibiotic use (n; %)			
YEAR 1	**386 (47.6%)**	**504 (54.1%)****	340 (52.7%)
YEAR 2	**413 (41.6%)**	520 (44.9%)	**374 (47.3%)***
YEAR 3	**313 (32.4%)**	396 (36.4%)	**281 (37.5%)***
YEAR 4	**260 (27.6%)**	331 (30.1%)	**247 (32.2%)***
YEAR 6	**227 (18.9%)**	**339 (23.0%)***	**266 (23.9%)****
EVER (at age 6 years)	**830 (85.0%)**	**1059 (89.4%)****	**774 (90.0%)****
Age at focus visit			
(Median ± range; years)	6.0 (5.0-8.8)	**6.0 (5.0-9.0)*****	**6.1 (4.9-9.1)*****
BMI for age (mean; sd; SDS)	0.20 (0.85)	0.25 (0.92)	**0.39 (0.95)*****

Values are means (SD), absolute numbers (percentages) or #medians (90% range). Based on observed dataset. Significantly different from 0 infections*P<0.05) ** (P<0.01) *** (P<0.001) Maternal atopy: eczema, house dust mite allergy, hay-fever,

any reported autoimmune disease includes diabetes mellitus, systemic lupus erythematosus, arthritis, multiple sclerosis, thyroid disorder or celiac disease

Part III

Consequences of Celiac Disease Autoimmunity

Chapter II

Growth Trajectories and Bone Mineral Density in Anti-Tissue Transglutaminase Antibody-positive Children: The Generation R Study

Michelle A.E. Jansen, Jessica C. Kiefte-de Jong, Romy Gaillard, Johanna C. Escher, Albert Hofman, Vincent W.V. Jaddoe, Herbert Hoijkaas, Henriette A. Moll

Clin Gastroenterol Hepatol 2015; May;13(5):913-20.

ABSTRACT

Background and aims

Increased levels of anti-tissue transglutaminase (tTG) levels have been associated with reduced weight and bone mineral density (BMD) in symptomatic patients with celiac disease. Little is known about the effects of these antibodies patients with subclinical or other forms of celiac disease. We examined associations between anti-tTG positivity and growth and BMD.

Methods

In a population-based prospective cohort study, serum samples were collected from children (median age, 6 years, n=4442) and analysed for anti-tTG levels. All children were born between April 2002 and January 2006 and were not previously diagnosed with celiac disease. Children were categorized as anti-tTGnegative (< 7 U/ml, *n=4249*), or anti-tTG positive (≥7 U/ml, *n=57*). Children's levels of anti-tTG positive levels were further categorized into 2 categories on the basis of ≥10 times upper limit of normal (70 U/ml). Height, weight and body mass index (BMI) age- and sex-adjusted standard deviation scores (SDS) ((observed value- mean)/SD) were obtained using Dutch reference growth charts. BMD was measured by dual-energy x-ray absorptiometry . Multivariable linear regression and linear mixed models were performed.

Results

Children who tested positive for anti-tTG had reduced growth in weight SDS/year (reduction of 0.05; 95% CI, reductions of 0.09,0.01) and BMI SDS/year (reduction of 0.10; 95% CI, reductions of 0.18,0.01) from 6 months until 6 years, compared with children without anti-tTG; they also tended to have reduced growth in height from 6 months until 6 years (reduction of 0.02 SDS/year; 95% CI reductions of 0.06, 0.02). Children who tested positive for anti-tTG were shorter (0.29 SDS less; 95% CI reductions of 0.55, 0.04), weighed less (0.38 SDS less; 95% CI , reductions of 0.64, 0.12), and had lower BMIs (0.26 SDS less; 95% CI, reductions of 0.49, 0.03) and BMDs (0.26 SDS less; 95% CI, reductions of 0.45, 0.08) at 6 years of age than anti-tTG negative children.

Conclusion

Anti-tTG positive children without gastrointestinal symptoms have lower BMDs and reduced growth trajectories until they are 6 years old. This suggests that subclinical or potential CD can affect BMD and growth.

INTRODUCTION

Celiac disease (CeD) is an autoimmune-mediated disease that is caused by ingestion to gluten in genetic predisposed individuals. Screening studies have shown that the prevalence of CeD is approximately 1%, and increases over time.[1-3] However, CeD is underdiagnosed, because clinical symptoms in childhood are often minor, atypical or even absent.[4] Screening on anti-tissue transglutaminase (tTG) concentrations has the potential to detect these forms of otherwise undetected subclinical and potential CeD.[5,6] However, it remains unclear whether these forms of CeD should be treated, because the consequences of untreated subclinical and potential CeD in childhood remain unclear. Consequences of classic CeD include a decreased length, weight, body mass index (BMI), and bone mineral density (BMD), which completely recover by using a gluten free diet (GFD) in childhood. However, by the time CeD is diagnosed in adulthood, complete bone recovery is not always achieved. It is estimated that one-third of adult patients have osteoporosis, one-third have osteopenia, and the remaining one-third have a normal BMD.[7,8] In addition, screening studies showed a decreased BMD even in asymptomatic adults who had positive concentrations of celiac antibodies,[7,9] but no villous atrophy,[10] indicating that merely the presence of positive serology in adulthood is sufficient to have adverse effects on BMD. Moreover, persisting osteopenia and osteoporosis in adults, despite a GFD, may be the result of subclinical CeD in childhood.[10,11] Because the effects on BMD in childhood are independent of clinical symptoms, [12] adverse effects might also be present in children with undetected forms of CeD. Although children with positive immunoglobulin A-endomysial antibody (EMA) serology are suggested to be shorter and lighter,[3] effects on growth trajectories and bone development are not known.

Therefore, the aim of this study was to assess whether anti-tTG positivity was associated with decreased height, weight, BMI and BMD in a population-based study of children from 6 months until 6 years of age.

METHODS

Study Design

This study was embedded within a population-based prospective cohort study.[13] All children were born between 2002 and 2006. From the age of 6 years, 6690 visited the research center (median age, 6.0 years).[13] of whom serum anti-tTG levels were available in 4442 (66%). Of these, we excluded twins, children with a questionnaire reported CeD diagnosis and a GFD. The resulting population for analysis consisted of 4306 children [Supplemental Figure 1].

Anti–tissue transglutaminase concentrations

Anti-tTG immunoglobulin A concentrations were measured in venous blood serum samples by using a fluorescence enzyme immunoassay (FEIA Phadia ImmunoCAP 250; EliA IgA, Phadia AB, Uppsala Sweden).[14] Concentration of anti-tTG were categorized into 2 groups: group 1, anti-tTG negative (< 7 U/ml) and group 2, anti-tTG positive (≥7 U/ml) [Supplemental Figure 1]. In addition, anti-tTG positive concentrations were categorized into 2 categories on the basis of below or above the ≥ 10 times upper limit of normal concentrations of the test kit (≥70 U/ml). [5]

HLA DQ2 DQ8

A genome-wide association scan (Illumia 610K) of child DNA was taken from 5908 cord-blood samples. A tag single nucleotide polymorphisms approach was used to capture whether children carried the HLA-DQ risk type DQ2 or DQ8.[15,16] Children were genotyped for HLA-DQ2 (rs2187668, rs2395182, rs4713586 and rs7775228) and DQ8 (rs7454108) by using genome wide Illumina 610 Quad Array (San Diego, CA, USA). Genotype and allele frequencies were in Hardy Weinberg equilibrium (rs2187668, p=0.88; rs2395182, p=0.78; rs4713586, p=0.95; rs7775228, p=0.86; rs7454108, p=0.90).

Growth measurements

Child anthropometrics were obtained by measurements at 6, 14, 24, 36 and 48 months and at 6 years of age.[17] We obtained age-adjusted standard deviation scores (SDS) by using Dutch reference growth curves (Growth Analyzer 3.0; Dutch Growth Research Foundation, Rotterdam, Netherlands). Definitions of overweight (>1.1 to 2.3 SDS) and obesity (BMI > 2.3 SDS) were based on internationally established age- and sex-adjusted BMI distributions.[17,18]

Gastrointestinal symptoms

Gastrointestinal symptoms were assessed by parental reported questionnaires (median age, 6 years). Functional constipation was defined if at least the following symptoms of Rome III [19] were reported in the past year: (1) defecation frequency <3 times a week, (2) predominantly hard faeces for the majority of stools, (3) ≥1 episode of fecal incontinence per week. The stool pattern of the child was assessed using questionnaires at the age of 6 years. Mothers were asked whether their child had abdominal complaints during the last 3 months (y/n), if their child was feeling sick or nauseous while or after having eaten (y/n), and how their child's faeces usually looked over the past 3 months (answer options included: 'very hard' to 'very soft/slushy' or 'watery'). Diarrhea was defined when 'very soft/slushy' or 'watery' faeces was indicated.

Bone Mineral Density, Bone Mineral Content and Bone Area

BMD (g/cm^2), bone mineral content (BMC) (g), and bone area (BA (cm^2) were measured by a dual-energy X-ray absorptiometry (DXA) scan (iDXA, General Electrics-Lunar, 2008,

Madison, WI) (median age, 6 years).[20] All scans were performed by well-trained research assistants and by using the same device and software (enCORE, version 13; GE Healthcare, Little Chalfont, Buckinghalshire, United Kingdom). Total body less head BMD, BMC, and BA were used to analyze total body bone mass. [21]

Covariates

Maternal anti-tTG concentrations were measured during pregnancy and maternal height and weight were measured at enrollment. Fetal gender, gestational age and birth weight were obtained from obstetric records.[13] Ethnicity was determined according to the classification of Statistics Netherlands.[22] Data on other sociodemographic and lifestyle factors, including educational level[23] and household income,[24] were obtained by a combination of pre- and postnatal questionnaires completed by both parents. Postnatal questionnaires also included information on feeding behavioure, physical activity, and family history of CeD.

Statistical methods

First, independent T test and Chi-square tests were used to test for differences in characteristics between the anti-tTG negative and positive groups. Then, to assess whether growth characteristics were different between these groups, 2 different types of analyses were performed. First, to longitudinally assess associations between anti-tTG and growth, repeated-measurement analyses were performed by using linear mixed models with an unstructured symmetry covariance structure (the Statistical Analysis System). These models take the correlations between repeated measurements within the same subject into account. The between-measurement correlations for height and weight decreased with increasing age, from 0.81 to 0.58 for height, and 0.84 to 0.52 for weight. Goodness-of-fit indices (AIC's) of models with a CS, UN and AR1 symmetry covariance structure were compared. The UN model provided the smallest AIC and was therefore applied. Second, we performed cross-sectional analyses to examine associations between anti-tTG positivity with height, weight, BMI and BMD at 6 years of age in 2 linear regression models: (1) a basic model and (2) a confounder model, which also adjusted for covariates that showed a change in effect estimate of $\geq$10%.[25] We constructed age-adjusted standard SDS ([observed value- mean]/SD) for childhood outcomes to enable comparison of effect estimates. In addition, we assessed whether HLA DQ2-DQ8 status, ethnicity, gender, breastfeeding and timing of gluten introduction modified the association, but no significant interactions were observed. Binary logistic regression models were performed to assess the associations between anti-tTG and gastrointestinal complaints. To reduce bias associated with missing data, multiple imputation (n=10) of the covariates was performed.[26] Regression coefficients were pooled by taking the average of the 10 imputed data sets. The pooled standard error was calculated by using the rules of Rubin and Schenker.[27] Last, sensitivity analyses were performed to test whether the associations between anti-tTG and height, weight, BMI and

BMD were different before and after excluding HLA DQ2/DQ8 negative children. Longitudinal mixed model analyses were performed in Statistical Analysis System 9.3 (SAS Institute Inc., Cary, NC). Imputation procedures and cross-sectional analyses were performed in SPSS 20.0 for Windows (SPSS Inc., Chicago, Il),

RESULTS

Subject characteristics

Mean child's age was 6.2 (range 4.9, 9.1) years. Out of 4306 children, 57 (1.3%) had anti-tTG positive levels at 6 years of age, and 4249 (98.7%) had anti-tTG negative concentrations [Table 1, Supplemental Figure 1, Supplementary Table 1]. Of children with anti-tTG positive concentrations, 91% carried the HLA-DQ2 or -DQ8 molecule [Table 1]. Higher income, female gender and Western Ethnicity were related to positive anti-tTG concentrations. Breastfeeding duration and timing of gluten introduction were not different between anti-tTG positive and negative children.[14]

Gastrointestinal symptoms

No significant differences were observed in doctor's diagnosis of (functional) constipation according to the ROME III criteria (adjusted odds ratio 0.57; 95% CI 0.22,1.46), abdominal pain (aOR 1.50; 95% CI 0.82, 2.72), a soft/watery stool pattern (aOR 1.32; 95% CI 0.55, 3.18), or a sick or nauseous feeling while/after having eaten (aOR 0.61; 95% CI 0.18,2.06) [Table 2].

Longitudinal growth outcomes

Longitudinal analysis on height, weight and BMI from 6 months until 6 years of age derived from linear mixed models are illustrated in Figure 1. Children who tested positive for anti-tTG tended to be shorter (0.12 SDS less; 95% CI reductions of 0.39, 0.14 SDS) and to weigh less (0.14 SDS less; 95% CI reductions of 0.41, 0.14) at 6 months than anti-tTG negative children, although these effects were not significant. However, anti-tTG positive children had reduced growth in weight SDS/year (reductions of 0.05; 95% CI reductions of 0.09, 0.01, p=0.01), and BMI SDS/year (reductions of 0.10; 95% reductions of 0.18, 0.01, p=0.03) from 6 months until 6 years, compared with children without ant-tTG; they also tended to have reduced growth in height (reduction of 0.02 SDS/year; 95% CI reductions of 0.06, 0.02) although this effect was non-significant. The differences in height, weight and BMI between children with and without anti-tTG positive levels became stronger with increasing age and were mostly noticeable at 6 years of age. Anti-tTG positive children tended to be shorter(0.24 SDS shorter; 95% CI reductions of 0.52, 0.04, p=0.09) and weighed less (0.41 SDS less; 95% CI reductions of 0.69, 0.14, p=0.003) at 6 years of age than anti-tTG negative children [**Figure 1**].

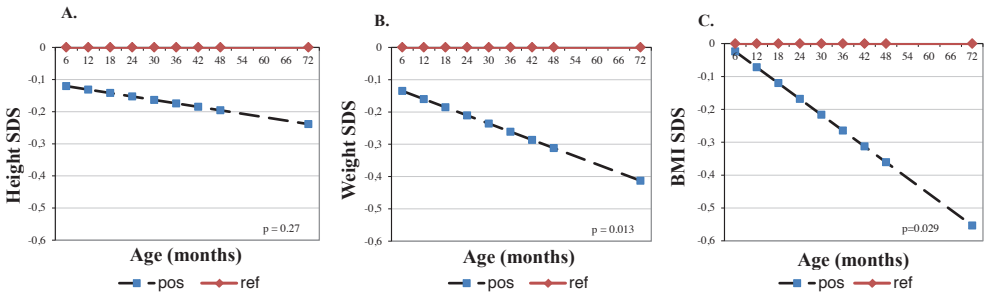


Figure 1. Longitudinal growth differences between tTG levels positive children relative to tTG negative children.

Estimated regression lines derived from longitudinal linear mixed model analysis reflect difference in height (A), weight (B) and BMI (C) SDS from 6 months until 6 years in in anti-tTG positive anti-tTG levels relative to anti-tTG negative children.

Á. No significant difference in height over time (p=0.27). Borderline difference (-0.24 SDS) in height at 6 years of age (p=0.09).

B. Significant difference in weight over time (p=0.013). Significant difference (-0.41 SDS) at 6 years of age (p=0.003)

C. Significant difference in BMI over time (p=0.029). Significant difference (-0.60 SDS) at 6 years of age (p=0.021)

Adjusted for timing of introduction of gluten, gender, ethnicity, income, breastfeeding, birth weight and maternal tTG level during pregnancy.

Growth outcomes at 6 years of age

Cross-sectional analysis on height, weight, and BMI are shown in Table 3 and Supplemental Table 2. Anti-tTG positive children were shorter (0.29 SDS shorter; 95% CI reductions of 0.55, 0.04), weighed less (0.38 SDS less; 95% CI reductions of 0.64, 0.12) and had lower BMI (0.26 SDS less; 95% CI reductions of 0.49, 0.03) at 6 years of age than anti-tTG negative children. Results were most eminent in children with high concentrations (≥ 70 U/ml). Children who tested positive for anti-tTG concentrations ≥ 70 U/ml weighed less (0.45 SDS less; 95% CI reductions of 0.80, 0.10) and had lower BMI (0.33 SDS less; 95% CI, reductions of 0.65, 0.02) at 6 years if age than anti-tTG negative children, but no differences in weight and BMI were observed in those with positive concentrations below 10 times upper limit of normal (7-70 U/ml), (respectively 0.28 SDS less; 95% CI reductions of 0.65, 0.09), (0.19 SDS less; 95% CI reductions of 0.53, 0.14). Associations were independent of gastrointestinal symptoms, including constipation, abdominal pain, soft, slushy or watery faeces and a sick or nauseous feeling while or after having eaten.

Bone Mineral Density outcome at 6 years of age

Anti-tTG positive children had lower BMDs of total body and spine (respectively, 0.26 SDS less; 95% CI reductions 0.47, 0.05) (0.35 SDS less; 95% CI reductions of 0.56, 0.13) at 6 years of age than anti-tTG negative children [Table 4]. In addition, inverse linear associations were observed between anti-tTG concentrations, weight, BMI, and BMD of total body and spine [Supplemental Table 3]. Results were most eminent in children with anti-tTG concentrations ≥ 10 times upper limit of normal [Table 4]. Children who teste positive for

anti-tTG concentrations ≥ 70 U/ml had lower total BMDs and spinal BMDs (respectively, 0.40 SDS less; 95% CI reductions of 0.70, 0.11) (0.55 SDS less; 95% CI reducetions of 0.85, 0.25) at 6 years of age than anti-tTG negative children, but no differences in total BMDs and spinal BMDs were observed in those with positive concentrations below the 10 times upper limit of normal (7-70 U/ml), (respectively, 0.11 SDS less; 95% CI reductions of 0.41, 0.20), (0.13 SDS less; 95% CI reductions of 0.44, 0.18). Associations were independent of gastrointestinal symptoms.

Sensitivity analyses

After excluding children who were HLA DQ2/DQ8 negative, sensitivity analysis confirmed the associations between anti-tTG positivity and decreased height, weight, BMI and BMD SDS. In fact, effect sizes were stronger in the HLA DQ2/DQ8 positive children. Within the population of HLA DQ2/DQ8 carriers, anti-tTG positive children were shorter (0.45 SDS shorter; 95% CI reductions of 0.77, 0.14), weighed less (0.51 SDS less; 95% CI reductions of 0.83, 0.20), had lower BMIs (0.31 SDS lower; 95% CI reductions of 0.60, 0.02), lower total body BMDs (0.29 SDS less; 95% CI reductions of 0.55, 0.003) and lower BMDs of the spine (0.46 SDS less; 95% CI reductions of 0.72, 0.20) than anti-tTG negative children. Differences in height weight, total BMDs, and spinal BMDs were more prominent in children who tested positive for anti-tTG ≥ 70 U/ml [Supplemental Table 4].

DISCUSSION

Interpretation of main findings

This prospective observational study showed that anti-tTG positivity was associated with a deviation in height, weight, BMI, and BMD in young children. The effects of positive anti-tTG were mostly present in children with anti-tTG levels ≥10 times upper limit normal (≥ 70 U/ml).

This is a population-based study that assessed associations between anti-tTG levels and growth longitudinally and BMD in anti-tTG positive children without gastrointestinal symptoms. Previous studies did not assess BMD[3] or growth longitudinally,[3] but they reported decreased weight and BMD at the time of CeD diagnosis in symptomatic children.[28-33] Our results indicate that reduced weight and BMD even occurred in screening identified anti-tTG positive children. All anti-tTG positive children did not receive a previous CeD diagnosis, because we excluded children with a GFD or CeD diagnosis. Interestingly, children carrying anti-tTG antibodies in our study did not report more often abdominal pain, functional constipation, a soft/watery stool pattern or a sick/nauseous feeling while or after having eaten. Apart from diarrhea, this is in line with Bingley et al,[3] who found that in antibody positive children, who were detected by active anti-EMA screening, gastro-intestinal symptoms such as vomiting, abdominal pain and constipation, were not prominent. Comparable to

our results they found that these antibody positive children were 0.76 SDS shorter and 0.54 SDS lighter than antibody negative children at 7 years of age. We found that children with anti-tTG concentrations ≥ 10 times upper limit of normal were 0.31 SDS shorter and 0.45 SDS lighter. This equates to about 6-9 months growth (approximately 2 kg weight gain) in an average child around this age. These associations were independent of gastrointestinal symptoms, which are in line with Bingley et al, who presumed that the effects on height and weight are unrelated to malabsorption. Our results can be explained by 2 mechanisms: First, anti-tTG positive concentrations might lead to villous atrophy and nutrient deficiency, such as calcium, and vitamin D deficiency, which consequently may have adverse effects on BMD and growth. Second, anti-tTG positive concentrations might induce inflammation, which influences bone development and growth.[32] A deviation of -0.55 SDS in BMD (total body less head) equates to about 1 year developmental delay in bone mass density at the age of 6 years.[33] A decreased BMD in childhood is a strong predictor of low peak bone mass in young adulthood, which influences the risk of osteoporosis in later life.[34]

In adult CeD patients, BMD has been shown to be independent of clinical symptoms at diagnosis.[9] Although a GFD improves the BMD in clinical CeD, it does not always restore osteopenia or osteoporosis completely. However, a GFD started in early childhood in patients with proven CeD reaches complete recovery of BMD within 1 year. [6,12] Once the GFD starts, most children catch up their height-weight growth and bone mineralization rate.[10] The younger the children start with a GFD, especially before puberty [31], the better the BMD.[12] Therefore, persisting osteopenia and osteoporosis in adults, despite a GFD, might reflect subclinical and asymptomatic CeD in childhood, which underlies failure to reach peak bone mass during bone-building years.[11] This highlights the importance of diagnosing CeD as early as possible, and it suggests that these patients would benefit from screening to start early with treatment. However, the effects of a GFD on growth and BMD in children with subclinical or potential CeD remain unclear. Until the effectiveness of a GFD in these children has been shown, it would be premature to recommend screening. In addition, follow-up is necessary to determine to what extent children with anti-tTG positivity will develop CeD later in life.

Methodological considerations

Detailed information on height, weight and BMI from birth until 6 years was collected prospectively, which enabled us to assess childhood growth longitudinally. In addition, information on a broad range of potential confounders was available.

However, some limitations should be taken into account. First, although the sensitivity and specificity of anti-tTG is high,[5] we did not have biopsy specimens and anti-EMA levels, thus final conclusions concerning CeD diagnosis should be made with caution. According to the ESPGHAN criteria[5], only when anti-tTG levels are increased ≥10 times upper limit of normal in accordance with clinical symptoms and EMA positive concentrations, CeD can be diagnosed without duodenal biopsy. However, even if childrenwith positive anti-tTG con-

centrations lack criteria for final CeD diagnosis, the association between anti-tTG positivity and decreased BMD remains valid. Another point of consideration is that HLA DQ2/DQ8 negative children were included which might have diluted the effect estimates. However, sensitivity analysis confirmed the observed associations between height, weight, BMI and BMD. In fact, effect estimates were stronger in the HLA DQ2/DQ8 positive population, and most pronounced in children with anti-tTG concentrations ≥ 70U/ml, suggesting that the results are driven by children who are likely to have CeD. Second, symptom assessment was performed by parental reported questionnaires, thus underreporting of symptoms cannot be totally excluded. However, questionnaires were filled in before anti-tTG measurement, which makes recall bias less likely. Third, we had 33% missing data on HLA DQ2 DQ8 status. Nevertheless, because the genetic risk types were defined from cordblood samples, availability did not depend on CeD diagnosis, and missing data on HLA are not likely to have influenced the observed associations. Fourth, we only determined anti-tTG concentrations at the age of 6 years. Therefore, it remains uncertain at what age children became positive. Last, missing blood samples at the 6-year visit were mainly due to non-consent of the parents or child. However, non-response would only lead to biased effect estimated if the association between anti-tTG concentration and height, weight and BMD was different between those included and those not included in the analysis. This seems unlikely, but cannot be excluded.

Conclusion

Screeningdetected anti-tTG positive children have lower BMDs and reduced growth trajectories until they are 6 years old. This suggests that subclinical or potential CeD can affect BMD and growth.

Table 1. Maternal, Paternal and Child characteristics according to anti-tTG levels multiple imputed data set (n=4306)

	Negative anti-tTG (<7 U/ml) *N=4249 (99%)*	**Positive anti-tTG (≥ 7 U/ml)** *N=57 (1%)*
	N (%)	*N* (%)
Maternal		
Age (mean ±SD; years)	30.7 (5.1)	31.7 (4.6)
Height (mean ±SD; cm)	167.7 (7.4)	168.1 (7.5)
Weight (mean ±SD; kg)	69.7 (13.0)	66.3 (11.0)
BMI before pregnancy (mean; SD: kg/m^2)	23.7(4.2)	23.0 (3.1)
Educational level (n;%)		
Low	2003 (47%)	22 (39%)
High	2246 (53%)	35 (61%)
Household income (n;%)		
< 2000 Euro	**1114 (26%)**	**11 (19%)***
≥ 2000 Euro	**3135 (74%)**	**46 (81%)**
Gastrointestinal disease (n;%)		
None or no reported	3208 (99%)	36 (97%)
Celiac disease	1 (0.0%)	1 (3%)
Lactose intolerance	1 (0.0%)	0 (0%)
Inflammatory Bowel disease	4 (0.1%)	0 (0%)
Bowel complaints without known organic cause	13 (0.4%)	0 (0%)
Any questionnaire reported autoimmune disease (diabetes mellitus, SLE, arthritis, MS or thyroid disorder) (n; %)	122 (3%)	2 (4%)
Smoking during pregnancy (n;%)	1102 (26%)	14 (25%)
Alcohol during pregnancy (n;%)	2313 (54%)	36 (63%)
Nulliparous (n;%)	2325 (55%)	31 (54%)
Paternal		
Age (mean ±SD; years)	33.3 (5.6)	33.4 (4.6)
Height (mean ±SD; cm)	182.3 (8.0)	181.9 (7.8)
Weight (mean ±SD; kg)	83.9 (13.0)	82.5 (13.6)
BMI (mean ±SD: kg/ m^2)	25.2 (3.4)	24.8 (3.8)
Child		
Male gender (n;%)	**2201 (52%)**	**19 ((33%)****
Western ethnicity (n; %)	**2738 (64%)**	**44 (77%)***
CeD in family (n; %)	626 (15%)	9 (16%)
Estimated fetal growth (mean ± SD: grams)		
Second trimester	384 ± 98	385 ± 82
Third trimester	1630 ± 264	1627 ± 250
Birth weight (mean ± SD: grams)	3436 ± 601	3466 ± 601
Gestational age (mean ± SD: weeks)	39.8 ± 1.4	39.9 ± 1.4

imputed data set (n=4306) (continued)

	Negative anti-tTG (<7 U/ml) *N=4249 (99%)*	**Positive anti-tTG (≥ 7 U/ml)** *N=57 (1%)*
Ever Breastfeeding (n;%)	3885 (91%)	55 (96%)
Timing of introduction of gluten <6 mo (n;%)	1819 (43%)	28 (49%)
HLA DQ2 or DQ8 present (n; %)	**1147 (40%)**	**38 (91%)*** **
Television watching ≥ 2 hours a day (n;%)	1002 (24%)	14 (25%)
Playing outside ≥ 2 hours a day (n;%)	1404 (33%)	15 (26%)
Suppletion of vitamin D, AD or multivitamin (n;%)	2118 (50%)	29 (51%)
Age child at 6 year visit (median ± range; years)	6.02 (4.9; 9.1)	6.04 (5.6; 8.2)
Weight child at 6 year visit (median ± range; kg)	**22.6 (13.0; 56.0)**	**21.0 (16.0; 35.6)****
Height child at 6 year visit (median ± range; cm)	119.3 (99.0; 148.3)	117.6 (1.06; 135.3)
BMI (body mass index (median ± range; kg/ m^2)	**15.9 (11.9; 29.2)**	**15.4 (13.0; 19.4)****
Overweight BMI >1.1-2.3 SDS (n;%)	307 (7.2%)	4 (7%)
Obesity BMI >2.3 SDS (n;%)	142 (3.3%)	0 (0%)

* P<0.05
** P<0.01
***P<0.001

Table 2. No significant association between tTG-IgA levels and constipation (A), abdominal pain (B), soft faeces (C) and feeling nauseous (D)

	A. Functional constipation according ROME III OR (95% CI) *N=3120*			B. Abdominal pain OR (95% CI) *N=3093*		
	N (%)	OR (95% CI)	aOR (95% CI)[a]	N (%)	OR (95% CI)	aOR (95% CI)[a]
Negative	572/3074 (19%)	Ref (1.0)	Ref (1.0)	1227/ 3048 (40%)	Ref (1.0)	Ref (1.0)
Positive	6/46 (13%)	0.66 (0.28,1.56)	0.57 (0.22,1.46)	24/45 (53%)	1.69 (0.94,3.06)	1.50 (0.82,2.72)
7-70	4/21 (19%)	1.03 (0.35, 3.07)	0.72 (0.21, 2.49)	11/20 (55%)	1.81 (0.75, 4.39)	1.48 (0.60, 3.68)
≥70	2/25 (8%)	0.62 (0.30, 1.27)	0.66 (0.32, 1.36)	13/25 (52%)	1.27(0.86, 1.88)	1.23 (0.83,1.82)

	C. Soft/slushy/watery faeces OR (95% CI) *N=3015*			D. Feeling sick /nauseous while/after eating OR (95% CI) *N=1562*		
	N (%)	OR (95% CI)	aOR (95% CI)[a]	N (%)	OR (95% CI)	aOR (95% CI)[a]
Negative	90/2971 (3%)	Ref (1.0)	Ref (1.0)	283/1535 (18%)	Ref (1.0)	Ref (1.0)
Positive	2/44 (5%)	2.26 (0.52, 3.03)	1.32 (0.55, 3.18)	4/27 (15%)	0.77 (0.26, 2.24)	0.61 (0.18, 2.06)
7-70	1/20 (5%)	1.69 (0.22, 12.7)	2.28 (0.29, 17.4)	2/13 (15%)	0.80 (0.18, 3.65)	0.39 (0.05, 3.06)
≥70	1/24 (4%)	1.18 (0.43, 3.23)	1.21 (0.44, 3.32)	2/14 (14%)	0.86 (0.41, 1.82)	0.91 (0.43, 1.94)

OR: Odds Ratio, aOR: adjusted Odds Ratio, 95% CI: 95% confidence Interval. [a] Adjusted for: ethnicity, income, sex

Table 3. Association between TG2A A level and Height (SDS), Weight (SDS) and BMI (SDS)

TG2A (U/ml)	Crude Model (β; 95% CI)	Multivariate Model ($_a$β; 95% CI)
	height (SDS)	**adjusted height**
Negative *(N=4249)*	Ref	ref
Positive (≥ 7)	-0.19 (-0.46,0.07)	**-0.29 (-0.55,-0.04)***
7-70 U/ml *(N=27)*	-0.14 (-0.52,0.25)	-0.22 (-0.59,0.14)
≥70 U/ml *(N=30)*	-0.25 (-0.61,0.12)	-0.31 (-0.65,0.04)
	weight (SDS)	**Adjusted weight (SDS)**
Negative *(N=4249)*	Ref	Ref
Positive (≥ 7)	**-0.35 (-0.62,-0.08)***	**-0.38 (-0.64,-0.12)****
7-70 U/ml *(N=27)*	-0.17 (-0.56,0.22)	-0.28 (-0.65,0.09)
≥70 U/ml *(N=30)*	**-0.51 (-0.88,-0.14)****	**-0.45 (-0.80,-0.10)***
	BMI	**adjusted BMI**
Negative *(N=4249)*	Ref	Ref
Positive (≥ 7)	**-0.30 (-0.54,-0.06)***	**-0.26 (-0.49,-0.03)***
7-70 U/ml *(N=27)*	-0.13 (-0.47,0.22)	-0.19 (-0.53, 0.14)
≥70 U/ml *(N=30)*	**-0.46 (-0.78,-0.13)****	**-0.33 (-0.65,-0.02)***

Values (β's) are based on linear regression models and reflect between group differences in mean height, weight and BMI SDS relative to reference group; 95% CI: 95% confidence interval. $_a$β; adjusted beta.
Multivariate Model: adjusted for child's age, sex, ethnicity, household income, birth weight, breastfeeding, timing of gluten introduction, and tv-watching. Models with weight as an outcome were additionally adjusted for height (SDS).
*P<0.05.
** P<0.01.

Table 4. Association between TG2A level and Bone Mass Density (SDS)

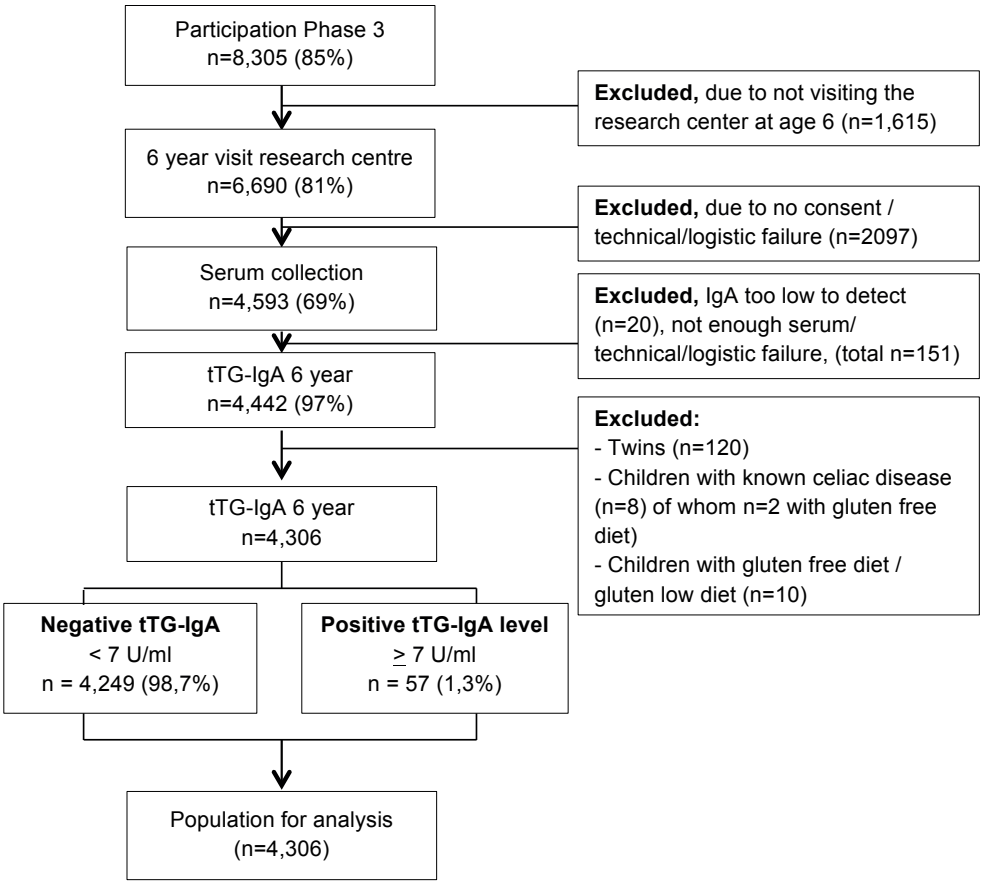
Anti-tTG (U/ml)	Crude Model (β; 95% CI)	Multivariate Model ($_a$β; 95% CI)
	Total BMD (SDS)	**Total BMD (SDS)**
Negative *(N=4249)*	Ref	Ref
Positive (≥ 7)	**-0.36 (-0.61,-0.11)****	**-0.26 (-0.47, -0.05)***
7-70 U/ml *(N=27)*	-0.14 (-0.49,0.22)	-0.11 (-0.41, 0.20)
≥70 U/ml *(N=30)*	**-0.56 (-0.91,-0.22)*****	**-0.40 (-0.70,-0.11)****
	BMD Spine (SDS)	**a BMD Spine (SDS)**
Negative *(N=4249)*	Ref	Ref
Positive (≥ 7)	**-0.41 (-0.66,-0.17)*****	**-0.35 (-0.56,-0.13)****
7-70 U/ml *(N=27)*	-0.11 (-0.45,0.24)	-0.13 (-0.44, 0.18)
≥70 U/ml *(N=30)*	**-0.70 (-1.04,-0.37)*****	**-0.55 (-0.85,-0.25)*****

Values (β's) are based on linear regression models and reflect between group differences in mean height, weight and BMI Standard Deviation Scores relative to reference group; 95% CI: 95% confidence interval. $_a$β; adjusted beta.
Multivariate Model: adjusted for child's age, sex, ethnicity (western/non-western), height (SDS), total body lean mass, total body fat mass, household income, breastfeeding, timing of gluten introduction, tv-watching, playing outside and vitamin D suppletion.
*P<0.05.
** P<0.01.
*** P<0.001.

REFERENCES

1. West J, Logan RF, Hill PG, Lloyd A, Lewis S, Hubbard R, Reader R, Holmes GK, Khaw KT. Seroprevalence, correlates, and characteristics of undetected coeliac disease in England. Gut 2003;52:960-5.
2. Maki M, Mustalahti K, Kokkonen J, Kulmala P, Haapalahti M, Karttunen T, Ilonen J, Laurila K, Dahlbom I, Hansson T, Hopfl P, Knip M. Prevalence of Celiac disease among children in Finland. N Engl J Med 2003;348:2517-24.
3. Bingley PJ, Williams AJ, Norcross AJ, Unsworth DJ, Lock RJ, Ness AR, Jones RW, Avon Longitudinal Study of P, Children Study T. Undiagnosed coeliac disease at age seven: population based prospective birth cohort study. BMJ 2004;328:322-3.
4. Green PH, Cellier C. Celiac disease. N Engl J Med 2007;357:1731-43.
5. Husby S, Koletzko S, Korponay-Szabo IR, Mearin ML, Phillips A, Shamir R, Troncone R, Giersiepen K, Branski D, Catassi C, Lelgeman M, Maki M, Ribes-Koninckx C, Ventura A, Zimmer KP, Diagnosis EWGoCD, Committee EG, European Society for Pediatric Gastroenterology H, Nutrition. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr 2012;54:136-60.
6. Ludvigsson JF, Leffler DA, Bai JC, Biagi F, Fasano A, Green PH, Hadjivassiliou M, Kaukinen K, Kelly CP, Leonard JN, Lundin KE, Murray JA, Sanders DS, Walker MM, Zingone F, Ciacci C. The Oslo definitions for coeliac disease and related terms. Gut 2013;62:43-52.
7. Cellier C, Flobert C, Cormier C, Roux C, Schmitz J. Severe osteopenia in symptom-free adults with a childhood diagnosis of coeliac disease. Lancet 2000;355:806.
8. Goddard CJ, Gillett HR. Complications of coeliac disease: are all patients at risk? Postgrad Med J 2006;82:705-12.
9. Mustalahti K, Collin P, Sievanen H, Salmi J, Maki M. Osteopenia in patients with clinically silent coeliac disease warrants screening. Lancet 1999;354:744-5.
10. Lucendo AJ, Garcia-Manzanares A. Bone mineral density in adult coeliac disease: An updated review. Rev Esp Enferm Dig 2013;105:154-162.
11. Mora S. Celiac disease: a bone perspective. J Pediatr Gastroenterol Nutr 2003;37:409-11.
12. Turner J, Pellerin G, Mager D. Prevalence of metabolic bone disease in children with celiac disease is independent of symptoms at diagnosis. J Pediatr Gastroenterol Nutr 2009;49:589-93.
13. Jaddoe VW, van Duijn CM, Franco OH, van der Heijden AJ, van Iizendoorn MH, de Jongste JC, van der Lugt A, Mackenbach JP, Moll HA, Raat H, Rivadeneira F, Steegers EA, Tiemeier H, Uitterlinden AG, Verhulst FC, Hofman A. The Generation R Study: design and cohort update 2012. Eur J Epidemiol 2012;27:739-56.
14. Jansen MAE, Tromp IIM, Kiefte-de Jong JC, Jaddoe VWV, Hofman A, Escher JC, Hooijkaas H, Moll HA. Infant feeding and anti-tissue transglutaminase antibody concentrations in the Generation R Study. The American Journal of Clinical Nutrition 2014; 100:1095-1101
15. Monsuur AJ, de Bakker PI, Zhernakova A, Pinto D, Verduijn W, Romanos J, Auricchio R, Lopez A, van Heel DA, Crusius JB, Wijmenga C. Effective detection of human leukocyte antigen risk alleles in celiac disease using tag single nucleotide polymorphisms. PLoS One 2008;3:e2270.
16. Kiefte-de Jong JC, Jaddoe VW, Uitterlinden AG, Steegers EA, Willemsen SP, Hofman A, Hooijkaas H, Moll HA. Levels of antibodies against tissue transglutaminase during pregnancy are associated with reduced fetal weight and birth weight. Gastroenterology 2013;144:726-735 e2.
17. Durmus B, Kruithof CJ, Gillman MH, Willemsen SP, Hofman A, Raat H, Eilers PH, Steegers EA, Jaddoe VW. Parental smoking during pregnancy, early growth, and risk of obesity in preschool children: the Generation R Study. Am J Clin Nutr 2011;94:164-71.
18. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ 2000;320:1240-3.

19. Rasquin A, Di Lorenzo C, Forbes D, Guiraldes E, Hyams JS, Staiano A, Walker LS. Childhood functional gastrointestinal disorders: child/adolescent. Gastroenterology 2006;130:1527-37.
20. Heppe DH,Medina-Gomez C, Hofman A, et al. Maternal first-trimester diet and childhood bone mass: the Generation R study. Am J Clin Nutr 2013;98:224-232
21. Leib ES, Lewiecki EM, Binkley N, Hamdy RC, International Society for Clinical D. Official positions of the International Society for Clinical Densitometry. J Clin Densitom 2004;7:1-6.
22. Swertz O, Duimelaar P, Thijssen J. Migrants in the Netherlands 2004. Voorburg/Heerlen: Statistics Netherlands, 2004.
23. Statistics Netherlands. Dutch Standard Classification of Education 2003. Voorburg/Heerlen: Statistics Netherlands, 2004.
24. Statistics Netherlands. Welfare in the Netherlands. Income, welfare and spending of households and individuals. Den Haag/Heerlen: Statistics Netherlands, 2012.
25. Greenland S, Mickey RM. Re: "The impact of confounder selection criteria on effect estimation. Am J Epidemiol 1989;130:1066.
26. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393.
27. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Stat Med 1991;10:585-98.
28. Balci TA, Koc ZP, Mitil HA. Bone mineral densitometry findings of children with newly diagnosed celiac disease. Mol Imaging Radionucl Ther 2011;20:59-62.
29. Jatla M, Zemel BS, Bierly P, Verma R. Bone mineral content deficits of the spine and whole body in children at time of diagnosis with celiac disease. J Pediatr Gastroenterol Nutr 2009;48:175-80.
30. Scotta MS, Salvatore S, Salvatoni A, De Amici M, Ghiringhelli D, Broggini M, Nespoli L. Bone mineralization and body composition in young patients with celiac disease. Am J Gastroenterol 1997;92:1331-4.
31. Tau C, Mautalen C, De Rosa S, Roca A, Valenzuela X. Bone mineral density in children with celiac disease. Effect of a Gluten-free diet. Eur J Clin Nutr 2006;60:358-63.
32. Krupa-Kozak U. Pathologic bone alterations in celiac disease: etiology, epidemiology, and treatment. Nutrition 2014;30:16-24
33. Zemel BS, Kalkwarf HJ, Gilsanz V, Lappe JM, Oberfield S, Shepherd JA, Frederick MM, Huang X, Lu M, Mahboubi S, Hangartner T, Winer KK. Revised reference curves for bone mineral content and areal bone mineral density according to age and sex for black and non-black children: results of the bone mineral density in childhood study. J Clin Endocrinol Metab 2011;96:3160-9
34. Wren TA, Kalkwarf HJ, Zemel BS, Lappe JM, Oberfield S, Shepherd JA, Winer KK, Gilsanz V, Bone Mineral Density in Childhood Study G. Longitudinal tracking of dual-energy X-ray absorptiometry bone measures over 6 years in children and adolescents: persistence of low bone mass to maturity. J Pediatr 2014;164:1280-5 e2.



Supplemental Figure 1. Flowchart of participants included in the analysis

Supplemental Table 1. Maternal, paternal and child characteristics according to anti-tTG levels observed data set (n=4306)

	Negative anti-tTG (<7 U/ml) n=4249 (99%)	Positive anti-tTG (≥ 7 U/ml) n=57 (1%)
Maternal		
Age (mean ±SD; years)	30.7 (5.1)	31.7 (4.6)
Missing (n; %)	0 (0%)	0 (0%)
Height (mean ±SD; cm)	167.7 (7.4)	168.1 (7.5)
Missing (n; %)	417 (10%)	11 (19%)
Weight **(**mean **±**SD; kg**)**	69.7 (13.0)	66.2 (11.0)
Missing (n; %)	417 (10%)	12 (21%)
BMI before pregnancy (mean; SD: kg/m^2)	23.6 (4.2)	22.3 (3.0)
Missing (n; %)	1093 (26%)	23 (40%)
Educational level (n;%)		
Low	1580 (43%)	17 (33%)
High	2094 (57%)	34 (67%)
Missing (n; %)	575 (14%)	6 (11%)
Household income (n;%)		
< 2000 Euro	828 (23%)	8 (16%)
≥ 2000 Euro	2836 (77%)	43 (84%)
Missing (n; %)	575 (14%)	6 (11%)
Gastrointestinal disease (n;%)		
None or no reported	3208 (99%)	**36 (97%)***
Celiac disease	1 (0%)	1 (3%)
Lactose intolerance	1 (0%)	0 (0%)
Inflammatory Bowel disease	4 (0.1%)	0 (0%)
Bowel complaints without known organic cause	13 (0.4%)	0 (0%)
Missing (n; %)	1011 (24%)	20 (35%)
Any reported autoimmune disease (diabetes mellitus, SLE, arthritis, MS or thyroid disorder) (n; %)	122 (3.0%)	2 (4.0%)
Unknown (n;%)	375 (9%)	8 (14%)
Smoking during pregnancy (n;%)	921 (25%)	11 (26%)
Missing (n; %)	559 (13%)	14 (25%)
Alcohol during pregnancy (n;%)	1871 (56%)	27 (69%)
Missing (n; %)	879 (21%)	21 (37%)
Nulliparous (n;%)	2253 (55%)	31 (55%)
Missing (n; %)	156 (4%)	1 (2%)
Paternal		
Age (mean ±SD; years)	33.3 (5.6)	33.4 (4.6)
Missing (n; %)	1336 (31%)	18 (32%)
Height (mean ±SD; cm)	182.2 (8.0)	181.9 (7.8)

Supplemental Table 1. Maternal, paternal and child characteristics according to anti-tTG levels observed data set (n=4306) (continued)

	Negative anti-tTG (<7 U/ml) n=4249 (99%)	Positive anti-tTG (≥ 7 U/ml) n=57 (1%)
Missing (n; %)	1341 (32%)	18 (32%)
Weight (mean ±SD; kg)	83.9 (13.0)	82.6 (13.6)
Missing (n; %)	1341 (32%)	19 (33%)
BMI before pregnancy (mean; SD: kg/m^2)	25.2 (3.4)	24.8 (3.8)
Missing (n; %)	1344 (32%)	19 (33%)
Child		
Male gender (n;%)	2201 (52%)	**19 (33%)****
Missing (n; %)	0 (0%)	0 (0%)
Western ethnicity (n; %)	2738 (64%)	**44 (77%)***
missing (n; %)	0 (0%)	0 (0%)
Celiac disease in family(n; %)	23 (0.8%)	**2 (4.5%)***
Missing (n;%)	1221 (29%)	13 (23%)
Estimated fetal growth (means ± SD: grams)		
Second trimester	383 ± 94	385 ± 95
missing	639 (15%)	12 (21%)
Third trimester	1631 ± 265	1627 ± 250
Missing (n; %)	519 (12%)	13 (23%)
Birth weight (mean ± SD: grams)	3437 ± 549	3466 ± 601.5
Missing (n; %)	8 (0.1%)	0 (0%)
Gestational age (mean ± SD: weeks)	39.84 ± 1.74	39.88 ± 1.41
Missing (n; %)	30 (0.7%)	1 (2%)
Ever Breastfeeding (n;%)	3091 (92%)	43 (98%)
Missing (n; %)	905 21%)	13 (23%)
Breastfeeding duration (median; range: months)	3.5 (0.1;30.0)	5.0 (0.5-12.0)
Missing (n; %)	188 (44%)	22 (39%)
Timing of introduction of gluten ≤6 months (n;%)	1240 (44%)	19 (48%)
Missing (n; %)	1412 (33%)	17 (30%)
HLA DQ2 or DQ8 present (n; %)	1147 (41%)	**38 (91%)*****
Missing (n; %)	1435 (34%)	15 (26.%)
Television watching ≥ 2 hours a day (n; %)	662 (20%)	11 (23%)
Missing	929 (22%)	10 (18%)
Playing outside ≥ 2 hours a day (n;%)	893 (29%)	10 (24%)
Missing	1208 (28%)	15 (26%)
Suppletion of vitamin D, AD or multivitamin (n;%)	928 (48%)	15 (58%)
Missing	2315 (54%)	31 (54%)
Age child at 6 year visit (median ± range; years)	6.02 (4.9; 9.1)	6.04 (5.6; 8.2)
missing (n; %)	0 (0%)	0 (0%)

Supplemental Table 1. Maternal, paternal and child characteristics according to anti-tTG levels observed data set (n=4306) (continued)

	Negative anti-tTG (<7 U/ml) n=4249 (99%)	Positive anti-tTG (≥ 7 U/ml) n=57 (1%)
Weight child at 6 year visit (median ± range; kg)	**22.6 (13.0; 56.0)**	**21.0 (16.0; 35.6)****
missing (n; %)	6 (<1%)	0 (0%)
Height child at 6 year visit (median ± range; cm)	119.3 (99.0; 148.3)	117.6 (1.06; 135.3)
missing (n; %)	6 (<1%)	**0 (0%)**
BMI (body mass index; median ± range; kg/ m^2)	**15.9 (11.9; 29.2)**	**15.4 (13.0; 19.4)****
missing (n; %)	6 (<1%)	0 (0%)
Overweight BMI >1.1-2.3 SDS (n;%)	307 (7.2%)	4 (7%)
missing (n; %)	6 (<1%)	0 (0%)
Obesity BMI >2.3 SDS (n;%)	142 (3.3%)	0 (0%)
missing (n; %)	6 (<1%)	0 (0%)

MS, multiple sclerosis; SLE systemic lupus erythematosus. *Significantly different from anti-tTG negative levels *p <0.05, **P <0.01, ***p<0.001

Supplemental Table 2. Linear association between tTG level and height, weight, BMI

	All tTG-IgA levels (SDS) 0.1 – 125 U/ml *(n=4306)*	Positive tTG levels (SDS) ≥ 7 U/ml (SDS) *(n=57)*
	(β; 95% CI)	**(β; 95% CI)**
Height		
Crude model	-0.02 (-0.06;0.01)	-0.02 (-0.07;0.04)
Adjusted Model	**-0.04 (-0.07;-0.00)***	-0.04 (-0.09;0.02)
Weight		
Crude model	**-0.05 (-0.08;-0.02)****	-0.05 (-0.10;0.01)
Adjusted Model	**-0.05 (-0.08;-0.02)****	**-0.05 (-0.09;-0.00)***
BMI		
Crude model	**-0.04 (-0.07;-0.01)****	**-0.04 (-0.09;0.00)***
Adjusted Model	**-0.04 (-0.07;-0.01)***	-0.04 (-0.09;0.02)

Note. Values are based on linear regression models and reflect differences and 95% CI in Height, weight and BMI (Standard Deviation Scores), for an increase in tTG-level (Standard Deviation Score). Models were adjusted for child's age, gender, ethnicity, Net household income, birth weight, breastfeeding, and tv-watching. Models with weight as an outcome were additionally adjusted for height (SDS). *p <0.05, **P <0.01

Supplemental Table 3. Linear association between tTG level and BMD, BMC, adjusted BMC and BA.

	All tTG-IgA levels (SDS) 0.1 – 125 U/ml *(n=4306)*	
	Model 1 (β; 95% CI)	**Model 2 (β; 95% CI)**
Total Body		
BMD	**-.05 (-.08; -.03)*** **	**-.04 (-.06;-.02)** **
BMC	**-.04 (-.07; -.02)*** **	**-0.02 (-0.03;-0.01)** **
Adjusted BMC	**-.02 (-.03;-.00)***	**-0.02(-0.03;-0.00)** **
BA	**-.03 (-.05;-.01)***	-0.01 (-0.02;0.00)
Spine		
BMD	**-.06 (-.09; -0.04)*** **	**-0.05 (-0.07;-0.02)*** **
BMC	**-.03 (-.05;-.00)***	-0.01 (-0.03;0.01)
Adjusted BMC	**-.03 (-.05;-.02)*** **	-0.01 (-0.02;-0.01)
BA	.00 (-.02;.03)	0.02 (-0.01; 0.04)

Note. BMD = BMC (g/cm)/=BA (g/cm^2) Values are based on linear regression models and reflect differences and 95% CI in Bone Mass Density, BMC and BA (SDS), for 1 SDS increase in tTG-level.
Model 1: adjusted for Height (SDS), models with aBMC as an outcome were additionally adjusted for BA (SDS).
Model 2: adjusted for child's age, gender, ethnicity, income, height at focus visit, total body lean mass, total body fat mass, birth weight, breastfeeding, timing of gluten introduction, vitamin D suppletion, tv-watching, playing outside. Models with adjusted BMC as an outcome were also adjusted for BA (SDS). *p <0.05, **P <0.01, ***p<0.001

Supplemental Table 4. Sensitivity analysis: Association between tTG-Immunloglobulin A level and height, weight, BMI and BMD in total population and in HLA DQ2/DQ8 population.

Total population (n=4306)		HLA DQ2/DQ8+ population (n=1185)	
Anti-tTG (U/ml)	**Multivariate Model ($_a\beta$; 95% CI)**	**Anti-tTG (U/ml)**	**Multivariate Model ($_a\beta$; 95% CI)**
	a Height		**a Height**
Negative *(N=4249)*	Ref	Negative *(N=1147)*	Ref
Positive (≥ 7)	**-0.29 (-0.55,-0.04)***	Positive (≥ 7)	**-0.45 (-0.77,-0.14)****
7-70 U/ml *(N=27)*	-0.22 (-0.59,0.14)	7-70 U/ml *(N=15)*	-0.27 (-0.77, 0.22)
≥70 U/ml *(N=30)*	-0.31 (-0.65,0.04)	≥70 U/ml *(N=23)*	**-0.57 (-0.98,-0.17)****
	a Weight (SDS)		**a Weight (SDS)**
Negative *(N=4249)*	Ref	Negative *(N=1147)*	Ref
Positive (≥ 7)	**-0.38 (-0.64,-0.12)****	Positive (≥ 7)	**-0.51 (-0.83,-0.20)****
7-70 U/ml *(N=27)*	-0.28 (-0.65,0.09)	7-70 U/ml *(N=15)*	-0.33 (-0.82,0.17)
≥70 U/ml *(N=30)*	**-0.45 (-0.80,-0.10)***	≥70 U/ml *(N=23)*	**-0.64 (-1.04, -0.23)****
	a BMI		**a BMI**
Negative *(N=4249)*	Ref	Negative *(N=1147)*	Ref
Positive (≥ 7)	**-0.26 (-0.49,-0.03)***	Positive (≥ 7)	**-0.31 (-0.60,-0.02)***
7-70 U/ml *(N=27)*	-0.19 (-0.53, 0.14)	7-70 U/ml *(N=15)*	-0.22 (-0.68,0.23)
≥70 U/ml *(N=30)*	**-0.33 (-0.65,-0.02)***	≥70 U/ml *(N=23)*	-0.36 (-0.73,0.01)
	BMD Total		**BMD Total**
Negative *(N=4249)*	Ref	Negative *(N=1147)*	Ref
Positive (≥ 7)	**-0.26 (-0.47, -0.05)***	Positive (≥ 7)	**-0.29 (-0.55,-0.03)***
7-70 U/ml *(N=27)*	-0.11 (-0.41, 0.20)	7-70 U/ml *(N=15)*	-0.15 (-0.55, 0.26)
≥70 U/ml *(N=30)*	**-0.40 (-0.70,-0.11)****	≥70 U/ml *(N=23)*	**-0.39 (-0.73, -0.06)***
	BMD Spine		**BMD Spine**
Negative *(N=4249)*	Ref	Negative *(N=1147)*	Ref
Positive (≥ 7)	**-0.35 (-0.56,-0.13)****	Positive (≥ 7)	**-0.46 (-0.72,-0.20)*****
7-70 U/ml *(N=27)*	-0.13 (-0.44, 0.18)	7-70 U/ml *(N=15)*	-0.34 (-0.74, 0.06)
≥70 U/ml *(N=30)*	**-0.55 (-0.85,-0.25)*****	≥70 U/ml *(N=23)*	**-0.55 (-0.88, -0.21)*****

Note. Values (β's) are based on linear regression models and reflect between group differences in mean height, weight, BMI and BMD Standard Deviation Scores relative to reference group; 95% CI: 95% confidence interval. $_a\beta$; adjusted beta.

Multivariate Models for height weight and BMI were adjusted for child's age, gender, ethnicity, household income, birth weight, breastfeeding, timing of gluten introduction, and tv-watching. Models with weight as an outcome were also adjusted for height (SDS). Multivariate Models for BMD were adjusted for child's age, gender, ethnicity, income, height at focus visit, total body lean mass, total body fat mass, birth weight, breastfeeding, timing of gluten introduction, vitamin D suppletion, tv-watching, playing outside. *P<0.05, ** p<0.01, ***p<0.001

Chapter 12

The Identification of Celiac Disease in Asymptomatic Children:

The Generation R Study

Michelle A.E. Jansen, Menno C. van Zelm, Michael Groeneweg, Vincent W.V. Jaddoe, Wim A. Dik, Marco W.J. Schreurs, Herbert Hooijkaas, Henriette A. Moll, Johanna C. Escher

J Gastroenterol. 2017 June 6.

ABSTRACT

Background

The objective of our study was to assess whether TG2A levels in the healthy childhood population can be predictive for subclinical CeD.

Methods

A total of 4442 children (median age 6.0 yrs) participating in a population-based prospective cohort study were screened on serum TG2A. Those with positive TG2A (≥7 U/ml; n=60, 1.4%) were invited for clinical evaluation (median age 9.0 yrs). Medical history, physical examination, serum TG2A and IgA-endomysium (EMA) were assessed, as well as HLA DQ 2.2/2.5/8 typing. Patients with positive serologies and genetic risk types underwent duodenal biopsies. TG2A levels at the time of biopsy were compared with the degree of enteropathy.

Results

Fifty-one TG2A-positive children were included in follow-up: 31 (60.8%) children had CeD, ten (19.6%) did not have CeD, and 10 (19.6%) were considered potential CeD cases because of inconclusive serologies. Duodenal biopsies were performed in 26/31 children. CeD with Marsh 3a/b enteropathy was observed in 75% (15/20) of children having TG2A levels ≥10ULN at 6 years of age, as well as in 75% (6/8) of children having a positive TG2A <10 ULN (OR 1.00; 95% CI 0.15-6.64). CeD cases had a lower BMI SDS (mean -0.49, SD 0.92) than children without CeD (mean 0.47, SD 1.37; p=0.02). No differences were observed in gastrointestinal symptoms.

Conclusions

Serum TG2A screening at 6 years of age in the healthy childhood population has a positive predictive value of 61% to detect subclinical CeD. We did not find a positive correlation between serum TG2A levels and the degree of enteropathy.

INTRODUCTION

Celiac Disease (CeD) is one of the most common, but largely underdiagnosed chronic diseases in childhood, and associated with excess morbidity and mortality.[1] The prevalence of CeD is 0.5-3% and increasing over time.[2-5] Despite increased awareness among physicians, active case-finding and screening strategies,[6-9] the majority of CeD patients still remain unrecognized in childhood. Diagnosis is difficult as recent insights have shown CeD can be present in asymptomatic children that do have a positive serology as well as enteropathy.[10]

CeD can be characterized by the presence of serum IgA against transglutaminase type 2 (TG2A), IgA against endomysium (EMA), genetic carriership for HLA DQ2.2, DQ2.5 and/or DQ8, and a gluten induced enteropathy, [11] which is graded according to the Marsh-Oberhuber criteria.[12-14] Serum TG2A and EMA positivities have sensitivities and specificities of >90% for detection of small intestinal enteropathy.[15] Screening of healthy populations frequently detects seropositive subjects without symptoms and with mild enteropathy (i.e., subclinical CeD), or seropositive subjects with normal small bowel mucosa (i.e., potential CeD). Thus, serology testing is only the first step in the diagnosis.[6] Nevertheless, several studies have shown that high TG2A levels ($\geq$10 times upper limit normal of the test; ULN) have a high specificity for severe enteropathy (Marsh 3) in symptomatic patients.[16-18] Therefore, the ESPGHAN guideline recommends that biopsies may be omitted when TG2A levels $\geq$10 ULN in symptomatic patients. It is still recommended to evaluate biopsies in asymptomatic or screening-identified individuals, irrespective of TG2A levels,[19] although some studies argue the need in subclinical and screening-detected patients with TG2A levels $\geq$10 ULN values.[20] Recently, one study found that TG2A levels correlate with the severity of mucosal lesions in both asymptomatic and symptomatic children.[21] Still, studies validating the correlation between TG2A levels and the degree of enteropathy in subclinical CeD are scarce.

Therefore, the primary aim of this study was to study the positive predictive value of TG2A screening to detect subclinical CeD in the healthy childhood population. A second aim was to study whether serum TG2A levels in subclinical CeD correlate with the degree of enteropathy.

MATERIALS AND METHODS

Design and screening strategy

This study was embedded within the Generation R study, a prospective population based cohort study from fetal life until young adulthood, described in detail previously. [22, 23] Children were born between April 2002 and January 2006 in Rotterdam, The Netherlands. At the age of 6 years, 6690 children visited the research center. During this visit, serum samples were collected from 4593 (69%) children and subsequently stored over a time period of 2.5 years from.[22] In 2013, after the inclusion of the whole cohort was completed

and children were 9 years old, these samples were thawed and analyzed for TG2A levels. Children and parents were not aware of TG2A determination. We excluded 20 children in whom total IgA concentrations were below the detection limit, possibly indicating IgA deficiency. Finally, 4442 children provided data on TG2A levels at 6 years of age.[24, 25] Between November and December 2013, when children were 9 years old, parents of 60 children were informed about a positive TG2A screening result at 6 years of age, and invited for diagnostic follow-up. Of these, eight children were lost to follow-up and one child was diagnosed with CeD prior to the initial serology screening at 5 years of age, and therefore excluded. Of the remaining 51 children, two children were on a gluten free diet (GFD), and had been diagnosed with CeD prior to the follow up (between 6 and 9 years of age). Therefore, these children were excluded from clinical retesting (but not from the final analyses). In total, 49 children received diagnostic follow-up within 3 months after initial notification of the elevated TG2A level at 9 years of age (January-April 2014). None of them consumed a low- or gluten free diet (GFD).

Written informed consent was obtained from all participants. Approval for the study was obtained from the Medical Ethical Committee of Erasmus MC, University Medical Centre Rotterdam, the Netherlands.

Screening on serum TG2A

TG2A serum levels were measured using a fluorescence enzyme immunoassay (EliA Celikey IgA, Phadia ImmunoCAP 250, Phadia AB, Uppsala, Sweden). The intra- and interassay coefficient of variation (CV) was below 10% and 15% respectively. Sera with a TG2A level of 7 U/ml or higher were considered to be positive per manufacturer's instructions. We further subdivided positive TG2A levels into two categories on the basis of below or above the 10 upper limit normal (ULN) of the test kit (>70 U/ml). [19, 24, 25]

Diagnostic Follow–up

Of 51 children included, two were diagnosed with CeD between 6 and 9 years of age, and one child received follow-up elsewhere [Figure 1]. To assess whether clinical symptoms were different between the CeD groups [Table 1], these three children were excluded from analyses (but not from the final analyses). From the remaining 48 children visiting the Sophia Children's Hospital at the median age of 9.9 years, a standardized medical and family history was taken by a pediatric gastroenterologist. The presence of gastrointestinal complaints was assessed systematically after parents were informed about the TG2A positivity at 6 years of age. Information on gluten intake was assessed by diet questions recorded at 6 years of age prior to TG2A screening, and at the moment of follow-up at 9 years of age, by asking whether their child had been on a low- or GFD. If not, the child was considered to be exposed to a normal amount of gluten. Height and weight were measured without shoes and heavy clothing, and body mass index (BMI; kg/m^2) was calculated. Age- and sex-adjusted standard deviation scores (SDS) were obtained using Dutch

reference growth curves.[26] Prior data on height and weight as recorded in the community health centers were included. A delay in linear growth over time was defined as decrease of >0.5 SD height-for-age relative to the normal SD line. The second TG2A test at 9 years of age was performed using the same fluorescence enzyme immunoassay (exact same kit). In addition, anti-endomysial antibody (anti-EMA) levels were measured, and children were genotyped for HLA DQ2.2, DQ2.5 and/or DQ8. Serological criteria for recommending an upper gastrointestinal (GI) endoscopy are summarized in Figure 1. All GI endoscopies were performed by two experienced pediatric gastroenterologists (MG or JCE) between March and December 2014. Mucosal biopsies were taken from both the proximal (including the bulb) and distal (2nd or 3rd) part of the duodenum as recommended.[27-29] Enteropathy was graded according to the Marsh-Oberhuber criteria,[12, 13] which can vary from intraepithelial lymphocytosis (IEL; ≥30 lymphocytes/100 enterocytes; Grade 1) to more extensive lesions including crypt hyperplasia (Grade 2), and various degrees of villous atrophy (Grade 3a partial; Grade 3b subtotal, Grade 3c total). Enteropathy was defined as having Marsh II or greater.[19]

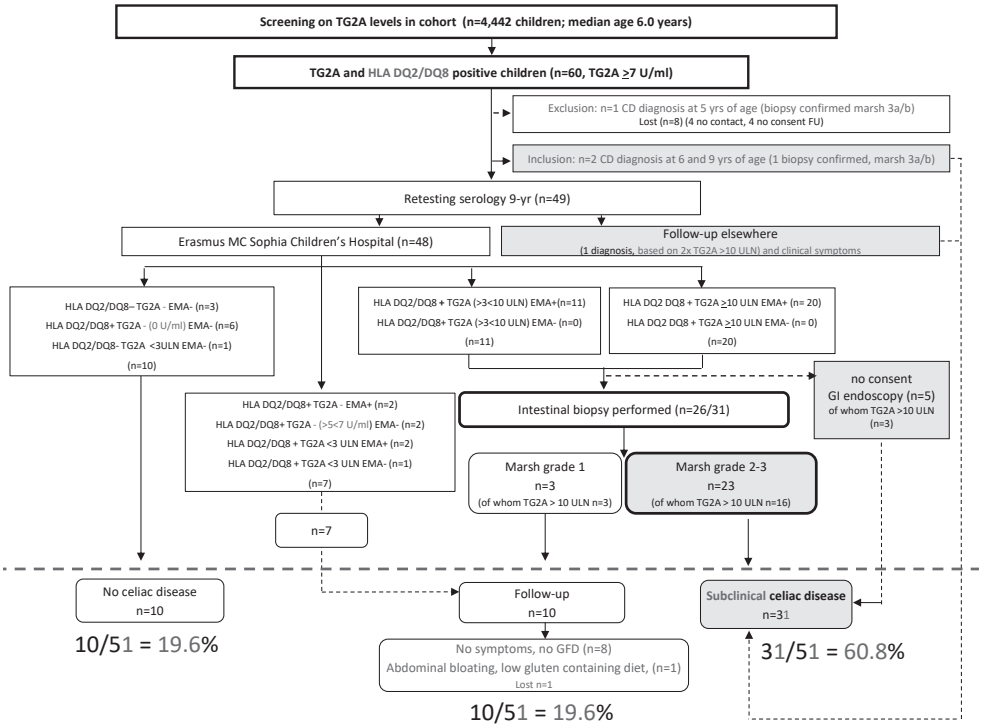


Figure 1. Outcome of screening for anti-tissue transglutaminase antibodies (TG2A) in a population based prospective cohort study.

Sixty children had a positive TG2A test result at 6 years of age. 31/51 (60.8%) children were considered to have CeD, 10/51 (19.6%) children needed follow-up, and 10/51 (19.6%) children did not have CeD.

Table 1. Characteristics according to Celiac Disease diagnosis (in tekst invoegen)

Characteristics	Diagnosis			p-value[b]
	No CeD (n=10)	Potential CeD (n=10)	Definitive CeD (n=28)	
Age at outpatient center (median; range; yrs)	**10.6 (9.9-11.4)**	9.1 (8.4-10.3)	**9.8 (8.6-11.2)**	**<0.01**
Female gender (n;%)	8 (80%)	7 (70%)	19 (68%)	0.48
Medical History				
Asymptomatic*	1 (10%)	3 (30%)	9 (32%)	0.17
No GI symptoms	1 (10%)	3 (30%)	10 (36%)	0.12
Abdominal pain (n;%)	7 (70%)	5 (50%)	16 (57%)	0.88
Constipation (n;%)	4 (40%)	1 (10%)	12 (43%)	0.88
Diarrhea (n; %)	4 (40%)	0 (0%)	6 (21%)	0.25
Nausea (n; %)	3 (30%)	1 (10%)	8 (29%)	0.93
Vomiting (n;%)	2 (20%)	0 (0%)	1 (4%)	0.10
≥2 gastrointestinal symptoms (n;%)	6 (60%)	1 (10%)	11 (39%)	0.26
≥3 gastrointestinal symptoms (n;%)	3 (30%)	0 (0%)	9 (32%)	0.90
Fatigue (n;%)	4 (40%)	1 (10%)	3 (11%)	0.04
GP visit for abdominal complaints (n;%)	2 (20%)	1 (10%)	0 (0%)	0.02
Family with CD (n;%)				
1st degree	0 (0%)	2 (20%)	1 (4%)	0.48
2nd degree	0 (0%)	1 (10%)	4 (14%)	
3rd degree	1 (10%)	0 (0%)	1 (4%)	
Physical examination				
Delayed linear growth curve 0-9 yrs (n;%)	1 (10%)	1 (10%)	3 (11%)	0.47
Missing data (n;%)	*1 (10%)*	*2 (20%)*	*8 (29%)*	
Height 9 years (median; range; cm)	147.2 (132.2-158.2)	135.0 (131.0- 153.5)	138.6 (105.0-164.2)	0.32
Weight 9 years (median; range; kg)	**36.8 (25.9-69.8)**	28.6 (27.1-50.4)	**29.3 (19.2-61.6)**	**0.04**
BMI (median; range; kg/m^2)	**17.5 (14.4-27.9)**	15.5 (15.1-21.4)	**15.4 (12.8-22.8)**	**0.04**
Height for age SDS (mean; SD)	-0.45 (1.30)	-0.09 (0.78)	-0.24 (1.05)	0.61
Weight for age SDS (mean; SD)	0.06 (1.68)	-0.20 (0.78)	-0.45 (0.99)	0.26
BMI for age SDS (mean; SD)	**0.47 (1.37)**	-0.11 (0.66)	**-0.49 (0.92)**	**0.02**

Abbreviations; CeD celiac disease; GI gastrointestinal; GP general practitioner; BMI Body Mass index; SDS standard deviation score adjusted for sex and age; TG2A Tissue transglutaminase type 2 antibody (IgA); ULN upper limit normal. Values represent means (SD's), medians (range), or numbers (percentages).

a Delayed linear growth was defined as: > -0.75-1.5 SDS decrease over time from 0-9 years of age.

b p-value reflects differences between biopsy proven (definitive CeD) and potential CeD group versus 'No CeD = reference' group (*Mann–Whitney U test*-tests were used for non-normally distributed variables, and $\chi 2$ tests were used to test for differences in proportions between groups).

*'Asymptomatic' refers to no GI symptoms, nor anorexia, fatigue, or irritability. None of the children was diagnosed with an autoimmune disease, including diabetes mellitus 1, or autoimmune thyroid disease. Thirty-one children were classified as definitive CeD cases, but only 28 were included in the analyses, because two children were diagnosed at an earlier age because of symptoms, and one child was diagnosed in another hospital (diagnosis was based on 2x TG2A >10 ULN in accordance with clinical symptoms), thus assessment of medical histories and physical examination may be different from the 28 asymptomatic children included in the analyses.

Endomysial antibodies

Endomysial antibodies (EmA) of IgA isotype were determined by indirect immunofluorescence using commercial monkey oesophagus slides, according to the manufacturer's instructions (Inova Diagnostics, San Diego, CA) [S1].

Genotyping of HLA DQ2.2, 2.5 and DQ8

Presence of CeD-associated HLA-DQ haplotypes DQ2.2 (DQA1*02/DQB1*02), DQ2.5 (DQA1*05/DQB1*02) and DQ8 (DQA1*0301/DQB1*0302) was determined by EUROArray, according to the manufacturer's instructions (Euroimmun AG, Lübeck, Germany) [Supplemental 1].[30]

Statistical analysis

Chi-square tests were used to test whether gastro-intestinal complaints were different between the children with and without final CeD diagnosis. The variables 'more than two gastrointestinal complaints' and 'more than three gastrointestinal complaints' were calculated by taking the sum of the following complaints: abdominal pain (y/n), constipation (y/n), diarrhoea (y/n), nausea (y/n) and vomiting (y/n), resulting in two dichotomous variables. Non-parametric Mann-Whitney U tests were used to test for differences in non-normally distributed variables, including median age, height, weight and BMI between children with and without CeD diagnosis. Chi-square tests were used to test whether clinical complaints and the degree of enteropathy were related to the TG2A categories. Binary logistic regression analyses were performed to test whether TG2A levels were related to CeD diagnosis (CeD versus no CeD, and potential CeD versus no CeD) and enteropathy (Marsh 3 versus Marsh 0, 1, or 2).

RESULTS

Subject characteristics

Of 4442 screened children (median age 6.0 years), 60 children (1.4%) had increased TG2A levels (≥7 U/ml), of whom 31 children had TG2A levels above 10 ULN (>70 U/ml). Of 60 children, eight children were lost to follow-up, and one child was diagnosed with CeD at 5 years of age (prior to the initial serology screening) and therefore excluded from analyses. Two children (4%) had received a CeD diagnosis in the period preceding retesting (between 6 and 9 years of age) and consumed a gluten free diet (GFD). Here, CeD was detected based on gastrointestinal symptoms by routine clinical care in the Netherlands, of whom one diagnosis was confirmed by biopsy (Marsh 3a/3b) [Figure 1].

Outcome of Follow-up

Of 48 children who were retested in the Sophia Children's Hospital, 31 (65%) were positive for TG2A and EMA, carried HLA DQ2.2, DQ2.5 or DQ8, and were advised to undergo gastrointestinal endoscopy. Of these, 20 had TG2A levels ≥10 ULN [Figure 1]. Of 26 children who underwent GI endoscopy, 20 children had Marsh grade 3 ((3a n=13), 3b (n=7)); three children had Marsh grade 2; and three children had Marsh grade 1. Of 23 children with Marsh grade 2-3 lesions, 16 children (70%) had TG2A levels ≥10 ULN. All three children with Marsh grade 1 lesions had TG2A levels ≥10 ULN and were advised to continue on a gluten containing diet, and to consider a second serology and/or intestinal biopsy when clinical complaints compatible with CeD occur. A total of 10 children were HLA DQ2.2, DQ2.5 and DQ8 negative, or carried the genetic risk type but lacked EMA and TG2A positivity, and were therefore considered not to have CeD. The second serology was not conclusive in seven children: they carried the genetic risk type, but had TG2A concentrations <3 ULN, or TG2A levels were negative in accordance with a positive EMA, whereas the initial screening TG2A test result at 6 years of age was positive. Furthermore, these children mentioned gastrointestinal complaints (abdominal pain), and had a first-degree family history of CeD (in contrast to the 10 children who were considered not to have CeD). Hence, these children were advised to be serologically retested in 6-12 months. During the period of follow-up (January 2014 - April 2017), none of the ten children received a CeD diagnosis; only one child occasionally mentioned complaints of abdominal bloating, and was on a low gluten containing diet (but not gluten free). The remaining children did not mention CeD associated symptoms while consuming a gluten containing diet.

In total, 31 of 51 (60.8%) children were considered to have CeD, of whom two children had developed symptomatic CeD within 3 years after the screening test was performed; 10/51 (19.6%) needed follow-up because of inconclusive serology (n=7) or negative intestinal biopsies (n=3, of whom all three had levels exceeding 10 ULN); and 10/51 (19.6%) children were considered to not have CeD [Figure 1]. Thus, the positive predictive value of TG2A screening was 61%, (95% CI 49-75), to detect subclinical CeD and 81%, (95% CI 70-91) to detect both potential and subclinical CeD.

Clinical symptoms

All 48 children consumed a normal gluten containing diet at 6 years of age, as well as at the visit to the pediatric gastroenterologist at 9-10 years. Of 28 CeD cases who provided data on medical history and physical examination at 9 years of age, nine (32%) children were truly asymptomatic relative to one (10%) of the children without CeD diagnosis (p=0.17). Abdominal pain was reported most frequently (57%), followed by constipation (43%), nausea (29%) and diarrhea (21%), but no significant differences in gastrointestinal symptoms were observed between children with and without CeD diagnosis (p=0.12) [Table 1]. In addition, no significant differences were observed in the presence of anorexia, irritability, food allergy, lactose intolerance, eczema, and absenteeism from school between children

with and without CeD diagnosis (data not shown), nor in family history of CeD [Table 1]. However, CeD cases weighed less (median 29.3 kg; range 19.2-61.6, p=0.04) than children who did not have CeD (median 36.8 kg; range 25.9-69.8). In addition, their BMI was lower (mean BMI SDS -0.49, SD 0.92, p=0.04) compared to children without CeD (mean BMI SDS 0.47, SD 1.37) No significant differences in height were observed between children with and without CeD diagnosis (p>0.32) [Table 1]. Associations were not substantially different in children without gastrointestinal symptoms [Supplemental Table 2].

In addition, no significant differences in gastrointestinal symptoms were observed between children who were TG2A negative (<7 U/ml) and strongly positive (≥10 ULN), but children having TG2A levels exceeding 10 ULN weighed less (median 29.1 kg, range 19.2-44.0) and had a lower BMI (mean BMI SDS -0.51, SD 0.85), than children who were TG2A negative (median 35.3 kg, range 25.9-69.8 and SDS BMI 0.21, SD 1.16 resp.) [Supplemental Table 3].

Development of CeD over time according to CeD diagnosis

We observed a high variability in serum TG2A levels at 6 and 9 years of age [Figure 2]. Of 29 CeD cases, 20 children (69%) had high TG2A levels (≥10 ULN) at 6 years of age, which

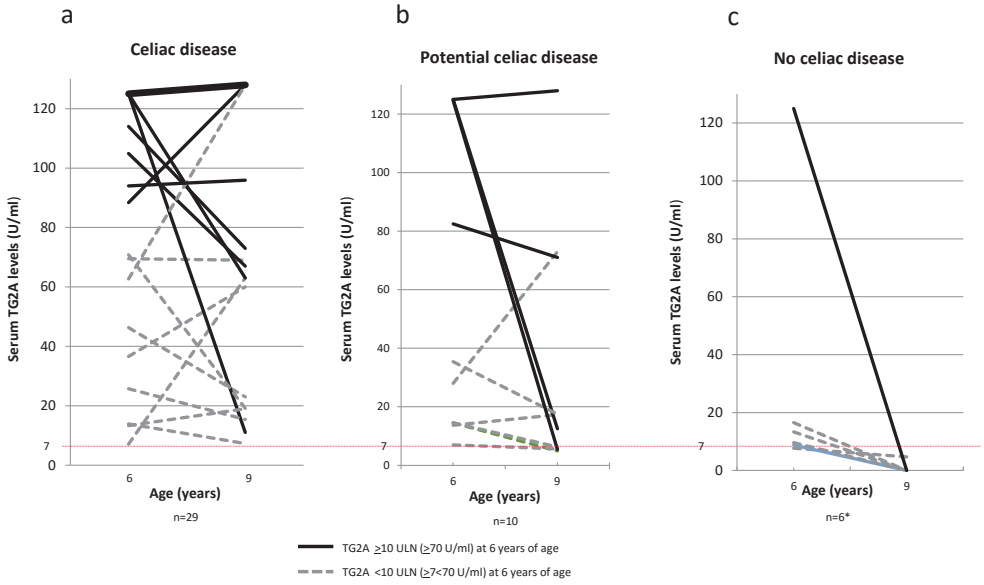


Figure 2. Serum TG2A concentrations at 6 and 9 years of age according to celiac disease diagnosis.
Thin dotted line indicates clinical cutoff for IgA-TG2A positivity (≥7 U/ml).
a, TG2A concentrations between 6-9 years in CeD group, bold line reflects n=14 CeD cases who had strong positive IgA-TG2A concentrations (≥125 U/ml at 6 years of age, and ≥128 U/ml at 9 years of age).
b, TG2A concentrations between 6-9 years in potential CeD group.
c, TG2A concentrations between 6-9 years in children lackes criteria for CeD diagnosis.
Four children were excluded because of negative genetic risk type. In addition, three children who received a CeD diagnosis between 6 and 9 years of age were excluded from this figure.

increased to a higher level in the majority (16/20) of cases at 9 years of age. Of ten children who needed follow-up at 9 years of age because of negative serologies, four children (40%) had TG2A concentrations exceeding 10 ULN levels at 6 years of age. Of 6 children who carried the genetic risk type, but did not have CeD, one child had TG2A levels exceeding 10 ULN at 6 years of age [**Figure 2**].

Association between level of TG2A concentration, CeD diagnosis and degree of enteropathy

A positive linear tendency was observed for TG2A levels in predicting CeD diagnosis: children having high TG2A levels (≥10 ULN) at 6 and 9 years of age were more frequently diagnosed with CeD than children having TG2A positive concentrations below <10 ULN ((OR 7.7; 95% CI 2.2-27.4) and (OR 10.0 (1.07-93.4) resp.) [Supplemental Table 4]. Second, we observed that TG2A levels at 6 or 9 years of age were not related to the presence of severe intestinal enteropathy (Marsh 3) [Supplemental Table 5; Figure 3].

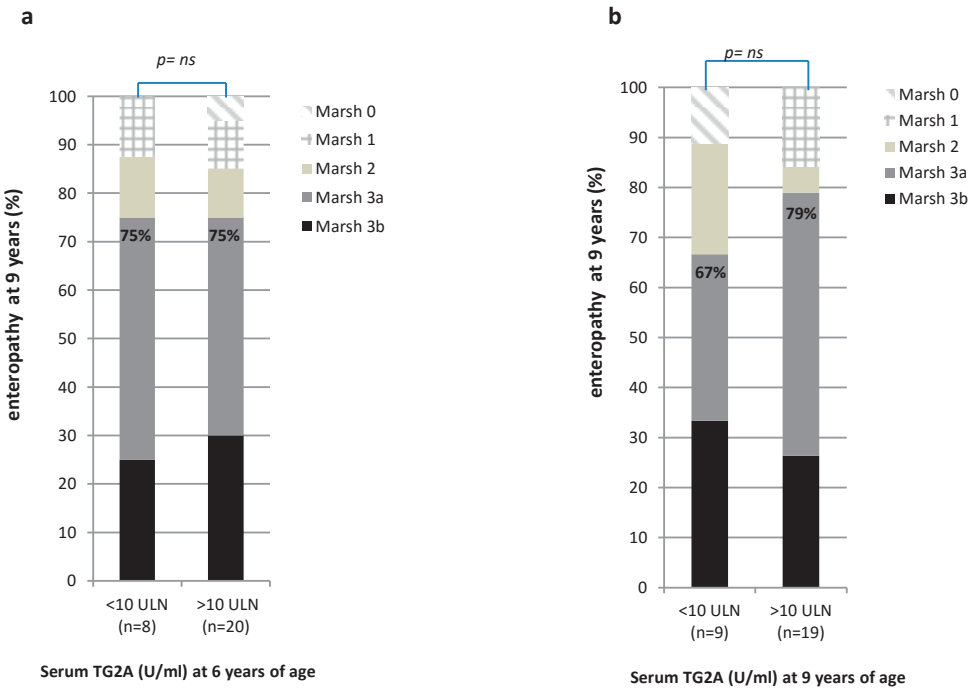


Figure 3. Association between serum IgA-TG2A concentrations and the degree of enteropathy in children with undiagnosed celiac disease.

a, TG2A concentrations at 6 years of age according to Marsh-Oberhuber classification.

b, TG2A concentrations at 9 years of age to Marsh-Oberhuber classification.

The three children who received a CeD diagnosis between 6 and 9 years of age were excluded from this figure.

Incidence of Celiac Disease in TG2A negative cohort

At the child's age of 9 years, parents were asked by questionnaire whether their child had ever received a doctors diagnosis of CeD. Of 4382 TG2A negative children at 6 years of age, nine (0.2%) received a doctors diagnosis of CeD between 6 and 9 years of age.

DISCUSSION

This population based prospective cohort study shows that: 1) Serum TG2A screening at 6 years of age in the healthy childhood population has a positive predictive value of 61% to detect subclinical CeD. 2) The level of serum TG2A in subclinical CeD is not predictive for the severity of enteropathy.

Although we followed the ESPGHAN guideline for CeD diagnosis in the majority of asymptomatic children, the PPV of TG2A screening may still have been influenced by misclassification of the outcome. According to the Oslo definitions,[10] *potential* CeD refers to children who are at increased risk of developing CeD as indicated by positive serology, but with a *normal* small intestinal bowel mucosa. In our study, 23 children were considered to have *subclinical* CeD on the basis of the combination of positive TG2A and the presence of enteropathy. However, five CeD cases did not consent for GI endoscopy, thus we lacked formal proof for diagnosis of CeD. Still, three of them had elevated levels >10 ULN, and they were all EmA and HLA DQ2 or 8 positive, suggesting that enteropathy is highly likely. Therefore, excluding these children would likely result in underdiagnosis. Second, loss to follow-up might have affected the precision of the PPV of TG2A screening. If children who were lost to follow-up (n=8) would not have CeD, then the initial PPV of screening on TG2A would have been 31/59 (53%). In contrast, if all children who were lost would be CeD cases, then the PPV would have been 39/59 (66%). Thus, the positive predictive value of serum TG2A screening to detect subclinical CeD in the general childhood population would be maximally 66%.

In contrast with several other studies, we did not find a positive association between serum TG2A levels and the degree of enteropathy.[16, 21, 31-33] However, these previous studies concerned children with gastrointestinal symptoms that triggered CeD testing, and contrast our cases of subclinical CeD, complicating a direct comparison between studies. In fact, it has been shown that TG2A levels are more strongly correlated with the severity of intestinal lesions in symptomatic patients, than in asymptomatic children.[21] Hence, the association between TG2A levels and the degree of enteropathy might be differentially influenced by the presence -or lack- of symptoms. In our study, the majority of children with high TG2A levels mentioned gastrointestinal symptoms. According to the ESPGHAN guideline, biopsies could have been omitted in this group. However, on the basis of the present study, it needs to be considered that children with TG2A ≥10 ULN in most cases needed a biopsy to confirm CeD. Moreover, our results confirm that the presence and

severity of gastrointestinal symptoms, such as abdominal pain, are insufficient to discriminate subclinical CeD cases from healthy children.[34] The majority (90%) of children who did not develop CeD mentioned gastrointestinal symptoms as well, such as abdominal pain, constipation, diarrhea, nausea or fatigue. Gastrointestinal symptoms were also not related to the degree of enteropathy, nor TG2A levels. Nonetheless, consistent with our previous findings,[25] subclinical children weighed less than healthy children. Interestingly, the TEDDY study demonstrated that the majority of subclinical children had normal weight and height growth at age 4,[21] which may suggest that the deviation in height and weight growth becomes more prominent between 4-6 years of age, and appears before gastrointestinal symptoms occur. Hence, increased clinical vigilance focusing on weight growth parameters, especially BMI,[35] or systematic growth monitoring in the general pediatric population, might improve early detection of CeD.[36] Nevertheless, the effects of early diagnosis, and the benefit of treatment with a gluten free diet of subclinical patients, needs to be further studied. In the short run, symptoms that may -or may not- have been recognized upon the start of a GFD may disappear, serving as a proof to parents,[37] but evidence in children is lacking. In addition, the benefits and cost-effectiveness of screening in asymptomatic individuals remains controversial.[38-40] Introducing a screening test in the general pediatric population with a positive predictive value of 61% to diagnose a disease with lifelong consequences (such as a gluten free diet, regular health care visits, screening for other diseases, screening and investigating family members, fear of complications, and economic burden from the gluten free diet) may not be reasonable. Therefore, it would be too premature to recommend screening in the general pediatric population.

Methodological considerations

The premise of the study was the identification of CeD in asymptomatic children. Three children in our study developed symptoms following screening, thereby triggering CeD diagnosis in routine clinical practice. One of the strengths of our study is that we used a prospective study design, with a similar protocol for the diagnostic process for all screened TG2A positive children. Furthermore, all screened TG2A positive children were of the same age, and analyzed 3 years later by the same laboratory and same test kit. Because all participants were unaware of TG2A determination, we were able to study the development of TG2A levels over 3 years of time. Ten children with negative TG2A did not have CeD, despite TG2A positivity at 6 years of age. Explanations for the contradictory test results at 6 and 9 years may be that TG2A levels were transiently high at 6 years, test results at 6 years were false positive (i.e. 1-specificity), or gluten intake decreased over time. Intriguingly, on top of the 1.4% of children with positive TG2A levels at 6 years of age (of whom 61% received a subclinical CeD diagnosis), 0.2% of TG2A negative children developed CeD within 3 years after the negative TG2A screening result at 6 years of age. The median age of retesting was 9.9 years, with a range between 8.4 and 11.4 years. This large age range is related to the inclusion period of 2.5 years. Another strength of our study is that TG2A measurements

were performed in combination with assessment of EMA, genetic risk types and biopsy specimens at the same time of second serology assessment. Last, all biopsies were evaluated by expert pathologists by using the Marsh-Oberhuber criteria to classify the degree of enteropathy. Nevertheless, several limitations should be taken into account as well. First of all, five children in the CeD group were considered to have subclinical CeD without biopsy results. Therefore, we did not fully meet the ESPGHAN criteria for CeD diagnosis. However, three of them had levels exceeding 10 ULN, they were all EMA positive and carried the genetic risk type. Hence, CeD diagnosis is highly likely. Second, symptoms were assessed only among those with TG2A positivity at 9 years of age, after parents were informed about the initial screening TG2A status which might have induced bias. However, the presence of gastrointestinal symptoms was evaluated systematically during routine clinical evaluation, preceding second serological (TG2A) assessment, additional EMA assessment, genetic risk typing, and gastrointestinal endoscopies at 9 years of age. Hence, symptom assessment was independent of the outcome, making bias less likely. Furthermore, symptoms were assessed by one experienced gastroenterologist minimizing inter-observer bias. We did not specifically ask for the presence of joint pains or neurological symptoms, but these complains were not reported on the general question " any complaints". Hence, the occurrence of joint pains and neurological symptoms seems unlikely.

Conclusion

Our results indicate that the majority of asymptomatic 6-year-old TG2A positive children developed subclinical CeD within 3 years following screening. We did not find a positive association between serum TG2A concentrations and the degree of enteropathy. Thus, TG2A screening at 6 years of age could substantially advance CeD diagnosis as children with subclinical CeD in our study only showed a small deviation in height and weight, and lacked gastrointestinal symptoms.

REFERENCES

1. Ludvigsson JF, Montgomery SM, Ekbom A, et al. Small-intestinal histopathology and mortality risk in celiac disease. JAMA. 2009;302(11):1171-8.
2. Ivarsson A, Myleus A, Norstrom F, et al. Prevalence of childhood celiac disease and changes in infant feeding. Pediatrics. 2013;131(3):e687-94..
3. Lohi S, Mustalahti K, Kaukinen K, et al. Increasing prevalence of coeliac disease over time. Aliment Pharmacol Ther. 2007;26(9):1217-25. 4. Rubio-Tapia A, Kyle RA, Kaplan EL, et al. Increased prevalence and mortality in undiagnosed celiac disease. Gastroenterology. 2009;137(1):88-93.
5. Ludvigsson JF, Rubio-Tapia A, van Dyke CT, et al. Increasing incidence of celiac disease in a North American population. Am J Gastroenterol. 2013;108(5):818-24.
6. Aggarwal S, Lebwohl B, Green PH. Screening for celiac disease in average-risk and high-risk populations. Therap Adv Gastroenterol. 2012;5(1):37-47.
7. Catassi C, Kryszak D, Louis-Jacques O, et al. Detection of Celiac disease in primary care: a multicenter case-finding study in North America. Am J Gastroenterol. 2007;102(7):1454-60.
8. Lebwohl B, Rubio-Tapia A, Assiri A, et al. Diagnosis of celiac disease. Gastrointest Endosc Clin N Am. 2012;22(4):661-77. Epub 2012/10/23.
9. Rubio-Tapia A, Hill ID, Kelly CP, et al. ACG clinical guidelines: diagnosis and management of celiac disease. Am J Gastroenterol. 2013;108(5):656-76; quiz 77.
10. Ludvigsson JF, Leffler DA, Bai JC, et al. The Oslo definitions for coeliac disease and related terms. Gut. 2013;62(1):43-52.
11. Green PH, Cellier C. Celiac disease. N Engl J Med. 2007;357(17):1731-43. Epub 2007/10/26.
12. Marsh MN. Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). Gastroenterology. 1992;102(1):330-54.
13. Oberhuber G, Granditsch G, Vogelsang H. The histopathology of coeliac disease: time for a standardized report scheme for pathologists. Eur J Gastroenterol Hepatol. 1999;11(10):1185-94.
14. Dickson BC, Streutker CJ, Chetty R. Coeliac disease: an update for pathologists. J Clin Pathol. 2006;59(10):1008-16.
15. Giersiepen K, Lelgemann M, Stuhldreher N, et al. Accuracy of diagnostic antibody tests for coeliac disease in children: summary of an evidence report. J Pediatr Gastroenterol Nutr. 2012;54(2):229-41.
16. Donaldson MR, Book LS, Leiferman KM, et al. Strongly positive tissue transglutaminase antibodies are associated with Marsh 3 histopathology in adult and pediatric celiac disease. J Clin Gastroenterol. 2008;42(3):256-60.
17. Mubarak A, Wolters VM, Gmelig-Meyling FH, et al. Tissue transglutaminase levels above 100 U/mL and celiac disease: a prospective study. World J Gastroenterol. 2012;18(32):4399-403.
18. Alessio MG, Tonutti E, Brusca I, et al. Correlation between IgA tissue transglutaminase antibody ratio and histological finding in celiac disease. J Pediatr Gastroenterol Nutr. 2012;55(1):44-9.
19. Husby S, Koletzko S, Korponay-Szabo IR, et al. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr. 2012;54(1):136-60.
20. Murray JA, Rashtak S, Rubio-Tapia A. Diagnosis of celiac disease in pediatric patients. Nat Rev Gastroenterol Hepatol. 2009;6(5):260-2.
21. Agardh D, Lee HS, Kurppa K, et al. Clinical features of celiac disease: a prospective birth cohort. Pediatrics. 2015;135(4):627-34.
22. Kruithof CJ, Kooijman MN, van Duijn CM, et al. The Generation R Study: Biobank update 2015. Eur J Epidemiol. 2014;29(12):911-27. Epub 2014/12/21.

23. Kooijman MN, Kruithof CJ, van Duijn CM, et al. The Generation R Study: design and cohort update 2017. Eur J Epidemiol. 2016;31(12):1243-64.
24. Jansen MA, Tromp, II, Kiefte-de Jong JC, et al. Infant feeding and anti-tissue transglutaminase antibody concentrations in the Generation R Study. Am J Clin Nutr. 2014;100(4):1095-101.
25. Jansen MA, Kiefte-de Jong JC, Gaillard R, et al. Growth Trajectories and Bone Mineral Density in Anti-Tissue Transglutaminase Antibody-positive Children: The Generation R Study. Clin Gastroenterol Hepatol. 2015;13(5):913-20 e5. Epub 2014/09/24.
26. Schonbeck Y, Talma H, van Dommelen P, et al. The world's tallest nation has stopped growing taller: the height of Dutch children from 1955 to 2009. Pediatr Res. 2013;73(3):371-7.
27. Rostom A, Murray JA, Kagnoff MF. American Gastroenterological Association (AGA) Institute technical review on the diagnosis and management of celiac disease. Gastroenterology. 2006;131(6):1981-2002. Epub 2006/11/08.
28. Hopper AD, Cross SS, Sanders DS. Patchy villous atrophy in adult patients with suspected gluten-sensitive enteropathy: is a multiple duodenal biopsy strategy appropriate? Endoscopy. 2008;40(3):219-24.
29. Latorre M, Lagana SM, Freedberg DE, et al. Endoscopic biopsy technique in the diagnosis of celiac disease: one bite or two? Gastrointest Endosc. 2015;81(5):1228-33. Epub 2015/02/02.
30. Mubarak A, Spierings E, Wolters V, et al. Human leukocyte antigen DQ2.2 and celiac disease. J Pediatr Gastroenterol Nutr. 2013;56(4):428-30.
31. Liu E, Bao F, Barriga K, et al. Fluctuating transglutaminase autoantibodies are related to histologic features of celiac disease. Clin Gastroenterol Hepatol. 2003;1(5):356-62.
32. Tursi A, Brandimarte G, Giorgetti GM. Prevalence of antitissue transglutaminase antibodies in different degrees of intestinal damage in celiac disease. J Clin Gastroenterol. 2003;36(3):219-21.
33. Webb C, Norstrom F, Myleus A, et al. Celiac Disease Can be Predicted by High Levels of Anti-Tissue Transglutaminase Antibodies in Population-Based Screening. J Pediatr Gastroenterol Nutr. 2015.
34. Rosen A, Sandstrom O, Carlsson A, et al. Usefulness of symptoms to screen for celiac disease. Pediatrics. 2014;133(2):211-8.
35. van Dommelen P, Grote FK, Oostdijk W, et al. Screening rules for growth to detect celiac disease: a case-control simulation study. BMC Pediatr. 2008;8:35..
36. Saari A, Harju S, Makitie O, et al. Systematic growth monitoring for the early detection of celiac disease in children. JAMA Pediatr. 2015;169(3):e1525.
37. Kurppa K, Paavola A, Collin P, et al. Benefits of a gluten-free diet for asymptomatic patients with serologic markers of celiac disease. Gastroenterology. 2014;147(3):610-7 e1.
38. Fasano A. Should we screen for coeliac disease? Yes. BMJ. 2009;339:b3592.
39. Korponay-Szabo IR, Szabados K, Pusztai J, et al. Population screening for coeliac disease in primary care by district nurses using a rapid antibody test: diagnostic accuracy and feasibility study. BMJ. 2007;335(7632):1244-7.
40. Evans KE, McAllister R, Sanders DS. Should we screen for coeliac disease? No. BMJ. 2009;339:b3674.

SUPPLEMENTAL METHODS

Endomysial antibodies

Endomysial antibodies (EmA) of IgA isotype were determined by indirect immunofluorescence using commercial monkey oesophagus slides, according to the manufacturer's instructions (Inova Diagnostics, San Diego, CA). Briefly, slides were incubated with 1:2 diluted human serum for 30 minutes at room temperature, followed by PBS wash and incubation with anti-human IgA-FITC (Inova Diagnostics) for 30 minutes at room temperature. Washed slides were embedded in glass cover slips and evaluated by fluorescence microscopy by two independent observers.

Genotyping of HLA DQ2.2, 2.5 and DQ8

Presence of CeD-associated HLA-DQ haplotypes DQ2.2 (DQA1*02/DQB1*02), DQ2.5 (DQA1*05/DQB1*02) and DQ8 (DQA1*0301/DQB1*0302) was determined by EUROArray, according to the manufacturer's instructions (Euroimmun AG, Lübeck, Germany).[30] Briefly, HLA-DQA1 and HLA-DQB1 gene fragments were PCR amplified in parallel from genomic DNA, isolated from whole blood using standard procedure. Fluorescent PCR products were subsequently hybridized on a microarray containing probe spots specific for different HLA-DQ variants (HLA-DQA1: *02, *02/*0301, *03, *0302/03 and *05; HLA-DQB1: *02 and *02/*0302). Detection of specifically hybridized PCR products was analyzed by the Euroimmun microarray scanner, and genotypes and diagnostic findings were automatically deduced by the EUROArrayScan software.

Supplemental Table 2. Growth characteristics according to diagnosis stratified on the presence of symptoms

	Diagnosis				
	No CeD (n=10)	Potential CeD (n=10)	CeD (n=28)		p-value[b]
Characteristics			No GI symptoms (n=10)	GI symptoms (n=18)	
Delayed linear growth 0-9 yrs (n;%)[a]	1 (10%)	1 (10%)	1 (10%)	2 (11%)	0.66
Missing data (n;%)	*1 (10%)*	*2 (20%)*	*1 (10%)*	*7 (39%)*	
Height 9 years (median; range; cm)	147.2 (132.2-158.2)	135.0 (131.0-153.5)	140.0 (128.0-152.2)	137.8 (105.0-164.2)	0.31
Weight 9 years (median; range; kg)	**36.8 (25.9-69.8)**	28.6 (27.1-50.4)	31.4 (25.0-36.8)	**28.9 (19.2-61.6)***	**0.04**
BMI (median; range; kg/m^2)	17.5 (14.4-27.9)	15.5 (15.1-21.4)	15.5 (14.4-18.5)	15.4 (12.8-22.8)	0.08
Height for age and sex SDS (mean; SD)	-0.45 (1.30)	-0.09 (0.78)	-0.47 (1.05)	-0.11 (1.05)	0.46
Weight for age and sex SDS (mean; SD)	0.06 (1.68)	-0.20 (0.78)	-0.58 (0.94)	-0.39 (1.05)	0.38
BMI for age and sex SDS (mean; SD)	**0.47 (1.37)**	-0.11 (0.66)	-0.51 (0.80)	**-0.48 (1.01)***	**0.04**

Abbreviations; CeD celiac disease; GI gastrointestinal; BMI Body Mass index; SDS standard deviation score adjusted for sex and age.

a Delayed linear growth was defined as: ≥ -0.75-1.5 SDS decrease over time from 0-9 years of age. Values represent means (SD's), medians (range), or numbers (percentages).

b p-value reflects differences between symptomatic CeD group versus no CeD group (MW-test were used for non-normally distributed variables, and $\chi 2$ tests were used to test for differences in proportions between groups) 'No GI symptoms' includes: No abdominal pain, constipation, diarrhea, nausea or vomiting. None of the children was diagnosed with an autoimmune disease, including diabetes mellitus 1, or autoimmune thyroid disease. n=32 children were classified as CeD cases, but only 28 were included in the analyses, because 4 children were diagnosed at an earlier age, or in other hospital, thus assessment of medical histories and physical examination might be different from the 28 children included in the analyses.

Supplemental Table 3. Characteristics according to TG2A levels

n=48 Characteristics	TG2A levels at 9 years of age			p-value[b]
	TG2A negative (<7 U/ml) (n=13)	TG2A <10 ULN (7<70 U/ml) (n=15)	TG2A ≥10 ULN (>70 U/ml) (n=20)	
Age at outpatient center (median; range; yrs)	10.3 (8.4-11.4)	9.7 (8.8-10.9)	9.9 (8.6-11.2)	0.30
Female gender (n;%)	10 (77%)	10 (67%)	14 (70%)	0.82
Medical History				
Asymptomatic*	2 (15%)	4 (27%)	7 (35%)	0.38
No GI symptoms	2 (15%)	4 (22%)	8 (38%)	0.49
Abdominal pain (n;%)	8 (62%)	10 (67%)	10 (50%)	0.59
Constipation (n;%)	3 (23%)	7 (47%)	7 (35%)	0.43
Diarrhea (n; %)	4 (31%)	3 (20%)	3 (15%)	0.55
Nausea (n; %)	3 (23%)	3 (20%)	6 (30%)	0.78
Vomiting (n;%)	2 (15%)	0 (0%)	1 (5%)	0.23
≥2 gastrointestinal symptoms (n;%)	5 (39%)	6 (40%)	7 (35%)	0.95
≥3 gastrointestinal symptoms (n;%)	3 (23%)	4 (27%)	5 (25%)	0.98
Food allergy (n;%)	0 (0%)	1 (7%)	1 (5%)	0.66
Anorexia (n;%)	1 (8%)	1 (7%)	2 (20%)	0.93
Fatigue (n;%)	4 (31%)	3 (20%)	1 (5%)	0.28
Irritability (n;%)	0 (0%)	0 (0%)	0 (0%)	NA
Lactose intolerance (n;%)	0 (0%)	0 (0%)	1 (5%)	0.49
Eczema (n;%)	1 (8%)	3 (20%)	1 (5%)	0.33
GP visit for abdominal complaints (n;%)	3 (23%)	0 (0%)	0 (0%)	0.01
Absenteeism from school (n;%)	1 (8%)	0 (0%)	1 (5%)	0.55
Family with CeD (n;%)				
1st degree	2 (15%)	1 (7%)	0 (0%)	
2nd degree	1 (8%)	1 (7%)	3 (15%)	0.61
3rd degree	1 (8%)	0 (0%)	1 (5%)	
Physical examination				
Delayed height growth curve 0-9 yrs (n;%) *Missing data*	1 (8%) *2 (15%)*	3 (20%) *3 (20%)*	1 (5%) *6 (30%)*	0.54
Height 9 years (median; range; cm)	144.9 (132.2-158.2)	137.8 (131.0-164.2)	138.2 (105.0-162.8)	0.32
Weight 9 years (median; range; kg)	**35.3 (25.9-69.8)**	29.6 (27.1-61.6)	**29.1 (19.2-44.0)**	**0.05**
BMI (median range; kg/m^2)	**16.6 (14.4-27.9)**	15.5 (13.3-24.2)	**15.4 (12.8-17.5)**	**0.08**
Height for age SDS (mean; SD)	-0.24 (1.26)	-0.03 (0.85)	-0.43 (1.03)	0.63
Weight for age SDS (mean; SD)	0.00 (1.46)	-0.16 (1.03)	-0.59 (0.92)	0.17
BMI for age SDS (mean; SD)	**0.21 (1.16)**	-0.18 (1.10)	**-0.51 (0.85)**	**0.047**

Abbreviations; CeD celiac disease; GI gastrointestinal; GP general practitioner; BMI Body Mass index; SDS standard deviation score adjusted for sex and age; TG2A Tissue transglutaminase type 2 antibody (IgA); ULN upper limit normal. Values represent means (SD's), medians (range), or numbers (percentages).
a Delayed linear growth was defined as: > -0.75-1.5 SDS decrease over time from 0-9 years of age.
b p-value reflects differences between TG2A >10 ULN group versus TG2A negative group (reference-group). *Mann–Whitney U* tests were used for non-normally distributed variables, and $\chi 2$ tests were used to test for differences in proportions between the 2 groups. None of the children was diagnosed with an autoimmune disease, including diabetes mellitus 1, or autoimmune thyroid disease.
*'Asymptomatic' refers to no GI symptoms, nor anorexia, fatigue, or irritability.

Supplemental Table 4. Association between TG2A levels and celiac disease diagnosis.

n=32	Number of children per TG2A group	CeD diagnosis n (%)	OR for CeD diagnosis (95% CI)	p-value
TG2A at 6 years of age				
<10 ULN (<70 U/ml)	29	9/24 (38%)	*reference*	
≥10 ULN (≥70 U/ml)	31	23/28 82%)	**7.7 (2.2; 27.4)**	**0.002**
Per Unit/ml increase	60	32	**1.03 (1.01; 1.04)**	***<0.001***
TG2A at 9 years of age				
Negative (TG2A <7U/ml)	13	-	-	
<10 ULN (<70 U/ml)	18	12/18 (67%)	*reference*	
≥10 ULN (≥70 U/ml)	21	20/21 (95%)	**10.0 (1.07; 93.4)**	***0.04***
Per Unit/ml increase	48	32/52 (62%)	**1.05 (1.02; 1.08)**	***<0.001***

Abbreviations; CeD celiac disease; TG2A Tissue transglutaminase type 2 antibody (IgA); ULN upper limit normal. OR: Odds ratio; 95% CI: 95% confidence interval. OR's are derived from binary logistic regression models, with CeD diagnosis as dependent variable (y/n), and TG2A levels as (dichotomous and continuous) independent variable.

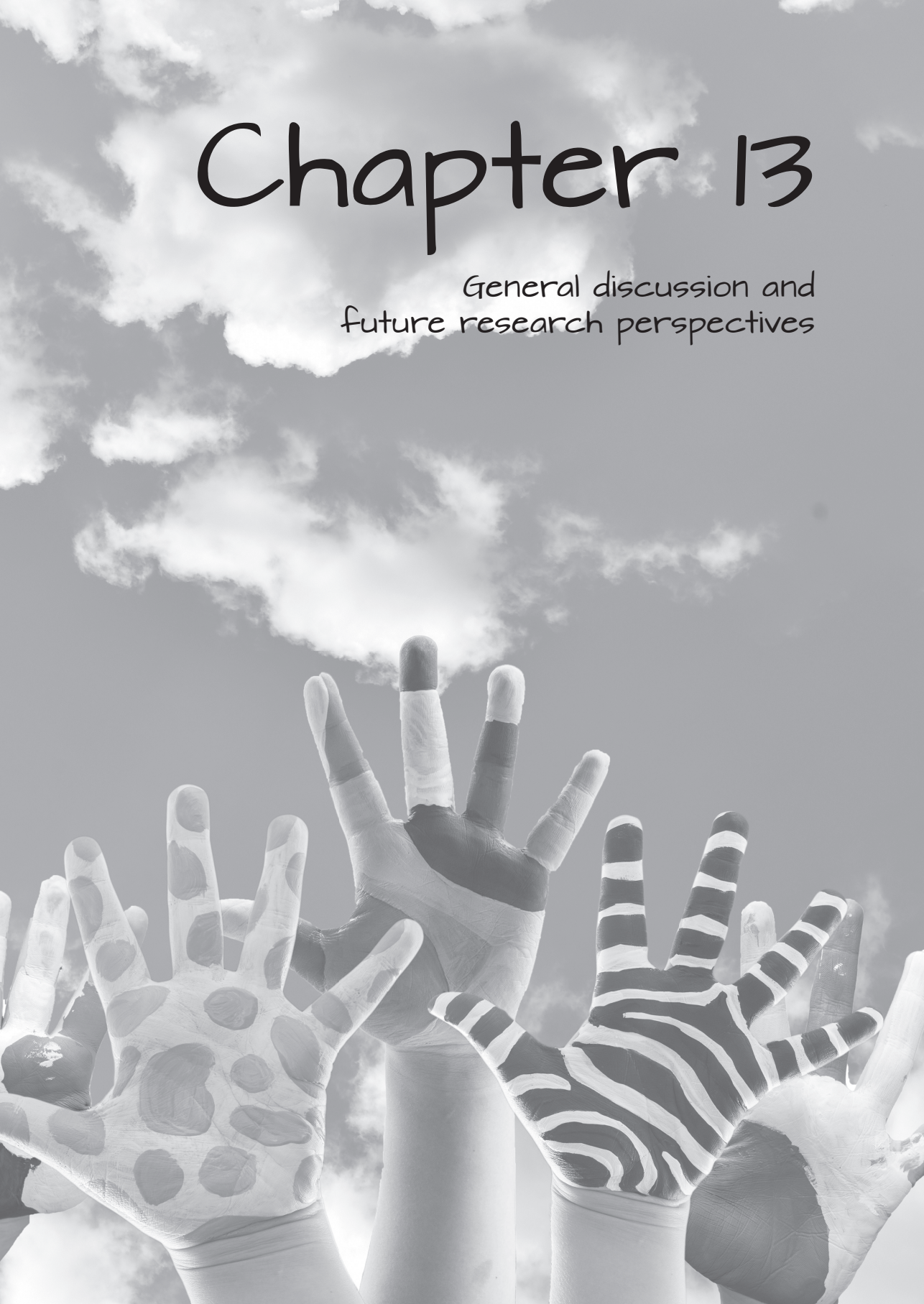
Supplemental Table 5. Association between TG2A levels and Marsh 3 enteropathy.

n=28	Number of biopsies per TG2A group n (%)	Marsh 3 biopsies n (%)	OR for Marsh 3 enteropathy (95% CI)	p-value
TG2A at 6 years of age				
<10 ULN (<70 U/ml)	8/29	6/8 (75%)	*reference*	
≥10 ULN (≥70 U/ml)	20/31	15/20 (75%)	1.00 (0.15; 6.64)	***0.99***
Per Unit/ml increase	28/60	21/28	1.00 (0.98; 1.02)	***0.98***
TG2A at 9 years of age				
<10 ULN (<70 U/ml)	9/18	6/9 (67%)	*reference*	
≥10 ULN (≥70 U/ml)	19/20	15/19 (79%)	3.33 (0.56, 19.59)	***0.18***
Per Unit/ml increase	28/48	21/28	1.02 (1.00; 1.04)	***0.09***

Abbreviations; TG2A Tissue transglutaminase type 2 antibody (IgA); ULN upper limit normal. OR: Odds ratio; 95% CI: 95% confidence interval. ORs are derived from binary logistic regression models, with Marsh 3 enteropathy relative to Marsh 2, 1 or 0) as dependent variable, and TG2A levels as (dichotomous and continuous) independent variable.

Chapter 13

General discussion and future research perspectives



GENERAL DISCUSSION

A. RATIONALE FOR STUDYING ACQUISITION OF IMMUNE COMPETENCE FROM A CLINICAL PERSPECTIVE

Celiac disease (CeD) is a multisystem immune-mediated disorder, characterized by an excessive inflammatory "Th1" response against dietary gluten in genetically susceptible individuals (i.e. HLA-DQ2/ HLA-DQ8 carriers).[1-3] The prevalence of CeD is 1-3% in the general population, while the HLA-DQ2/ HLA-DQ8 haplotypes are present in >25% of the general population, suggesting that, besides genetic predisposition, additional factors must contribute to disease initiation.[1, 2, 4-6] Still, it remains unclear why exposure to dietary gluten in one genetically susceptible person results in an excessive immune reaction, yet not in another exposed genetically susceptible person. Apart from gluten and the role of dietary factors early in life, such as the duration and exclusiveness of breastfeeding, the search for environmental determinants of CeD has reached limited success.[1, 7] Furthermore, the search has been complicated by the broad heterogeneity of the disease, including subclinical and potential forms,[8, 9] raising further questions, such as: What is the clinical course of these undiagnosed forms? Are these clinically relevant? Do these forms cause detrimental health effects, and should these forms be treated to avoid deleterious effects?

Understanding normal acquisition of immune competence in childhood, and its main driving forces, may help to better understand disorders such as CeD, as well as other forms of immune dysfunction (either unwanted responses (i.e. autoimmunity, atopic diseases) or insufficient responses, i.e. immunodeficiencies or premature immunosenescence). Many research focuses on how and when aberration results in disease, and on the mechanisms that operate to prevent these unenviable aberrations.[10] Therefore, it may be helpful to better understand normal processes of immune maturation. This vision captures the principles of the Barker's hypothesis, or Developmental Origin of Health and Disease (DOHad)[11] that posits that environmental exposures in early life (pregnancy, neonatal period, and early childhood) may present a "window of immune vulnerability" for dysregulation and development of immune-mediated diseases later in life. During early childhood, exposure to infectious and non-infectious particles induces immune responses that shape immune homeostasis and induce the formation of long-term immunity. However, which determinants define this early childhood immune homeostasis? And to which extent do these early life exposures leave a lasting "fingerprint" on the immune system? And is this "fingerprint" associated with beneficial or potentially detrimental effects later in life?

This thesis is composed of research conducted in a population-based prospective cohort study with two main aims: 1. Identification of determinants of the maturing immune system in childhood, focusing predominantly on adaptive immunity; 2. To study determinants and consequences of celiac disease autoimmunity. In this General Discussion chapter, the

main findings of the research are placed in a broader context, and methodological aspects are discussed. In the following paragraphs, the effects of various pre-, peri- and postnatal determinants on immune maturation will be discussed, with a predominant focus on breastfeeding and herpesvirus seropositivity. Based on the results presented in this thesis, the crucial role of timing of these external determinants will be discussed; i.e. whether a window of "immune vulnerability" exists, whether immune effects are dynamic, temporary or persistent, and which factors are associated with immune-mediated diseases, such as celiac disease autoimmunity (CDA).

B. WHAT IS "NORMAL" IMMUNE MATURATION DURING EARLY CHILDHOOD?

B.1 "Normal" immune maturation is a broad concept influenced by age, sex and various external factors

Which factors drive immune homeostasis? This question remains difficult to explain as blood and tissue cell numbers are affected by many factors that are connected, and still not entirely elucidated.[12, 13] For example, the numbers of circulating leukocytes depend on their production from primary lymphoid tissue (bone marrow, thymus), the in- and efflux from circulation to lymphoid and peripheral tissues, such as intestines, lungs and skin, proliferation in immune responses and cell death (Figure 1).[13-15] Therefore, direct translation to biological functions such as immune homeostasis, production, recirculation,

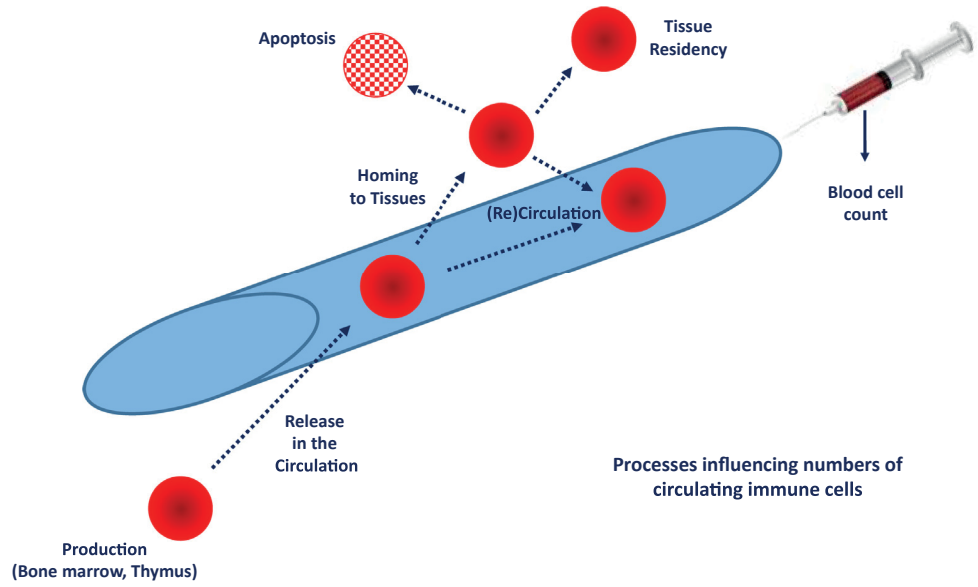


Figure 1. Schematic representation of the different mechanims that influence blood immune cell counts. Adapted from van Lier et al., J Allergy Clin Immunol 2017; 139:1793-4

extravascular homing, persistence and cell death remains difficult.[13] These functional and connecting immunological processes may in part be delineated by using mouse models. However, mouse models have limitations as well, therefore some have advocated convincingly that we should return to concentrate on human studies.[15-17] We performed our studies within the framework of a large-scale cohort of healthy children, which provided us with the unique opportunity to observe populations prospectively, and the ability to delineate the factors that drive immune maturation over a large period of time. The studies described in **Part I** were performed during early childhood (birth to 6 years) when the immune cell composition is most dynamic.[18-21] This unique approach enabled assessment of the relative contributions of various pre-, peri and postnatal determinants (**Chapter 2**). Our results were in line with previous studies estimating that non-genetic factors account for ~50-75% of inter-individual immunological variation,[16, 22-24] of which the most profound effect can be attributed to aging.[16, 25] In early childhood, the strongest fluctuation in cell numbers is observed shortly after birth, in the first 6 months of life, followed by a relative stabilization after 5-6 years of age (**Chapter 2**). Which factors drive these robust *age-associated changes* in immune equilibrium?

First, transient immunological fluctuations shortly after birth reflect the newborns sudden, and robust adjustment from the relatively protective uterus to the extra-uterine environment, a period of quickly changing exposures and colonization by various harmful and non-harmful pathogens. Newborns are more vulnerable to infections than adults; this is not entirely understood, but suspected to be in part explained by the lack of acquired immunity, as well as by the prolonged "feto-maternal tolerance"[26] that may persist after delivery.[10, 27-29] However, this immunological priming may not be detrimental, it may also challenge the neonate's immature immune system to 'learn' to balance between tolerating non-harmful antigens and minimizing potential harmful inflammation.[10] Furthermore, in the first months of life, maternal derived antibodies (esp. sIgA from breastmilk) are likely to provide passive humoral protection, thereby supporting the child's innate immunity to provide protection from invading pathogens (**Chapter 3,** Supplemental figure 1). Breastfeeding induces a transient skewing towards cellular immunity, whilst the formation of memory B cells is reduced (**Chapter 3**), possibly in part through the passive protection of maternal immunoglobulins. Hence, a longer breastfeeding duration may not only affect local B cell, but also systemic B-cell responses. Furthermore, breastfeeding may prime the T cell compartment (**Chapter 3**), which may further explain why breastfed children respond better to vaccinations than non-breastfed children.[30] Furthermore, T cell maturation, at the expense of B cell maturation, may prove beneficial for infections with persistent cytomegalovirus (CMV) that require memory T cell responses to achieve long-term suppression (**Chapters 2, 3 and 4**).

A second determinant that drives age-related changes in immune cell composition is postnatal infections. Various pathogens profoundly determine the number and differentiation state of memory cells.[13, 31-34] Of all pre-, peri- and postnatal determinants, chronic

infections with cytomegalovirus (CMV) and Epstein Barr virus (EBV) were the strongest determinants for CD4+ and CD8+ T-cell maturation (**Chapters 2 and 4**). Our observations are in line with previous thoughts that variation in immune cell composition between individuals is largely driven by non-heritable factors, of which the most profound effect has been attributed to aging and CMV.[23] It remains unclear though, why CMV infection in children was predominantly associated with enlargement of late CD45RA+ memory T cells, whereas EBV was predominantly associated with enlargement of early CD45RO+ effector memory T cells. Explanations may include that these distinct viruses need different antiviral responses to clear and suppress each virus,[33] and functional properties of phenotypically distinct memory CD8+ T cells are different with respect to proliferative potential, cytotoxicity, polyfunctionality, avidity, and molecule production upon stimulation.[15, 33-39] It may also be that the dosage and timing of the initial anti-viral response defines the phenotype and functional characteristics of the CD8+ T cells. In **Chapter 4** we observed higher CMV than EBV seroprevalence in 2-year-old children, which may suggest that CMV infection occurs at a relatively younger age than EBV infection due to vertical transmission. Hence, differences in the inflammatory environment at the timing of infection, and route of primary herpesvirus transmission may further explain our findings.[15, 40] For example, growth factors and cytokines (for example IL-15) in breastmilk may stimulate CD8+ T cell functions, reduce intestinal permeability, or activate the local intestinal epithelial immune response against infectious pathogens.[41-43] Intriguingly, the major factor that explains age-dependent inter-individual immunological variation (and convergence) in adults was recently attributed to "parenting": Parents raising a child together seemingly have more similar immune profiles than randomized couples; a robust effect that even extended the effect of acute and untreated gastroenteritis.[16] Various explanations were provided, including the process of "spousal concordance", which relates to shared lifestyle behaviors, such as diet, smoking, alcohol intake, exercise, sleep deprivation, chronic diseases and convergent microbiomes in couples.[16, 44] Furthermore, explanations may be sought in shared transmission of childhood viral pathogens, including EBV and CMV.[16, 23] Hence, childhood herpesvirus infection may also shift the immunological landscape of their parents towards a longstanding, closer equilibrium.

Third, age-related differences may in part be explained by sex, with girls showing a mild skewing of humoral and early differentiated CD4+ T-cell responses over cellular cytotoxic responses in contrast to boys (**Chapters 2 and 5**). Our observations are consistent with previous reports showing subtle effects of sex on CD4+ T cells.[16, 45] Furthermore, sex may interact with CMV infection, as CMV-infected pre-elderly men had lower numbers of naive and central memory T cells within the CD4 and CD8 lineages, as well as memory B cells than CMV-infected females.[46] Notably, this interaction only holds in adults, not in children (**Chapter 5**). The potential difference between these populations could be age-related, and affected by differences in hormone balances, as well as by genetic differences.[47] If this hypothesis holds, then interactions between sex and CMV become noticeable in teenagers

and in young adults. These insights may be valuable considering that autoimmune diseases are more prevalent in premenopausal females than males.[47, 48] Furthermore, sex may differentially affect vaccine responses, as responses may be more pronounced in women,[49] although inconsistent results have been reported.[49-51]

Last, age-related changes may be explained by (epi)genetics. It has been estimated that 80% of the age-related changes in the immune system can be attributed to age and genetics, and the remaining variability by age-genetic interactions.[16] However, the influences of (epi)genetics are beyond the scope of this thesis and will therefore not be discussed. In conclusion, age and CMV may be considered the most profound non-genetic determinants explaining age-dependent inter-individual variability in the immunological profile in childhood. These observations are important considering that immune function in elderly, such as susceptibility to infections, vaccination responses and cancer immunosurveillance, declines with age.[16]

B.2. Prenatal determinants: a critical window of immune–vulnerability?

The strongest prenatal effect on innate leukocyte development was observed for maternal educational level, (**Chapter 2**) which was specifically associated with a reduction in patterns of eosinophils and classical monocytes over time. Importantly, the magnitude of this "lifestyle" effect was comparable to the substantial positive effect size of postnatal CMV and EBV infection on CD8+ TemRO and TemRA lymphocyte dynamics over time, which suggests that prenatal maternal lifestyle may shape innate immunity as much as postnatal determinants such as CMV and EBV shape adaptive immunity. In line, maternal educational level was found to be the strongest confounder in the association between breastfeeding and immune cell subsets at six months of age (**Chapter 3**). How can these lifestyle effects be explained? As we included cord blood immune cells in the longitudinal analyses, prenatal maternal lifestyle or maternal physiology may affect fetal immune maturation. The effect of maternal educational level could not be explained by prenatal maternal autoimmune measures, such as maternal atopy, maternal TG2A and anti-TPO levels, nor by measures related to prenatal maternal lifestyle such as maternal Body Mass index, smoking, and alcohol consumption (**Chapters 2 and 3**). Hence, the mediator(s) of maternal educational level still need to be elucidated, but might be related to other (unmeasured) environmental factors that may be linked to maternal lifestyle and socioeconomic position, such as nutritional status, vitamin D exposures, or other infectious agents.[52, 53] Fetal regulatory T cells have been shown to respond to non-genetic maternal antigens such as infectious agents and food antigens,[29, 54] and this effect seems to persist at least until early adulthood.[29, 55] Hence, early development of the immune system may relate to significant biological consequences, such as development of peripheral tolerance and responses against pathogens that cross the placenta. These functions may develop specifically during a phase of rapidly changing peripheral environments,[55] which may link to Barker's principles[11]. In conclusion, considering the longitudinal effect of maternal

prenatal lifestyle on immune cell dynamics during early childhood, prenatal factors seem as much as important as postnatal viral exposures (**Chapter 2**).

B.3. Temporary or persistent effects on adaptive immunity

Various pre-, peri and postnatal factors influence the longitudinal kinetics of quantitative adaptive lymphocyte cells from birth until 6 years of age (**chapter 2**), and maybe longer, which suggests that most external factors have persistent, or longer-lasting priming effects on adaptive immunity in childhood, rather than transient or temporary effects. However, our analyses were limited up to the age of 6 years, and lacked information thereafter. A limitation might be that the strong longitudinal effects of bacterial and viral exposure, such as of herpesvirus infections were mainly driven by the moment IgG serology against these viruses was determined (in the majority of children only measured at 6 years of age). Most of CMV-seropositive children were infected before the age of 2 years **(chapter 4),** suggesting a longer virus-induced persistence of immune kinetics over time. Still, to enable conclusions on the persistence of these kinetics in childhood and adolescence, detailed longitudinal follow-up is necessary. On the other hand, the combined results from **Chapters 2 and 3** may provide evidence for temporary effects. Based on these chapters it could be speculated that breastfeeding transiently reduces the induction of systemic B cell responses during the period of breastfeeding and not thereafter, whereas breastfeeding may have a long-lasting reducing effect on the local CD27-IgA+ cells after a longer period of breastfeeding, or effects appear in children older than 6 months of age.[56] In line, the effects on CD8+ TCM cells are likely to be transient, as effects were no longer detectable at older ages. Though speculative, these transient effects may explain the beneficial effects of breastfeeding on the reduction of gastro-intestinal and respiratory tract infections. One of the major postnatal factors that drive immune maturation are herpesviruses (**Chapter 2** and **4**). The results in **Chapter 4** suggest that EBV infection within the first 2 years of life induces a transient decline of memory B-cells around 2 years of age that normalizes before 6 years of age, whereas vaccination responses against measles and tetanus seemed normal. Together, our results imply that the majority of environmental determinants induce temporary responses, but we also found evidence for persistent effects on the development of lymphocyte kinetics from birth until 6 years of age. Longitudinal follow-up may determine to what extent the effects on the longitudinal kinetics in early childhood remain present during adolescence.

B.4. Age at primary herpesvirus infection, immune plasticity and associations with clinical disease

We observed that CMV and EBV infection before 2 years of age induced smaller T cell-expansions than infection between 2-6 years of age (**Chapter 4**). These observations may be important considering the age-dependent associations between timing of herpesvirus acquisition and the prevalence of atopic and allergic diseases. EBV infection before 2 years

of age has been associated with beneficial effects in relation to atopic and allergic disease; whereas infection after 2 years of age was associated with opposite effects.[57] Although effector memory T cell populations did not increase in early CMV and EBV infected children between 2-6 years of life, long-term follow-up of these children is needed to assess whether the observed mild changes prevail, or eventually result in detrimental immunological effects. In contrast to early childhood, CMV infection in elderly has been associated with accumulation of CD8+ late effector memory T cells, and a subsequent profound shrinkage of the T-cell receptor (TCR) diversity and loss of naive cells.[56, 58] Furthermore, CMV-specific T cells in children were mainly CD45RA- (**Chapter 4**) rather than CD45RA+, which is in line with previous reports in children,[38] but opposes observations in elderly showing that CMV specific T cells are mainly CD45RA+ (TemRA).[56] These CD45RA+ cells in elderly may have more effector cell properties, are likely to increase with age, and thus may be a consequence of repeated antigenic stimulation.[33, 56, 59] Notably, accumulations of TemRA may also be associated with loss of immunity, as these cells show signs of replicative senescence.[56, 60, 61] Hence, differences in timing of herpesvirus acquisition may in part explain whether the immune response is mild or symptomatic, and to what extent the immune system is able to return to its steady state. For example, most primary EBV infections in children are completely asymptomatic, whereas primary EBV infection in adolescence may manifest as acute infectious mononucleosis (AIM).[62] These symptoms in AIM have been suggested to be caused by CD8 overexpression,[63] whereas carriership of EBV infection in young children is associated with an accumulation, rather than over-expansion, of virus specific CD8+ T cells (**Chapter 4**).[63] Additionally, delayed EBV and CMV acquisition, esp. in high socioeconomic groups, (**Chapter 10**) has been associated with a 2-3 times increased risk of malignancies,[62] cardiovascular disease risk and mortality.[64, 65] The onset and course of these clinical outcomes may depend on interactions between environmental factors, genetics and CMV that together may determine the speed of immunosenescence.[58] This hypothesis is in line with the observation in **Chapter 6**, showing that CMV infection modifies associations between CeD and CD57+ exhausted T cells, which may suggest that replicative senescence may be involved in disease onset or deterioration.[61] Notably, alterations in blood Vδ1 and CD57+ expressing T-cell numbers were already present in children with subclinical CeD (**Chapter 6**), suggesting that early recruitment of Vδ1+ T cells from blood to the intestinal mucosa potentially occurs prior to the onset of intestinal inflammation and appearance of symptoms. Although we need to be careful with conclusions on causal mechanisms (**Chapter 8)**, future studies assessing blood lymphocyte markers in (subclinical) CeD patients should adjust for confounding elicited by herpesviruses. In conclusion, age at primary herpesvirus acquisition may contribute to explain the onset and course of CMV and EBV associated disease epidemiology. Altogether, the studies in **Part I** of this thesis illustrate the dynamics of childhood immune maturation, the factors that drive immune maturation, and the plasticity of the childhood immune system upon exposure to various external immunogenic factors.

C. HERPESVIRUS INFECTIONS IN CHILDHOOD

C.1. The changing epidemiology of primary herpesvirus infections in childhood

Infections with CMV, EBV, HSV-1 and VZV are common, and reach 30-100% seropositivity in adults.[66-74] Though, a large variation in seroprevalence between various ethnic and geographical groups exists worldwide, with substantial differences in timing of seroconversion. We hypothesized that age at primary infection of herpesviruses is dependent on behavior that facilitates transmission, which may differ according to ethnic background. **Chapter 10** describes the basic epidemiology of three herpesviruses among children living in modern multi-ethnic urban societies. We observed substantial racial/ethnic differences in antibody prevalence, with native-Dutch 6-year old children having lower seroprevalences than five distinct non-native Dutch ethnic groups living in Rotterdam. These observations are in line with previous literature showing that primary infection occurs early in non-industrialized countries and in low socioeconomic groups, whereas in wealthy countries primary infection is often delayed until adolescence.[62] Socioeconomic position and factors related to lifestyle explained up to 39% of the ethnic differences in EBV seroprevalences, of which socioeconomic position was the most important contributor. In contrast, these factors did not explain differences in CMV antibody prevalence, suggesting important roles for vertical transmission such as breastfeeding. Hence, despite the occurrence of substantial environmental changes in Western societies during the last half-decade, the routes of childhood HSV-1 and EBV transmission still seem comparable (**Chapter 10**).[75] Nowadays, low family net household income is still one of the most predominant risk factors for transmission of both HSV-1 and EBV (**Chapter 10**). Similarly, in the 1960s and 1970s, childhood herpesvirus infections were prevalent among lower income families with children living under crowded conditions.[75] Between 1950-1970 transmission of herpesviruses decreased in higher income households with only 1 or 2 children sleeping in separate bedrooms.[76] Concurrently, breastfeeding and contact with saliva (pre-chewing baby's food, or kissing) became less common in these families. Therefore, primary EBV acquisition moved from early childhood to late adolescence resulting in symptomatic acute infectious mononucleosis (AIM).[75, 77] HSV-1 infection showed a similar pattern of early acquisition in children in poor families. Nowadays, the trend of delayed primary herpesvirus acquisition seems to persist: only 13% of native Dutch 6-year old children were seropositive for HSV-1, whereas seropositivities reach 69% in Dutch pregnant women.[78] Notably, delayed primary HSV-1 acquisition during pregnancy may increase the prevalence of maternal genital herpes and congenital herpes; similar to what was observed for delayed EBV acquisition resulting in AIM.[75] This is in line with observations of increasing genital herpes incidence in the Netherlands, the UK, Ireland and the USA between 2002-2014.[79-83] Explanations could be an increase in unsafe sexual behavior, better and faster diagnosis, but also a delay in HSV-1 acquisition. If the rise in primary genital HSV-1 infection in young women continues, it may be hypothesized that neonatal herpes infections will occur more frequently and severely in the next decade.[81]

C.2. Herpesvirus infections and autoimmunity

Herpesviruses are ubiquitous and highly contagious, but the age at primary infection is delayed in high Western countries compared to non-western countries,[84, 85] concurrent with an increase in autoimmune diseases such as CeD.[86-89] Together with the observations that herpesviruses are less common among higher socioeconomic groups with higher incidence of CeD (**Chapters 9 and 10**), it is tempting to speculate that the herpesvirus epidemiology explains some of the ethnic/geographic variation in incidence of autoimmune diseases, such as CeD. In **Chapter 8,** we observed inverse associations between CMV seropositivity and CDA. Similar protective effects of herpesviruses have previously been described for allergic diseases.[56, 57, 90] Importantly, the protective effect was limited to infection before the second year of life.[56, 57] In line with these observations, we observed smaller T-cell expansions in children who were infected with CMV and EBV before 2 years of age (**Chapter 4**). Additionally, the protective role may not be limited to CMV single infection, but may also include infection with multiple herpesviruses, as children with single herpesvirus infections were more often seropositive for other herpesviruses (**Chapter 10**). Moreover, the association with CDA was stronger for multiple herpesvirus infections than for single CMV infection. Potential causal mechanisms have been described in **Chapter 8**. Alternatively, herpesvirus seropositivity may be a proxy for the amount of other infections in early life, which is in line with the hygiene hypothesis, or old friend's theory.[40] Increasing hygienic living conditions, smaller family sizes with fewer siblings, and decreasing rates of breastfeeding, in combination with fewer early life pathogen exposures and concomitant changes in the diversity of the microbiota may contribute to the increasing incidence of allergic and autoimmune diseases.[40, 56] Hence, these environmental changes may also contribute to explain ethnic differences in autoimmunity, such as CeD. In **Chapter 9** we confirmed the presence of substantial ethnic variation in CeD incidence,[86, 88, 89, 91-95] and observed that up to 47% of ethnic differences in CDA could be explained by differences in daycare attendance, socioeconomic position and CMV acquisition in childhood. The nature of early life herpesvirus infection is often asymptomatic, and combined with the observations that these viruses induce only small T cell expansions when acquired <2 years of age, in contrast to infections later in life, suggests that it may be favorable to acquire these viruses early in life. Early life infection has been shown to protect against the development of CDA and allergies, particularly if children are infected prior to 2 years of age.[57] Therefore, it may be questioned whether immunization against these viruses may prove beneficial in children that were not infected during the first few years in life.[56] Considering the broad adverse effects of herpesvirus infections in pregnant women and newborns, as well as complications of primary herpesvirus infections in elderly or immunocompromised or -suppressed individuals, one may argue for primary prevention in the form of universal vaccination. Although some argue in favor of universal vaccination,[73, 96, 97] the effectiveness and safety of current vaccines are still discussed. Nevertheless, future vaccination pro-

grammes in typical Western multiethnic urbanizations such as Rotterdam, should at least include geographic and socioeconomic information.

D. CELIAC DISEASE: DEFINITIONS, DIAGNOSIS AND TREATMENT

Celiac disease (CeD) is a common chronic immune-mediated disorder, but largely unrecognized by clinicians due to the heterogeneity of disease presentation.[98] Patients with *silent CeD* (the patient meets the diagnostic criteria for CeD, but lacks symptoms), *subclinical CeD* (symptoms are below the CeD testing threshold and biopsies showing villous atrophy) or *potential CeD* (positive serology without histologic confirmation) may experience health effects, however this has been extensively debated.[99] Primary prevention is defined as preventing the development of CeD, and secondary prevention as diagnosing CeD in its earliest stage through active case finding, targeted screening in at risk populations, or population mass-screening. However, should we search for CeD, diagnose and treat patients with minimal or no symptoms associated with CeD? The general discussion focuses on improvement of early diagnosis, and reduction of the high underdiagnosis rate of CeD, taking long-term consequences of delayed diagnosis, complications, and early treatment with a gluten free diet (GFD) into account.

D.1. PRIMARY PREVENTION OF CELIAC DISEASE: ENVIRONMENTAL DETERMINANTS

D.1.1. A focus on early nutrition

Until 2014-2015, a special emphasis was placed on the timing and mode of gluten introduction to prevent the development of CeD. Results from previous retrospective observational studies suggested that the introduction of gluten between 4 and 6 months of age represents 'a window of opportunity' to influence the risk of developing CeD.[100, 101] Therefore, in 2012, Dutch and European guidelines recommended to introduce gluten gradually between 4-6 months of age (not earlier or later), preferably while the infant is still being breastfed.[102] However, we showed that the risk of CDA is not influenced by introducing gluten after the age of 6 months, nor by breastfeeding (**Chapter 7**). Our results are in line with results from other prospective studies.[103-105] The final answer was given by the PREVENT CD Study,[1] a randomized, double-blinded placebo controlled trial comparing infants from high risk families, whose children were randomized in either placebo or gluten intervention at the age of 3 months and underwent repeated CeD screening up to 3 years of age: Introducing gluten in small quantities between 4-6 months of age while breastfeeding was not associated with CeD risk.[1, 104] Recently, neither high amounts of gluten, nor gluten consumption patterns were associated with altered CeD risk.[106] Furthermore, maintenance of breastfeeding at the time of gluten introduction, and breastfeeding in general did not reduce the risk of CeD. In line, a systematic review[107] concluded that breastfeeding and

timing of gluten introduction did not ameliorate CeD risk, obligating for an update of the European guidelines in 2016.[108] These adjusted guidelines recommend introducing dietary gluten between 4 and 12 months of age, and state that breastfeeding and the timing of gluten introduction do not influence the risk of developing CeD in childhood.[108]

D.1.2. Geographical and ethnic differences in CeD

Genetic variation in the HLA DQ2.2, 2.5 and DQ8 haplotypes, and exposure to dietary gluten are the two undeniable factors required to develop CeD. However, most of the individuals carrying HLA-DQ2 and/or HLA-DQ8 who are exposed to gluten do not develop the disease. 40% of the European population carries these predisposing haplotypes, while CeD occurs in only a small fraction (1%).[1, 109] Hence, besides gluten, it is conceivable that additional genetic and environmental factors play a role in disease initiation.

Epidemiological studies observed substantial differences in CeD prevalence according to geographic location and ethnic background: CeD is common in Europe, US, North Africa, the middle east, sub-Saharan Africa and east-Asia, and more common among whites than among individuals from other ethnicities.[110-112] Also across different European countries, CeD prevalence varies remarkably (8-fold).[112] Furthermore, the occurrence of CeD seems to follow a north-south gradient in the US population, which was found to be independent of race/ethnicity, socioeconomic status, or body mass index.[110] Notably, this north-south gradient was also observed in occurrence of several other auto-immune diseases, which may suggest a role for vitamin D, sunlight, or ultraviolet B radiation in disease pathogenesis.[110, 113]

Intriguingly, not only CeD, but also the prevalence of CDA varies according to ethnicity: we found CDA to be more common among Western vs non-Western children living in urban Rotterdam (**Chapter 9**). Our results are in line with previous observations in the USA, where CDA is estimated to be 4-8 times greater among non-Hispanic whites than among other ethnic minorities.[114] Aside from differences in HLA-DQ2/DQ8 carriership among different ethnic groups, and differences in gluten consumption, variations in other environmental factors may explain some of the ethnic and geographic variations in CeD prevalence. Some studies suggested socioeconomic variation as an explanatory factor, as children with a higher socioeconomic position (SEP) may face different infectious agents during early childhood, consequently affecting the gut microbiota and immune programming (**Part I**). Indeed, of environmental factors, SEP contributed mostly to explain the ethnic variation in CDA prevalence, followed by daycare and CMV seropositivity. Furthermore, we found that CMV seropositivity is inversely related to the development of CDA (**Chapter 8**). This may reflect a causal association; however, as effects on CDA were stronger for the combination of EBV, HSV-1 and CMV (**Chapter 8**), it is much more likely that herpesvirus seropositivity may serve as a proxy for other, undefined, determinants, of which the combination of these different types of environmental factors may eventually induce villous atrophy. Children who acquire CMV early in life come from a lower socioeconomic environmental back-

ground (e.g. they are longer breastfed, attend daycare more often, and experience more often respiratory and gastrointestinal tract infections in their first year of life, than children primarily infected at a relatively later age (**Chapter 10**). Hence, numerous factors may contribute. For example, reovirus infections have recently been found to trigger development of TH1 immunity to gluten.[115] Still, it is likely that additional events are necessary for induction of TG2A antibodies and villous atrophy, and it may be reasonable to assume that the combination of various types of other, in part still undefined environmental stimuli, such as members of the microbiota and intestinal dysbiosis may eventually cause villous atrophy.[115] An obvious counter-argument against the partial explanations by SEP, daycare and CMV (**Chapter 9**) may be that early infant feeding practices differ between Western and non-Western children. Nonetheless, as we adjusted for breastfeeding duration and timing of gluten introduction (**Chapter 9),** these factors are less likely to contribute. Besides, early feeding practices have been shown that they cannot prevent the onset of CeD.[1, 104] A second counter-argument may be that genetic predisposition may vary among different ethnic groups, as we did not adjust for HLA DQ2/DQ8 carriership. Still, associations between ethnic background and CDA remain valid even in individuals who lack carriership for these associated haplotypes, and thus lack a final CeD diagnosis. Furthermore, the presence of other (protective) genes may contribute to prevent disease onset. To date, it has been estimated that up to ~50% of CeD heritability can be explained by genetic variation.[116] These are promising findings that leave a role for other CeD associated loci and environmental factors that may serve as novel disease biomarkers in future.

In conclusion, primary prevention of CeD (prevention of the development of CeD) is not possible at this moment. Therefore, secondary prevention strategies represent the only way to improve the high underdiagnosis rate.

D.2. SECONDARY PREVENTION: IS IT TIME TO SCREEN FOR CELIAC DISEASE?

D.2.1. Active–case finding, and targeted screening in populations at risk: when and how?

To diagnose CeD in its earliest stage, "active-case finding" is recommended,[102] which means an active approach by clinicians to seek for CeD-associated symptoms and related diseases. However, due to the wide variety of disease presentation (often subtle, atypical, or absence of symptoms), the majority of screening-detected CeD cases cannot be diagnosed by case-finding.[117] Therefore, current European[102] and American guidelines [118, 119] recommend to screen asymptomatic children and adolescents with an increased risk for CeD, provided that they have consumed an adequate, gluten-containing diet for at least 1 year prior to testing. At risk populations include first- or second- degree relatives, or patients with other autoimmune diseases, such as type 1 diabetes mellitus (T1DM). First-degree relatives have a 5-20% higher risk of developing CeD. However, the risk is differentially affected by sex (F/M= 2-3:1 at age 3 yrs.) and HLA-genotype. Homozygous HLA-DQ2 increases the risk to 14.9%, followed by heterozygous HLA-DQ2 (3.9%) and HLA DQ8 hetero- and homozygous (0.9%).[1] HLA-testing, followed by IgA-TG2A determination, is recommended to start around

3 years of age.[119] Approximately 50% of children with a CeD diagnosis and a family history of CeD have been shown to develop CeD antibodies by 3 years of age,[1] suggesting that the potential forms start early in life, and testing may start prior to 3 years of age. Even in very young children, TG2A testing shows good sensitivities and specificities.[1] If IgA-TG2A is negative and IgA deficiency is excluded, then CeD is unlikely. However, as the disease may still develop later in life, serological testing should be repeated at regular intervals; i.e. every 2 to 3 years. Notably, targeted screening in the Generation R cohort would not have led to early identification of CeD, as only a few children with positive TG2A levels had a first or second-degree relative with CeD (**Chapters 10 and 11**).

D.2.2. Population mass screening: yes or no? If yes, when and how?

Mass-screening in the general population may be another approach to improve early diagnosis of CeD, which means that all individuals regardless of symptoms undergo serological testing, followed by esophagogastroduodenoscopy (EGD) with duodenal biopsy in those with positive TG2A levels.[102] Currently, biopsies can only be avoided in *symptomatic* children presenting with 1 symptom, TG2A levels exceeding 10 ULN, and positive EmA twice.[120] These criteria have recently been validated by a large, multinational, prospective study,[120] showing that burden, risks and costs of endoscopy can be prevented for >50% of CeD patients worldwide. Moreover, HLA testing seems not to improve accuracy of CeD diagnosis in symptomatic patients, and may therefore be omitted.[120] Notably, the guideline is currently discussed, as the diagnostic ESPGHAN criteria for symptomatic children may also be applied to asymptomatic screening identified children. Swedish observations[121] and previous observations from the Prevent CD study[1] indicate that TG2A levels >10 ULN show good diagnostic accuracy in screened children. In line, all children with TG2A levels >10 ULN in the Generation R cohort and in the Prevent CD Study[1] had severe enteropathy, and the positive predictive value of TG2A did not depend on symptoms (**Chapters 10 and 11**),[1] necessitating for an update of the diagnostic ESPGHAN criteria. Hence, as burden, risks and costs of endoscopy may be avoided, the screening debate would be placed in a totally different context. To assess the harms and benefits of mass-screening for CeD, the criteria according to Wilson and Jungner[122] can be applied (**Table 1).** Several recent epidemiological studies, including ours (**Chapters 10 and 11**), made efforts to reopen the screening debate. For example, 3 US studies focusing on the natural course of CeD found that 0% to 15% of patients with positive serology (without histologic abnormalities) developed symptoms after 10 to 45 years.[86, 87, 123] Still, some questions remain unanswered. In the following paragraphs, we aimed to describe recent considerations to assess the net harm and benefits of mass-screening.

Table 1. Wilson and Jungner criteria for mass-screening applied to CeD

Criterium	CeD situation	References
1. Important health problem	1% of general population:33.600 children in The Netherlands	124-126
2. Accepted treatment	GFD, good adherence in screening detected cases in prospective studies	127, 128
3. Facilities diagnosis	Presence of CeD specific antibodies	8, 102
4. Recognizable latent/ early symptomatic stage	Presence of CeD specific antibodies	8, 102
5. Suitable test detection	Presence of CeD specific antibodies	8, 102
6. Test acceptable for the population	Depends, data from prospective studies, PPV's in subclinical disease discussed	86, 87, 123, 129
7. Natural history understood	Data from prospective studies showing potential and subclinical forms	86, 87, 123, 127, 130, 131
8. Agreed policy of whom to treat	Depends, data from prospective studies showing adverse health effects in subclinical disease, and recovery upon treatment	125, 132-136 127, 131, 137, 138
9. Cost-case finding economically balanced	Limited data	139, 140
10. Case finding continuous process	No data available	

D.2.3. Potential benefits of screening in asymptomatic populations

Symptomatic CeD is associated with a wide variety of symptoms of malabsorption (which may appear several years prior to diagnosis), excess mortality, intestinal adenocarcinoma, and lymphoma.[86] Screening may enable early diagnosis and treatment before overt symptoms develop, relief mild but unacknowledged symptoms, and prevent malabsorption and associated nutritional deficiencies.[99] However, it remains unclear whether subclinical CeD is associated with the same risks as reported for symptomatic CeD. Screening identified children with subclinical CeD in the Generation R cohort had reduced bone mineral densities and growth trajectories from 6 months onwards (**Chapter 10**). Our observations are in line with others,[132-134] showing that growth falters in most children with CeD, even in anti-EMA positive children without biopsies.[134] Recent Norwegian observations[132] showed that growth retardation in children commonly starts at 12 months of age, and precedes clinical symptoms that usually bring the suspicion of CeD. Notably, the average age of CeD diagnosis is between 40-60 years of age;[99, 141, 142] often too late to reverse detrimental effects on bone health, such as reduced mineral density, osteopenia/osteoporosis. Hence, if considering mass-screening, screening should start early in life. Importantly, screening studies showed a decreased BMD even in asymptomatic adults who had positive concentrations of celiac antibodies,[135, 136] but no villous atrophy,[143] indicating that the presence of positive serology may induce adverse effects on BMD.

In contrast to the relatively high PPV of 61% in our study (**Chapter 11**), studies among US adults reported rather lower PPV values, i.e. from 0% to 15%,[86, 87, 123] suggesting that the PPV of TG2A to detect (subclinical) CeD may be substantially different among children

and adults. Varying results may be in part explained by differences in disease prevalence among children and adults,[144] differences in time of follow-up, the absence or presence of histologic confirmation, [99, 145] and the spectrum of CeD. For example, in untreated CeD, high sensitivities (TG2A IgA 85-95%; EMA IgA 80-90%) and specificities (TG2A- IgA 95-99% ; EMA IgA 95-100%) have been reported, whereas in refractory CeD patients, sensitivities reach barely 50%.[146] These patients had normal levels of TG2A while mucosal atrophy was not recovered upon a GFD. Hence, sensitivities and specificities of CeD specific antibody tests may be different for screening-identified 'subclinical' CeD than for symptomatic CeD.

D2.4. Effectiveness of early treatment in asymptomatic populations

Several studies have shown that screening and early treatment with a GFD among screen detected CeD children and adults results in significant health improvements.[127, 131, 137, 138] On the short run, a GFD has been shown to improve symptoms of diarrhea, abdominal pain, constipation, fatigue, malabsorption, irritability, oral aphthous ulcers and growth retardation in Dutch screened children.[127] In adult Finish CeD patients, a GFD improved chronic health problems as gastroesophageal reflux, indigestion and anxiety.[131] Furthermore, dietary adherence seems excellent after screening; CeD-related QOL was not different between screen-detected and symptom-detected patients.[127, 131, 137] Of screening identified children with CeD in the Generation R cohort (**Chapter 11**), treatment with a GFD in the ~first 6-9 months following diagnosis resulted in the majority of children in improvement of abdominal pain and fatigue (which were present in 11/18 children). CeD associated symptoms were subtle and only recognized after initiation of a GFD, suggesting that mass-screening enables identification of patients who are symptomatic in retrospect. Nonetheless, these 'subtle' symptoms did not result in visitation of the general practitioner (**Chapters 10 and 11**). Hence, a GFD seems to alleviate even mild, but otherwise unrecognized symptoms that are below the awareness of the patient as a clinical symptom. Furthermore, a GFD resulted in improvement of growth parameters at 13 years of age. On the long run, screening may prevent complications such as osteoporosis, reduced height, subfertility and perinatal problems. Osteopenia affects 26-72% of all untreated CeD patients, resulting in increased fracture risk.[147, 148] Furthermore, we showed that screening identified children with CeD experience growth retardation (**Chapter 10**), which is in line with others.[132, 134] Generally, osteopenia/osteoporosis and growth failure recover completely within 1-2 years upon start of the GFD in childhood, in contrast to diagnosis after 25 years of age, underscoring the need for early diagnosis and treatment.[149-151]

D.2.5. Potential harms of screening in asymptomatic populations

False-positive, or inconclusive serology and biopsy results are potential harms of screening, resulting in unnecessary burden or complications. In high risk individuals, the diagnostic accuracy of screening tests is considered to be good with sensitivities and specificities reaching >90%, (**Chapter 1**) and PPV's of approx. 40-80%.[99] However, evidence in asymp-

tomatic pediatric populations is scarce. Only one study reported that TG2A levels >10 times ULN has, concurrent with a positive EMA IgA test result, a sensitivity and specificity of 83% and 67% to detect enteropathy (i.e. Marsh stage 2 or 3).[129] We found the PPV of IgA-TG2A testing for CeD in 6-year old children to be relatively high with 61% (**Chapter 11**). These observations are in line with observations in the general European adult population.[112] Of 29,000 screened individuals, 147 individuals had both positive TG2A as well as positive EMA, of whom 100 had enteropathy, suggesting a combined PPV for CeD of 68%. Assuming a PPV of 61%, still, more than 1/3 of children should currently undergo biopsies, whereas they do not have CeD. On the other hand, a PPV of 60-70% may be appreciable, considering that colorectal cancer in the general population (age >50 years) is accepted in most countries of the world, based on a PPV of guaiac fecal occult blood (gFOC) of 9%, and together with detection of adenoma's, a PPV of 30%.[152]

Second, it remains unclear how many asymptomatic individuals who screen positive, will undergo invasive gastrointestinal endoscopies. Though speculative, because of the absence of symptoms, these individuals may be more likely to refuse biopsies, thereby complicating diagnosis. In our study, 16% (5/31) of parents of children who were suspected to have CeD, refused duodenal biopsies, of whom 3 had levels exceeding 10 ULN (**Chapter 11**). Considering that CeD diagnosis in screening identified children can be made without biopsies, the screening debate would be placed in a totally different view, as complications, patient burden, anxiety and costs of time-consuming endoscopies may be avoided.

Third, the natural course of asymptomatic and subclinical CeD/CDA is understudied in the general population. Only 5 long-term follow-up studies in children and adults,[86, 87, 123, 127, 130] have shown that some screening identified CeD cases never develop symptoms. Together, these observations suggest that over-diagnosis may be an important concern. On the other hand, adverse effects on health seem to be present irrespective of the presence of clinical symptoms (**Chapter 10**).[117, 132] However, most individuals are probably not truly asymptomatic, and only become aware of the presence of (subtle) symptoms after they have been placed on a GFD.[131]

Fourth, dietary adherence may be questionable, as the GFD may have a major impact on daily life, is expensive, and difficult to maintain, especially in screening identified, asymptomatic, or minimally symptomatic individuals who do not expect to gain health as much as symptomatic patients. However, dietary adherence has been found excellent in Dutch and 12-year old Swedish children screened positive.[127, 128] Second, in screened Dutch children and Finnish adults, a GFD has been shown effective in terms of alleviating early and chronic CeD associated symptoms (including symptoms of diarrhea, abdominal pain, constipation, fatigue, irritability, oral aphthous ulcers, growth failure, indigestion, gastro-esophageal reflux and anxiety).[127, 131] Furthermore, a GFD has been shown to improve serology and mucosal atrophy in adults.[131]

Last, the impact of CeD diagnosis and treatment of GFD on quality of life probably varies between countries, as changes in perceived health must be balanced against logistical,

economic and social burden.[153, 154] Despite these concerns, Swedish observations suggest that mass-screening is acceptable for most of CeD patients and their parents.[155] In fact, parents suggested it to be more beneficial if performed early in life.[155] In line, quality of life among Finnish, US, and Dutch patients seems high, and not materially different between screening detected patients and patients who are diagnosed based on symptoms.[127, 131, 137, 156]

To conclude is it time to screen for CeD instead of targeted screening? Considering the important detrimental health effects of unrecognized CeD, mass-screening may be justified. However, the benefits of screening (prevention or reduction of complications of undiagnosed CeD by early diagnosis and treatment) should outweigh the harms (false negative test results, anxiety, costs and burden of gastrointestinal biopsies and GFD). To assess this balance more adequately, some questions remain unanswered. Furthermore, cost-effectiveness analyses are missing in a large childhood population.

METHODOLOGICAL CONSIDERATIONS

The studies described in this thesis were performed within the Generation R Study; a population based prospective cohort study.[157] Cohort studies provide the unique opportunity to observe populations prospectively, thereby identifying factors associated with an outcome of interest. The studies presented in this thesis combine a unique combination of detailed immunological analyses on the origin, maturation and responsiveness of specific lymphocyte populations, and large-scale epidemiological analyses that include associations with herpesvirus acquisition and CDA. However, several methodological considerations need to be taken into account.

E. Outcome definitions / Information bias

E.1. Immune cell analyses

The mainly cross-sectional studies described in **Part I** include detailed assessment of immunological flowcytometric analyses at multiple time points. Flow cytometry has proven highly successful in characterizing blood immune cells in health and disease.[158] Population numbers included (n=230-377 from birth until 24m; and n=~1,100 at 6 years), are considered extremely large in the field of immunology, however, –from an epidemiological point of view- these groups comprise relative small numbers, thereby limiting statistical power (and the possibility to adjust for a wide range of potential confounders).

E.2. Misclassification of determinants

We gathered the majority of data on pre-, peri and postnatal determinants (such as maternal lifestyle, infant feeding practices (breastfeeding duration, gluten introduction)) retrospectively by questionnaires, which may be prone to misclassification.[159] However,

because any misclassification would be independent of our outcomes (as participants were not aware of determining laboratory findings such as immune cell numbers, CDA, and herpesvirus persistence), it is unlikely that these have affected our outcomes.

E.3. Selection bias, missing data and data imputation

Cohort studies may be prone to selection bias do to either loss of study participants, or loss to follow-up. Initially, this may occur if a proportion of the target study population does not participate, and is systematically different from the original study population with regard to exposure and outcomes. Participating mothers in the Generation R Study more often come from a higher socioeconomic background than non-participating mothers, which may have affected the external validity. However, only if associations between determinants and outcomes are different for participants and non-participants, selection bias could have influenced our findings. In non-immunological chapters, we dealt with missing data by using multiple imputation techniques. In the field of epidemiology, multiple imputations are recommended to adjust for potential attrition bias and loss of information that may occur when performing complete-case analyses. Notably, results of multiple imputations are valid under the assumption that data are missing at random (MAR).[160, 161] However, in this thesis it was not possible to distinguish between data being MAR or missing not at random (MNAR). When data was MNAR this could have affected the prediction of the missing values and therefore the uncertainty of the effect estimates observed.

E.4. Confounding and reverse causation

Prospective cohort studies do not conclusively prove causal association, because first, associations between exposure and outcome may be distorted by confounders (i.e. factors that are related to both determinant and outcome, and are not intermediates in the pathway from exposure to outcome). We tended to adjusted our studies for multiple potentially confounding variables. Still, residual confounding cannot be fully excluded. This may be due to the fact that some potential confounders were not (sufficiently) measured during the process of data gathering (e.g. timing, duration and dose of infections, or because measurement of variables contained errors. Second, reverse causality cannot be fully excluded, for example in the association between herpesvirus seropositivity and CDA, as both herpesvirus seropositivity and TG2A were determined cross-sectional at 6 years of age.

CONSIDERATIONS FOR FUTURE RESEARCH

Non–genetic determinants that shape immunity

The results presented in **Part I** may be essential to begin to understand the gradual and ordered acquisition of immune competence in childhood.[13] However, our observations

raise further questions, such as: "Which mechanisms operate to regulate circulating, (and lymphoid tissue) resident immune cells?" How do circulating immune cell numbers relate to functions such as production, homeostasis and inflammation.[12, 13] Furthermore, future studies should include information on the diversity of antigen receptor repertoires to determine the immunologic flexibility to respond adequately to pathogens.[13] Many postnatal determinants have been studied in cross-sectional studies including only a small numbers of participants, but the crucial role of timing of these environmental factors, i.e. whether a critical or vulnerable time period exists, needs to be further studied in longitudinal studies. For example, future studies should evaluate the timing of herpesvirus infection on immune plasticity/ TCR diversity. These studies should include not only postnatal, but also pre- and perinatal factors in a multivariable model to assess their relative contributions. The Generation R Study provides a unique opportunity to longitudinally assess immunological patterns, and to provide information on the quantification and maturation of, for example, IgE producing B cells that are strongly linked to atopic diseases. Questions include: "Which pre- peri- and postnatal determinants are associated with IgE-producing B cells at 9 years of age? What is the role for CD27-IgA+ cells in gut homeostasis, and to what extent are CD27-IgG+ cells related to CD(A)?" Furthermore, the intestinal microbiome should be linked to immune maturation.

Determinants and consequences of Celiac Disease Autoimmunity

The overall feasibility of mass-screening needs to be assessed by studying the acceptance-rate of mass-screening in the general population, among parents and clinicians, the psychological impact, compliance to diagnostic procedures and treatment in those screened positive, cost-benefit analyses, and the ethical aspects involved. Furthermore, the optimal age-range to screen asymptomatic children should be assessed, as well as the need for repeated testing.[144] Furthermore, a cost-effectiveness analysis of screening and treatment (compared with the total healthcare costs) should be performed in the general pediatric population that includes all of the aforementioned aspects of screening, and also the long-term CeD associated complications. In the young adult general population, mass-screening for CeD was found to be cost-effective if the time delay to diagnosis was longer than 6 years,[140] with a cost-effectiveness ratio of 48.960 USD per QALY. One of the aims is to further explore the cost-effectiveness balance in Dutch children participating in the Generation R Study. Furthermore, the course of CDA in the general population remains unknown. The Generation R cohort provides the unique opportunity to longitudinally evaluate the course of TG2A seropositivity/subclinical CeD at 6 and 9 years of age. This would provide information on the number of new onset CeD or CDA cases if screening is repeated.

CONCLUDING REMARKS

To conclude,

1. "Healthy" or "normal" maturation of the immune system in childhood is characterized by a wide range of distinct innate and adaptive immune cell patterns, which differ greatly between children, and depend highly on age. This immunological variation between individuals is largely driven by non-heritable (environmental) factors, of which postnatal herpesvirus acquisition has been considered one of the most profound determinants. Furthermore, the results presented in this thesis imply that the "window of immune vulnerability" starts in utero, of which pre- and perinatal determinants may be just as important as postnatal determinants that leave a longer-lasting fingerprint on the maturing immune system. Furthermore, our results imply that early herpesvirus acquisition does not adversely affect the plasticity of the immune system, which, together with observations from others, may prove beneficial in terms of modifying "clinical disease risk".
2. Currently, primary prevention strategies cannot prevent the onset of CeD. Other environmental determinants, such as infections, may play a role in disease onset, but are relatively understudied, and poorly understood. Therefore, the only method to optimize early diagnosis of CeD is secondary prevention. Because active-case finding and targeted screening cannot resolve the high underdiagnosis rate, mass-screening may be a reasonable and realistic alternative.

REFERENCES

1. Vriezinga SL, Auricchio R, Bravi E, Castillejo G, Chmielewska A, Crespo Escobar P, et al. Randomized feeding intervention in infants at high risk for celiac disease. N Engl J Med. 2014;371:1304-15.
2. Vriezinga SL, Schweizer JJ, Koning F, Mearin ML. Coeliac disease and gluten-related disorders in childhood. Nat Rev Gastroenterol Hepatol. 2015;12:527-36.
3. Jabri B, Sollid LM. T Cells in Celiac Disease. J Immunol. 2017;198:3005-14.
4. Monsuur AJ, de Bakker PI, Zhernakova A, Pinto D, Verduijn W, Romanos J, et al. Effective detection of human leukocyte antigen risk alleles in celiac disease using tag single nucleotide polymorphisms. PLoS One. 2008;3:e2270.
5. Mearin ML. Celiac disease among children and adolescents. Curr Probl Pediatr Adolesc Health Care. 2007;37:86-105.
6. Romanos J, Rosen A, Kumar V, Trynka G, Franke L, Szperl A, et al. Improving coeliac disease risk prediction by testing non-HLA variants additional to HLA variants. Gut. 2014;63:415-22.
7. van Bergen J, Mulder CJ, Mearin ML, Koning F. Local communication among mucosal immune cells in patients with celiac disease. Gastroenterology. 2015;148:1187-94.
8. Ludvigsson JF, Leffler DA, Bai JC, Biagi F, Fasano A, Green PH, et al. The Oslo definitions for coeliac disease and related terms. Gut. 2013;62:43-52.
9. Maki M, Collin P. Coeliac disease. Lancet. 1997;349:1755-9.
10. MacGillivray DM, Kollmann TR. The role of environmental factors in modulating immune responses in early life. Front Immunol. 2014;5:434.
11. Barker DJ. The origins of the developmental origins theory. J Intern Med. 2007;261:412-7.
12. Mandala WL, Ananworanich J, Apornpong T, Kerr SJ, MacLennan JM, Hanson C, et al. Control lymphocyte subsets: can one country's values serve for another's? J Allergy Clin Immunol. 2014;134:759-61 e8.
13. van Lier RAW, Kuijpers TW. The cellular immune system comes of age. J Allergy Clin Immunol. 2017;139:1793-4.
14. van Zelm MC, van der Burg M, van Dongen JJ. Homeostatic and maturation-associated proliferation in the peripheral B-cell compartment. Cell Cycle. 2007;6:2890-5.
15. Farber DL, Yudanin NA, Restifo NP. Human memory T cells: generation, compartmentalization and homeostasis. Nat Rev Immunol. 2014;14:24-35.
16. Carr EJ, Dooley J, Garcia-Perez JE, Lagou V, Lee JC, Wouters C, et al. The cellular composition of the human immune system is shaped by age and cohabitation. Nat Immunol. 2016;17:461-8.
17. Davis MM. Immunology taught by humans. Sci Transl Med. 2012;4:117fs2.
18. Morbach H, Eichhorn EM, Liese JG, Girschick HJ. Reference values for B cell subpopulations from infancy to adulthood. Clin Exp Immunol. 2010;162:271-9.
19. Piatosa B, Wolska-Kusnierz B, Pac M, Siewiera K, Galkowska E, Bernatowska E. B cell subsets in healthy children: reference values for evaluation of B cell maturation process in peripheral blood. Cytometry B Clin Cytom. 2010;78:372-81.
20. Duchamp M, Sterlin D, Diabate A, Uring-Lambert B, Guerin-El Khourouj V, Le Mauff B, et al. B-cell subpopulations in children: National reference values. Immun Inflamm Dis. 2014;2:131-40.
21. Tosato F, Bucciol G, Pantano G, Putti MC, Sanzari MC, Basso G, et al. Lymphocytes subsets reference values in childhood. Cytometry A. 2015;87:81-5.
22. Orru V, Steri M, Sole G, Sidore C, Virdis F, Dei M, et al. Genetic variants regulating immune cell levels in health and disease. Cell. 2013;155:242-56.
23. Brodin P, Jojic V, Gao T, Bhattacharya S, Angel CJ, Furman D, et al. Variation in the human immune system is largely driven by non-heritable influences. Cell. 2015;160:37-47.

24. De Jager PL, Hacohen N, Mathis D, Regev A, Stranger BE, Benoist C. ImmVar project: Insights and design considerations for future studies of "healthy" immune variation. Semin Immunol. 2015;27:51-7.
25. Shaw AC, Goldstein DR, Montgomery RR. Age-dependent dysregulation of innate immunity. Nat Rev Immunol. 2013;13:875-87.
26. PrabhuDas M, Bonney E, Caron K, Dey S, Erlebacher A, Fazleabas A, et al. Immune mechanisms at the maternal-fetal interface: perspectives and challenges. Nat Immunol. 2015;16:328-34.
27. Ghazal P, Dickinson P, Smith CL. Early life response to infection. Curr Opin Infect Dis. 2013;26:213-8.
28. Kollmann TR, Crabtree J, Rein-Weston A, Blimkie D, Thommai F, Wang XY, et al. Neonatal innate TLR-mediated responses are distinct from those of adults. J Immunol. 2009;183:7150-60.
29. Zhang X, Zhivaki D, Lo-Man R. Unique aspects of the perinatal immune system. Nat Rev Immunol. 2017.
30. Pabst HF, Godel J, Grace M, Cho H, Spady DW. Effect of breast-feeding on immune response to BCG vaccination. Lancet. 1989;1:295-7.
31. Kuijpers TW, Vossen MT, Gent MR, Davin JC, Roos MT, Wertheim-van Dillen PM, et al. Frequencies of circulating cytolytic, CD45RA+CD27-, CD8+ T lymphocytes depend on infection with CMV. J Immunol. 2003;170:4342-8.
32. Kuijpers TW, Baars PA, Dantin C, van den Burg M, van Lier RA, Roosnek E. Human NK cells can control CMV infection in the absence of T cells. Blood. 2008;112:914-5.
33. Appay V, Dunbar PR, Callan M, Klenerman P, Gillespie GM, Papagno L, et al. Memory CD8+ T cells vary in differentiation phenotype in different persistent virus infections. Nat Med. 2002;8:379-85.
34. Appay V, van Lier RA, Sallusto F, Roederer M. Phenotype and function of human T lymphocyte subsets: consensus and issues. Cytometry A. 2008;73:975-83.
35. Lachmann R, Bajwa M, Vita S, Smith H, Cheek E, Akbar A, et al. Polyfunctional T cells accumulate in large human cytomegalovirus-specific T cell responses. J Virol. 2012;86:1001-9.
36. Riddell NE, Griffiths SJ, Rivino L, King DC, Teo GH, Henson SM, et al. Multifunctional cytomegalovirus (CMV)-specific CD8(+) T cells are not restricted by telomere-related senescence in young or old adults. Immunology. 2015;144:549-60.
37. Riou C, Treurnicht F, Abrahams MR, Mlisana K, Liu MK, Goonetilleke N, et al. Increased memory differentiation is associated with decreased polyfunctionality for HIV but not for cytomegalovirus-specific CD8+ T cells. J Immunol. 2012;189:3838-47.
38. Griffiths SJ, Riddell NE, Masters J, Libri V, Henson SM, Wertheimer A, et al. Age-associated increase of low-avidity cytomegalovirus-specific CD8+ T cells that re-express CD45RA. J Immunol. 2013;190:5363-72.
39. Geginat J, Lanzavecchia A, Sallusto F. Proliferation and differentiation potential of human CD8+ memory T-cell subsets in response to antigen or homeostatic cytokines. Blood. 2003;101:4260-6.
40. Strachan DP. Hay fever, hygiene, and household size. BMJ. 1989;299:1259-60.
41. M'Rabet L, Vos AP, Boehm G, Garssen J. Breast-feeding and its role in early development of the immune system in infants: consequences for health later in life. J Nutr. 2008;138:1782S-90S.
42. Mincheva-Nilsson L, Hammarstrom ML, Juto P, Hammarstrom S. Human milk contains proteins that stimulate and suppress T lymphocyte proliferation. Clin Exp Immunol. 1990;79:463-9.
43. Brugman S, Perdijk O, van Neerven RJ, Savelkoul HF. Mucosal Immune Development in Early Life: Setting the Stage. Arch Immunol Ther Exp (Warsz). 2015;63:251-68.
44. Lax S, Smith DP, Hampton-Marcell J, Owens SM, Handley KM, Scott NM, et al. Longitudinal analysis of microbial interaction between humans and the indoor environment. Science. 2014;345:1048-52.
45. Amadori A, Zamarchi R, De Silvestro G, Forza G, Cavatton G, Danieli GA, et al. Genetic control of the CD4/CD8 T-cell ratio in humans. Nat Med. 1995;1:1279-83.
46. van der Heiden M, van Zelm MC, Bartol SJ, de Rond LG, Berbers GA, Boots AM, et al. Differential effects of Cytomegalovirus carriage on the immune phenotype of middle-aged males and females. Sci Rep. 2016;6:26892.

47. Ober C, Loisel DA, Gilad Y. Sex-specific genetic architecture of human disease. Nat Rev Genet. 2008;9:911-22.
48. Whitacre CC, Reingold SC, O'Looney PA. A gender gap in autoimmunity. Science. 1999;283:1277-8.
49. Giefing-Kroll C, Berger P, Lepperdinger G, Grubeck-Loebenstein B. How sex and age affect immune responses, susceptibility to infections, and response to vaccination. Aging Cell. 2015;14:309-21.
50. Furman D, Hejblum BP, Simon N, Jojic V, Dekker CL, Thiebaut R, et al. Systems analysis of sex differences reveals an immunosuppressive role for testosterone in the response to influenza vaccination. Proc Natl Acad Sci U S A. 2014;111:869-74.
51. Tsang JS, Schwartzberg PL, Kotliarov Y, Biancotto A, Xie Z, Germain RN, et al. Global analyses of human immune variation reveal baseline predictors of postvaccination responses. Cell. 2014;157:499-513.
52. Clark CE, Fay MP, Chico ME, Sandoval CA, Vaca MG, Boyd A, et al. Maternal Helminth Infection Is Associated With Higher Infant Immunoglobulin A Titers to Antigen in Orally Administered Vaccines. J Infect Dis. 2016;213:1996-2004.
53. Dauby N, Goetghebuer T, Kollmann TR, Levy J, Marchant A. Uninfected but not unaffected: chronic maternal infections during pregnancy, fetal immunity, and susceptibility to postnatal infections. Lancet Infect Dis. 2012;12:330-40.
54. Mold JE, Michaelsson J, Burt TD, Muench MO, Beckerman KP, Busch MP, et al. Maternal alloantigens promote the development of tolerogenic fetal regulatory T cells in utero. Science. 2008;322:1562-5.
55. Mold JE, McCune JM. Immunological tolerance during fetal development: from mouse to man. Adv Immunol. 2012;115:73-111.
56. Van den Heuvel D. The Immune System out of Shape? Shaping of adaptive immunity by persistent viral infections in young children. Rotterdam: Erasmus University Rotterdam; 2015.
57. Saghafian-Hedengren S, Sverremark-Ekstrom E, Linde A, Lilja G, Nilsson C. Early-life EBV infection protects against persistent IgE sensitization. J Allergy Clin Immunol. 2010;125:433-8.
58. Bauer ME, Wieck A, Petersen LE, Baptista TS. Neuroendocrine and viral correlates of premature immunosenescence. Ann N Y Acad Sci. 2015;1351:11-21.
59. Gamadia LE, van Leeuwen EM, Remmerswaal EB, Yong SL, Surachno S, Wertheim-van Dillen PM, et al. The size and phenotype of virus-specific T cell populations is determined by repetitive antigenic stimulation and environmental cytokines. J Immunol. 2004;172:6107-14.
60. Appay V, Rowland-Jones SL. Lessons from the study of T-cell differentiation in persistent human virus infection. Semin Immunol. 2004;16:205-12.
61. Brenchley JM, Karandikar NJ, Betts MR, Ambrozak DR, Hill BJ, Crotty LE, et al. Expression of CD57 defines replicative senescence and antigen-induced apoptotic death of CD8+ T cells. Blood. 2003;101:2711-20.
62. Macsween KF, Crawford DH. Epstein-Barr virus-recent advances. Lancet Infect Dis. 2003;3:131-40.
63. Jayasooriya S, de Silva TI, Njie-jobe J, Sanyang C, Leese AM, Bell AI, et al. Early virological and immunological events in asymptomatic Epstein-Barr virus infection in African children. PLoS Pathog. 2015;11:e1004746.
64. Simanek AM, Dowd JB, Pawelec G, Melzer D, Dutta A, Aiello AE. Seropositivity to cytomegalovirus, inflammation, all-cause and cardiovascular disease-related mortality in the United States. PLoS One. 2011;6:e16103.
65. Muhlestein JB, Horne BD, Carlquist JF, Madsen TE, Bair TL, Pearson RR, et al. Cytomegalovirus seropositivity and C-reactive protein have independent and combined predictive value for mortality in patients with angiographically demonstrated coronary artery disease. Circulation. 2000;102:1917-23.
66. Staras SA, Dollard SC, Radford KW, Flanders WD, Pass RF, Cannon MJ. Seroprevalence of cytomegalovirus infection in the United States, 1988-1994. Clin Infect Dis. 2006;43:1143-51.
67. Cannon MJ, Schmid DS, Hyde TB. Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. Rev Med Virol. 2010;20:202-13.

68. Bradley H, Markowitz LE, Gibson T, McQuillan GM. Seroprevalence of herpes simplex virus types 1 and 2--United States, 1999-2010. J Infect Dis. 2014;209:325-33.
69. Korndewal MJ, Mollema L, Tcherniaeva I, van der Klis F, Kroes AC, Oudesluys-Murphy AM, et al. Cytomegalovirus infection in the Netherlands: seroprevalence, risk factors, and implications. J Clin Virol. 2015;63:53-8.
70. Gaytant MA, Galama JM, Semmekrot BA, Melchers WJ, Sporken JM, Oosterbaan HP, et al. The incidence of congenital cytomegalovirus infections in The Netherlands. J Med Virol. 2005;76:71-5.
71. Dowd JB, Palermo T, Brite J, McDade TW, Aiello A. Seroprevalence of Epstein-Barr virus infection in U.S. children ages 6-19, 2003-2010. PLoS One. 2013;8:e64921.
72. Smith JS, Robinson NJ. Age-specific prevalence of infection with herpes simplex virus types 2 and 1: a global review. J Infect Dis. 2002;186 Suppl 1:S3-28.
73. Pembrey L, Raynor P, Griffiths P, Chaytor S, Wright J, Hall AJ. Seroprevalence of cytomegalovirus, Epstein Barr virus and varicella zoster virus among pregnant women in Bradford: a cohort study. PLoS One. 2013;8:e81881.
74. van Lier A, Smits G, Mollema L, Waaijenborg S, Berbers G, van der Klis F, et al. Varicella zoster virus infection occurs at a relatively young age in The Netherlands. Vaccine. 2013;31:5127-33.
75. Grose C, Johanson DC. Transmission of Cytomegalovirus, Epstein-Barr Virus, and Herpes Simplex Virus Infections: From the Lucy Australopithecus Epoch to Modern-Day Netherlands. J Pediatr. 2016;170:9-10.
76. Stagno S, Pass RF, Cloud G, Britt WJ, Henderson RE, Walton PD, et al. Primary cytomegalovirus infection in pregnancy. Incidence, transmission to fetus, and clinical outcome. JAMA. 1986;256:1904-8.
77. Hoagland RJ. The clinical manifestations of infectious mononucleosis: a report of two hundred cases. Am J Med Sci. 1960;240:55-63.
78. Gaytant MA, Steegers EA, van Laere M, Semmekrot BA, Groen J, Weel JF, et al. Seroprevalences of herpes simplex virus type 1 and type 2 among pregnant women in the Netherlands. Sex Transm Dis. 2002;29:710-4.
79. Bernstein DI, Bellamy AR, Hook EW, 3rd, Levin MJ, Wald A, Ewell MG, et al. Epidemiology, clinical presentation, and antibody response to primary infection with herpes simplex virus type 1 and type 2 in young women. Clin Infect Dis. 2013;56:344-51.
80. Xu F, Sternberg MR, Kottiri BJ, McQuillan GM, Lee FK, Nahmias AJ, et al. Trends in herpes simplex virus type 1 and type 2 seroprevalence in the United States. JAMA. 2006;296:964-73.
81. Kimberlin DW. The scarlet H. J Infect Dis. 2014;209:315-7.
82. Aar F van (RIVM) NMvdR. Herpes genitalis: Zijn er verschillen tussen Nederland en andere landen? In: Volksgezondheid Toekomst Verkenning, Nationaal Kompas Volksgezondheid Bilthoven: RIVM 2 juli 2014.
83. WHO. CISID database 7022 Genital herpes simplex incidence rate, acces date 18-7-2017
84. Nilsson C, Larsson Sigfrinius AK, Montgomery SM, Sverremark-Ekstrom E, Linde A, Lilja G, et al. Epstein-Barr virus and cytomegalovirus are differentially associated with numbers of cytokine-producing cells and early atopy. Clin Exp Allergy. 2009;39:509-17.
85. Aarnisalo J, Ilonen J, Vainionpaa R, Volanen I, Kaitosaari T, Simell O. Development of antibodies against cytomegalovirus, varicella-zoster virus and herpes simplex virus in Finland during the first eight years of life: a prospective study. Scand J Infect Dis. 2003;35:750-3.
86. Rubio-Tapia A, Kyle RA, Kaplan EL, Johnson DR, Page W, Erdtmann F, et al. Increased prevalence and mortality in undiagnosed celiac disease. Gastroenterology. 2009;137:88-93.
87. Catassi C, Kryszak D, Bhatti B, Sturgeon C, Helzlsouer K, Clipp SL, et al. Natural history of celiac disease autoimmunity in a USA cohort followed since 1974. Ann Med. 2010;42:530-8.
88. Lohi S, Mustalahti K, Kaukinen K, Laurila K, Collin P, Rissanen H, et al. Increasing prevalence of coeliac disease over time. Aliment Pharmacol Ther. 2007;26:1217-25.

89. White LE, Merrick VM, Bannerman E, Russell RK, Basude D, Henderson P, et al. The rising incidence of celiac disease in Scotland. Pediatrics. 2013;132:e924-31.
90. Silverberg JI, Norowitz KB, Kleiman E, Silverberg NB, Durkin HG, Joks R, et al. Association between varicella zoster virus infection and atopic dermatitis in early and late childhood: a case-control study. J Allergy Clin Immunol. 2010;126:300-5.
91. Burger JP, Roovers EA, Drenth JP, Meijer JW, Wahab PJ. Rising incidence of celiac disease in the Netherlands; an analysis of temporal trends from 1995 to 2010. Scand J Gastroenterol. 2014;49:933-41.
92. Kang JY, Kang AH, Green A, Gwee KA, Ho KY. Systematic review: worldwide variation in the frequency of coeliac disease and changes over time. Aliment Pharmacol Ther. 2013;38:226-45.
93. Choung RS, Ditah IC, Nadeau AM, Rubio-Tapia A, Marietta EV, Brantner TL, et al. Trends and racial/ethnic disparities in gluten-sensitive problems in the United States: findings from the National Health and Nutrition Examination Surveys from 1988 to 2012. Am J Gastroenterol. 2015;110:455-61.
94. Riddle MS, Murray JA, Porter CK. The incidence and risk of celiac disease in a healthy US adult population. Am J Gastroenterol. 2012;107:1248-55.
95. West J, Fleming KM, Tata LJ, Card TR, Crooks CJ. Incidence and prevalence of celiac disease and dermatitis herpetiformis in the UK over two decades: population-based study. Am J Gastroenterol. 2014;109:757-68.
96. Griffiths PD. Burden of disease associated with human cytomegalovirus and prospects for elimination by universal immunisation. Lancet Infect Dis. 2012;12:790-8.
97. Plotkin SA. Vaccination against cytomegalovirus, the changeling demon. Pediatr Infect Dis J. 1999;18:313-25; quiz 26.
98. Csizmadia CG, Mearin ML, von Blomberg BM, Brand R, Verloove-Vanhorick SP. An iceberg of childhood coeliac disease in the Netherlands. Lancet. 1999;353:813-4.
99. Bibbins-Domingo K, Grossman DC, Curry SJ, Barry MJ, Davidson KW, Doubeni CA, et al. Screening for Celiac Disease: US Preventive Services Task Force Recommendation Statement. JAMA. 2017;317:1252-7.
100. Ivarsson A, Myleus A, Norstrom F, van der Pals M, Rosen A, Hogberg L, et al. Prevalence of childhood celiac disease and changes in infant feeding. Pediatrics. 2013;131:e687-94.
101. Norris JM, Barriga K, Hoffenberg EJ, Taki I, Miao D, Haas JE, et al. Risk of celiac disease autoimmunity and timing of gluten introduction in the diet of infants at increased risk of disease. JAMA. 2005;293:2343-51.
102. Husby S, Koletzko S, Korponay-Szabo IR, Mearin ML, Phillips A, Shamir R, et al. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr. 2012;54:136-60.
103. Aronsson CA, Lee HS, Liu E, Uusitalo U, Hummel S, Yang J, et al. Age at gluten introduction and risk of celiac disease. Pediatrics. 2015;135:239-45.
104. Lionetti E, Castellaneta S, Francavilla R, Pulvirenti A, Tonutti E, Amarri S, et al. Introduction of gluten, HLA status, and the risk of celiac disease in children. N Engl J Med. 2014;371:1295-303.
105. Stordal K, White RA, Eggesbo M. Early feeding and risk of celiac disease in a prospective birth cohort. Pediatrics. 2013;132:e1202-9.
106. Crespo-Escobar P, Mearin ML, Hervas D, Auricchio R, Castillejo G, Gyimesi J, et al. The role of gluten consumption at an early age in celiac disease development: a further analysis of the prospective PreventCD cohort study. Am J Clin Nutr. 2017;105:890-6.
107. Szajewska H, Shamir R, Chmielewska A, Piescik-Lech M, Auricchio R, Ivarsson A, et al. Systematic review with meta-analysis: early infant feeding and coeliac disease--update 2015. Aliment Pharmacol Ther. 2015;41:1038-54.
108. Szajewska H, Shamir R, Mearin L, Ribes-Koninckx C, Catassi C, Domellof M, et al. Gluten Introduction and the Risk of Coeliac Disease: A Position Paper by the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition. J Pediatr Gastroenterol Nutr. 2016;62:507-13.

109. Liu E, Rewers M, Eisenbarth GS. Genetic testing: who should do the testing and what is the role of genetic testing in the setting of celiac disease? Gastroenterology. 2005;128:S33-7.
110. Unalp-Arida A, Ruhl CE, Choung RS, Brantner TL, Murray JA. Lower Prevalence of Celiac Disease and Gluten-Related Disorders in Persons Living in Southern vs Northern Latitudes of the United States. Gastroenterology. 2017;152:1922-32 e2.
111. Liu E, Lee HS, Agardh D. Risk of celiac disease according to HLA haplotype and country. N Engl J Med. 2014;371:1074.
112. Mustalahti K, Catassi C, Reunanen A, Fabiani E, Heier M, McMillan S, et al. The prevalence of celiac disease in Europe: results of a centralized, international mass screening project. Ann Med. 2010;42:587-95.
113. Arnson Y, Amital H, Shoenfeld Y. Vitamin D and autoimmunity: new aetiological and therapeutic considerations. Ann Rheum Dis. 2007;66:1137-42.
114. Mardini HE, Westgate P, Grigorian AY. Racial Differences in the Prevalence of Celiac Disease in the US Population: National Health and Nutrition Examination Survey (NHANES) 2009-2012. Dig Dis Sci. 2015;60:1738-42.
115. Bouziat R, Hinterleitner R, Brown JJ, Stencel-Baerenwald JE, Ikizler M, Mayassi T, et al. Reovirus infection triggers inflammatory responses to dietary antigens and development of celiac disease. Science. 2017;356:44-50.
116. Withoff S, Li Y, Jonkers I, Wijmenga C. Understanding Celiac Disease by Genomics. Trends Genet. 2016;32:295-308.
117. Rosen A, Sandstrom O, Carlsson A, Hogberg L, Olen O, Stenlund H, et al. Usefulness of symptoms to screen for celiac disease. Pediatrics. 2014;133:211-8.
118. Rubio-Tapia A, Hill ID, Kelly CP, Calderwood AH, Murray JA. ACG clinical guidelines: diagnosis and management of celiac disease. Am J Gastroenterol. 2013;108:656-76; quiz 77.
119. Hill ID, Dirks MH, Liptak GS, Colletti RB, Fasano A, Guandalini S, et al. Guideline for the diagnosis and treatment of celiac disease in children: recommendations of the North American Society for Pediatric Gastroenterology, Hepatology and Nutrition. J Pediatr Gastroenterol Nutr. 2005;40:1-19.
120. Werkstetter KJ, Korponay-Szabo IR, Popp A, Villanacci V, Salemme M, Heilig G, et al. Accuracy in Diagnosis of Celiac Disease Without Biopsies in Clinical Practice. Gastroenterology. 2017.
121. Webb C, Norstrom F, Myleus A, Ivarsson A, Halvarsson B, Hogberg L, et al. Celiac disease can be predicted by high levels of anti-tissue transglutaminase antibodies in population-based screening. J Pediatr Gastroenterol Nutr. 2015;60:787-91.
122. Wilson JM, Jungner YG. [Principles and practice of mass screening for disease]. Bol Oficina Sanit Panam. 1968;65:281-393.
123. Godfrey JD, Brantner TL, Brinjikji W, Christensen KN, Brogan DL, Van Dyke CT, et al. Morbidity and mortality among older individuals with undiagnosed celiac disease. Gastroenterology. 2010;139:763-9.
124. Leonard MM, Sapone A, Catassi C, Fasano A. Celiac Disease and Nonceliac Gluten Sensitivity: A Review. JAMA. 2017;318:647-56.
125. Jansen M, van Zelm M, Groeneweg M, Jaddoe V, Dik W, Schreurs M, et al. The identification of celiac disease in asymptomatic children: the Generation R Study. J Gastroenterol. 2017.
126. Lebwohl B, Sanders DS, Green PHR. Coeliac disease. Lancet. 2017.
127. van Koppen EJ, Schweizer JJ, Csizmadia CG, Krom Y, Hylkema HB, van Geel AM, et al. Long-term health and quality-of-life consequences of mass screening for childhood celiac disease: a 10-year follow-up study. Pediatrics. 2009;123:e582-8.
128. Webb C, Myleus A, Norstrom F, Hammarroth S, Hogberg L, Lagerqvist C, et al. High adherence to a gluten-free diet in adolescents with screening-detected celiac disease. J Pediatr Gastroenterol Nutr. 2015;60:54-9.

129. Nevoral J, Kotalova R, Hradsky O, Valtrova V, Zarubova K, Lastovicka J, et al. Symptom positivity is essential for omitting biopsy in children with suspected celiac disease according to the new ESPGHAN guidelines. Eur J Pediatr. 2013.
130. Maki M, Mustalahti K, Kokkonen J, Kulmala P, Haapalahti M, Karttunen T, et al. Prevalence of Celiac disease among children in Finland. N Engl J Med. 2003;348:2517-24.
131. Kurppa K, Paavola A, Collin P, Sievanen H, Laurila K, Huhtala H, et al. Benefits of a gluten-free diet for asymptomatic patients with serologic markers of celiac disease. Gastroenterology. 2014;147:610-7 e1.
132. Kahrs CR, Magnus MC, Stigum H, Lundin KEA, Stordal K. Early growth in children with coeliac disease: a cohort study. Arch Dis Child. 2017.
133. Saari A, Harju S, Makitie O, Saha MT, Dunkel L, Sankilampi U. Systematic growth monitoring for the early detection of celiac disease in children. JAMA Pediatr. 2015;169:e1525.
134. Bingley PJ, Williams AJ, Norcross AJ, Unsworth DJ, Lock RJ, Ness AR, et al. Undiagnosed coeliac disease at age seven: population based prospective birth cohort study. BMJ. 2004;328:322-3.
135. Mustalahti K, Collin P, Sievanen H, Salmi J, Maki M. Osteopenia in patients with clinically silent coeliac disease warrants screening. Lancet. 1999;354:744-5.
136. Cellier C, Flobert C, Cormier C, Roux C, Schmitz J. Severe osteopenia in symptom-free adults with a childhood diagnosis of coeliac disease. Lancet. 2000;355:806.
137. Mahadev S, Gardner R, Lewis SK, Lebwohl B, Green PH. Quality of Life in Screen-detected Celiac Disease Patients in the United States. J Clin Gastroenterol. 2016;50:393-7.
138. Marine M, Fernandez-Banares F, Alsina M, Farre C, Cortijo M, Santaolalla R, et al. Impact of mass screening for gluten-sensitive enteropathy in working population. World J Gastroenterol. 2009;15:1331-8.
139. Shamir R, Hernell O, Leshno M. Cost-effectiveness analysis of screening for celiac disease in the adult population. Med Decis Making. 2006;26:282-93.
140. Hershcovici T, Leshno M, Goldin E, Shamir R, Israeli E. Cost effectiveness of mass screening for coeliac disease is determined by time-delay to diagnosis and quality of life on a gluten-free diet. Aliment Pharmacol Ther. 2010;31:901-10.
141. Rampertab SD, Pooran N, Brar P, Singh P, Green PH. Trends in the presentation of celiac disease. Am J Med. 2006;119:355 e9-14.
142. van Gils T, Rootsaert B, Bouma G, Mulder CJ. Celiac Disease in The Netherlands: Demographic Data of Members of the Dutch Celiac Society. J Gastrointestin Liver Dis. 2016;25:441-5.
143. Lucendo AJ, Garcia-Manzanares A. Bone mineral density in adult coeliac disease: an updated review. Rev Esp Enferm Dig. 2013;105:154-62.
144. Liu E, Dong F, Baron AE, Taki I, Norris JM, Frohnert BI, et al. High Incidence of Celiac Disease in a Long-term Study of Adolescents With Susceptibility Genotypes. Gastroenterology. 2017;152:1329-36 e1.
145. Chou R, Bougatsos C, Blazina I, Mackey K, Grusing S, Selph S. Screening for Celiac Disease: Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. 2017;317:1258-68.
146. Silvester JA, Kurada S, Szwajcer A, Kelly CP, Leffler DA, Duerksen DR. Tests for Serum Transglutaminase and Endomysial Antibodies Do Not Detect Most Patients With Celiac Disease and Persistent Villous Atrophy on Gluten-free Diets: a Meta-analysis. Gastroenterology. 2017.
147. Heikkila K, Pearce J, Maki M, Kaukinen K. Celiac disease and bone fractures: a systematic review and meta-analysis. J Clin Endocrinol Metab. 2015;100:25-34.
148. Larussa T, Suraci E, Nazionale I, Abenavoli L, Imeneo M, Luzza F. Bone mineralization in celiac disease. Gastroenterol Res Pract. 2012;2012:198025.
149. Tau C, Mautalen C, De Rosa S, Roca A, Valenzuela X. Bone mineral density in children with celiac disease. Effect of a Gluten-free diet. Eur J Clin Nutr. 2006;60:358-63.

150. Ciacci C, Maurelli L, Klain M, Savino G, Salvatore M, Mazzacca G, et al. Effects of dietary treatment on bone mineral density in adults with celiac disease: factors predicting response. Am J Gastroenterol. 1997;92:992-6.
151. Troncone R, Kosova R. Short stature and catch-up growth in celiac disease. J Pediatr Gastroenterol Nutr. 2010;51 Suppl 3:S137-8.
152. Frazier AL, Colditz GA, Fuchs CS, Kuntz KM. Cost-effectiveness of screening for colorectal cancer in the general population. JAMA. 2000;284:1954-61.
153. Rosen A, Ivarsson A, Nordyke K, Karlsson E, Carlsson A, Danielsson L, et al. Balancing health benefits and social sacrifices: a qualitative study of how screening-detected celiac disease impacts adolescents' quality of life. BMC Pediatr. 2011;11:32.
154. Leffler DA, Kelly CP. The cost of a loaf of bread in symptomless celiac disease. Gastroenterology. 2014;147:557-9.
155. Rosen A, Emmelin M, Carlsson A, Hammarroth S, Karlsson E, Ivarsson A. Mass screening for celiac disease from the perspective of newly diagnosed adolescents and their parents: a mixed-method study. BMC Public Health. 2011;11:822.
156. Nordyke K, Norstrom F, Lindholm L, Carlsson A, Danielsson L, Emmelin M, et al. Health-related quality-of-life in children with coeliac disease, measured prior to receiving their diagnosis through screening. J Med Screen. 2011;18:187-92.
157. Kooijman MN, Kruithof CJ, van Duijn CM, Duijts L, Franco OH, van IMH, et al. The Generation R Study: design and cohort update 2017. Eur J Epidemiol. 2016;31:1243-64.
158. Tarnok A. Flow cytometry and immune disorders. Cytometry A. 2012;81:819-20.
159. Li R, Scanlon KS, Serdula MK. The validity and reliability of maternal recall of breastfeeding practice. Nutr Rev. 2005;63:103-10.
160. Spratt M, Carpenter J, Sterne JA, Carlin JB, Heron J, Henderson J, et al. Strategies for multiple imputation in longitudinal studies. Am J Epidemiol. 2010;172:478-87.
161. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.

Chapter 14

Summary/Samenvatting

SUMMARY

The early childhood period may reflect a vulnerable time period, during which various determinants 'shape' the immune system, and possibly affect the development of immune-mediated diseases. A competent immune system provides long-lasting protection against harmful antigens, without losing control, which requires a delicate balance between 'too little and too much' tolerance. Aberrant immune protection, however, results in clinical disease, such as autoimmune diseases with an immune response directed against self-antigens, or allergies/hypersensitivities with an excessive reaction to harmless foreign particles. The development of celiac disease autoimmunity (CDA) is an example of such an excessive reaction against dietary gluten. Although considerable knowledge on the etiology of celiac disease (CeD) has been gained with studies focusing on early childhood nutrition, the etiology of CeD is still not fully understood. 40% of the general European population carries the HLA DQ2/DQ8 risk alleles, but only 1% develops CeD, suggesting that additional factors must contribute to pathogenesis. However, besides dietary gluten and the role of breastfeeding, other environmental determinants are relatively understudied.

The first aim of this thesis was to identify determinants of the developing immune system in childhood, focusing predominantly on the adaptive immune compartment. The second aim was to study determinants and the consequences of celiac disease autoimmunity within a population based prospective cohort. These aims were explored within the framework of The Generation R Study, a population-based prospective cohort study from fetal life onwards. The studies in this thesis particularly referred to the childhood period from birth until 6 years of age.

Part I is devoted to determinants that affect the shaping of adaptive immunity in childhood. In **Chapter 2** we studied the dynamics of leukocyte subsets within the innate, naive and memory T- and B-cell lineages in childhood from birth until the age of 6 years, and we studied which environmental determinants affect these dynamics over time. We identified that patterns of leukocyte subsets could be divided in 4 major groups, with functionally similar leukocyte populations following similar kinetics over time. Various non-genetic determinants were identified that influence these patterns, including factors that are related to prenatal maternal lifestyle, maternal immune-mediated diseases, birth characteristics, and postnatal bacterial and viral exposures. Persistent viral infections, esp. cytomegalovirus (CMV) and Epstein Barr Virus (EBV), contributed mostly to the kinetics of memory B and T cells over time. Our data confirms previously described effects of these viral infections on the memory T and B cell compartments, and identified new determinants, including inverse associations between herpes simplex virus 1 (HSV-1) seropositivity and B cell numbers. Our large-scale analyses of all leukocyte subsets and various non-genetic determinants shed new light on the shaping of the adaptive immune system in the first 6 years of life, and enabled identification of various pre- and postnatal determinants that contribute to this shaping.

The effects of breastfeeding on the immune compartments were studied in more detail in **Chapter 3**. Breastfed versus bottle-fed children showed increased cytotoxic CD8+ central memory T(cm) cells at 6 months of age, which was irrespective of the duration of breastfeeding, whereas the numbers of T-cell dependent CD27- IgG+, CD27+ IgM+ and CD27+ IgA+ memory B cell subsets at 6 months of age were inversely related to the duration of breastfeeding. Nevertheless, these effects disappeared after discontinuation of breastfeeding. The transient increase in cytotoxic T cell numbers may suggest a lasting priming of CD8+ T-cell immunity via breastfeeding. Furthermore, breastfeeding may enable stepwise immune maturation, by reducing the need for humoral immune responses, while priming cellular immune responses.

Chapter 4 reveals how CMV and EBV infections affect the memory B and T cell compartments in 6-year-old children. Infected adults were previously reported to have expansions of memory T and B cells, at the expense of naive and Tcm cells. We observed similar virus-associated memory T-cell expansions in children. However, these expansions did not result in loss of naive or Tcm cells, nor in loss of responses to the non-related Measles or Tetanus vaccines. Moreover, children infected before 2 years of age showed a stable control of these persistent viruses with only moderate expansions of effector memory T cells, which is suggestive of better viral suppression with less effector memory T cells upon childhood infection with CMV or EBV. Timing of CMV or EBV infection may be crucial to determine to what extent these herpesviruses have immunomodulatory properties. This is important for our understanding of herpesvirus-associated immunosenescence in elderly. Furthermore, previous literature suggests that persistent infections with CMV differentially affects the host immune phenotype in middle-aged males and females, however whether these differential effects also exist in children remains unclear.

In **Chapter 5** we studied whether sex modifies associations between CMV and multiple B and T cell subsets in 6-year-old children. We observed that sex and CMV independently impacted on multiple B and T cell subsets, however there was no interaction between the two. Importantly, the effects of sex and CMV were in part explained by age and infections with other herpes viruses. Thus, immune aging is likely to be more complex, with involvement of hormonal changes, socioeconomic status, birth characteristics and pathogen exposure.

The associations between specific peripheral T cells, persistent viral infections and the development of subclinical celiac disease (CeD) are discussed in **Chapter 6**. CeD is a digestive and autoimmune disorder driven by an immune response to modified gluten peptides. Affected intestines show infiltrates of various T-cell and NK-cell subsets, but it remains unclear if systemic abnormalities in immune cells are already present in children with subclinical CeD, and when present, whether CMV infection modifies this association. We observed abnormalities in CD57+ cytotoxic T cells and Vδ1+γδ T cells in 6-year old children with subclinical CeD, which were affected by cytomegalovirus seropositivity. Children with subclinical CeD had fewer peripheral Vδ1+ T cells which was predominantly observed

in CMV seronegative children, and more CD57+ T cells than healthy genetic susceptible controls. Hence, children with subclinical CeD have alterations in specific blood T cell subsets that are linked to viral pathology.

In **Part II** we identified determinants that are associated with CDA and persistent herpes virus seropositivity. The timing of gluten introduction and breastfeeding duration in relation to the development of CDA at the age of 6 years is discussed in **Chapter 7.** A delayed introduction of gluten was not associated with an increased risk of CDA. In addition, breastfeeding did not decrease the risk of CDA. Hence, other environmental factors are likely to contribute to the increasing incidence and geographic variability of celiac disease incidence worldwide. Persistent viral infections have been implicated in the etiology of autoimmune diseases in adulthood, but it is not known whether persistent viral infections in childhood are associated with CeD. In **Chapter 8** we described inverse associations between herpesvirus infections and the presence of Transglutaminase type 2 Antibodies. CMV single infection, and combined CMV, EBV and HSV-1 infections were inversely associated with transglutaminase type 2 antibody positivity in children at 6 years of age. Hence, persistent viral infections may have a protective effect in the pathogenesis of CDA.

Ethnic differences in CDA in childhood were described in **Chapter 9.** Within our cohort, we observed a higher incidence of CDA in Western than in non-Western children, and we examined how these differences may be explained by sociodemographic and environmental factors. Following correction for potential differences in the timing of gluten introduction and duration of breastfeeding, we observed that socioeconomic position, daycare attendance and CMV seropositivity in part explained these differences, which may serve as targets for future prevention strategies.

Large differences in infection seroprevalences for CMV, EBV and HSV-1 in children with different ethnic origins were described in **Chapter 10**. The ethnic differences in EBV and HSV-1 could in part (up to 39%) be ascribed to differences in socioeconomic position, and other factors related to lifestyle, which suggests an important role for horizontal transmission for EBV and HSV-1. In contrast, these factors did not explain ethnic differences in CMV seroprevalences in childhood. Together with the observation that breastfeeding was a strong determinant for CMV seropositivity, this suggests an important role for vertical transmission from mother to child.

The consequences of CDA (*TG2A has only been measured on one occasion and was positive*) are described in **Part III**. Screening identified TG2A positive children without gastrointestinal symptoms have lower bone mineral densities, and reduced growth trajectories until they are 6 years old (**Chapter 11**). This suggests that CDA is associated with reduced bone development and growth retardation. Hence, there is a need for early identification of children at risk, which could be made possible through implementation of a screening program for TG2A in young children between 6-10 years of age.

Furthermore, in **Chapter 12** we aimed to describe to what extent CDA in the healthy childhood population can be predictive for the identification of (potential or subclinical)

CeD**.** We observed that serum TG2A screening at 6 years of age in the healthy pediatric population has a positive predictive value of 61% to detect subclinical CeD. We did not find a positive correlation between serum TG2A levels and the degree of enteropathy.

Finally, in **Chapter 13**, a general discussion on the results of this thesis has been described related to recent published studies. In addition, recommendations for future research are provided.

Combined, the studies described in **Part I** of this thesis illustrate the complex, and large age-dependent inter-individual variability of immune maturation during childhood. This variability is largely driven by non-heritable (environmental) factors, of which postnatal herpesvirus acquisition may be considered one of the most profound determinants. Our results stress the ability of the immune system in childhood to suppress a persistent virus, such as herpesviruses, without the negative effects that have been observed in elderly. Follow-up of this cohort in young adults is recommended to observe whether immune modulatory effects are persistent. These studies should include comprehensive analyses on the role of the intestinal microbiome to further unravel the complex interconnecting pathways involved in immune maturation. Furthermore, we provide evidence that the epidemiology of herpesvirus infections has not changed substantially during the last half-decade, and we identified several high risk groups that may serve as targets for future prevention strategies of herpesvirus associated diseases.

The studies described in **Part II** of this thesis underscore that current primary prevention strategies cannot prevent the onset of CeD. Other environmental determinants, such as infections, may play a role in disease onset, but are relatively understudied, and poorly understood. Therefore, the only method to optimize early diagnosis of CeD is secondary prevention. Because active-case finding and targeted screening cannot resolve the high underdiagnosis rate, mass-screening may be a reasonable and realistic alternative. Future follow-up of the Generation R cohort is recommended to identify new onset CeD cases, and to analyze cost-benefit ratios taking short- and long term complications into account. Furthermore, the overall feasibility of mass-screening should be assessed, by studying the acceptance-rate of mass-screening in the general population, among parents and clinicians, the psychological impact, compliance to diagnostic procedures and treatment in those screened positive with inclusion of ethical aspects involved.

SAMENVATTING

De vroege kinderleeftijd reflecteert een gevoelige periode waarin diverse determinanten het immuunsysteem beïnvloeden, en mogelijk ook de ontwikkeling van immuun-gemedieerde aandoeningen later in het leven. Een gezond immuunsysteem biedt langdurige bescherming tegen schadelijke antigenen, zonder dat het immuunsysteem de controle verliest. Dit vereist een gezonde balans tussen "te weinig" en "te veel" tolerantie. Afwijkende immuunreacties daarentegen kunnen resulteren in verschillende klinische ziektebeelden, zoals auto-immuunziektes waarbij de afweer is gericht tegen antigenen van het eigen lichaam, of allergieën/ hypersensitiviteiten waarbij de afweer is gericht tegen niet-eigen onschadelijke allergenen. De ontwikkeling van coeliakie auto-immuniteit is een voorbeeld van een multifactoriële aandoening, waarbij een excessieve reactie bestaat tegen gluten uit het dieet. Hoewel aanzienlijke kennis over de etiologie van coeliakie is bewerkstelligd met onderzoek dat zich richt op de rol van voeding op de vroege leeftijd, is de etiologie van coeliakie nog steeds niet helemaal opgehelderd. 40% van de algemene Europese populatie draagt een van de HLA DQ2/ DQ8 allelen, maar slechts 1% ontwikkelt coeliakie. Deze bevindingen suggereren dat aanvullende factoren een rol moeten spelen in de pathogenese van coeliakie. Echter, naast voeding in het vroege leven, zoals de introductie van gluten en borstvoeding, zijn andere mogelijke omgevingsdeterminanten relatief weinig bestudeerd.

Het eerste doel van dit proefschrift was het in kaart brengen van determinanten van het ontwikkelende immuunsysteem op de kinderleeftijd. We waren hoofdzakelijk gericht op het beschrijven van de normale ontwikkeling van het immuunsysteem, waarbij onze aandacht uitging naar het adaptieve, ofwel verworven, immuunsysteem. Ons doel was om te beschrijven hoe de normale immuun-ontwikkeling op de kinderleeftijd verloopt als reactie op diverse antigenen die het immuunsysteem na de geboorte tegenkomt, met specifieke aandacht voor het effect van borstvoeding en persisterende herpesvirus infecties. Deze herpesvirussen blijven na infectie latent in het lichaam aanwezig, en oefenen grote druk uit op het immuunsysteem. Het tweede doel was om de determinanten van coeliakie auto-immuniteit te bestuderen, met specifieke aandacht voor de rol van voeding in het eerste levensjaar, en het doormaken van persisterende herpesvirus infecties. Daarnaast was ons doel om de consequenties van ongediagnosticeerde coeliakie auto-immuniteit op de kinderleeftijd te bestuderen. De doelstellingen van dit proefschrift werden onderzocht binnen de Generation R Studie, een populatie-gebaseerd prospectief bevolkingsonderzoek in Rotterdam waarin kinderen gevolgd worden vanaf het foetale leven. De studies beschreven in dit proefschrift richten zich voornamelijk op de vroege kinderleeftijd; de periode vanaf de geboorte tot op 6-jarige leeftijd.

Deel I is toegewijd aan determinanten die het adaptieve immuunsysteem vorm geven. In **Hoofdstuk 2** hebben we bestudeerd hoe de aantallen van diverse leukocyten subtypes in het bloed binnen de aangeboren, naïeve en geheugen B- en T cellijnen veranderen vanaf de geboorte tot op 6-jarige leeftijd. Daarnaast hebben we bestudeerd hoe omgevingsfac-

toren deze dynamische patronen over de tijd beïnvloeden. Op basis van de dynamiek in cel aantallen over de tijd hebben we 4 verschillende patronen geïdentificeerd, waarbij functioneel gelijke leukocyten populaties dezelfde kinetiek volgden. We toonden aan dat diverse niet-genetische determinanten deze patronen beïnvloeden, inclusief factoren die gerelateerd zijn aan de levensstijl van moeder tijdens de zwangerschap, maternale immuun-gemedieerde ziektebeelden, perinatale factoren, en postnatale blootstelling aan bacteriële en virale ziekteverwekkers. Persisterende virale infecties, met name cytomegalovirus (CMV) en Epstein Barr Virus (EBV), waren de voornaamste factoren die de kinetiek van geheugen B en T cellen op de kinderleeftijd beïnvloeden. Onze data bevestigen voorgaande onderzoeken die de effecten van deze virale infecties beschrijven op de geheugen B en T cellen. Daarnaast toonden we nieuwe associaties aan, zoals (inverse) relaties tussen herpes simplex virus 1 (HSV-1) seropositiviteit en naïeve B cel aantallen. Onze grootschalige analyse van alle leukocyten populaties, waarbij we de invloed onderzochten van vele non-genetische determinanten, biedt inzicht in hoe de normale ontwikkeling van het adaptieve immuunsysteem verloopt onder invloed van vele omgevingsfactoren, en brengt in kaart welke omgevingsfactoren de meeste invloed hebben op een gezonde immunologische ontwikkeling.

Het effect van borstvoeding op het immuunsysteem werd in meer detail onderzocht in **Hoofdstuk 3**. Borst-gevoede kinderen hadden ten opzicht van kinderen die kunstvoeding kregen een vergrote populatie van cytotoxische CD8+ centrale memory cellen T(cm) op de leeftijd van 6 maanden. Dit effect was onafhankelijk van de duur van borstvoeding. Daarentegen werden T-cel afhankelijke B-geheugencel populaties (CD27- IgG+, CD27+ IgM+ and CD27+ IgA+), kleiner naarmate de duur van borstvoeding toenam. Desondanks verdwenen deze effecten na het stoppen van borstvoeding. De tijdelijke vergroting in cytotoxische T cel populaties suggereert een langdurige ”priming” van CD8+ T cellen via borstvoeding. Daarnaast, suggereren deze bevindingen dat borstvoeding een stapsgewijze immunologische ontwikkeling mogelijk maakt doordat er tijdelijk minder humorale immuniteit gevormd hoeft te worden, waardoor de vorming van cellulaire immuniteit wordt gestimuleerd.

Hoofdstuk 4 onthult in meer detail hoe CMV en EBV infectie aantallen van T en B geheugencellen beïnvloedt in 6-jarige kinderen. In voorgaande studies is aangetoond dat in een selectie van CMV- en EBV geïnfecteerde ouderen een expansie optreedt van infectie specifieke T en B geheugencellen, ten koste van naïeve en Tcm cellen, en ten koste van de afweerrespons tegen vaccinaties. We hebben aangetoond dat CMV en EBV infectie werden geassocieerd met een toename van verschillende typen geheugen T-cellen op 6-jarige leeftijd. Echter, in tegenstelling tot in ouderen, resulteerden deze expansies in kinderen niet in verlies van naïeve of Tcm cellen, of in verlies van vaccinatie responsen tegen de niet-gerelateerde ziekteverwekkers mazelen of tetanus. Bovendien lieten kinderen die voor het 2^{e} levensjaar waren geïnfecteerd met CMV of EBV een kleinere expansie zien van geheugen effector T cellen, dan kinderen die na het 2^{e} levensjaar waren geïnfecteerd. Deze bevindingen suggereren dat vroege infectie er voor kan zorgen dat er minder effector geheugen T cellen nodig zijn, en er daardoor alleen

een milde verschuiving in het T cel repertoire ontstaat. Het tijdstip van CMV en EBV infectie kan daarom cruciaal zijn voor de mate waarin deze herpesvirussen immuunmodulatoire effecten bewerkstelligen. Deze bevindingen zijn belangrijk voor het begrijpen van herpesvirus geassocieerde immuunveroudering in ouderen.

In de literatuur is beschreven dat persisterende infecties met CMV het immunologische fenotype verschillend beïnvloeden in mannen en vrouwen op middelbare leeftijd. Het is echter niet bekend of dit verschil ook aanwezig is op de kinderleeftijd. In **Hoofdstuk 5** hebben we daarom bestudeerd of geslacht de associaties tussen CMV en aantallen van meerdere typen T en B cellen modificeert in 6-jarige kinderen. We vonden dat geslacht en CMV onafhankelijk van elkaar de T en B cel aantallen beïnvloedden, maar dat er geen interactief effect aanwezig was tussen CMV en geslacht. De onafhankelijke effecten van geslacht en CMV werden gedeeltelijk verklaard door leeftijd en infecties met andere herpesvirussen. Het is dus aannemelijk dat immuunveroudering complexer verloopt dan gedacht. Ook hormonale factoren, socio-economische status, perinatale factoren en postnatale blootstelling aan verscheidene pathogenen zijn hoogstwaarschijnlijk van invloed op dit proces.

De associaties tussen specifieke perifere T cellen, persisterende virale infecties en de ontwikkeling van subklinische coeliakie auto-immuniteit worden bediscussieerd in **Hoofdstuk 6**. Coeliakie, ofwel glutenenteropathie, is een auto-immuun gemedieerd ziektebeeld met inflammatoire kenmerken ter hoogte van de mucosa van het duodenum. Dit geeft in typische gevallen aanleiding tot het ontstaan van de kenmerkende villeuze atrofie en hyperplasie van de crypten. Deze worden gedreven door een immuunrespons tegen gemodificeerde glutenpeptides. De aangedane mucosa bevat infiltraten van vele typen T- en NK-cellen, maar het is onbekend of deze aantallen ook afwijkend zijn in het bloed van kinderen met subklinische coeliakie. Daarnaast is het onbekend in hoeverre CMV infectie deze associaties beïnvloeden. We vonden afwijkende aantallen van CD57+ cytotoxische T cellen en Vδ1+γδ T cellen in 6-jarige kinderen met subklinische coeliakie, welke beïnvloed werden door cytomegalovirus seropositiviteit. Kinderen met subklinische coeliakie hadden minder Vδ1+ T cellen in het bloed, en dit effect was voornamelijk zichtbaar in CMV seronegatieve kinderen. Daarentegen waren er meer CD57+ T cellen aanwezig in het bloed in vergelijking met gezonde kinderen die de genetische predispositie voor coeliakie hebben. Deze bevindingen suggereren dat kinderen met subklinische coeliakie verschuivingen in specifieke aantallen T cellen in het bloed laten zien die geassocieerd kunnen worden met virale pathologie.

In **Deel II** hebben we determinanten geïdentificeerd die geassocieerd zijn met coeliakie auto-immuniteit en herpesvirus seropositiviteit. Het tijdstip van gluten introductie en de duur van borstvoeding in relatie tot de ontwikkeling van coeliakie auto-immuniteit worden bediscussieerd in **Hoofdstuk 7.** We toonden aan dat een late introductie van gluten (>6 maanden) niet geassocieerd was met een hoger risico op coeliakie auto-immuniteit. Daarnaast toonden we aan dat borstvoeding het risico op de ontwikkeling van coeliakie auto-immuniteit niet verlaagt. Het is daarom aannemelijk dat andere omgevingsfactoren

bijdragen aan de stijgende incidentie en geografische variabiliteit in de incidentie van coeliakie wereldwijd.

Persisterende virale infecties spelen mogelijk een rol in de etiologie van verschillende auto-immuunziekten in volwassenen, maar het is onduidelijk of persisterende virale infecties op de kinderleeftijd zijn geassocieerd met coeliakie. In **Hoofdstuk 8** beschrijven we een omgekeerde associatie tussen herpesvirus seropositiviteit en de aanwezigheid van antistoffen tegen transglutaminase type 2. CMV infectie, en de combinatie van CMV met EBV en HSV-1, waren omgekeerd geassocieerd met de aanwezigheid van antistoffen tegen transglutaminase type 2 in kinderen op 6-jarige leeftijd. Dit suggereert dat persisterende virale infecties een beschermende rol kunnen spelen in de pathogenese van coeliakie auto-immuniteit.

Etnische verschillen in coeliakie auto-immuniteit zijn beschreven in **Hoofdstuk 9.** We observeerden een hogere incidentie van coeliakie auto-immuniteit in Westerse dan in niet-Westerse kinderen in ons cohort, en we onderzochten hoe deze verschillen konden worden verklaard door socio-demografische en omgevingsfactoren. Na correctie voor potentiele verschillen in het tijdstip van gluten introductie en de duur van borstvoeding, vonden we dat etnische verschillen gedeeltelijk werden verklaard door sociaal economische verschillen, verschillen in bezoek aan kinderdagverblijf, en CMV seropositiviteit. Deze factoren kunnen mogelijk dienen als aangrijpingspunten in toekomstige preventiestrategieën.

In **Hoofdstuk 10** toonden we grote verschillen aan in seroprevalenties voor CMV, EBV en HSV-1 tussen kinderen van verschillende etnische achtergronden. De etnische verschillen in EBV en HSV-1 konden gedeeltelijk (tot en met 39%) toegeschreven worden aan verschillen in socio-economische positie, en andere factoren die gerelateerd zijn aan de levensstijl. Deze bevinden suggereren dat er een belangrijke rol is weggelegd voor horizontale transmission van EBV en HSV-1, terwijl deze factoren de etnische verschillen in CMV seroprevalentie op de kinderleeftijd niet verklaarden. Dit suggereert, gezamenlijk met de bevinding dat borstvoeding een sterke determinant was voor CMV seropositiviteit, een belangrijke rol voor verticale transmissie van moeder op kind.

De consequenties van coeliakie auto-immuniteit (gedefinieerd als *eenmaal TG2A positiviteit*) zijn beschreven in **Deel III**. TG2A positieve kinderen zonder gastro-intestinale symptomen die door middel van screening werden geïdentificeerd, hadden een verminderde botdichtheid, en lieten retrospectief een vertraagde groei zien tot op de leeftijd van 6 jaar (**Hoofdstuk 11**). Deze bevindingen suggereren dat CDA is geassocieerd met groeiretardatie en een verminderde botontwikkeling, en laten zien dat er een noodzaak bestaat voor vroege identificatie van kinderen die een verhoogd risico hebben op het ontwikkelen van coeliakie. Dit kan mogelijk gemaakt worden door implementatie van een screeningsprogramma in kinderen van 6-10 jaar oud.

Daarnaast was ons doel om te beschrijven in hoeverre coeliakie auto-immuniteit in de algemene kinderpopulatie voorspellend kan zijn voor het identificeren van kinderen met (potentiele of subklinische) coeliakie (**Hoofdstuk 12)**. We toonden aan dat serum TG2A

screening een positief voorspellende waarde heeft van 61% voor het detecteren van subklinische coeliakie in een gezonde populatie van 6-jarige kinderen. We vonden echter geen positieve correlatie tussen serum TG2A waarden en de ernst van de enteropathie.

Tenslotte beschrijven we in **Hoofdstuk 13** een algemene discussie over de resultaten in dit proefschrift in relatie tot andere studies die recent gepubliceerd zijn. Ook worden aanbevelingen gegeven voor vervolgonderzoek.

Gezamenlijk illustreren de studies in **Deel I** van dit proefschrift de complexe, en leeftijdsafhankelijke –interindividuele variabiliteit van het ontwikkelende immuunsysteem op de jonge kinderleeftijd. Deze variabiliteit wordt met name gedreven door niet-genetische (omgevings-) factoren, in het bijzonder door postnatale herpesvirus infecties. Onze resultaten benadrukken het vermogen van het immuunsysteem van kinderen om een infectie met een persisterende virus, zoals de herpesvirussen onder controle te houden, zonder dat dit gepaard gaat met negatieve effecten zoals die beschreven worden in ouderen. We adviseren vervolgonderzoek in adolescenten om te bestuderen of de geobserveerde immuunmodulatoire effecten persisteren. Deze studies moeten idealiter uitgebreide analyses bevatten die ook de effecten van het intestinale microbioom omvatten, zodat de samenhangende compartimenten van het ontwikkelende immuunsysteem beter in kaart gebracht kunnen worden. Daarnaast suggereert onze data dat de epidemiologie van herpesvirusinfecties niet wezenlijk veranderd is gedurende de afgelopen 50 jaar, en we identificeren risicogroepen voor primaire preventie van herpesvirus geassocieerde aandoeningen.

De studies in **Deel II** van dit proefschrift benadrukken dat de huidige primaire preventieve strategieën niet toereikend zijn om de ontwikkeling van coeliakie te voorkomen. Andere omgevingsdeterminanten, zoals infecties, kunnen een rol spelen in het ontstaan van coeliakie, maar deze zijn onduidelijk en relatief onder bestudeerd. De enige manier om vroege diagnose van coeliakie te optimaliseren is daarom secundaire preventie. Omdat actieve opsporing van coeliakie, en gerichte screening van risicogroepen, de hoge onder diagnose cijfers niet kan terugdringen, is algehele populatiescreening wellicht een redelijk en realistisch alternatief. Vervolgonderzoek van het Generation R cohort wordt aanbevolen om het aantal nieuwe coeliakie diagnoses in kaart te brengen, en om een kosteneffectiviteitsanalyse uit te voeren, met inachtneming van gevolgen op zowel de korte als lange termijn. Daarnaast moet de algemene haalbaarheid van populatiescreening worden onderzocht, door te inventariseren hoe acceptabel screening is onder ouders en clinici, de psychologische impact van screening, de acceptatie om diagnostische procedures te ondergaan, de compliance van het glutenvrije dieet in positief gescreende kinderen, en de ethische aspecten hiervan.

Chapter 15

Appendices

I. List of abbreviations

II. Author's affiliations

III. List of Publications

IV. About the author

V. PhD portfolio

VI. Dankwoord

I LIST OF ABBREVIATIONS

AAP	American Association of Pediatrics
Anti-TPO	Anti-thyroid peroxidase IgG
aOR	Adjusted odds ratio
APC	Antigen presenting cell
AD	Atopic dermatitis
β	Beta; regression coefficient
BCR	B-cell receptor
BMC	Bone mineral content
BMD	Bone mineral density
BMI	Body Mass Index
CCR	Chemokine receptor
CI	Confidence Interval
CeD	Celiac disease
CDA	Celiac disease autoimmunity
CMV	Cytomegalovirus
CR	Complement receptor
CSR	Class switch recombination
CV	Coefficient of variation
DC	Dendritic cell
DGPA	Deamidated gliadin peptide antibody
DNA	Desoxyribonucleic acid
DXA	Dual-energy x-ray absorptiometry
EAACI	European Society of Allergy and Clinical Immunology
EBV	Epstein Barr Virus
EmA	Endomysial antibody
ESPGHAN	European Society for Pediatric Gastroenterology, Hepatology and Nutrition
FSC	Forward scatter
FFQ	Food frequency questionnaire
g	Gram
GA	Gestational age
GEE	Generalized Estimating Equations
GI	Gastrointestinal tract infections
GLMM	Generalized linear mixed model
H. Influenzae	Haemophilus Influenzae
HLA	Human leukocyte antigen
HSV-1	Herpes simplex virus -1
Ig	Immunoglobulin
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IgM	Immunoglobulin M

IL	Interleukin
INF	interferon
IQR	Interquartile range
ISAAC	International Study of Asthma and Allergy in Childhood
kg	Kilogram
L	Liter
LRTI	Lower respiratory tract infections
m	Meter
MAR	Missing at random
MCAR	Missing completely at random
MCMC	Markov Chain Monte Carlo
MHC	Major histocompatibility complex
MNAR	Missing not at random
M. cattarhalis	Moraxella catarrhalis
MEC	Medical Ethical Committee
mg	Milligram
ml	Milliliter
mm	millimeter
MRI	Magnetic resonance imaging
n	Number (sample size)
NK cell	Natural killer cell
nmol	Nanomol
OR	Odds ratio
PCA	Principal component analyses
RR	Relative risk
RTI	Respiratory tract infections
S. aureus	Staphyococcus Aureus
SD	Standard deviation
SDS	Standard deviation score
SES	Socioeconomic status
SEP	Socioeconomic position
SGA	Small for gestational age
SHM	Somatic hypermutation
S. pneumoniae	Streptococcus pneumoniae
SPSS	Statistical package social sciences
SSC	Side scatter
Tc	Cytotoxic T cell; CD8+ T cell
Tcm	Central memory T cell
TCR	T-cell receptor
TD	T cell-dependent
Tem	Effector memory T cell
TemRA	CD45RO- effector memory T cells

TemRO	CD45RO+ effector memory T cells
TG2A	Tissue transglutaminase type 2 antibody (IgA)
Th	Helper T cell; CD4+ T cell
TI	T cell-independent
TNF	Tumour necrosis factor
Treg	Regulatory T cell
TV	Television
URTI	Upper respiratory tract infections
VCA	Viral capsid antigen
VZV	Varicella Zoster Virus
WFH	Weight for height
WHO	World Health Organization

II AUTHOR'S AFFILIATIONS

The Generation R Study Group, Erasmus Medical Center, Rotterdam, the Netherlands.
Ilse I.M. Tromp, Vincent W.V. Jaddoe, Albert Hofman, Hein Raat, Selma H. Bouthoorn, Jessica C. Kiefte-de Jong, Sytske A. Beth, Romy Gaillard, Johan C. de Jongste, Henriëtte A. Moll.

Department of General Pediatrics, Erasmus Medical Center, Rotterdam, the Netherlands.
Vincent W.V. Jaddoe, Henriëtte A. Moll

Department of Immunology, Erasmus Medical Center, Rotterdam, the Netherlands
Menno C. van Zelm, Diana van den Heuvel, Herbert Hooijkaas, Jacques J.M.van Dongen, Wim A. Dik, Marco W.J. Schreurs, Halima Bouallouch-Charif, E.G. van Lochem, L.E. Bakker-Jonges.

Department of Immunology and Pathology, Central Clinical School, Monash University, Melbourne, Victoria, Australia
Menno C. van Zelm

Department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands.
Vincent W.V. Jaddoe, Albert Hofman, Jessica C. Kiefte-de Jong, Romy Gaillard

Department of Pediatric Gastroenterology, Erasmus Medical Center, Rotterdam, the Netherlands.
Johanna C. Escher

Department of Pediatric Gastroenterology, Maasstad Hospital, Rotterdam, the Netherlands.
Michael Groeneweg

Department of Pediatric Pulmonology, Erasmus Medical Center, Rotterdam, the Netherlands.
Johan C. de Jongste

Department of Public Health, Erasmus Medical Center, Rotterdam, the Netherlands.
Hein Raat, Selma H. Bouthoorn

Department of Biostatistics, Erasmus Medical Center, Rotterdam, the Netherlands.
Kazem Nasserinejad

III LIST OF PUBLICATIONS

1) **Jansen MAE**, Tromp IIM, Kiefte-de Jong JC, Jaddoe VWV, Hofman A, Escher JC, Hooijkaas H, Moll H.A. Infant feeding and anti-tissue transglutaminase antibody concentrations in the Generation R Study. Am J Clin Nutr 2014 Oct;100(4):1095-101.

2) **Jansen MAE**, Kiefte-de Jong JC, Gaillard R, Escher JC, Hofman A, Jaddoe VWV, Hooijkaas H, Moll HA. Growth Trajectories and Bone Mineral Density in Anti-Tissue Transglutaminase Antibody-positive Children: The Generation R Study. Clin Gastroenterol Hepatol 2015 May;13(5):913-20.e5.

3) **Jansen MAE**, van den Heuvel D, van Zelm MC, Jaddoe VWV, Hofman A, de Jongste JC, Hooijkaas H, Moll HA. Decreased memory B cells and increased CD8 memory T cells in blood of breastfed children: The Generation R Study. PLoS One 2015; May 18;10(5)**:**e0126019.

4) Van den Heuvel D, **Jansen MAE**, Dik WA, Bouallouch-Charif H, Zhao D, van Kester K, Smits-te Nijenhuis MAW, Kolijn-Couwenberg MJ, Jaddoe VWV, Arens R, van Dongen JJM, Moll HA, van Zelm MC. Cytomegalovirus- and Epstein-Barr Virus-Induced T-cell Expansions in Young Children Do Not Impair Naive T-cell Populations or Vaccination Responses: The Generation R Study. J Infect Dis. 2016 Jan 15;213(2):233-42.

5) **Jansen MAE**, van den Heuvel D, Bouthoorn SH, Jaddoe VWV, Hooijkaas H, Raat H, Fraaij PLA, van Zelm MC, Moll HA. Determinants of Ethnic differences in Cytomegalovirus, Epstein-Barr Virus, and Herpes Simplex Virus Type 1 Seroprevalence in Childhood. J Pediatr. 2016 Mar;170:126-34.e1-6.

6) **Jansen MAE**, van den Heuvel D, van der Zwet KVM, Jaddoe VWV, Hofman A, Escher JC, Fraaij PLA, Hooijkaas H, van Zelm MC, Moll HA. Herpesvirus infections and Transglutaminase Type 2 Antibody Positivity in Childhood. The Generation R Study. J Pediatr Gastroenterol Nutr. 2016 Oct;63(4):423-30.

7) Beth SA, **Jansen MAE**, Elfrink MEC, Kiefte-de Jong JC, Wolvius EB, Jaddoe VWV, van Zelm MC, Moll HA. Generation R birth cohort study shows that specific enamel defects were not associated with elevated serum transglutaminase type 2 antibodies. Acta Paediatr. 2016 Oct;105(10):e485-91.

8) Gervin K, Page CM, Aass HCD, **Jansen MAE**, Fjeldstad HE, Andreassen BK, Duijts L, van Meurs JB, van Zelm MC, Jaddoe VWV, Nordeng H, Knudsen GP, Magnus P, Nystad W, Staff AC, Felix JF, Lyle R. Cell-type specific DNA methylation in cord blood: A 450K reference data set and cell count-based validation of estimated cell-type composition. Epigenetics. 2016 Sep;11(9):690-698.

9) Van den Heuvel D, **Jansen MAE**, Bell A, Rickinson AB, Jaddoe VWV, van Dongen JJM, Moll HA, van Zelm MC. Transient reduction in IgA+ and IgG+ memory B-cell numbers in young EBV-Seropositive children. The Generation R study. J Leukoc Biol. 2017 Apr;101(4):949-956.

10) Van den Heuvel D, **Jansen MAE**, K Nasserinejad, Dik WA, van Lochem EG, Bakker-Jonges LE, Bouallouch-Charif H, Jaddoe VWV, Hooijkaas H, van Dongen JJM, Moll HA, van Zelm MC. Effects of nongenetic factors on immune cell dynamincs in early childhood: The Generation R Study. J Allergy Clin Immunol. 2017;139(6): 1923-1934.e17.

11) **Jansen MAE**, Beth SA, Kiefte-de Jong JC, Raat H, Jaddoe VWV, van Zelm MC, Moll HA. Ethnic differences in coeliac disease autoimmunity in childhood. The Generation R Study. Arch Dis Child. 2017 Jun; 102(6):529-534

12) **Jansen MAE**, van den Heuvel D, Jaddoe VWV, Moll HA, van Zelm MC. No Interactive Effects of Sex and Persistent Cytomegalovirus on Immune Phenotypes in Young Children. The Generation R Study. J Infect Dis. 2017 Mar 15;215(6):883-888**.**

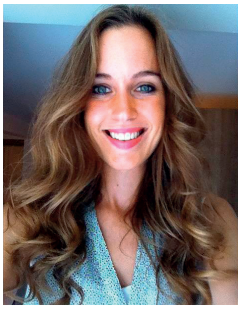
13) **Jansen MAE**, van den Heuvel D, Jaddoe VWV, van Zelm MC, Moll HA. Abnormalities in CD57+ cytotoxic T cells and Vδ1+γδ T cells in subclinical Celiac Disease in Childhood are affected by Cytomegalovirus. Clin Immunol. 2017 Apr 26. pii: S1521-6616(17)30278-4.

14. **Jansen MAE**, van Zelm MC, Groeneweg M, Jaddoe VWV, Dik WA, Schreurs MWJ, Hooijkaas H, Moll HA, Escher JC. The Identification of Celiac Disease in Asymptomatic Children. J. Gastroenterol. 2016 June 6. [Epub ahead of print].

15. Looman KIM, **Jansen MAE**, Voortman T, van den Heuvel D, Jaddoe VWV, Franco OH, van Zelm MC, Moll HA. The Role of Vitamin D on circulating Memory T cells in Children: The Generation R Study. Pediatr Allergy Immunol. 2017 Sep;28(6):579-587

16. Looman KIM, **Jansen MAE**, Voortman T, van den Heuvel D, Jaddoe VWV, Franco OH, van Zelm MC, Moll HA. Fatty acids and the Maturation of Memory T cells in Children: The Generation R Study. In progress

IV ABOUT THE AUTHOR



Michelle Anne Elisa Jansen was born on april 14th 1985 in Deventer, the Netherlands. In 2003 she completed secondary school at the Carmel College Salland in Raalte, Overijssel. In the same year, she started her study Biomedical Sciences at the University of Utrech, and in 2004, she started her Medicine Study at the Erasmus University of Rotterdam. During her study, she participated in several research projects in the Sophia Children's Hospital. In 2008, she received, as one of two selected students of her year, a 6 months research scholarship at Johns Hopkins Medical Hospital, in Baltimore, USA. Under supervision of prof. Dr. Marilee C. Allen, she worked on her MSc research topic 'preterm brain injury in neonates'. During her internships (2009-2011) she spent 7 weeks at the neonatology and pediatric intensive care unit (NICU, PICU) in the Auckland City Hospital in New Zealand (supervisor Dr. Jane Alsweiler). In april 2011 she obtained her medical docter (MD) degree (cum laude) at the Erasmus University Rotterdam. From april 2011 she started working as a medical doctor (MD) and PhD student at the Generation R Study, department of general pediatrics, Immunology and Epidemiology (supervisor Prof. Dr. Henriëtte A. Moll and Dr. Menno C. van Zelm). Her research was focused on early determinants shaping adaptive immunity and celiac disease autoimmunity in childhood, of which the results are presented in this dissertation. In 2013 she obtained a MSc degree in Clinical Epidemiology (Netherlands Institute for Health Sciences, Erasmus University Rotterdam) and she worked as clinical epidemiologist and MD onwards. In 2013 and 2014, she was awarded a fellowship from the Dutch society of gastroenterology and hepatology for attendance of the international conference in Jerusalam, she received the first price of the European Pediatric young investigator award from the European Society for Pediatrics in Barcelona, and she obtained a research funding from the Dutch Organization NutsOhra. From April-September 2015 she worked as a pediatric resident (ANIOS) at the Maasstad Hospital in Rotterdam (Supervisor Dr. M. Groeneweg), and from September 2015 she returned back to her PhD project to complete her research while focusing fulltime on the project. In 2016 and 2017 she worked as a resident (ANIOS) at the department of pediatric psychiatry, Lucertis, Rotterdam, and at the outdoor department eating disorders of Rintveld, Altrecht, Zeist). From March 2018 onwards she hopes to start with the general practitioner training at the University Medical Center Utrecht (UMCU).

In her spare time, Michelle likes to sport, travel, read, golf- and kitesurf, and spend time with family and friends. Michelle is married to Paul Didden (1981), a gastroenterologist working at the University Medical Center Utrecht (UMCU). They have one daughter Velize Mila (October 2016).

V PHD PORTFOLIO

Name PhD student:	Michelle Anne Elisa Jansen
Erasmus MC Department:	Pediatrics, Immunology, Generation R
Research School:	Netherlands Institute for Health Sciences
Promotor:	Prof. dr. H.A. Moll
Copromotor:	Dr. Menno C van Zelm
PhD period:	April 2011 - December 2017

1.PhD training	**Year**	**Workload (ECTS)**
Master's degree Health Sciences, specialization		
Clinical Epidemiology, NIHES, Erasmus University	2011-2013	120
General courses		
Principles of Research in Medicine and Epidemiology	2011	0.7
Clinical Decision Analysis	2011	0.7
Methods of Public Health Research	2011	0.7
Topics in Meta-analysis	2012	0.7
Pharmaco-epidemiology	2011	0.7
Markers and Prognostic Research	2011	0.7
The Practice of Epidemiologic Analysis	2011	0.7
Conceptual Foundation of Epidemiologic Study Design	2012	0.7
Cohort Studies	2012	0.7
Case-control studies	2012	0.7
Introduction to Global Public Health	2012	0.7
Social Epidemiology	2012	0.7
Core courses		
Biostatistical Methods: Basic Principles	2011	5.7
Biostatistical Methods II: Popular Regression Models	2011	4.3
Study Design	2012	4.3
Clinical Epidemiology	2012	5.7
Methodologic Topics in Epidemiologic Research	2012	1.4
Advanced courses		
Epidemiology of Infectious Diseases	2012	1.4
Repeated Measurements in Clinical Studies	2012	1.4
Courses for the Quantitative Researcher	2012	1.4
Women's Health	2013	0.9

General academic skills		
English Language	2011	1.4
Introduction to Medical Writing	2011	1.1
In depth courses		
Biomedical English Writing and communication	2012	4.0
Advanced Molecular Immunology, Postgrad course, Molmed Erasmus MC	2013	1.4
Seminars, workshops and symposia		
Generation R research meetings	2011-2015	2.6
Seminars department of Epidemiology	2011-2015	1.0
TULIPS Grant writing and presenting weekend, NVK	2013	0.6
(Inter)national presentations at conferences and congresses		
-General pediatrics research meeting - poster presentation	2012	0.4
-Sophia Children's hospital research day - poster presentation	2012	0.4
-Developmental Origins of Health and Disease (DOHaD), Rotterdam, poster presentation	2012	0.4
-European Society for pediatrics and infectious diseases (ESPID), Thessaloniki, Greece, poster presentation	2012	0.4
-European Academy of Pediatrics (EAP), Lyon, France, 2 oral presentations	2013	2.8
-European Society for Pediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN) Jerusalem, Israel, oral presentation	2014	1.4
-European Academy of Pediatric Society (EAPS) Barcelona, 2 oral presentations	2014	2.8
-Pediatric Gastroenterology Meeting (SMDL-NVK), Utrecht oral presentation	2014	1.4
-European Society for Pediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN), poster presentation	2015	0.4
-Dutch society for pediatrics - oral presentation	2015	1.4
Obtained fundings		
Fonds NutsOhra 'Celiac Disease; the tip of the iceberg'	2013	1.4
M*race* Doelmatigheid 'Coeliakie, een gemiste diagnose?'	2014	0.4
Other (o.a. Stichting Coolsingel, Nederlandse coeliakie vereniging)	2014-2015	1.4
Other		
Peer review of articles for scientific journals *(JAMA pediatrics, Journal of Pediatric Gastroenterology and Nutrition)*	2014-2016	1.0

2.Teaching activities	**Year**	**Workload (ECTS)**
Supervising Master's theses medical students:		
-'Herpes virusinfections and childhood celiac disease'	2013	2.0
Mw. Kirsten van der Zwet, MSc.		
-"Herpes virusinfections and childhood asthma symptoms'	2013	2.0
Mw. Noortje Haenen, MSc.		
- 'Vitamin D and childhood immune phenotypes' and	2016-	4.0
Maternal prenatal fatty acids and childhood immune phenotypes	2017	
Mw. Kirsten I.M. Looman, MSc.		

VI DANKWOORD

Tot slot wil ik graag de belangrijkste personen bedanken zonder wie dit proefschrift nooit tot stand gekomen zou zijn. Allereerst, alle kleine deelnemers en ouders: zonder jullie zou dit onderzoek er nooit zijn geweest.

Ten tweede, prof.dr. H.A. Moll, promotor van dit proefschrift en hoogleraar kindergeneeskunde, lieve Henriette, bedankt voor je vertrouwen in mij, de vrijheid die je me gaf om mijn wensen en ideeën uit te werken, en om waar nodig bij te sturen en van een kritische noot te voorzien. Je kan een ingewikkeld concept helder samenvatten en direct de vertaalslag maken naar de klinische praktijk, waardoor ik altijd vol inspiratie verder kon werken. Als 'mama' van jouw onderzoeksgroep reik jij alle handvaten aan die een bloeiende onderzoeksgroep nodig heeft: inspiratie, energie en vertrouwen. Tijdens onze vele overlegmomenten was er niet alleen ruimte voor wetenschappelijk inhoudelijke onderwerpen, ook was er ruimte voor gezelligheid en wist je de juiste vragen te stellen waardoor ik me geen fijnere promotor kon wensen. Veel dank hiervoor.

Ten derde, dr. Van Zelm, Beste Menno, wat was het fijn om met je te kunnen 'sparren'. Je nam de tijd voor nieuwe ideeën en dacht vanuit een hele andere, verfrissende hoek, waardoor je onze manuscripten naar een hoger niveau tilde. Ik vond het inspirerend om te merken dat jij vaak linksom dacht en ik rechtsom (of andersom), waardoor ik getriggerd werd om op een andere manier te denken. Ik heb veel van je geleerd, zowel van de 'diehard' immunologische onderwerpen als op tekstueel gebied. Er zijn maar weinig onderzoekers die zo snel denken en zo goed kunnen schrijven als jij, wat mede maakt dat ik veel bewondering voor je heb als onderzoeker. Sorry dat ik zoveel eisen stelde, en af en toe mijn ongeduld iets te veel liet blijken, zeker op het laatst..! Je hield me terecht een spiegel voor, dankjewel hiervoor. Succes 'down under', geniet van het mooie leven daar; maar met wat ik heb gezien via skype komt dat zeker goed! Wellicht tot over een paar jaar!

Ten vierde gaat mijn dank uit naar de experts in de kleine commissie, prof. dr. Escher, prof. Dr. Hein Raat, en prof Dr. Frits Koning. Bedankt voor het kritisch beoordelen van mijn proefschrift. Beste Hankje, dankjewel dat je niet alleen secretaris van de promotiecommissie wilt zijn, maar me ook nuttige klinische feedback hebt gegeven over de coeliakie stukken. Dankzij jouw goede zorg en alle scopieën die je hebt uitgevoerd samen met Michael, heb je het onderzoek klinische betekenis gegeven. Ook kijk ik met veel plezier terug naar onze gezellige tijd in Israël. Daarnaast wil ik ook graag prof. dr. Harry de Koning, Dr. Anton Langerak en Dr. Pieter Fraaij bedanken voor het zitting nemen in de commissie. Pieter, dank voor je laagdrempelige contact en je nuttige feedback op de manuscripten.

Ten vijfde, mijn cruciale schakel naar de immunologie: Diana van den Heuvel. Lieve, Diana, inmidels ook doctor, wat waren het bewogen jaren, met af en toe wat downs, maar bovenal veel ups! Zonder jou was dit proefschrift er nooit geweest. Dankzij jouw enorme hoeveelheid werk hebben we beiden ons proefschrift kunnen afmaken. Bedankt voor onze vele gezellige klets- en koffie momentjes. Inmiddels ben je postdoc in Leiden, volgens mij

zit je goed op je plek! Met jouw enorme doorzettingsvermogen weet ik zeker dat je een mooie toekomst tegemoet gaat! Ten zesde, wil ik graag een speciaal woord richten aan Prof. Dr. Hooijkaas: Beste Herbert, veel dank dat ik, ook dankzij jouw vertrouwen in mij, aan dit project kon starten. Ik heb met veel plezier met je samengewerkt. De immunologie was soms lastig, maar door jouw uitleg werd het toch enigszins behapbaar, en dat is een kunst. Door een noodlottige gebeurtenis, ik was ontdaan bij het horen van dit bericht, kon je je werkzaamheden helaas niet voortzetten, een gemis voor de afdeling. Ik wens je heel veel sterkte toe, en hoop dat je met een goed gevoel terugkijkt op de mooie dingen die we hebben bereikt. Ook wil ik graag alle co-auteurs van de manuscripten bedanken: Prof. dr. Vincent Jaddoe, beste Vincent, dank voor je snelle blik op de manuscripten. Je leverde altijd nuttige en helder commentaar waar ik van kon leren, dank voor de fijne en gezellige samenwerking. Prof. dr. A. Hofman, prof. dr. H. Raat, Prof. dr. O. Franco, en prof. dr. J.C. de Jongste: bedankt voor het kritisch beoordelen van mijn manuscripten.

Een speciaal woord van dank wil ik richten richten aan mijn dierbare vriendin Sanne Beth. Lieve Sanne, bedankt dat je mijn lieve vriendin bent. Je bent een bijzonder warm, sterk en lief persoon. Dat je daarnaast ook nog eens mijn collega'tje werd, was natuurlijk een extra fijne en leuke bijkomstigheid. Maar bovenal wil ik je graag laten weten dat jouw vriendschap heel veel voor me betekent. Je hebt het niet altijd makkelijk: weet dat ik veel bewondering voor je heb en veel waarde hecht aan onze vriendschap. Je bent een fantastische moeder van een lief, klein ventje: jouw grootste en liefste wens. Zo bijzonder dat onze kleintjes samen opgroeien, en dat we ook deze periode in ons leven kunnen delen met elkaar. Ik hoop op nog heel veel jaren samen!

Daarnaast de vele collega's van Generation R. Het gaat te ver om alle mensen op te noemen die afgelopen jaren voor mij tot een feestje maakten, maar ik wil graag enkele personen noemen die me in het bijzonder hebben geraakt. Lieve Sel en An; wat was het gezellig met jullie, we waren al snel vriendinnen, door jullie voelde GenR als thuis. Marjolein en Zoe; fijn om op maandag eerst een lekkere doppio met jullie te doen! Dank voor de gezelligheid! Ankie, door jou werd ik nieuwsgierig naar een plekje bij Henriette, dankjewel voor je tip, en de vele gezellige borrels tijdens onze coschappen inmiddels heel wat jaren geleden. Volgens mij worden we nu echt oud..!;) Lieve Jes, wat een geluk dat ik als groentje bij jou op de kamer kwam. Je bent de meest slimme analyticus ever! Dankjewel dat ik zoveel van je mocht leren. Lieve Zoe, heerlijke positieveling, jij maakte me altijd aan het lachen! Inmiddels ben je in opleiding tot gynaecoloog, wat ben je een goede dokter, en niet op de laatste plaats ook moeder van een lief meisje met een - net zo - vrolijke lach: zo moeder, zo dochter!:) Lieve Ilse, inmiddels ook doctor! Wat fijn dat je samen met mij een stuk wilde schrijven, dat ging van een "leien dakje", dankjewel voor de gezellige momentjes! Marjolein en Claudia, de immuno datasets waren soms "verschrikkelijk" maar altijd reageerden jullie razendsnel op mijn verzoeken; het leek nooit te veel waar ik om vroeg; veel dank hiervoor! Zonder jullie hulp zou ik nergens zijn! Mijn studenten Kirsten, Noortje en Kirsten: we hebben een aantal mooie manuscripten kunnen schrijven: bedankt voor

jullie inzet! Kirsten, je gaat nu zelf ook aan de slag: veel succes, je bent een enorm attente en supersnelle leerling! Was fijn om met je te werken! Ook wil ik graag de vele personen bedanken die me hebben geholpen om de flowcytometrie data in het lab te verzamelen: Marja, Marion, Nicole, Halima en Kevin, en ongewtijfeld vergeet ik nog een heleboel meer personen: Sorry dat alles -maar dan ook echt alles- overnieuw moest. Ik heb bewondering voor de enorme hoeveelheid werk die jullie verzet hebben om dag- in- dag uit- weer aan de slag te gaan met die complexe analyses, net zo lang tot onze ogen vierkant werden van het scherm. Bedankt voor jullie onuitputtelijke energie, zonder jullie hulp waren die analyses nog steeds niet af! Diana Dufour, bedankt dat je alle TG2A levels zo snel hebt bepaald! Marco Schreurs en Wim Dik: bedankt voor jullie immuno-methodologische bijdrage aan onze manuscripten. Tevens gaat mijn dank uit naar de collega's in het Maasstad ziekenhuis die mijn tijd op de kindergeneeskunde bijzonder hebben gemaakt. Carsten, dankjewel dat je me in contact bracht met Henriette. Michael, dankjewel voor je betrokkenheid, en uiteraard voor al je werk om alle scopieën uit te voeren. "The intuitive mind is a sacred gift, and the rational mind a faithful servant." Dat had je goed aangevoeld! Dank ook aan Jennifer voor je adviezen, verhalen en luisterend oor! Tevens dank voor alle gezellige (afdelings-) uitjes met onze focusdames, medepromovendi uit het Sophia, Epidemiologie en Immunologie: helaas te veel om bij naam te noemen. Astrid, dank dat je altijd alle afspraken en documenten tot in de puntjes voor me hebt geregeld. Rosa, Pety en Naomi; jullie hebben gezorgd voor een fijne plek bij Lucertis! Dank voor jullie betrokkenheid, ook tijdens mijn zwangerschap. Mijn nieuwe collega's Linne, Oriana en Maaike: Ik werk nog maar kort bij jullie, maar dankzij jullie met veel plezier!

Ten slotte, mijn lieve vriendinnen Anna, Jorie, Sanne, Margaux, Wendy, Wendelien, Lalini en Margreeth: jullie maakten mijn tijd in Rotterdam, en de vele jaren erna, extra warm. Het blijft altijd een gevoel van 'thuiskomen' als ik weer samen met jullie ben. Lieve Margaux, wat fijn dat je terugkomt uit Curaçao: aftellen geblazen; ik heb je gemist! En je kleine knappe hartendief(jes) ook! Lieve Melis, wat hebben we een gezellige tijd samen gehad (Baltimore, en iets met een emmer..;)) ons contact is me nog steeds heel dierbaar! Lieve Manon, wat straalde je onlangs: prachtige bruid! Al op de peuterspeelzaal en basisschool deden we volgens mij lippenstift op die we kregen van..?;) Lieve Loet, bijzonder dat we alle ontwikkelingen in ons leven samen delen. Je bent een prachtige, sterke vrouw, fantastische moeder, crea bea, optimist, en altijd lieve, oprechte en geïntereseerde vriendin: aan een half woord hebben we genoeg. Ik zie onze meisjes volgend jaar al samen zandtorentjes bouwen in het zand bij Rapa Nui..J

Lieve (schoon)familie The-Didden; bedankt voor jullie zorg en betrokkenheid, en de vele lekkere indische verassingen die jullie altijd meebrengen. Lieve Im en "brabo G": heerlijke wereldreizigers! Gelukkig was ik, dankzij jullie verhalen, ook weer een beetje op ontdekkingsreis. Maar nog veel fijner dat jullie weer thuis zijn! Bedankt dat jullie op Vtje hebben gepast, voor de momenten dat jullie hebben getimmerd en gezaagd in ons huis en voor de nodige likjes verf her en der!;) Lieve sis, ik weet niet wat ik zonder jou had gemoeten. Be-

dankt voor je adviezen, gezelligheid, luisterend oor, en warme kaartjes de afgelopen jaren. Ik kan me geen liever zusje wensen. Lieve pap en mam, jullie zijn van onschatbare waarde, Vtje is zo ontzettend op jullie gesteld, klusjes oppakken zonder dat we er om vragen; niets is jullie te veel. Maar nog meer ben ik jullie dankbaar voor alle mogelijkheden die jullie me hebben aangereikt. Jullie onvoorwaardelijke en onuitputtelijke steun en liefde draag ik altijd bij me, ik kan me geen fijnere ouders wensen.

Als laatste gaat mijn dank uit naar mijn man Paul. Lieve Paul, met jou naast me is het zoveel leuker! Je bent mijn 'rock', je hebt ons een prachtige dochter geschonken, we zijn inmiddels 3 huizen verder, een paar steden verder, een paar banen verder; we did it! Ik hoop op nog veel meer mooie avonturen met jou samen! Lieve Velize, ik heb het laatste gedeelte van dit proefschrift geschreven in de uurtjes dat je lag te slapen. Stiekem hoop ik dat je niet te snel groeit, want ik geniet zo van jouw kleine en vrolijke aanwezigheid. Ik hoop dat je altijd zult voelen dat je het meest prachtige geschenk bent dat ik me kan wensen!