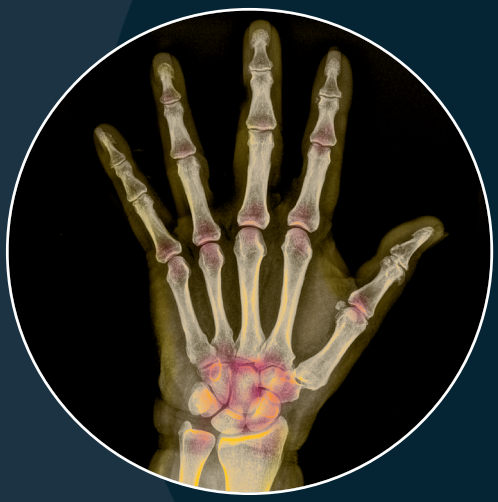
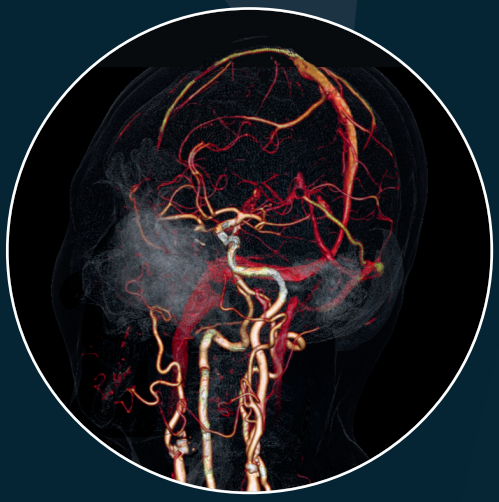




LEANDER R. BUISMAN

EARLY COST-EFFECTIVENESS OF MEDICAL TESTS IN RHEUMATOID ARTHRITIS AND ISCHEMIC STROKE

Early cost-effectiveness of medical tests in rheumatoid arthritis and ischemic stroke

Leander R. Buisman

Cover design and layout: SuvaalOntwerp, Rotterdam
Cover photo (right): Centre Cardio-Thoracique de Monaco, Monaco
Printed by Optima, Rotterdam

ISBN 9789070116743

Early Cost-Effectiveness of Medical Tests in Rheumatoid Arthritis and Ischemic Stroke

Vroege kosteneffectiviteit van medische testen
in reumatoïde artritis en ischemische beroerte

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

donderdag 30 juni 2016 om 13.30 uur

door

Leander Robert Buisman
geboren te Schiedam

Erasmus University Rotterdam

PROMOTIECOMMISSIE

PROMOTOR

Prof.dr. M.P.M.H. Rutten-van Mölken

OVERIGE LEDEN

Prof.dr. J.L. Severens
Prof.dr. J.W.J. Bijlsma
Prof.dr. M.J.A.P. Daemen

COPROMOTOR

Dr. W.K. Redekop

CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

On January 30, 2015, President Obama of the United States launched the Precision Medicine Initiative in his State of the Union speech.[1] He introduced precision medicine as 'one of the greatest opportunities for new medical breakthroughs that we have ever seen'[1] and emphasized that the promise of precision medicine involves 'delivering the right treatments, at the right time, every time to the right person'.[1] The goal of precision medicine is to enable healthcare providers to identify people to individualized treatment and prevention strategies based on people's unique characteristics, including genetics, microbiome composition, lifestyle, health history, and environment factors.[1-3] Many new medical tests are currently being developed, such as tests and biomarkers in genomics, cytomics, metabolomics, imaging, and biomechanical analysis based on imaging information to improve the identification of patients for appropriate and optimal treatments. The aim of these tests is to classify individuals or patients into different subpopulations that differ in the risk or prognosis of a specific disease, or response to treatment. Preventive or therapeutic interventions can then be targeted to the subpopulations that will benefit most from these interventions. This is what precision medicine entails.[1-3]

The value of a medical test is dependent on the qualities of available treatments. Effective prevention or treatment should be available to prevent patients from getting a specific disease or to slow down disease progression (i.e., a new medical test with exceptional diagnostic accuracy can be considered useless if the available treatments are not effective in any patient subpopulation). On the other hand, the availability of a relatively cheap and sufficiently effective preventive or therapeutic intervention can render a test useless since all patients could receive it without being tested. In general, the value of a test is greatest when available treatments are effective but also have important shortcomings (e.g., high costs, high risks, or variation in effectiveness between patients). If the test's sensitivity and specificity are not 100%, the test (depending on the application of the test) diagnoses,

identifies the prognosis, or predicts treatment response for a selection of patients incorrectly, resulting in inappropriate treatments. Inappropriate treatments may result in poorer health outcomes for patients (e.g., complications due to unnecessary treatments or poorer health status as a consequence of forgone treatment). In addition, a medical test itself can also be invasive, resulting in a disutility and burden to patients. Overall, the use of new tests with improved test characteristics improves the use of treatments.

The effectiveness and cost-effectiveness of new test strategies should be assessed before these strategies can be used in daily clinical practice. When assessing the impact of new test strategies on quality of life over time, the effectiveness is usually measured in terms of quality-adjusted life years (QALYs). The QALY combines the number of life years with the level of health-related quality of life (i.e., utilities) in those years.[4] Since the budget to spend on healthcare is limited, motivated choices are required before investing the scarce resources. Therefore, new medical tests should demonstrate good value for money. A cost-effectiveness analysis (CEA) helps to identify how the scarce resources available for healthcare can be allocated optimally given the health outcomes and costs of alternative test strategies.[5] The results of a CEA provide evidence of the relative impact of one strategy versus another.

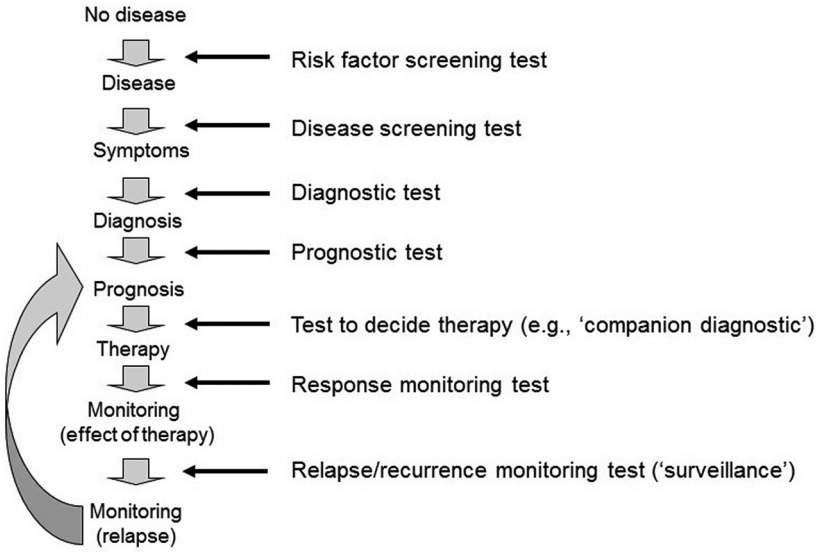
In healthcare systems worldwide, CEAs are mostly performed in the late stages of medical test development, mainly by medical test manufacturers/developers (referred to as test developers) to demonstrate to payers (i.e., healthcare insurers, governments, managed care organizations) that a test is good value for money and offers substantial return on investment.[6-10]

To test developers, adequate reimbursement of new tests is important for wide implementation in clinical practice, which improves return on investment. Medical tests are most often reimbursed as part of a Diagnosis Related Group (DRG) payment or fee-for-service. Hence, those who actually decide about reimbursement include opinion leaders among the clinicians, hospital managers, government authorities and healthcare insurers. Given the serious consequences of negative reimbursement decisions, test developers take a considerable risk when conducting cost-effectiveness analyses only in the late stages of medical test development, when large R&D investments have already been made. Therefore, evaluation of new tests in the early stages of development (early-CEA) has attracted increasing attention.

Early-CEAs guide the further development process of new medical tests, and set realistic performance-price goals.[9] Over the past two decades, most early-CEAs focused on new drug therapy.[11] The increasing development of medical tests for use in various phases of disease prevention and treatment as part of the promise of precision medicine, and the increasing use of companion diagnostics to personalized medicine call for early-CEAs to assess whether these tests have the potential to improve health outcomes and re-

duce overall costs. Early-CEAs are iterative by nature and should be performed throughout the different stages of test development, since new data and ideas may emerge regarding its potential application, target population, test performance (i.e., sensitivity and specificity), and costs. The continuous integration of these new insights and evidence that arise through feedback during the test development may convince developers to return to earlier development steps, if necessary, and will result in more informed decisions by test developers about its potential application in the healthcare system and target population. Early-CEAs are also highly relevant to other stakeholders. To clinicians, such evaluations can identify which specific patient subpopulations will benefit in terms of improved health outcomes and reduced costs from using a new test to guide treatment decisions; as noted above, this is part of precision medicine. Early-CEAs can therefore help to signal promising new tests early resulting in faster adoption of tests. Furthermore, the involvement of patients as stakeholders in the development process guarantees that patients have full benefits of new and useful tests as well as faster access to these tests. Moreover, hospital managers can use early-CEAs to investigate the budgetary possibilities to implement new tests. In addition, the demonstration of the potential cost-effectiveness of using new tests in clinical practice to government authorities and healthcare insurers will guide informed decisions about the reimbursement of these tests.

Figure 1.1: Various applications of medical tests in the healthcare system



The vertical arrows represent the different phases of disease and its treatment (based on Redekop and Uyl-de Groot[12]).

A new medical test can have a variety of applications. Figure 1.1 shows the various applications of medical tests and can be used to explore where along the patient care pathway a test could be used. Redekop and Uyl-de Groot[12] describe these various applications, following the sequence of screening/case finding, diagnosis, disease progression and treatment.

This thesis consists of two case studies of medical tests for which the application of the tests differs in both case studies. The first case study focusses on new diagnostic tests to detect early rheumatoid arthritis (RA). The second case study focusses on new prognostic tests to assess the risk of a recurrent ischemic stroke event in patients with a recent transient ischemic attack (TIA) or minor ischemic stroke. These two case studies will be introduced now.

1.1 Case study 1: Diagnostic tests for rheumatoid arthritis

Epidemiology

Rheumatoid arthritis (RA) is a chronic inflammatory disease characterized by structural irreversible joint damage and chronic synovitis leading to severe disability, serious loss of quality of life, and premature mortality.[13-18] Heritability for RA is estimated at around 50%-60%. It is regarded as a complex disease in which several genes play a role.[19] Lifestyle factors also contribute to the risk of developing RA, including smoking, alcohol intake, coffee intake, and vitamin D status.[20-21] The prevalence of RA increases with age and is highest in women older than 65 years.[22] The global prevalence of RA was estimated to be 0.24% in the Global Burden of Disease 2010 study with an almost three times higher prevalence among women than men (0.35% versus 0.13%, respectively).[23] Geographical differences in RA frequency were also found, with the disease more common in high-income regions where the prevalence among adults is 0.5-1.0% with a maximum incidence between ages 35 to 55.[24]

Burden of disease

RA presents a substantial burden to patients in terms of reduced quality of life, premature mortality, decreased work participation, and income loss.[24-34] RA has also an impact on society with reduced total health, productivity losses, and increased healthcare utilization.[35]

Quality of life

From a patient's perspective, RA has an impact on physical functioning and to a lesser extent on mental functioning, which both affect the participation of patients in daily activities.[25] When quality of life is expressed in utilities, RA contributes to an average loss of 0.1-0.3 in quality of life compared to the general population.[26]

Mortality

Several studies have demonstrated increased mortality in RA patients compared to the general population. The standardized mortality ratios vary between 1.28 and 2.98 per 100,000 life years depending on age, gender, geographical location, and disease activity status.[27-34] The increased mortality is caused by a 3 times higher risk of myocardial infarction and 6 times higher risk of silent myocardial infarction than the general population.[24] Furthermore, cancer is the second most common cause of mortality amongst RA patients. Infections and tuberculosis are also comorbidities of RA patients which may result from the drug treatment (e.g., an increased risk of tuberculosis is found in patients using Anti-TNF therapy[33,34]).

Costs

The costs for RA patients include direct healthcare and non-healthcare costs and indirect costs of productivity losses. Franke et al.[36] calculated the mean annual costs per patient based on 26 cost-of-illness studies. Most of these studies were conducted in Western Europe and the costs were annualized and converted to 2006 PPP Euros. The annual healthcare costs were on average €4,170 per patient (including drug costs) in which outpatient costs contribute for more than two thirds to the costs. The annual costs of informal care are on average €2,284 per patient.

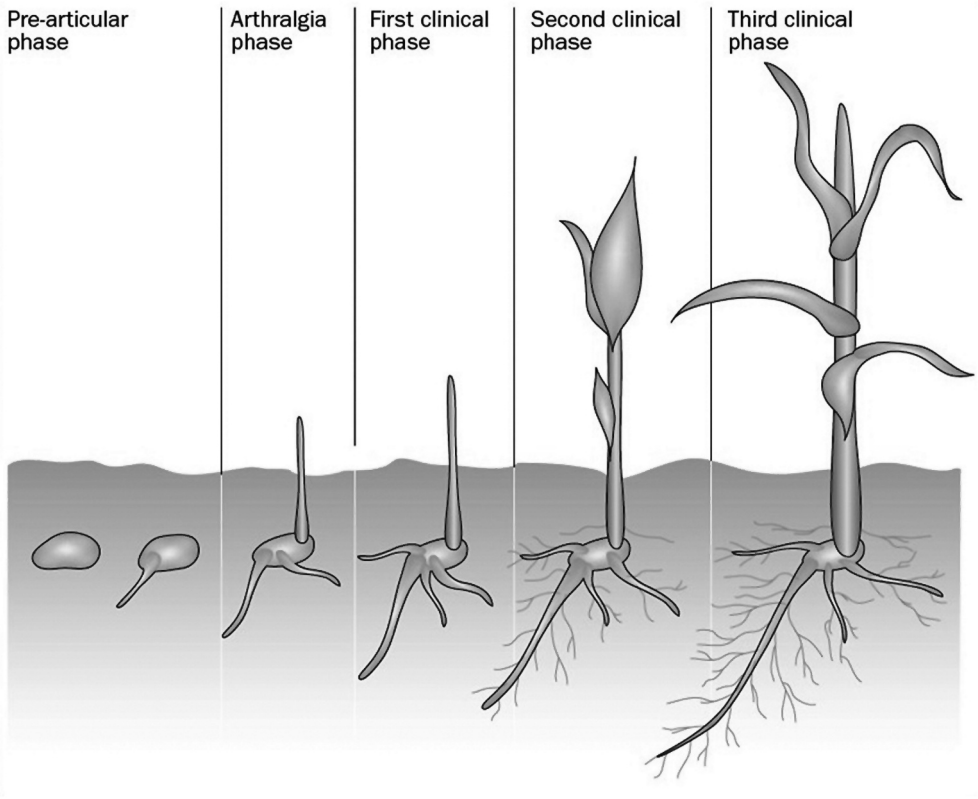
Since RA affects the working-age population, patients with RA have high productivity losses due to absence from paid work and presenteeism. Presenteeism is defined as being present at work, but not able to function optimally. The productivity losses result in huge societal productivity costs or indirect costs. Based on a review of 26 studies, the mean annual productivity costs per patient were estimated using two different measurement approaches: €1,441 (friction cost approach) and €8,452 (human capital approach).[36] The friction cost approach only considers short term losses with a maximum friction period of 160 days.[37] The assumption is that after 160 days the employee is substituted by another. On the other hand, the human capital approach accounts for all productivity losses till the retirement age, since the assumption is that each person in society has the potential to contribute.

Development of RA from pre-clinical to clinical disease

The development of RA from its earlier stages can be characterized by a metaphor of a growing plant from its seed to a mature plant described by Hazes and Luime[38] and shown in Figure 1.2. In the pre-articular stage, the genetic factors for the development of RA are present, but symptoms of RA are not. From this stage, patients can develop arthralgia in the following stage in which clinically objective synovitis is not present. In the first clinical stage, the first signs of synovitis and symptoms indicate the presence of in-

flammatory arthritis but the type of arthritis might be unclassifiable and hard to distinguish from other inflammatory diseases. RA develops in the second clinical stage and the arthritis is classifiable as RA in the last clinical stage.

Figure 1.2: Development of RA characterized by a metaphor of a growing plant



This figure is reproduced from Hazes and Luime[38] in agreement with the journal.

Diagnosis of RA

The American College of Rheumatology (ACR) and the European League Against Rheumatism (EULAR) developed classification criteria for RA. Although these ACR/EULAR 2010 RA classification criteria were initially developed for the classification of RA, they perform well in a diagnostic setting for early inflammatory arthritis patients (i.e., first clinical stage in Figure 1.2).[39] The ACR/EULAR 2010 RA classification criteria should be applied to patients who have at least one swollen joint which is not explained by another disease.[18] The application of the criteria provides a score between

0 to 10, with a score of 6 or higher indicating the presence RA. Patients with a score below 6 do not fulfil the ACR/EULAR 2010 RA classification criteria, but might be diagnosed as having RA in the future. Patients classify as intermediate-risk with a score of 3-5 and low-risk with a score of 0-2 on ACR/EULAR 2010 RA classification criteria. The score is based on the combination of (i) joint evaluation, (ii) a history of symptom duration, (iii) at least one serological test [i.e., rheumatoid factor (RF) and anti-citrullinated protein antibody (ACPA)], and (iv) one acute-phase response measure [i.e., erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP)].

Potential value of new diagnostic tests

For each patient, the onset, course of inflammation, rate of progression of joint damage, and response to therapy are different, making treatment of RA difficult. Therefore, it is generally acknowledged that patient tailored treatment is needed. However, optimal tests for diagnosis of RA in specific patient subpopulations (e.g., intermediate-risk patients) are not yet available. This case study focusses on diagnostic tests to detect early RA. As a result of early detection with the new diagnostic tests, patients at risk of having RA start early treatment and will have an improved prognosis. Therefore, several diagnostic tests are currently being developed to improve the identification of patient subpopulations at risk of having RA in inflammatory arthritis patients. By use of (a combination of) these tests, prediction rules for cost-effective treatment should improve individualized treatment, preventing unnecessary delay of treatment, side effects, and high costs of biologic treatment.

New diagnostic tests could be used as a replacement of the current diagnostic test strategy or as an add-on to the current test strategy for specific patient subpopulations. Based on REACH data*, the current diagnostic test strategy classifies around 45% of all tested inflammatory arthritis patients as high-risk of having RA, about 50% as intermediate-risk, and the rest as low-risk. Of the intermediate-risk patients, a large proportion might develop RA in the near future. Intermediate-risk patients would therefore benefit from more accurate new diagnostic tests that are able to select patients for early start of optimal treatment.

1.2 Case study 2: Prognostic tests for recurrent ischemic stroke

Epidemiology

Stoke is a collective noun for all strokes due to either ischemia or hemorrhage.[40] A hemorrhagic stroke is defined as a bleeding within the closed cranial cavity, resulting in insufficient supply of oxygen and nutrients to the brain. An ischemic stroke is defined as a focal neurological deficit of sudden

* REACH is the Rotterdam Early Arthritis Cohort with 552 inflammatory arthritis who were suspected of having RA. Details about the cohort can be found in Alves et al.[39]

onset of presumed vascular origin, lasting at least 24 hours, and may have typical signs of brain infarction with brain imaging. Patients with an ischemic stroke might have had a TIA which is defined as a focal neurologic deficit of sudden onset lasting less than 24 hours and with no signs of recent infarction based on CT or MRI scans. Globally, the proportion of strokes caused by ischemia is 68% (versus 32% due to hemorrhage).[41] In the United States, these proportions are 87% ischemic and 13% hemorrhage strokes, reflecting a proportionally higher incidence of ischemic strokes in high-income countries.[42] The remainder of this case study description focuses on *ischemic stroke*.

From 1990 to 2010, the absolute number of ischemic stroke events increased by 37%.[41] However, due to a larger increase in the world's population than the absolute number of ischemic stroke events since 1990, the incidence of ischemic stroke has not changed much (nonsignificant decrease by 2%). In 2010, 11.5 million people (176 per 100,000 life years) had an incident ischemic stroke worldwide compared to 7.2 million people (181 per 100,000 life years) in 1990.[41] Furthermore, geographical differences have been observed with a significant (13%) reduction in incidence in high-income countries, compared to a possible (though nonsignificant) increase in incidence of 6% in low- and middle-income countries.[41] The incidence of ischemic stroke increases significantly with age.[41] In addition, men have a higher incidence than women at younger ages; the pattern is different amongst people aged≥75, where the incidence is higher among women.[42] Other important risk factors of ischemic stroke are diabetes mellitus, hypertension, high cholesterol, heart disease, smoking, physical inactivity, family history and genetics, and a previous TIA or ischemic stroke event.[42]

Burden of disease

Stroke is the second most common cause of death and the third most common cause of disability worldwide.[43] Advances in prevention and treatment of TIA and ischemic stroke during the past two decades have resulted in significant changes in medical practice, such as the organization of stroke care into integrated stroke services, use of intravenous thrombolysis, and advances in secondary prevention.[44-49] These advances have contributed to reduced stroke-related mortality and increased survival rates. However, survivors of an ischemic stroke can experience severe health loss, increased healthcare utilization, decreased work participation, and income loss which results in substantial healthcare and societal costs.

Quality of life

The impact of an ischemic stroke on quality of life is dependent on the severity of the ischemic stroke and complications due to the ischemic stroke event. Patients with an ischemic stroke can have medical complications[50]

(e.g., urinary tract infection, constipation, or pneumonia), neurological complications,[51] or post-stroke depression.[52] Several studies have expressed the quality of life by severity of the ischemic stroke in utilities. For example, Hallan et al.[53] report utilities of patients after a TIA, minor ischemic stroke, and major ischemic stroke of 0.88, 0.71, and 0.31, respectively.

Mortality

Worldwide, the absolute annual number of people who died from an ischemic stroke in 2010 (2.8 million people) has increased by 21% since 1990.[41] Despite this absolute increase, the mortality ratios for ischemic stroke have actually decreased worldwide from 1990 to 2010 (from 58 to 42 per 100,000 life years) due to an increased world population.[41] In addition, the mortality ratios for ischemic stroke increase with age; people aged≥75 have a more than 5-fold higher mortality risk than people aged 65-74.[41]

Costs

The costs for patients who experienced an ischemic stroke include direct healthcare, non-healthcare and indirect costs of productivity losses. The direct healthcare costs include costs of stroke care in hospitals, rehabilitation centers, long-term care facilities, nursing home care, and general practitioners.

Several studies have found total stroke-related costs (including hospital costs, other healthcare costs, and productivity costs) to be substantially higher in the first year after ischemic stroke onset than in subsequent years.[54,55] Costs of patients with a minor and major ischemic stroke in the Netherlands from these old studies have been converted to 2014 Euros using the consumer price index.[56] Costs of minor ischemic stroke were €8,731 in the first year and €1,494 in subsequent years. Costs of major ischemic stroke were €49,790 in the first year and €29,072 in subsequent years. However, these cost studies were performed with resource use data from more than 10 years ago before major improvements such as integrated stroke services were implemented in stroke care.

Prognosis of ischemic stroke

From 12 hours to 7 days after ischemic stroke onset, patients without complications experience moderate improvement in neurological impairments.[57] In the 3 to 6 months after ischemic stroke onset, patients experience the greatest improvement, with some further improvement up to 18 months.[58]

The prognosis of patients with an ischemic stroke is dependent on a number of determinants, including stroke severity,[58] age,[59,60] stroke-related complications, comorbidities, and infarct volume.[61] Furthermore, medical interventions such as thrombolysis, integrated stroke services and rehabilitation help to improve the outcome of ischemic stroke.

Patients with a recent TIA or minor ischemic stroke are at increased risk of a recurrent ischemic stroke. A meta-analysis of 11 observational studies found that the risk of recurrent stroke at 2 days, 30 days, and 90 days after TIA was 3.5%, 8.0%, and 9.2%, respectively.[62] In addition, a prospective cohort study of 2,447 patients found that the risk of recurrent ischemic stroke was highest shortly after TIA or minor ischemic stroke (about 5%), decreased to about 1% at three years, and then progressively increased over the remainder of the 10-year follow-up (1-2%).[63]

Potential value of new prognostic tests

Patients with a recent TIA or minor ischemic stroke are at increased risk of a recurrent ischemic stroke, which may be caused by a plaque rupture in the carotid artery. Surgery (carotid endarterectomy) can reduce this risk, but the procedure is invasive and can lead to death and morbidity. A systematic review and meta-analysis showed that the estimated risk of death was 1.1% and risk of non-fatal ischemic stroke 6.6% for symptomatic patients undergoing CEA.[64] Better stroke risk prediction would therefore help to determine which patients should undergo carotid endarterectomy. At present, there is a limited ability to predict plaque rupture in the carotid artery and subsequently select patients for carotid endarterectomy based on rupture risk. Advances in non-invasive molecular imaging tests, such as contrast enhanced magnetic resonance imaging (CE-MRI), duplex ultrasonography (DUS),[65] CT angiography (CTA), and biomechanical analysis based on imaging information[66] may help to improve the ability to predict future strokes in patients with a recent TIA or minor ischemic stroke. The improved stroke risk prediction can help to select patients for carotid endarterectomy.

1.3 Thesis aims

The aims of this thesis are three-fold. This first aim is to evaluate the health effects and costs of current and new diagnostic test strategies in the early diagnosis of RA to select patients for early start of optimal treatment. First, the tests that are requested by rheumatologists, the test costs, and diagnoses considered in current clinical care are examined. Then, the cost-effectiveness of different new diagnostic test strategies is compared to the current diagnostic test strategy using decision models with 1-year and 5-year time horizon.

The second aim is to evaluate the health effects and costs of current and new prognostic test strategies to assess the risk of a recurrent ischemic stroke event. To assess the methodological challenges of performing cost-effectiveness analyses when practice variation in current clinical care exists, this thesis illustrates the importance of investigating practice variation and to use multiple comparators when important practice variation exists. Then, a cost

analysis shows the effect of the major improvements of stroke care on hospital costs. Lastly, the minimum prognostic sensitivity and specificity of a new test were assessed for different patient subpopulations in order to be cost-effective versus current test strategies.

The third aim is to provide a framework describing the general steps of conducting early cost-effectiveness analyses of new medical tests. This framework is applied to the early cost-effectiveness analyses in the two presented case studies.

1.4 Thesis outline

Chapter 2 provides a framework describing the general steps of an early cost-effectiveness analysis of new medical tests and illustrating the application of this approach to two case studies: a diagnostic test to detect early RA and a prognostic test to assess the risk of a recurrent ischemic stroke. The following chapters of this thesis are categorized into two parts according to disease area. The first part addresses the studies related to RA (Chapters 3-5) and the second part concerns the studies related to ischemic stroke (Chapters 6-8). **Chapter 3** provides insight into the costs of diagnostic tests, which tests are requested, what differential diagnoses are considered, and what strategies are used by rheumatologists in patients with early inflammatory arthritis at risk of having RA. **Chapter 4** presents an early cost-effectiveness model for evaluating new diagnostic tests in the work-up of patients with inflammatory arthritis at risk of having RA. This model was used to assess the short term cost-effectiveness of 4 diagnostic tests with assumed test characteristics and costs based on literature and expert opinion, compared to the ACR/EULAR 2010 RA classification criteria. **Chapter 5** presents an early cost-effectiveness model with a five-year time horizon for inflammatory arthritis patients who were suspected of having RA. This model was used to analyze the cost-effectiveness of different new test strategies compared to the ACR/EULAR 2010 RA classification criteria over a five-year time horizon. **Chapter 6** illustrates the importance of investigating possible clinical practice variation and deviation from national guidelines, and the need to perform hospital-level CEAs, which incorporate local hospital conditions, when important practice variation exists. We specifically focused on diagnostic imaging tests for the assessment of carotid stenosis and criteria for treatment of patients with a recent TIA or minor ischemic stroke in the Netherlands. **Chapter 7** describes the hospital resource use and costs of ischemic stroke and TIA in the Netherlands for 2012. In addition, the associations between hospital costs of ischemic stroke and TIA and various patient and hospital characteristics are discussed. **Chapter 8** presents an early cost-effectiveness analysis that estimates the potential health and economic impact of new tests to identify which patients witch a recent TIA or minor ischemic stroke will benefit more from surgery (i.e., carotid en-

darterectomy) than optimal medical treatment alone. This model was used to estimate the minimum prognostic performance (i.e., sensitivity and specificity) that a new test must have in order for it to be cost-effective versus currently available test strategies. **Chapter 9** provides a general discussion of the main findings, its implications for the different stakeholders, and the methodological challenges of performing early-CEAs.

1.5 References

1 Obama B. Remarks by the President on Precision Medicine. Washington, DC: The White House. State of the Union speech, January 30, 2015. Available from: https://www.whitehouse.gov/the-press-office/2015/01/30/remarks-president-precision-medicine. Accessed August 30, 2015.

2 National Research Council. Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease. Washington, DC: National Academies Press; 2011.

3 Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015;372(9):793–5.

4 Pliskin JS, Shepard D, Weinstein MC. Utility functions for life years and health status. *Operations Research* 1980;28(1):206–24.

5 Drummond MF, Sculpher MJ, Torrance GW, et al. Methods for the Economic Evaluation of Health Care Programmes. 3rd ed. Oxford: Oxford University Press; 2005.

6 Drummond, MF. Economic evaluation of pharmaceuticals. Science or marketing? *Pharmacoeconomics* 1992;1(1):8-13.

7 Johannesson, M. Economic evaluation of drugs and its potential uses in policy-making. *Pharmacoeconomics* 1995;8(3):190-8.

8 Annemans, L, Genesté B, Jolain B. Early Modelling for Assessing Health and Economic Outcomes of Drug Therapy. *Value Health* 2000;3(6):427-34.

9 Pietzsch JB, Paté-Cornell ME. Early technology assessment of new medical devices. *Int J Technol Assess Health Care* 2008;24(1):36-44.

10 IJzerman MJ, Steuten LM. Early assessment of medical technologies to inform product development and market access: a review of methods and applications. *Appl Health Econ Health Policy* 2011;9(5):331-47.

11 Hartz S, John J. Contribution of economic evaluation to decision making in early phases of product development: A methodological and empirical review. *Int J Technol Assess Health Care* 2008;24(4):465-72.

12 Redekop K, Uyl-de Groot C. Diagnostiek en economische evaluatie. In: Rutten-van Mölken M, Uyl-de Groot C, Rutten F, editors. Van kosten tot effecten: Een handleiding voor economische evaluatie studies in de gezondheidszorg. 2nd ed. Amsterdam: Elsevier Gezondheidszorg; 2010. p. 157-76.

13 Pincus T, Callahan LF, Sale WG, et al. Severe functional declines, work disability, and increased mortality in seventy-five rheumatoid arthritis patients studied over nine years. *Arthritis Rheum* 1984;27(8):864–72.
14 Mitchell DM, Spitz PW, Young DY, et al. Survival, prognosis, and causes of death in rheumatoid arthritis. *Arthritis Rheum* 1986;29(6):706–14.
15 Scott DL, Symmons DP, Coulton BL, et al. Long-term outcome of treating rheumatoid arthritis: results after 20 years. *Lancet* 1987;1(8542):1108–11.
16 Isomäki H. Long-term outcome of rheumatoid arthritis. *Scand J Rheumatol Suppl* 1992;95:3–8.
17 Wolfe F. The natural history of rheumatoid arthritis. *J Rheumatol Suppl* 1996;44:13–22.
18 Aletaha D, Neogi T, Silman AJ, et al. Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;69(9):1580–8.
19 Yarwood A, Huizinga TW, Worthington J. The genetics of rheumatoid arthritis: risk and protection in different stages of the evolution of RA. *Rheumatology (Oxford)* 2016;55(2):199–209.
20 Carlens C, Hergens MP, Grunewald J, et al. Smoking, use of moist snuff, and risk of chronic inflammatory diseases. *Am J Respir Crit Care Med* 2010;181:1217–22.
21 Liao KP, Alfredsson L, Karlson EW. Environmental influences on risk for rheumatoid arthritis. *Curr Opin Rheumatol* 2009;21(3):279–83.
22 Symmons D, Turner G, Webb R, et al. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology (Oxford)* 2002; 41: 793–800.
23 Cross M, Smith E, Hoy D, et al. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014;73(7):1316–22.
24 Gabriel SE, Michaud K. Epidemiological studies in incidence, prevalence, mortality, and comorbidity of the rheumatic diseases. *Arthritis Res Ther* 2009;11(3):229.
25 Chorus AM, Miedema HS, Boonen A, et al. Quality of life and work in patients with rheumatoid arthritis and ankylosing spondylitis of working age. *Ann Rheum Dis* 2003;62(12):1178–84.
26 Bansback N, Harrison M, Brazier J, et al. Health state utility values: a description of their development and application for rheumatic diseases. *Arthritis Rheum* 2008;59(7):1018–26.
27 Kvalvik AG, Jones MA, Symmons DP. Mortality in a cohort of Norwegian patients with rheumatoid arthritis followed from 1977 to 1992. *Scand J Rheumatol* 2000;29(1):29–37.
28 Krause D, Schleusser B, Herborn G, et al. Response to methotrexate treatment is associated with reduced mortality in patients with severe rheumatoid arthritis. *Arthritis Rheum* 2000;43(1):14–21.

29 Riise T, Jacobsen BK, Gran JT, et al. Total mortality is increased in rheumatoid arthritis. A 17-year prospective study. *Clin Rheumatol* 2001;20(2):123–7.
30 Doran MF, Pond GR, Crowson CS, et al. Trends in incidence and mortality in rheumatoid arthritis in Rochester, Minnesota, over a forty-year period. *Arthritis Rheum* 2002;46(3):625–31.
31 Young A, Koduri G, Batley M, et al. Mortality in rheumatoid arthritis. Increased in the early course of disease, in ischaemic heart disease and in pulmonary fibrosis. *Rheumatology (Oxford)* 2007;46(2):350–7.
32 Ziadé N, Jougla E, Coste J. Population-level influence of rheumatoid arthritis on mortality and recent trends: a multiple cause-of-death analysis in France, 1970-2002. *J Rheumatol* 2008;35(10):1950–7.
33 Sichletidis L, Settas L, Spyratos D, et al. Tuberculosis in patients receiving anti-TNF agents despite chemoprophylaxis. *Int J Tuberc Lung Dis* 2006;10(10):1127–32.
34 Gómez-Reino JJ, Carmona L, Angel Descalzo M, et al. Risk of tuberculosis in patients treated with tumor necrosis factor antagonists due to incomplete prevention of reactivation of latent infection. *Arthritis Rheum* 2007;57(5):756–61.
35 Boonen A, Severens JL. The burden of illness of rheumatoid arthritis. *Clin Rheumatol* 2011;30(1):S3–8.
36 Franke LC, Ament AJ, Van de Laar MA, et al. Cost-of-illness of rheumatoid arthritis and ankylosing spondylitis. *Clin Exp Rheumatol* 2009;27(4 Suppl 55):S118–23.
37 Brouwer WB, Koopmanschap MA. The friction-cost method: replacement for nothing and leisure for free? *Pharmacoeconomics* 2005;23(2):105–11.
38 Hazes JM, Luime JJ. The epidemiology of early inflammatory arthritis. *Nat Rev Rheumatol* 2011;7(7):381–90.
39 Alves C, Luime JJ, Van Zeben D, et al. Diagnostic performance of the ACR/EULAR 2010 criteria for rheumatoid arthritis and two diagnostic algorithms in an early arthritis clinic (REACH). *Ann Rheum Dis* 2011;70(9):1645–7.
40 Caplan LR. Basic pathology, anatomy, and pathopysiology of stroke. In: Caplan's Stroke: A Clinical Approach. 4th ed. Philadelphia: Saunders Elsevier; 2009.
41 Krishnamurthi RV, Feigin VL, Forouzanfar MH, et al. Global and regional burden of first-ever ischaemic and haemorrhagic stroke during 1990-2010: findings from the Global Burden of Disease Study 2010. *Lancet Glob Health* 2013;1(5):e259–81.
42 Go AS, Mozaffarian D, Roger VL, et al. Heart disease and stroke statistics—2014 update: a report from the American Heart Association. *Circulation* 2014;129(3):e28–e292.
43 Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;380(9859):2095–128.
44 Minkman MM, Schouten LM, Huijsman R, et al. Integrated care for patients with a stroke in the Netherlands: results and experiences from a national Breakthrough Collaborative Improvement project. *Int J Integr Care* 2005;5:e14.

45 Saver JL, Fonarow GC, Smith EE, et al. Time to treatment with intravenous tissue plasminogen activator and outcome from acute ischemic stroke. *JAMA* 2013;309:2480–8.
46 Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis. North American Symptomatic Carotid Endarterectomy Trial Collaborators. *N Engl J Med* 1991;325:445–53.
47 MRC European Carotid Surgery Trial: interim results for symptomatic patients with severe (70–99%) or with mild (0–29%) carotid stenosis. European Carotid Surgery Trialists' Collaborative Group. *Lancet* 1991;337:1235–43.
48 Randomised trial of a perindopril-based blood pressure lowering regimen among 6,105 patients with prior stroke or transient ischaemic attack. PROGRESS Collaborative Group. *Lancet* 2001;358:1033–41.
49 Amarenco P, Bogousslavsky J, Callahan A, et al. High-dose atorvastatin after stroke or transient ischemic attack. *N Engl J Med* 2006;355:549–59.
50 Ingeman A, Andersen G, Hundborg HH, et al. In-hospital medical complications, length of stay, and mortality among stroke unit patients. *Stroke* 2011;42(11):3214–8.
51 Thanvi B, Treadwell S, Robinson T. Early neurological deterioration in acute ischaemic stroke: predictors, mechanisms and management. *Postgrad Med J* 2008;84(994):412–7.
52 Pohjasvaara T, Vataja R, Leppävuori A, et al. Cognitive functions and depression as predictors of poor outcome 15 months after stroke. *Cerebrovasc Dis* 2002;14(3-4):228–33.
53 Hallan S, Asberg A, Indredavik B, et al. Quality of life after cerebrovascular stroke: a systematic study of patients' preferences for different functional outcomes. *J Intern Med* 1999;246:309–16.
54 Bergman L, Van der Meulen J, Limburg M, et al. Cost of medical care after first-ever stroke in the Netherlands. *Stroke* 1995;26:1830–6.
55 Evers SM, Engel GL, Ament AJ. Cost of stroke in the Netherlands from a societal perspective. *Stroke* 1997;28:1375–81.
56 Statistics Netherlands (in Dutch: Centraal Bureau voor de Statistiek, CBS). CBS Statline. Available from: http://statline.cbs.nl. Accessed February 1, 2014.
57 Ward NS, Brown MM, Thompson AJ, et al. Neural correlates of motor recovery after stroke: a longitudinal fMRI study. *Brain* 2003;126(11):2476–96.
58 Hankey GJ, Spiesser J, Hakimi Z, et al. Rate, degree, and predictors of recovery from disability following ischemic stroke. *Neurology* 2007;68(19):1583–7.
59 Weimar C, König IR, Kraywinkel K, et al. Age and National Institutes of Health Stroke Scale Score within 6 hours after onset are accurate predictors of outcome after cerebral ischemia: development and external validation of prognostic models. *Stroke* 2004;35(1):158.
60 Andersen KK, Andersen ZJ, Olsen TS. Predictors of early and late case-fatality in a nationwide Danish study of 26,818 patients with first-ever ischemic stroke. *Stroke* 2011;42(10):2806–12.

61 Schiemanck SK, Kwakkel G, Post MW, et al. Predictive value of ischemic lesion volume assessed with magnetic resonance imaging for neurological deficits and functional outcome poststroke: A critical review of the literature. *Neurorehabil Neural Repair* 2006;20(4):492–502.

62 Wu CM, McLaughlin K, Lorenzetti DL, et al. Early risk of stroke after transient ischemic attack: a systematic review and meta-analysis. *Arch Intern Med* 2007;167(22):2417–22.

63 Van Wijk I, Kappelle LJ, Van Gijn J, et al. Long-term survival and vascular event risk after transient ischaemic attack or minor ischaemic stroke: a cohort study. *Lancet* 2005;365(9477):2098–104.

64 Rothwell PM, Eliasziw M, Gutnikov SA, et al. Analysis of pooled data from the randomised controlled trials of endarterectomy for symptomatic carotid stenosis. *Lancet* 2003;361(9352):107–16.

65 Van Engelen A, Wannarong T, Parraga G, et al. Three-dimensional carotid ultrasound plaque texture predicts vascular events. *Stroke* 2014;45(9):2695-701.

66 Chai CK, Akyildiz AC, Speelman L, et al. Local axial compressive mechanical properties of human carotid atherosclerotic plaques-characterisation by indentation test and inverse finite element analysis. *J Biomech* 2013;46(10):1759-66.

CHAPTER 2

THE EARLY BIRD CATCHES THE WORM: EARLY COST-EFFECTIVENESS ANALYSIS OF NEW MEDICAL TESTS

Leander R. Buisman[1,2], Maureen P.M.H. Rutten-van Mölken[1,2], Douwe Postmus[3], Jolanda J. Luime[4], Carin A. Uyl-de Groot[1,2], William K. Redekop[1,2]

[1] Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands
[2] Institute for Medical Technology Assessment, Erasmus University Rotterdam, Rotterdam, The Netherlands
[3] Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
[4] Department of Rheumatology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

Published in *International Journal of Technology Assessment in Health Care* 2016;32(1):1–8. Epub 2016 Mar 22. doi:10.1017/S0266462316000064.

ABSTRACT

Objectives: There is little specific guidance on performing an early cost-effectiveness analysis (CEA) of medical tests. We develop a framework with general steps and apply it to two cases.

Methods: Step 1 is to narrow down the scope of analysis by defining the test's application, target population, outcome measures, and investigating current test strategies and test strategies if the new test was available. Step 2 is to collect evidence on the current test strategy. Step 3 is to develop a conceptual model of the current and new test strategies. Step 4 is to conduct the early-CEA by evaluating the potential (cost-)effectiveness of the new test in clinical practice. Step 5 involves a decision about the further development of the test.

Results: The first case illustrated the impact of varying the test performance on the headroom (maximum possible price) of an add-on test for patients with an intermediate-risk of having rheumatoid arthritis. Analyses showed that the headroom is particularly dependent on test performance. The second case estimated the minimum performance of a confirmatory imaging test to predict individual stroke risk. Different combinations of sensitivity and specificity were found to be cost-effective; if these combinations are attainable, the medical test developer can feel more confident about the value of further development of the test.

Conclusions: A well-designed early-CEA methodology can improve the ability to develop (cost-)effective medical tests in an efficient manner. Early-CEAs should continuously integrate insights and evidence that arise through feedback, which may convince developers to return to earlier steps.

Keywords: early cost-effectiveness analysis, medical test, decision support, research and development, manufacturer, test developer.

2.1 Introduction

In healthcare systems worldwide, comparative cost-effectiveness research is mostly done in the late stages of medical test* development. Such evaluation is mainly used by medical test manufacturers/developers (referred to as test developers) to demonstrate to payers (i.e., healthcare insurers, governments, managed care organizations) that a test is good value for money.[2-6] To test developers, adequate reimbursement of new tests is important for wide implementation in clinical practice, which improves return on investment. Medical tests are most often reimbursed as part of a Diagnosis Related Group (DRG) payment or fee-for-service. Hence, those who actually decide about reimbursement include opinion leaders among the clinicians, hospital managers, government authorities and healthcare insurers. Given the serious consequences of negative reimbursement decisions, test developers take a considerable risk when conducting cost-effectiveness analyses only in the late stages of medical test development, when large R&D investments have already been made. Therefore, evaluation of new tests in the early stages of development has attracted increasing attention.

Early cost-effectiveness analysis (CEA) helps test developers to decide about further development of medical tests, set realistic performance-price goals, and design and manage reimbursement strategies.[5] Early-CEAs may guide the resources invested in the development process. However, they require close cooperation between test developers and the researchers performing early-CEAs. Over the past two decades, most early-CEAs focused on new drug therapy.[7] The increasing use of medical tests in various phases of disease prevention and treatment, including companion diagnostics to 'personalize' medicine, calls for early-CEAs to assess how much these tests could really improve health outcomes and healthcare efficiency. However, there is little specific guidance on performing early-CEAs of medical tests. Some general steps of early-CEAs were described by O'Prinsen et al.[8] In this paper, we developed a framework with the general steps of early-CEAs of new medical tests and applied it to two cases. The development of the framework was an iterative process because in the cases we applied the methods in our framework but used the experience and insights gained from the cases to refine the framework.

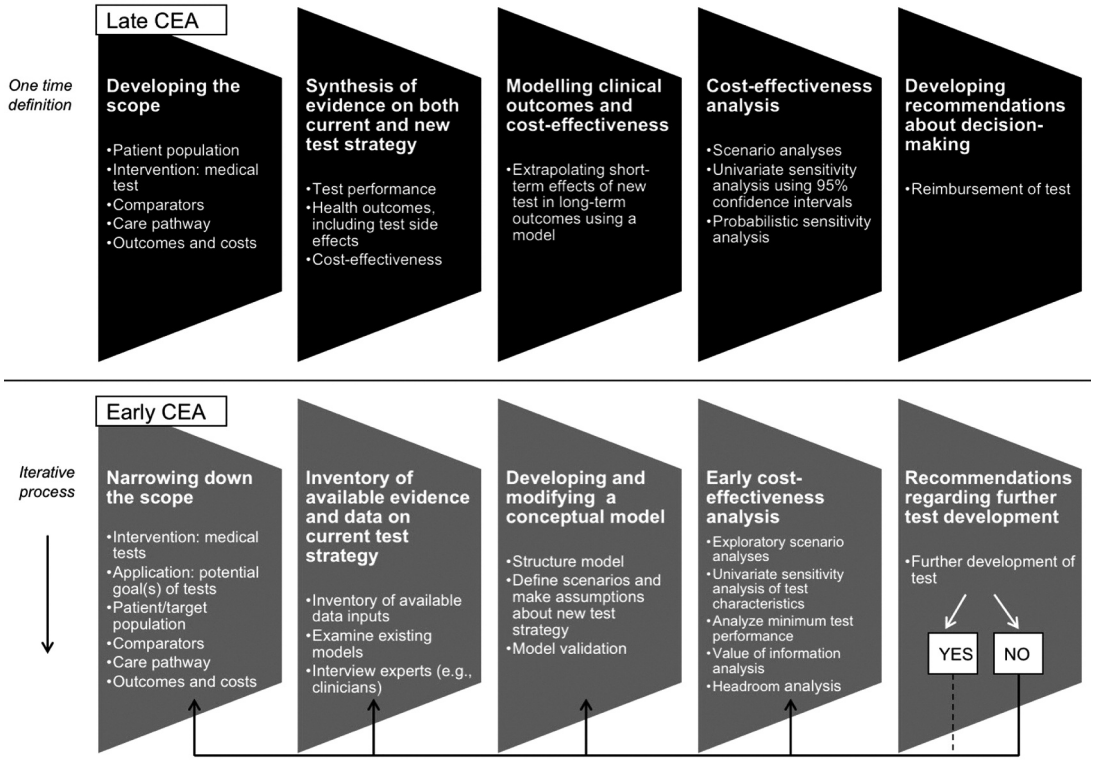
2.2 General steps of early-CEAs

Before presenting the general steps of early-CEAs, we describe the differences between early and late-CEAs of tests. Figure 2.1 shows the five general steps of late and early-CEAs of medical tests and the main differences between them.

* Medical tests are used 'to determine the presence or absence of a definite disease or of some substance in any of the fluids, tissues, or excretions of the body, or to determine the presence or degree of a psychological or behavioural trait'.[1] An early-CEA can be important for any type of device, including the medical tests addressed in this paper.

The general steps of *late-CEAs* of medical tests are based on the Diagnostic Assessment Programme (DAP) Manual from NICE.[9] The DAP manual promotes the consistent and rapid adoption of clinically innovative and cost-effective medical tests in the United Kingdom. There are at least four main differences between early and late-CEAs of medical tests. Firstly, the first step of a late-CEA of medical tests is to **define** the scope, while the first step of an early-CEA is **to narrow down** the scope. Secondly, early-CEAs are much more iterative than late-CEAs. Throughout the development process of a test, new data and ideas may emerge, which may convince developers to return to earlier steps. Thirdly, much less data are available for early-CEAs than for late-CEAs, making sensitivity analyses of early-CEAs much more exploratory. Fourthly, late-CEAs are mainly used by payers in reimbursement decisions, while early-CEAs are used by test developers for internal decision-making about further development of a test and setting realistic performance-price goals. The following section describes the five steps of an early-CEA as shown in Figure 2.1.

Figure 2.1: Differences and similarities of late and early-CEAs of medical tests

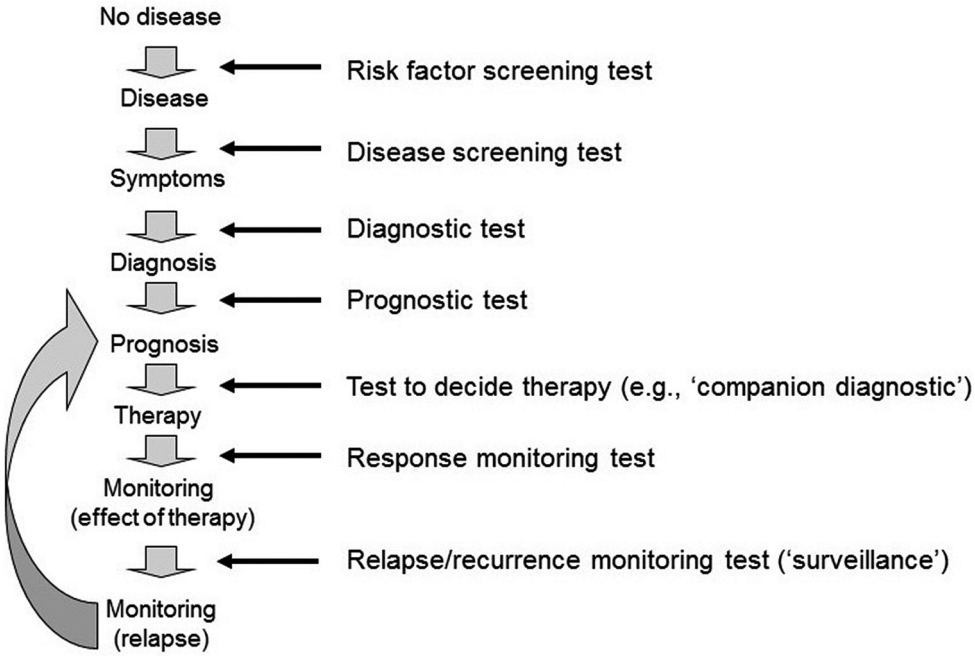


Steps of late-CEA based on NICE Diagnostic Assessment Programme manual;[9]
Steps of early-CEA were defined by the authors.

Step 1: Narrowing down the scope

The test developer needs to consider where in the healthcare system the new test will be used. Figure 2.2 shows a variety of applications of medical tests following the sequence of screening/case finding, diagnosis, disease progression and treatment.

Figure 2.2: Various applications of medical tests in the healthcare system



The vertical arrows represent the different phases of disease and its treatment (based on Redekop and Uyl-de Groot[10]).

One way to start exploring where a test could be used in patient care is to apply the Patient Population, Intervention, Comparator, and Outcomes (PICO) method.[11] This way, one can systematically compare different potential areas of application in the healthcare system. If the application of the test is clear, the target population may also be clear. However, if the application is not clear, literature research and discussions with test developers and clinicians on areas of highest unmet need and greatest added benefit would help to determine the application in the healthcare system and define the target population. Given the context of an early-CEA, we extend the PICO method and re-order the different elements into an 'APCOI' (Appli-

cation, Patient Population, Comparator, Outcomes, Intervention) method, resulting in the following questions:

1 Application (in healthcare): What is the anticipated application of the test in the healthcare system?
2 Patient Population (participants): What is the target population?
3 Comparator: How are current test strategies (i.e., current tests and treatment options based on the test result), anticipated future comparators, and resulting clinical care organized?
4 Outcomes: How will effectiveness be defined and which costs need to be taken into account?
5 Intervention: What will clinical care look like with the new medical test?

Discussions with different stakeholders are essential during test development. For example, a test can be cost-effective, but if clinicians are not convinced that the test improves patient care, the new test will not be used in clinical practice and will not yield a sufficient return on investment.

Step 2: Inventory of available evidence and data on current test strategy

A proper examination of the strengths and weaknesses of current care is invaluable in estimating the added value of the new test. For example, if current care is already highly effective, it may be difficult –if not impossible– to improve upon it (although the new test can still be considered cost-effective if it reduces costs). Researchers should examine how current test strategies, anticipated future comparators, and clinical care are organized and which evidence and data are available on the costs and health outcomes. In addition, potentially relevant existing models of the disease and target population should be reviewed.

Step 3: Developing and modifying a conceptual model

A conceptual model of both the current and new test strategies should be developed in Step 3 and informed by Step 2. Since little is known about the impact of the new test strategy, various scenarios should be defined, which could include varying subpopulations as part of the target population (e.g., specific age-sex groups), the prevalence of the disease of interest, applications of the test, costs, test performance, and health improvements. Due to the stronger iterative nature of early-CEAs as compared to late-CEAs, conceptual models should be revised as new insights and evidence arise during test development. The model's validity should be scrutinized on its validity similar to late-stage modelling (face, cross-model, external, predictive, and verification validity), to the extent that is possible in the early stages of test development.[12]

Step 4: Early cost-effectiveness analysis

An early-CEA can be conducted to evaluate the potential impact of the new test in clinical practice when all parameters and their values have been determined. Normally, initial estimates of the model parameters may have to be derived from expert opinion, observational studies or small clinical trials. Therefore, exploratory scenario analyses can help to set a benchmark for the minimum performance that is required for a test to become cost-effective compared to current practice. For example, one could derive the minimum sensitivity and specificity at which the test becomes an attractive alternative from an effectiveness standpoint. These scenario analyses can be reapplied and modified throughout the entire development process. Moreover, an early-CEA should at least include a univariate sensitivity analysis in which a range of parameters are varied to identify which of them have the most impact on incremental costs, effectiveness and cost-effectiveness. When uncertainty can be quantified, a value of information analysis is recommended in the early stages of test development to decide whether additional research is needed to decide which test scenario should be chosen.[13-15]

Although early-CEAs provide valuable insights into the clinical and economic value of new medical tests, they do not directly provide an answer to the question of whether the test developer should continue developing the test. To address this question, the results must be translated into an estimate of the test's maximum sales price. This price can be derived from the cost-effectiveness model, for any given combination of parameters. This price can be further substantiated by applying the principle of value-based pricing, in which the price is largely driven by the innovative nature of the test and the extent to which it addresses unmet needs, and by making assumptions about how the volume of sales is affected by the maximum sales price.[16] The maximum sales price can then be fed into an appropriate product-investment evaluation method, such as the headroom method.[16] The headroom (or potential profit) method assesses the maximum additional cost at which the medical test is still likely to be cost-effective at a given willingness-to-pay threshold.[17,18]

Step 5: Developing recommendations regarding further test development

The results of early-CEAs can help test developers of medical tests to decide about the further development of the test (go/no-go decision). If an early-CEA shows that the test is unfeasible or unlikely to be cost-effective, it is unlikely that the test will be reimbursed. Therefore, test developers may decide to return to earlier steps of the early-CEA. Even if an expensive test is found to be cost-effective, it may be difficult for clinicians to use the test until the higher test costs have been incorporated into a higher DRG payment.

2.3 Cases of early-CEAs of medical tests

This section applies the steps of early-CEAs of medical tests to two cases: a diagnostic test to detect early rheumatoid arthritis (RA) and a prognostic test to assess the risk of a recurrent ischemic stroke.

Case 1: diagnostic test for RA

RA is a chronic inflammatory disease characterized by structural irreversible joint damage, leading to severe disability and premature death.[19-24] Early detection is important because early treatment with synthetic and biologic disease modifying antirheumatic drugs (DMARDs) has been shown to slow down disease progression.[25-29]

When applying the general steps of an early-CEA, we started to narrow down the scope of the analysis to tests that help to diagnose RA in an early stage (Step 1). The current diagnostic standard for RA, and thus the comparator in the CEA, is the ACR/EULAR 2010 RA classification criteria (referred to as RA-2010 criteria).[24,30,31] Use of the RA-2010 criteria at baseline as a risk prediction tool usually results in a considerable proportion of early arthritis patients being classified as having an intermediate-risk to develop RA in the near future (3-5 points). What is needed is an additional test to reclassify these intermediate-risk patients into high-risk and low-risk ones. A B-cell test is a candidate for this purpose.[32]

As part of Step 2 of our early-CEA, we examined how current clinical care is organized using the RA-2010 criteria. Moreover, we investigated which tests are used to diagnose RA by conducting interviews with rheumatologists and analyzing resource use data. We found that a variable number of diagnostic tests (average: 32 diagnostic tests per patient) is requested by rheumatologists during the initial outpatient visit to exclude differential diagnoses of RA (Benner et al. 2015, unpublished data). The mean diagnostic costs per patient were €422 (SD: €168). In addition, interviews were conducted with the test developer of the B-cell test about the performance and costs of the test to diagnose RA in a population of early arthritis patients. The B-cell test has been studied in a cohort of seropositive arthralgia patients (i.e., subsample of early arthritis patients) and a sensitivity of 60% and specificity of 90% were found.[32] We assumed that this test performance was applicable for all early arthritis patients.

As Step 3, a conceptual decision model with a five-year time horizon was developed for the diagnosis and treatment of patients with early arthritis based on evidence obtained in Step 2 (see Appendix 2.1 for the model). Interviews with rheumatologists and test developer were conducted about the potential use of B-cell test in the early diagnosis of RA to formulate test scenarios. In the new test strategy, a B-cell test was used as add-on for patients with an intermediate-risk according to the RA-2010 criteria. In this strategy,

intermediate-risk patients can be reclassified as high-risk or low-risk. At a sensitivity of 60% and specificity of 90%, 29% (75/263) would be reclassified as high-risk and 71% (188/263) as low-risk. Patients were classified as true positive (TP) if they had a positive test result (scored ≥6 points on the RA-2010 criteria or were B-cell positive) at baseline and used MTX at 12 months. Patients were considered as true negative (TN) if they had a negative test result (scored <6 points on the RA-2010 criteria or were B-cell negative) at baseline and did not use MTX at 12 months. False positive (FP) patients were patients who scored ≥6 points on the RA-2010 criteria or were B-cell positive at baseline but did not use MTX at 12 months. Patients were classified as false negative (FN) if they scored <6 points on the RA-2010 criteria or were B-cell negative at baseline but used MTX at 12 months.

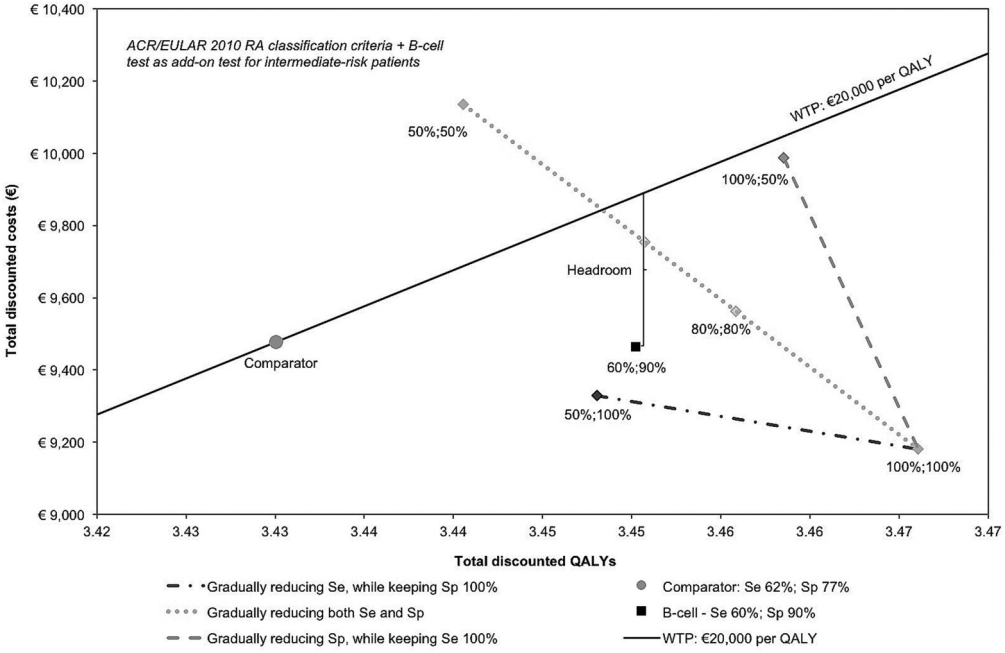
Patients classified as TP or FN at 12 months were defined as RA patients and entered a patient-level state transition model at 12 months in which the disease activity (DAS28) and treatment course were simulated based on data from two cohorts (REACH[30] and tREACH[31,33]) and published data[34]. Patients could switch from MTX to more expensive biologic DMARDs if they had a DAS28≥3.2 with a swollen joint count>0 and no comorbidity. Treatment with a biologic DMARD incurs additional treatment costs but also improves health-related quality of life (EQ-5D utilities). Patients classified as TN or FP at 12 months entered a background model in which they stayed for the remaining four years, assuming no change in utilities, biologic DMARD costs for 10% of FPs in the first year after diagnosis, and otherwise no RA-related costs.

The decision model was populated with data of patients in the REACH cohort.[30] The prevalence of RA among these patients is 54% at 12 months using the RA-2010 criteria. The quality of life (EQ-5D) at time of diagnosis and follow-up was obtained from the literature and the REACH cohort. Direct medical costs were based on the Dutch Manual of Costing.[35] Costs of blood tests were based on tariffs provided by the Dutch Healthcare Authority,[36] and costs of MTX and biologic DMARDs were obtained from the National Health Care Institute.[37] All costs were adjusted to 2014 Euros using the general price index from the Dutch Central Bureau of Statistics,[38] and a healthcare perspective was used.

In Step 4, exploratory scenario analyses were performed using different potential test scenarios of the B-cell test. In this paper, we present the results of the B-cell test as *add-on test for patients with an intermediate-risk* based on the RA-2010 criteria. The sensitivity and specificity of these criteria in the REACH data were estimated to be 62% and 77%, respectively. Probabilistic sensitivity analysis was performed and the maximum cost of a B-cell test required for the incremental cost-effectiveness ratio (ICER) of the new test strategy to stay below €20,000 per quality-adjusted life year (QALY)

(i.e., the headroom) was calculated (i.e., headroom). Figure 2.3 shows the impact of varying the sensitivity and specificity between 50% and 100% on the headroom of a B-cell test.

Figure 2.3: Cost-effectiveness plane for a B-cell test as add-on for patients with an intermediate-risk based on the ACR/EULAR 2010 RA classification criteria



The figure presents the sensitivity and specificity of an add-on test for intermediate-risk patients in addition to the ACR/EULAR 2010 RA classification criteria; Se = sensitivity; Sp = specificity; WTP = willingness-to-pay; QALY = quality-adjusted life year.

An add-on B-cell test with 100% sensitivity and specificity dominates the RA-2010 criteria since it reduces costs (by €296) and increases health (by 0.036 QALYs). Similarly, a B-cell test with 60% sensitivity and 90% specificity dominates the comparator since it also reduces costs (by €14) and increases health (by 0.020 QALYs). The headroom of the B-cell test (€417) is shown in Figure 2.3 by the difference in costs between the willingness-to-pay threshold line of €20,000 per QALY gained and the cost-effectiveness of the new test strategy (where the test is assumed to be free).

Case 2: prognostic test for recurrent stroke

Patients with a recent transient ischemic attack (TIA) or minor ischemic stroke are at risk of a recurrent ischemic stroke, which may be caused by a plaque rupture in the carotid artery. Surgery (carotid endarterectomy) can reduce this risk, but the procedure can lead to death and morbidity. Better stroke risk prediction would therefore help to determine which patients should undergo surgery. Non-invasive molecular imaging technologies, such as contrast enhanced magnetic resonance imaging, computed tomography angiography, and biomechanical analysis are technologies that can be used to improve risk prediction.

When applying the general steps of an early-CEA, we narrowed down the scope of the analysis using the APCOI method (Step 1). The application (and target population) of the new medical test was a *prognostic test* to predict the risk of a recurrent stroke caused by plaque rupture in *patients with recent TIA or minor ischemic stroke and 30-69% carotid stenosis*. We then examined how current test strategies and clinical care are organized and discovered a substantial amount of variation. For example, the national stroke guidelines were followed by 60% of Dutch hospitals while other hospitals used various other test combinations.[39] The comparators representing current clinical practice were therefore defined as 'guideline-based care'[40] and other current test strategies. For the new test strategy, interviews were conducted with test developers and clinicians about the optimal combination of tests and potential test performance, since many combinations are possible. The most likely application for a new test would be as a confirmatory imaging test for patients with a 30-69% carotid stenosis according to an initial duplex ultrasonography (sensitivity: 89% and specificity: 84%;[41] costs: €125[36]). We assumed that if the confirmatory imaging test identified patients as being at high-risk of a recurrent stroke, patients underwent surgery, while patients with a low-risk of recurrent stroke received medicines alone.

As Step 2, we examined the costs and health outcomes of current care using clinical stroke guidelines and literature. These findings were discussed with vascular neurologists and radiologists to assess the quality and relevance. Also, any existing ischemic stroke CEA models were reviewed including those for diagnostic tests (e.g., Tholen et al.[41]).

The results from Step 2 were used to develop a conceptual model in Step 3. The model consisted of 3 parts: prognostic testing, treatment, and health outcomes. The initial version focused on the use of a new confirmatory imaging test for patients with 30-69% carotid stenosis tested with an initial duplex ultrasonography.

At present, little is known about the performance of a test to predict the risk of plaque rupture. Therefore, decision modelling was used to estimate the minimum test performance (i.e., sensitivity and specificity) that a new

confirmatory imaging test must have in order to be cost-effective compared to current care (Step 4). Exploratory sensitivity analyses were performed to identify which combinations of test costs and performance resulted in acceptable cost-effectiveness ratios. All costs were adjusted to 2014 Euros using the general price index from the Dutch Central Bureau of Statistics,[38] and a societal perspective was used.

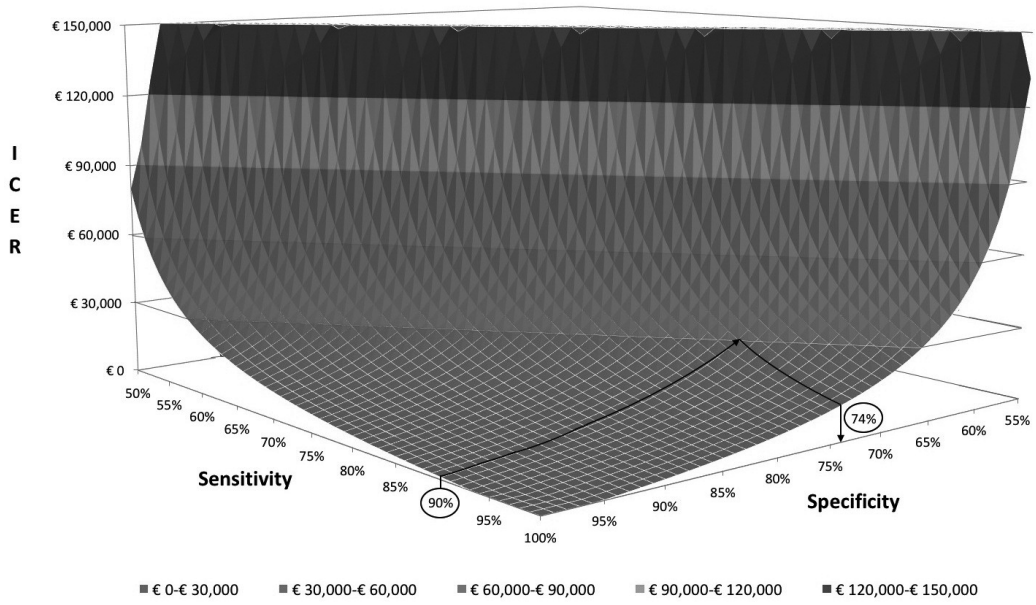
A perfect confirmatory imaging test (100% sensitivity and specificity) with a cost of €362 for 60-year-old men appears to be cost-effective compared to 'guideline-based care' using a life-time horizon. A perfect confirmatory imaging test dominates 'guideline-based care' since it reduces costs (by €110) and increases health (by 0.066 QALYs), because the test identifies all patients correctly and ensures that they all receive the appropriate treatment. Figure 2.4 shows the minimum values of sensitivity and specificity needed for a new confirmatory imaging test at different thresholds regarding the willingness-to-pay to gain one QALY. For example, if a test is 90% sensitive, it must have a specificity of at least 74% to be cost-effective given a willingness-to-pay threshold of €30,000 per QALY gained. This combination can be found in Figure 2.4 by starting from the sensitivity axis at 90%, following the gridline to the boundary between the blue and orange parts (i.e., willingness-to-pay threshold of €30,000 per QALY gained), following the gridline for specificity to the right-hand side, and then moving vertically down to the specificity axis.

2.4 Discussion

CEAs in the early stages of medical test development have important benefits. For test developers, they are useful in guiding further development of tests, for example by estimating the maximum cost of a new test and the minimum test performance required for the test to be cost-effectiveness. For clinicians, early-CEAs provide valuable information about the patient (sub)populations in which the test is potentially cost-effective. An early-CEA can convince clinicians of the potential improvements in patient care and health outcomes which may result in faster take-up of tests in clinical practice. Published early-CEA results can also help payers to identify promising new tests early, resulting in more timely decisions about reimbursement. A major advantage of early-CEA is that it generates optimal product development and pricing, although it might seem resource-intensive at first. This was illustrated by the two cases. We showed how the inevitable uncertainties in the early stage of the development can be addressed by applying different scenario and sensitivity analyses.

New tests can lead to significant improvements in health outcomes and efficiency only if effective treatments are available. In the RA case, effective treatment was available to slow down disease progression in patients diag-

Figure 2.4: Minimum sensitivity and specificity of a new confirmatory imaging test (test costs: €362)



The minimum values of sensitivity and specificity can be found by starting from a value of sensitivity at the sensitivity axis, moving vertically up to the corresponding gridline, following the gridline to a predefined threshold regarding the willingness-to-pay to gain one quality-adjusted life year (QALY), following the gridline for specificity to the right-hand side, and then moving vertically down to the specificity axis; ICER = incremental cost-effectiveness ratio.

nosed with a high-risk of having RA, while in the stroke case, patients with a high-risk of recurrent ischemic stroke underwent surgery.

We developed a framework with general steps of an early-CEA of new medical tests. Our framework is developed to evaluate the potential cost-effectiveness of new medical tests and is not completely applicable to new drugs (see Drummond et al.[42] for important differences in CEAs for tests and drugs). Some studies have previously been performed, which can be seen as examples of the steps in our framework. Postmus et al.[17] describe in more formal terms how Steps 2-5 can be performed for a risk factor screening test. Cao et al.[43] describe how to handle the use of expert opinion in an early-CEA model of medical tests in a probabilistic way. We recommend further research that applies our proposed framework of early-CEAs of medical tests to the assessment of specific tests.

Applying the framework to our example cases, we saw that varying the sensitivity and specificity influenced the headroom of an add-on test for in-

termediate-risk RA patients. To calculate the headroom, we used a fixed willingness-to-pay threshold per QALY gained, but different thresholds led to different estimates of the maximum sales price. If this maximum cost offers sufficient degree of headroom from a commercial standpoint, test developers may opt to continue developing the test as planned. In the second case, to predict individual stroke risk, different combinations of sensitivity and specificity of a new confirmatory imaging test were cost-effective at a given willingness-to-pay threshold. Decisions about further test development may therefore be dependent on the threshold used.

The framework has also some potential limitations. First of all, the general steps might be too general to make the framework applicable to all early-CEAs of different types of tests in all diseases; the necessary test-specific details will have to be developed and documented. Furthermore, the success of a new medical test depends on more factors than cost-effectiveness. Factors such as total revenue, future market with potential competitors, future investments during test development, costs of scaling-up the production, stakeholder preferences, and marketing among professionals should be considered by test developers in the early stages of development. A business case developed and refined during the early stages should incorporate these factors.[4,5] At the later stages payers will have important considerations, such as the safety of the test, and budget impact.

2.5 Conclusion

A well-designed early-CEA methodology as presented in this paper can improve the ability to develop effective and cost-effective medical tests in an efficient manner. The continuous integration of new insights and evidence that arise through feedback during the test development may convince developers to return to earlier development steps and will result in more informed decisions by test developers about its potential application in the healthcare system and target population.

2.6 Acknowledgments

This research was performed within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), projects PARISk (grant 01C-202) and TRACER (grant 04I-202), and supported by the Dutch Heart Foundation and the Dutch Arthritis Foundation.

2.7 References

1 Dirckx JH. Stedman's concise: Medical Dictionary for the Health Professions. 4th ed. Baltimore: Lippincott Williams & Wilkins; 2001.

2 Drummond MF. Economic evaluation of pharmaceuticals. Science or marketing? *Pharmacoeconomics* 1992;1(1):8-13.

3 Johannesson M. Economic evaluation of drugs and its potential uses in policy-making. *Pharmacoeconomics* 1995;8(3):190-8.

4 Annemans L, Genesté B, Jolain B. Early Modelling for Assessing Health and Economic Outcomes of Drug Therapy. *Value Health* 2000;3(6):427-34.

5 Pietzsch JB, Paté-Cornell ME. Early technology assessment of new medical devices. *Int J Technol Assess Health Care* 2008;24(1):36-44.

6 IJzerman MJ, Steuten LM. Early assessment of medical technologies to inform product development and market access: a review of methods and applications. *Appl Health Econ Health Policy* 2011;9(5):331-47.

7 Hartz S, John J. Contribution of economic evaluation to decision making in early phases of product development: A methodological and empirical review. *Int J Technol Assess Health Care* 2008;24(4):465-72.

8 O'Prinsen AC, Gaultney J, Redekop WK. Universal Steps for Conducting Early-Stage Medical Techonolgy Assessment. *ISPOR Connections* 2009;November/December:4-6.

9 National Institute for Health and Care Excellence (NICE). Diagnostic Assessment Programme Manual. London: National Institute for Health and Care Excellence; 2011.

10 Redekop K, Uyl-de Groot C. Diagnostiek en economische evaluatie. In: Rutten-van Mölken M, Uyl-de Groot C, Rutten F, editors. Van kosten tot effecten: Een handleiding voor economische evaluatie studies in de gezondheidszorg. 2nd ed. Amsterdam: Elsevier Gezondheidszorg; 2010. p. 157-76.

11 O'Connor D, Green S, Higgins JPT. Defining the review question and developing criteria for including studies. In: Higgins, JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration; 2011. Available from: http://www.cochrane-handbook.org. Accessed May 2, 2014.

12 Eddy DM, Hollingworth W, Caro JJ, et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Med Decis Making* 2012;32(5):733-43.

13 Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999;18(3):341-64.

14 Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty: when is there sufficient evidence? *Value Health* 2005;8(4):433-46.

15 Claxton KP, Sculpher MJ. Using value of information analysis to prioritise health research: some lessons from recent UK experience. *Pharmacoeconomics* 2006;24(11):1055-68.
16 Girling A, Young T, Brown C, et al. Early-stage valuation of medical devices: the role of developmental uncertainty. *Value Health* 2010;13(5):585-91.
17 Postmus D, De Graaf G, Hillege HL, et al. A method for the early health technology assessment of novel biomarker measurement in primary prevention programs. *Stat Med* 2012;31(23):2733-44.
18 Cosh E, Girling A, Lilford R, et al. Investing in new medical technologies: a decision framework. *J Commerc Biotechnol* 2007;13(4):263–71.
19 Pincus T, Callahan LF, Sale WG, et al. Severe functional declines, work disability, and increased mortality in seventy-five rheumatoid arthritis patients studied over nine years. *Arthritis Rheum* 1984;27(8):864–72.
20 Mitchell DM, Spitz PW, Young DY, et al. Survival, prognosis, and causes of death in rheumatoid arthritis. *Arthritis Rheum* 1986;29(6):706–14.
21 Scott DL, Symmons DP, Coulton BL, et al. Long-term outcome of treating rheumatoid arthritis: results after 20 years. *Lancet* 1987;1(8542):1108–11.
22 Isomäki H. Long-term outcome of rheumatoid arthritis. *Scand J Rheumatol Suppl* 1992;95:3–8.
23 Wolfe F. The natural history of rheumatoid arthritis. *J Rheumatol Suppl* 1996;44:13–22.
24 Aletaha D, Neogi T, Silman AJ, et al. Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;69(9):1580–8.
25 Bijlsma JW, Weinblatt ME. Optimal use of methotrexate: the advantages of tight control. *Ann Rheum Dis* 2007;66(11):1409–10.
26 Jacobs JW. Optimal use of non-biologic therapy in the treatment of rheumatoid arthritis. *Rheumatology (Oxford)* 2012;51(4):iv3–8.
27 Chen YF, Jobanputra P, Barton P, et al. A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness. *Health Technol Assess* 2006;10(42):iii–iv, xi–xiii, 1–229.
28 Donahue KE, Garthehner G, Jonas DE, et al. Comparative Effectiveness of Drug Therapy for Rheumatoid Arthritis and Psoriatic Arthritis in Adults. Rockville, MD: AHRQ Comparative Effectiveness Reviews; 2007.
29 Singh JA, Christensen R, Wells GA, et al. A network meta-analysis of randomized controlled trials of biologics for rheumatoid arthritis: a Cochrane overview. *CMAJ* 2009;181(11):787–96.
30 Alves C, Luime JJ, Van Zeben D, et al. Diagnostic performance of the ACR/EULAR 2010 criteria for rheumatoid arthritis and two diagnostic algorithms in an early arthritis clinic (REACH). *Ann Rheum Dis* 2011;70(9):1645–7.

31 Claessen SJ, Hazes JM, Huisman MA, et al. Use of risk stratification to target therapies in patients with recent onset arthritis; design of a prospective randomized multicenter controlled trial. *BMC Musculoskelet Disord* 2009;10:71.

32 Van Baarsen LG, Bos WH, Rustenburg F, et al. Gene expression profiling in autoantibody-positive patients with arthralgia predicts development of arthritis. *Arthritis Rheum* 2010;62(3):694–704.

33 De Jong PH, Hazes JM, Han HK, et al. Randomised comparison of initial triple DMARD therapy with methotrexate monotherapy in combination with low-dose glucocorticoid bridging therapy; 1-year data of the tREACH trial. *Ann Rheum Dis* 2014;73(7):1331–9.

34 Vermeer M, Kievit W, Kuper HH, et al. Treating to the target of remission in early rheumatoid arthritis is cost-effective: results of the DREAM registry. *BMC Musculoskelet Disord* 2013;14:350.

35 Tan SS, Bouwmans CA, Rutten FF, et al. Update of the Dutch Manual for Costing in Economic Evaluations. *Int J Technol Assess Health Care* 2012;28(2):152–8.

36 Dutch Healthcare Authority (in Dutch: Nederlandse Zorgauthoriteit, NZa). NZa Zorgproducten Tariefapplicatie. Available from: http://dbc-zorgproducten-tarieven.nza.nl/nzaZpTarief/ZoekfunctieDot.aspx. Accessed February 4, 2015.

37 National Health Care Institute. Diemen: National Health Care Institute. Available from: http://www.medicijnkosten.nl/default.asp. Accessed February 2, 2015.

38 Statistics Netherlands. The Hague: Statistics Netherlands. Available from: http://statline.cbs.nl. Accessed February 4, 2015.

39 Buisman LR, Rijnsburger AJ, Den Hertog HM, et al. Clinical Practice Variation Needs to be Considered in Cost-Effectiveness Analyses: A Case Study of Patients with a Recent Transient Ischemic Attack or Minor Ischemic Stroke. *Appl Health Econ Health Policy* 2016;14(1):67–75.

40 Dutch Institute for Healthcare Improvement CBO. Guideline: Diagnostics, treatment and healthcare for patients with a stroke (In Dutch: Richtlijn: Diagnostiek, behandeling en zorg voor patiënten met een beroerte). Utrecht: Dutch Institute for Healthcare Improvement CBO; 2008.

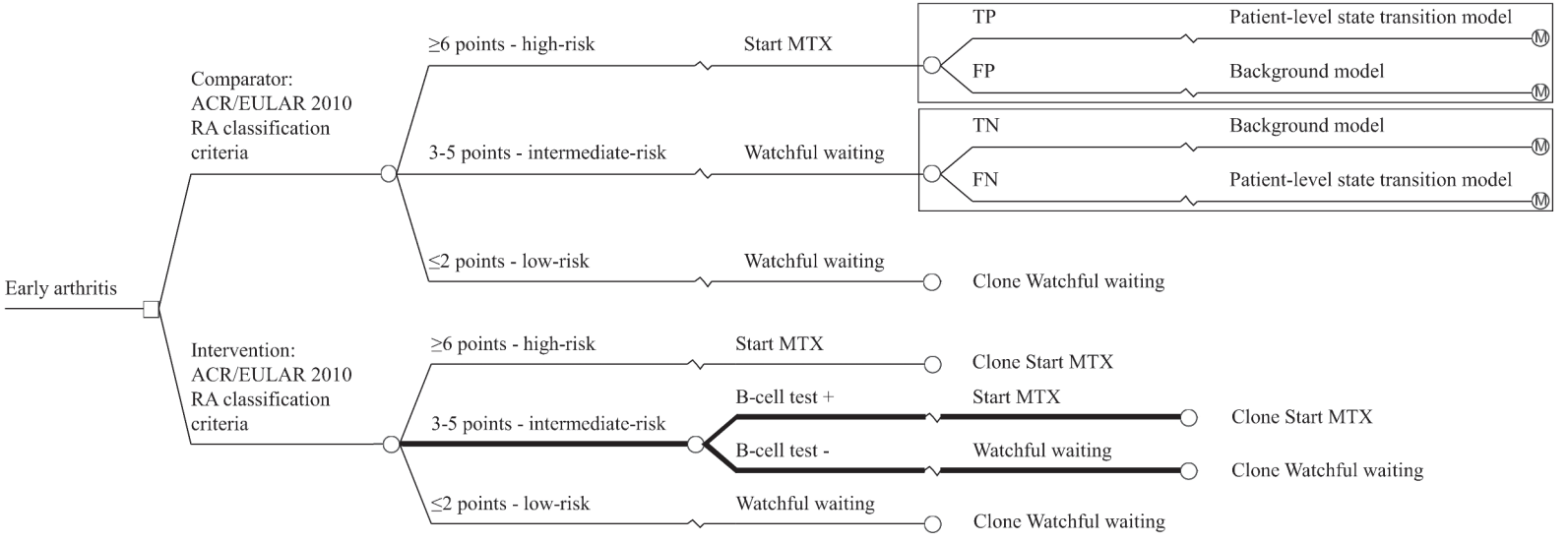
41 Tholen ATR, De Monyé C, Genders TSS, et al. Suspected Carotid Artery Stenosis: Cost-effectiveness of CT Angiography in Work-up of Patients with Recent TIA or Minor Ischemic Stroke. *Radiology* 2010;256:585–97.

42 Drummond M, Griffin A, Tarricone R. Economic evaluation for devices and drugs—same or different? *Value Health* 2009;12(4):402–4.

43 Cao Q, Postmus D, Hillege HL, et al. Probability elicitation to inform early health economic evaluations of new medical technologies: a case study in heart failure disease management. *Value Health* 2013;16(4):529–35.

2.8 Appendix

Appendix 2.1: Decision tree comparing the current diagnostic test strategy with the new diagnostic add-on test strategy for intermediate-risk patients



MTX = methotrexate; TP = true positive; FP = false positive; TN = true negative; FN = false negative.

CHAPTER 3

COSTS OF DIAGNOSTIC TESTS IN THE WORK-UP OF EARLY INFLAMMATORY ARTHRITIS PATIENTS: A MIXED-METHODS APPROACH

Bernadette J.M. Benner[1], Leander R. Buisman[2,3], Marijn Vis[1], Maureen P.M.H. Rutten-van Mölken[2,3], Johanna M.W. Hazes[1], Jolanda J. Luime[1]

[1] Department of Rheumatology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
[2] Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands
[3] Institute for Medical Technology Assessment, Erasmus University Rotterdam, Rotterdam, The Netherlands

Submitted for publication.

ABSTRACT

Objectives: To provide insight in the costs of diagnostic tests, which tests are requested, what differential diagnoses are considered, and what strategies are used by rheumatologists in patients with early inflammatory arthritis (EIA) at risk of having rheumatoid arthritis (RA).

Methods: A mixed-methods approach was used combining diagnostic test resource use data and its costs of 114 new EIA patients (50% RA) of one outpatient clinic with data from semi-structured telephone interviews with 11 rheumatologists to elicit views on the diagnostic process of EIA and requesting of tests.

Results: A median of 25 [IQR 20-35] tests were requested with mean costs of €422 (SD: €168) per patient. Imaging contributed to half of the costs and laboratory tests to the other half. Of the laboratory tests, specific antibody and detailed tests using ELISA techniques incurred the highest costs. Diagnostic tests for SLE, Sjögren, crystal arthropathies, and spondylartropathies were observed frequently, with ANA being the most frequent requested test outside the array of standard laboratory tests. Rheumatologists confirmed that no sequential ordering strategy was used. Patient characteristics and distribution of painful joints were the main reasons to rule in or rule out certain diagnoses. If a patient presented with an atypical joint distribution, more tests were requested to rule out diagnoses.

Conclusions: Rheumatologists spend about 25% of costs compensated by healthcare insurers (i.e., tariff) on diagnostic tests in EIA patients at risk of having RA. Most of these tests were used to rule out other diagnoses in which imaging modalities induce the highest cost per test.

Keywords: rheumatoid arthritis, diagnosis, costs.

3.1 Introduction

Rheumatoid arthritis (RA) is a chronic inflammatory, autoimmune disease predominantly characterized by arthritis of small joints potentially leading to irreversible joint destruction. If RA would be left untreated then it would lead to severe disability, co-morbidities and premature mortality.[1-3] Optimal patient outcome is achieved by early tight controlled disease modifying treatment.[2,4,5] This requires early detection of RA for which rheumatologists request diagnostic tests.

In order to confirm the presence of RA, a few diagnostic tests are needed: ESR, Rheumatoid Factor, ACPA, radiographs of hands and feet, and physical examination. However, differential diagnoses need to be ruled out before the ACR/EULAR 2010 RA classification criteria can be applied.[1] Which tests are requested depends on other likely diagnoses considered by the rheumatologist. Paper based scenarios evaluated by French rheumatologists suggest that there is no clear set of tests for alternative diagnoses and that under increasing uncertainty doctors tend to request more diagnostic tests.[6] No data is available for what rheumatologists do in daily practice and what the costs are of the diagnostic work-up of patients at risk of having RA. With pressure on hospital and healthcare budgets, rheumatologists are confronted with budget cuts and, therefore, need to choose which care to provide.

This paper aims to provide insight in the costs of diagnostic tests, which tests are requested, what differential diagnoses (DDx) are considered, and what strategies are used by rheumatologists in patients with early inflammatory arthritis (EIA) at risk of having RA.

3.2 Methods

This study uses a mixed-methods approach obtaining both quantitative and qualitative data on diagnostic tests. The quantitative data was collected from hospital records of newly diagnosed patients with EIA from one academic hospital in the Netherlands in 2010. The qualitative data was collected by semi-structured interviews with 11 randomly selected rheumatologists from 10 different hospitals across the Netherlands.

Quantitative data

Patient selection

Patients with EIA that visited the outpatient clinic for the first time in 2010 were identified using the 'diagnosis treatment combinations' (DBC; see Appendix 3.1 for the underlying categories) database of the hospital (retrospective cohort). In the Netherlands, healthcare insurers reimbursed the hospital for each patient using the fixed price of a DBC in 2010. Because patients were assigned to a particular DBC in the early stage of diagnostic

work-up, assigned DBCs are at risk of misspecification. Therefore, we checked each diagnosis against the patient's medical hospital record. Patients with RA, mono-arthritis, oligo-arthritis and polyarthritis were included. Second opinions and patients previously diagnosed by other rheumatologists were excluded. Diagnostic tests performed elsewhere were excluded from our analysis.

Diagnostic test data and costs

For each patient, all diagnostic test data and costs were extracted from the hospital database. The cost price included costs as specified by the hospital. The different diagnostic tests were grouped related to a DDx that was likely considered by the rheumatologists. This grouping was based on literature and discussions with experienced rheumatologists.[1,7-35] Details are presented in Appendix 3.2.

Data analyses

Descriptive statistics were used to describe the patterns of requesting diagnostic tests and the accompanying costs for each of the 5 diagnoses. STATA12 and Microsoft Excel were used to organize and analyze the data.

Qualitative data - Interviews

Selection of rheumatologists

To investigate practice variation between rheumatologists in the Netherlands on the diagnostic work-up of patients with EIA, we conducted semi-structured interviews with 11 rheumatologists. These rheumatologists were randomly selected from the Dutch Society for Rheumatology (NVR) members list. Forty rheumatologists received an invitation per email to participate in our study.

Data collection

We started each interview with the following question: "What differential diagnoses would you consider for a patient with complaints of early arthralgia at a first consultation, taken into account that the complaints might be explained by arthritis?" We asked the rheumatologists what their first thoughts were when being confronted with this hypothetical patient. Further pre-specified questions addressed characteristics of inflammatory joint disease and aspects of the differential diagnostic reasoning. We also queried about the role of diagnostic tests and their way of reasoning behind requesting these tests.

Analysis

ATLAS.ti 5.2 was used to code the quotations derived from the interviews. Two investigators read and coded two interview transcripts of the first

coder, independently. The three investigators discussed the codes derived from the interview content, by which consensus was achieved about the codes used.

3.3 Results

Quantitative analysis

In 2010, 863 new patients visited the outpatient clinic of which 335 had a DBC indicating an EIA diagnosis (gout excluded). After checking the patients' medical records for the clinical diagnosis, 114 patients had a clinical diagnosis relevant for our analysis: RA (n=49), probable RA (n=8), mono-arthritis (n=16), oligo-arthritis (n=23) and poly-arthritis (n=18).

Costs

The mean costs of diagnostic tests per patient were €422 (SD: €328) with a median of 25 (IQR 20-35) tests per patient (see Table 3.1). A cost breakdown of the requested tests is shown in Appendix 3.3. The costs of individual patients varied between €72 and €2,033. Three patients (2%) had costs less than €100 on diagnostics, while in 6 patients (5%) the costs were more than €1,000.

The total costs of 144 patients were about €48,000. Standard laboratory tests for internal medicine (leukocytes, CRP, ESR, hematocrit, differential count, granulocytes count, MCV, hemocytometry, ALAT, ASAT) resulted in €1,496 (3% of the total costs), while plain radiographs of hands and feet accounted for €9,975 (21% of the total costs). Although only few MRIs and other imaging modalities were used in all patients, the high costs of these techniques had substantial impact on the total diagnostic test costs (€24,400). Compared to specific blood markers (e.g., ACPA) the imaging modalities were about 5 to 10 times more expensive per test. Small differences in costs were observed between the different diagnoses as shown in Table 3.1.

Table 3.1: Costs and number of tests per diagnostic category

	Number of patients	Median number of diagnostic tests per patient (SD)	Mean costs per patient (SD)
(probable) RA	57	25 (16)	€424 (€27)
Mono-arthritis	16	23 (17)	€438 (€31)
Oligo-arthritis	23	25 (8)	€398 (€41)
Poly-arthritis	18	31 (18)	€454 (€23)
	Total	**25**	**€422**

SD = standard deviation

The costs related to DDx are presented in Table 3.2. Median costs resemble the total costs per patient in which the diagnosis was considered, rather than the specific test costs per DDx. The highest costs were observed among patients in which Polymyositis/dermatomyositis/Scleroderma and Vasculitis were considered, but this was diagnosed in only very few patients. Among more frequent considered DDx, intestinal illness resulted in the highest costs, which is mainly due to the use of imaging modalities, such as MRI and CT. The octreotide scan was the most expensive test (€787 per scan), but was only requested once to rule out sarcoidosis. It had little impact on the overall costs for patients in which sarcoidosis was considered.

In patients finally diagnosed with RA, only one extra diagnosis was often considered (mostly SLE/M. Sjögren). For mono-, oligo- and polyarthritis the DDx was more divers but with a mean of 2 considered diagnoses (according to our analysis of the requested diagnostic tests).

Table 3.2: Considered differential diagnosis and overall costs

		Median costs per patient	
Diagnostic tests related to:	**%[$]**	**€ (min-max)***	
SLE / M. Sjögren	42	€437	(€91-€2,033)
Crystal arthropathy-(pseudo)gout	22	€418	(€72-€1,783)
Septic arthritis[#]	15	€688	(€329-€2,033)
Intestinal illnesses	10	€790	(€87-€2,033)
Malignancy	12	€575	(€277-€1,783)
Polymyositis/dermatomyositis/Scleroderma	4	€802	(€219-€2,033)
Reactive arthritis	7	€680	(€335-€1,705)
Sarcoidosis	8	€423	(€293-€1,451)
Spondylarthropathy/M. Bechterew/IBD	17	€519	(€72-€1,451)
Vasculitis	2	€1,037	(€802-€,1273)
Viral arthritis	14	€541	(€335-€2,033)

[$] Number of patients in which the diagnosis was considered; [#] 'Septic arthritis' includes pathogenics like bacteria, mold, protozoa, Lyme infection, legionella infection and tuberculosis; * This presents the total costs made in those particular patients – not separated for the specific diagnostic tests.

Qualitative analysis

Eleven rheumatologists were available for interview. Nine of them were trained rheumatologists with a median practice experience of 14 years (range 1-29 years), and two were trainee rheumatologists. Five rheumatologists were working in an academic hospital, while the others were practicing in a general hospital.

The rheumatologists responded with similar strategies on what to do in case a patient with inflammatory arthralgia at risk of having RA presented

in their practices for the first time. All strategies included a thorough history taking, with special attention to duration of symptoms, number of joints involved and inflammatory characteristics (e.g., morning stiffness and swollen joints, followed by a physical examination). The combination of patient characteristics (sex, age, comorbidities, family history, and prior medical events), the number of joints involved, and accompanying symptoms (psoriasis, sunlight hypersensitivity, Raynaud's syndrome, sicca complaints and uveitis) were important in creating the eventual DDx.

After narrowing down the list of potential reasons for the observed arthritis, diagnostic tests were requested to rule in or out certain diagnoses. Independent of the DDx, a subset of diagnostic tests, in particular RF and ACPA, was requested by rheumatologists in almost all patients with EIA (independent of the likelihood of the RA diagnosis). Rheumatologists request all needed diagnostics during the first consultation of the patient. This approach was mainly based on practical than on medical reasons. The practical reasons were related to the organization of the healthcare system, time constraints and the need for a quick resolution of the diagnostic uncertainty of the patient.

Diagnostic uncertainty of the physician was a theme that emerged in relation to the number of diagnostic tests requested. On the one hand, if rheumatologists expected a benign course, a wait-and-see policy could be applied to see whether symptoms become more pronounced or would resolve. On the other hand, if the uncertainty was related to a potentially underlying threatening disease such as septic arthritis, rheumatologists aimed to rule out all possible causes before accepting that no clear cause explained the presence of the arthritis. This uncertainty increased the number of diagnostic tests, which we also found in our retrospective cohort for the patients suspected of septic arthritis among the mono-arthritis patients and viral diagnostics in the poly-arthritis patients.

3.4 Discussion

Rheumatologists spend about 25% of costs compensated by healthcare insurers (i.e., tariff) on diagnostic tests in EIA patients at risk of having RA (i.e., mean diagnostic test costs are €422 per patient). Whether the final diagnosis was RA or unclassified arthritis (i.e., mono/oligo/polyarthritis) had little impact on the costs, although differences were found in the DDx considered. Imaging contributed to half of the diagnostic test costs due the usual hand and feet x-rays, MRI and scintigraphy and laboratory tests to the other half. Standard laboratory tests accounted for 3% of the total costs, while plain radiographs of hands and feet contributed 21% to the total costs.

We found few studies which evaluated the costs of requested diagnostic tests for the DDx process in EIA. Using case vignettes (i.e., paper patients),

three studies from France found wide variation among rheumatologists in their requested diagnostic tests.[36,37] As we used retrospective data of patients of one hospital we were unable to investigate whether part of the costs were related to unnecessary requested tests. However, given the French data and the substantial diagnostic test costs per patient in our study, we recommend a critical appraisal of budgets spent on diagnostic tests in daily clinical practice. This may help rheumatologists in making informed decisions about requesting diagnostic tests given the confronted budget cuts. In addition, particular tests could be carried out by more than one technique, which vary substantially in costs. It was sometimes unclear why the laboratory used a particular technique if a cheaper alternative was available with similar test characteristics.

Limitations of our research may include the retrospective data analysis, patient selection and limited number of interviewed rheumatologists. The retrospective data analysis was performed using data in reverse chronological order and the diagnoses considered were based on the requested diagnostic tests. We could have assumed the wrong DDx if the same test was used for another DDx than we considered in our list of DDx. The amount of distortion of the results depends on how specific a certain test is for the underlying condition. For example, a Borellia test is specifically requested for Lyme's disease, while a colonoscopy can be used in diagnosing an inflammatory bowel disease or a malignancy. For the interpretation of the diagnostic costs, the reverse order of interpretation did not matter as we used the observed costs per patient.

The patient selection may be not generalizable to general rheumatologists. We selected patients from one academic hospital for which the data on type of test and test costs were available for analysis. A risk is the presence of a more severe case load in academic compared to general hospitals. Although we were not fully able to rule out this risk, we selected those patients that visited the outpatient clinic for the first time, excluding second and third opinions. Regarding patients that were included in other studies,[38] case load seems comparable to general hospitals for levels of disease activity. Therefore, we assume that the patients in our study are a cross-section of the EIA population in the Netherlands.

Some readers may regard the representative sample of 11 Dutch rheumatologists interviewed small. This is true if assessed from a quantitative research perspective, but in qualitative research, interviews are held until saturation occurs (i.e., no additional information could be elicited in subsequent interviews). However, there may be another limitation. During the interviews, the rheumatologists reflected on what to do in practice, rather than we observed what they really do. This concern was eliminated by comparing the answers with real life data of our quantitative analysis. Our quantitative data confirmed that a substantial part of the diagnostic tests were

requested after first consultation, and if arthritis had been objectified, RF and ACPA were requested in almost all patients.

In conclusion, rheumatologists spend about 25% of costs compensated by healthcare insurers (i.e., tariff) on diagnostic tests in EIA patients at risk of having RA. Most of these tests were used to rule out other diagnoses in which imaging modalities induce the highest cost per test. Standard laboratory tests and radiographs of hands and feet account for 24% of diagnostic test costs. Atypical patient presentation increased diagnostic test use, especially if the patient presented with acute symptoms.

3.5 Acknowledgments

This research was performed within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), and the Dutch Arthritis Foundation, project TRACER (grant O4I-202).

3.6 References

1 Aletaha D, Neogi T, Silman AJ, et al. Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* 2010;62(9):2569–81.

2 Hazes JM, Luime JJ. The epidemiology of early inflammatory arthritis. *Nat Rev Rheumatol* 2011;7(7):381–90.

3 Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. *Lancet* 2010;376(9746):1094–1108.

4 Aletaha D, Eberl G, Nell VP, et al. Attitudes to early rheumatoid arthritis: changing patterns. Results of a survey. *Ann Rheum Dis* 2004;63(10):1269–75.

5 Van Dongen H, Van Aken J, Lard LR, et al. Efficacy of methotrexate treatment in patients with probable rheumatoid arthritis: a double-blind, randomized, placebo-controlled trial. *Arthritis Rheum* 2007;56(5):1424–32.

6 Fautrel B, Saraux A, Maillefert JF, et al. Costs of workups for the diagnosis of early arthritis: results of a nationwide survey. *Arthritis Rheum* 2004;51(4):507–12.

7 Agarwal V, Dabra AK, Kaur R. Remitting seronegative symmetrical synovitis with pitting edema (RS3PE) syndrome: ultrasonography as a diagnostic tool. *Clin Rheumatol* 2005;24(5):476–9.

8 Alamanos Y, Tsifetaki N, Voulgari PV. Epidemiology of primary Sjogren's syndrome in north-west Greece, 1982-2003. *Rheumatology (Oxford)* 2006;45(2):187–91.

9 Bayetto K, Logan RM. Sjogren's syndrome: a review of aetiology, pathogenesis, diagnosis and management. *Aust Dent J* 2010;55(1):39–47.

10 Bertsias G, Ioannidis JP, Boletis J, et al. EULAR recommendations for the management of systemic lupus erythematosus. Report of a Task Force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics. *Ann Rheum Dis* 2008;67(2):195–205.

11 Bird HA, Leeb BF, Montecucco CM, et al. A comparison of the sensitivity of diagnostic criteria for polymyalgia rheumatica. *Ann Rheum Dis* 2005;64(4):626–9.

12 Bourikas LA, Papadakis KA. Musculoskeletal manifestations of inflammatory bowel disease. *Inflamm Bowel Dis* 2009;15(12):1915–24.

13 Choi HK, Atkinson K, Karlson EW, et al. Purine-rich foods, dairy and protein intake, and the risk of gout in men. *N Engl J Med* 2004;350(11):1093–103.

14 Dalakas MC, Hohlfeld R. Polymyositis and dermatomyositis. *Lancet* 2003;362(9388):971–82.

15 Fam AG. Paraneoplastic rheumatic syndromes. *Baillieres Best Pract Res Clin Rheumatol* 2000;14(3):515–33.

16 Goldenberg DL. Septic arthritis. *Lancet* 1998;351(9097):197–202.

17 Gunnarsson R, Molberg O, Gilboe IM, et al. The prevalence and incidence of mixed connective tissue disease: a national multicentre survey of Norwegian patients. *Ann Rheum Dis* 2011;70(6):1047–51.

18 Helmick CG, Felson DT, Lawrence RC, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I. *Arthritis Rheum* 2008;58(1):15–25.

19 Kassan SS, Moutsopoulos HM. Clinical manifestations and early diagnosis of Sjogren syndrome. *Arch Intern Med* 2004;164(12):1275–84.

20 Kortekaas MC, Kwok WY, Reijnierse M, et al. Pain in hand osteoarthritis is associated with inflammation: the value of ultrasound. *Ann Rheum Dis* 2010;69(7):1367–9.

21 Kwok WY, Kloppenburg M, Rosendaal FR, et al. Erosive hand osteoarthritis: its prevalence and clinical impact in the general population and symptomatic hand osteoarthritis. *Ann Rheum Dis* 2011;70(7):1238–42.

22 Mannoni A, Briganti MP, Di Bari M, et al. Epidemiological profile of symptomatic osteoarthritis in older adults: a population based study in Dicomano, Italy. *Ann Rheum Dis* 2003;62(6):576–8.

23 Minden K. Adult outcomes of patients with juvenile idiopathic arthritis. *Horm Res* 2009;72(1):20–5.

24 Nagaraju K, Lundberg IE. Polymyositis and dermatomyositis: pathophysiology. *Rheum Dis Clin North Am* 2011;37(2):159–71.

25 Neal-Boylan L. Update on rheumatology: part 1. *Home Healthc Nurse* 2009;27(5):286–96; quiz 297–8.

26 Nothnagl T, Leeb BF. Diagnosis, differential diagnosis and treatment of polymyalgia rheumatica. *Drugs Aging* 2006;23(5):391–402.

27 Petri M. Review of classification criteria for systemic lupus erythematosus. *Rheum Dis Clin North Am* 2005;31(2):245–54.

28 Roddy E, Doherty M. Epidemiology of gout. *Arthritis Res Ther* 2010;12(6):223.

29 Rudwaleit M, Van der Heijde D, Landewe R, et al. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. *Ann Rheum Dis* 2011;70(1):25–31.

30 Rugiene R, Dadoniene J, Aleknavicius E, et al. Prevalence of paraneoplastic rheumatic syndromes and their antibody profile among patients with solid tumours. *Clin Rheumatol* 2011;30(3):373–80.

31 Taylor W, Gladman D, Helliwell P, et al. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. *Arthritis Rheum* 2006;54(8):2665–73.

32 Wallace SL, Robinson H, Masi AT, et al. Preliminary criteria for the classification of the acute arthritis of primary gout. *Arthritis Rheum* 1977;20(3):895–900.

33 Zhang W, Doherty M, Bardin T, et al. EULAR evidence based recommendations for gout. Part II: Management. Report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis* 2006;65(10):1312–24.

34 Zhang W, Doherty M, Leeb BF, et al. EULAR evidence-based recommendations for the diagnosis of hand osteoarthritis: report of a task force of ESCISIT. *Ann Rheum Dis* 2009;68(1):8–17.

35 Zhang W, Doherty M, Pascual E, et al. EULAR evidence based recommendations for gout. Part I: Diagnosis. Report of a task force of the Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis* 2006;65(10):1301–11.

36 Saraux A, Maillefert JF, Fautrel B, et al. Laboratory and imaging studies used by French rheumatologists to determine the cause of recent onset polyarthritis without extra-articular manifestations. *Ann Rheum Dis* 2002;61(7):626–9.

37 Saraux A, Fautrel B, Maillefert JF, et al. Laboratory and imaging studies used by French rheumatologists to evaluate patients with early arthritis. *J Rheumatol* 2006;33(5):897–902.

38 Alves C, Luime JJ, Van Zeben D, et al. Diagnostic performance of the ACR/EULAR 2010 criteria for rheumatoid arthritis and two diagnostic algorithms in an early arthritis clinic (REACH). *Ann Rheum Dis* 2011;70(9):1645–7.

3.7 Appendices

Appendix 3.1: Selected DBC codes and names

DBC diagnosis codes	Name of DBC code	DBC diagnosis codes	Name of DBC code
101	Rheumatoid arthritis	305	Scleroderma
105	Reactive arthritis	307	Poly- dermatomyositis
108	Lyme arthritis-Borreliosis	308	M. Sjögren
109	Bacterial arthritis	309	PMR
110	Parvo virus infection	310	Temporal arteritis
113	Palindromic rheumatism	312	M. Wegener
115	Mono arthritis – unclassified	313	Sarcoidosis
116	Oligo arthritis – unclassified	316	Relapsing polychondritis
117	Poly arthritis – unclassified	399	Other vasculitis | Systemic diseases
118	Synovitis villonodularis pigmentosa	401	Axial (spine)
201	Ankylosing spondyli tis	701	Arthralgia and/or myalgia
202	Psoriatic arthritis (mainly axially)	702	Arthropathy in endocrine disease
203	Unclassified spondylartropathy	703	Carpal tunnel syndrome | entrapment
301	SLE	705	Dupuytren|other fibromatosis
303	Lupus like antiphospholipid s.	706	Enthesopathy
304	CREST syndrome	720	Psoriasis

Appendix 3.2: Differential diagnoses with their disease specific diagnostic test(s)

DIAGNOSIS	DIFFERENTIATING TESTS
SLE / M. Sjögren	Antinuclear antibodies (ANA)
	Antibody against ENA nucleair antigen
	Antibody DNA *
	Complement component C3 *
	Complement component C4 *
	Immuneglobulin G- IGG #
Crystal arthropathy-(pseudo)gout	Synovial fluid pyrophosphate crystals
	Synovial fluid uric acid crystals
	Uric acid
Septic arthritis	Legionella antigen rapid test ×
	Borrelia antibodies IGG α
	Protozoan cysts ‡
	Strongyloides stercoralis ‡
	Candida culture test @
	Mycological examination general @
	Toxoplasma -IGG
	Toxoplasma -IGM
	Tuberculosis provision -quantiferon-IFN-Gamma quantitative ∧
Intestinal illnesses	Colon- coloscopy including the ileum
	Small intestine -oesofagogastroduodenoscopy
	Culture -anaerobically
	Culture test -2-3 media bacteria
	Culture test - more than 3 media bacteria
	Determination micro-organisms bacteria
Malignancy	CT abdomen
	CT thorax
	CT upper abdomen
	Ultrasound abdomen
	CA-125 immuno-enzymes &
	Prostate specific antigen $
Polymyositis/dermatomyositis/ Scleroderma	Creatine phosphokinase
	Oesofagography-oral ~
Reactive arthritis	Anti-chlamydial IGA
	Anti-chlamydial IGG
	Anti-chlamydial IGM
	Antistreptolysin titer- serology
	Chlamydia trachomatis PCR

Appendix 3.2: Differential diagnoses with their disease specific diagnostic test(s) (continued)

DIAGNOSIS	DIFFERENTIATING TESTS
	Lues antibodies IGM
	Neisseria gonorrhoe
Sarcoidosis	Angiotensin-converting enzyme (ACE)
	Neuro-endocrine SCINT.COMP.DETAIL STAT. IN-111 OCTREO
Spondylarthropathy/M. Bechterew/IBD	HLA-B27
	Hip-thigh
	Pelvis-lying
Vasculitis	Cryoglobuline
Viral arthritis	Anti hepatitis-BC virus -HBC- IGT-ELISA
	Anti hepatitis-C virus -ELISA blood
	Anti-adenovirus IGA -ELISA blood
	Antibody against hepatitis-B surface antigen quantitative IA
	Anti-coxsackievirus IGA EIA quantitative
	Anti-coxsackievirus IGG EIA quantitative
	Anti-coxsackievirus IGM EIA quantitative
	Anti-cytomegalovirus IG of IGG-ELISA virology
	Anti-cytomegalovirus IGM-ELISA
	Anti-echovirus IGA -EIA quantitative
	Anti-echovirus IGG -EIA quantitative
	Anti-echovirus IGM -EIA quantitative
	Anti-enterovirus IGA -ELISA
	Anti-enterovirus IGM -ELISA
	Anti-epstein barr after IGG -ELISA serum
	Anti-epstein barr EA IGG -ELISA serum
	Anti-epstein barr VCA IGG -ELISA serum
	Anti-epstein barr VCA IGM -ELISA serum
	Anti-HAV IG of IGG-ELISA or RIA in a sample
	Anti-hepatitis-E-virus IGG -ELISA various sources
	Anti-hepatitis-E-virus IGM -ELISA various sources
	Anti-herpes-simplex-virus IGM ELISA
	Anti-herpes-simplex-virus-1 -HSV-1 IGG -ELISA
	Anti-influenza A-virus IGA -ELISA blood
	Anti-influenza B-virus IGA -ELISA blood
	Anti-LAV-HTLV-III IG-ELISA
	Anti-para-influenza-virus type 1 IGA -ELISA blood
	Anti-parvovirus B19 IGM -ELISA serology

Appendix 3.2: Differential diagnoses with their disease specific diagnostic test(s) (continued)

DIAGNOSIS	DIFFERENTIATING TESTS
	Anti-parvovirus IGG -ELISA various sources
	Anti-respiratory syncytiumvirus -RSV- IGA - ELISA blood
	Anti-rubella IGG-ELISA
	Anti-rubella IGM ELISA
	Anti-varicella-zoster-virus IGM ELISA
	Cytomegalovirus DNA-CMV -PCR
	Epstein barr virus DNA -quantitative PCR
	Epstein barr virus DNA-PCR
	Hepatitis-B surface antigen IA
	Herpes simplex virus DNA-HSV -PCR
	Parvo-DNA -PCR
	Varicella zoster virus genome detection -PCR
	Virological tests cell culture - less than 2 media
	Virus detection -IFT after cell culture

* Diagnostic tests specific in SLE
\# Diagnostic tests specific in M. Sjögren
× Diagnostic tests specific in Legionella
α Diagnostic tests specific in Lyme arthritis
‡ Diagnostic tests specific for protozoa
@ Diagnostic tests specific in mold infection
∧ Diagnostic tests specific in tuberculosis
& Diagnostic tests specific in ovarian carcinoma
$ Diagnostic tests specific in prostate carcinoma
~ Diagnostic tests specific in scleroderma

Appendix 3.3: Costs of tests grouped by application

Application	Units per patient	Mean test cost per unit	Mean test costs per patient
Imaging	5.21	€ 41.08	€ 214.07
Important diagnostics	5.10	€ 16.69	€ 85.06
Standard intern lab	7.26	€ 1.81	€ 13.12
General laboratory	0.30	€ 2.99	€ 0.89
Immune protein	0.01	€ 24.53	€ 0.22
Hematology/anemia	2.51	€ 3.28	€ 8.23
Deficiencies/osteoporosis	0.25	€ 1.08	€ 0.27
Coagulation	1.07	€ 5.88	€ 6.29
Liver chemistries	1.51	€ 1.19	€ 1.80
Kidney laboratory tests	2.53	€ 1.13	€ 2.85
Laboratory tests of digestive tract	0.06	€ 98.58	€ 6.05
Laboratory tests of respiratory tract	0.13	€ 25.76	€ 3.39
Musclefunction test	0.01	€ 0.20	€ 0.00
Tumor marker	0.01	€ 6.25	€ 0.05
Serology	0.54	€ 23.63	€ 12.65
Hormone laboratory tests	0.02	€ 13.79	€ 0.24
(Para) Thyroid laboratory tests	0.02	€ 12.12	€ 0.21
Cholesterol laboratory tests	1.03	€ 2.17	€ 2.23
Glucose laboratory tests	0.59	€ 2.16	€ 1.27
Pathogens-bacteria	0.89	€ 21.90	€ 19.59
Pathogens-mold	0.03	€ 16.11	€ 0.42
Pathogens-protozoa	0.02	€ 36.60	€ 0.64
Pathogens-virus	0.89	€ 24.26	€ 21.71
Tissue research	0.09	€ 67.59	€ 5.93
Flowchart	0.96	€ 12.43	€ 11.99
CSF (cerebral spine fluid) examination	0.11	€ 7.74	€ 0.81
Urinalysis	0.35	€ 5.19	€ 1.82
Total			**€ 421.82**

CHAPTER 4

A COST-EFFECTIVENESS MODEL FOR EVALUATING NEW DIAGNOSTIC TESTS IN THE WORK-UP OF PATIENTS WITH INFLAMMATORY ARTHRITIS AT RISK OF HAVING RHEUMATOID ARTHRITIS

Jolanda J. Luime[1], Leander R. Buisman[2,3], Mark Oppe[2,3,4], Johanna M.W. Hazes[1], Maureen P.M.H. Rutten-van Mölken[2,3]

[1] Department of Rheumatology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
[2] Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands
[3] Institute for Medical Technology Assessment, Erasmus University Rotterdam, Rotterdam, The Netherlands
[4] EuroQol Research Foundation, Rotterdam, The Netherlands

First authorship: JJL and LRB contributed equally to this work.

Published in *Arthritis Care & Research* 2015. Epub 2015 Nov 10.
doi:10.1002/acr.22776.

ABSTRACT

Background: New opportunities have emerged for early diagnosis with the arrival of new technologies that assess the impact of genomics, proteomics, metabolomics and cytomics on rheumatoid arthritis (RA) risk.

Objective: This early Health Technology Assessment study assesses the short term cost-effectiveness of 4 add-on diagnostic tests in early inflammatory arthritis patients at risk of RA.

Methods: We modeled four diagnostic add-on tests to the ACR/EULAR 2010 RA classification criteria covering the first year after diagnosis, using REACH data. Sensitivity (Se), specificity (Sp) and costs were assigned to the magnetic resonance imaging (MRI) of hands and feet (Se 0.90; Sp 0.60; €756), IL6 serum level test (Se 0.70; Sp 0.53; €50), B-cell related gene expression (Se 0.60; Sp 0.90; €150,) and gene assay for RA (Se 0.40;Sp 0.85;€750), based on literature and expert opinion. Outcomes were evaluated using the unweighted diagnostic Net Benefit (udNB) and the incremental costs per quality-adjusted life year gained (ICER) in all patients (n=552), intermediate risk patients (n=263) and seronegative patients (n=329).

Results: The highest udNB was found when using B-cell assay in intermediate risk patients (43%; ICER €5,314), while the IL6 test in seronegative patients resulted in the lowest udNB (-11.4%; ICER €7,650). If a threshold of €20,000 is applied, the B-cell assay would be preferred over the other alternatives with a probability of being cost-effective of 78% for intermediate risk patients, 57% for all patients and 73% for seronegative patients.

Conclusion: Diagnostic add-on tests favoring specificity over sensitivity with a headroom less than €370 per test are cost-effective, with the largest diagnostic benefit occurring in intermediate risk patients.

Keywords: early arthritis, cost-effectiveness, diagnosis, rheumatoid arthritis.

4.1 Introduction

In 13% to 32% of patients with inflammatory arthritis (IA), the underlying pathology is rheumatoid arthritis (RA).[1-4] Detection of RA is essential to reduce the probability of erosive disease and to optimize patient outcomes in terms of physical ability and social participation, including work.[5] Treatment with synthetic Disease Modifying Drugs (sDMARDs), and if necessary subsequent biological DMARDs, reduces radiological progression rates by 1-7 Sharp van der Heijde points per year.[6] This results in a 60% relative benefit compared to the control group, and a reduction in absenteeism of 50%.[7] The benefit of treatment, however, depends on the early commencement of treatment. The window of opportunity for treatment is regarded as the first 3 to 4 months after first symptoms start.[5,8]

With the arrival of new technologies to assess the impact of genomics, proteomics, metabolomics and cytomics on the development of RA, new opportunities have emerged for early diagnosis. In addition, existing imaging modalities such as MRI and ultrasound (US) may also be useful as they have proved to be effective tools in monitoring disease activity.[9,10] However, it is unclear which options would be valuable in addition to the current diagnostic work-up, taking into account both the metric characteristics and the costs of tests. The purpose of this early Health Technology Assessment (HTA) study was to assess the short term cost-effectiveness of 4 diagnostic tests with assumed test characteristics and costs based on literature and expert opinion, compared with the ACR/EULAR 2010 RA classification criteria, (henceforth RA 2010 criteria).[11] Particular attention was paid to the prior probability of having RA by selecting 3 target populations in which the prevalence of RA was different.

4.2 Materials and Methods

Study population

The target population comprised patients with IA who were suspected of having RA. These were patients presenting with at least one clinical synovitis who could not be classified as suffering from another inflammatory joint disease as suggested by the developers of the RA 2010 criteria.[11] Data was used from the Rotterdam Early Arthritis Cohort (REACH).[12] Details of the baseline characteristics of the patients used for this analysis are described in Appendix 4.1.

Decision tree

A decision tree covering the first year after diagnosis was developed to evaluate and compare diagnostic alternatives as add-on and replacement (see Figure 4.1). For ease of interpretation of the decision tree, only the test strat-

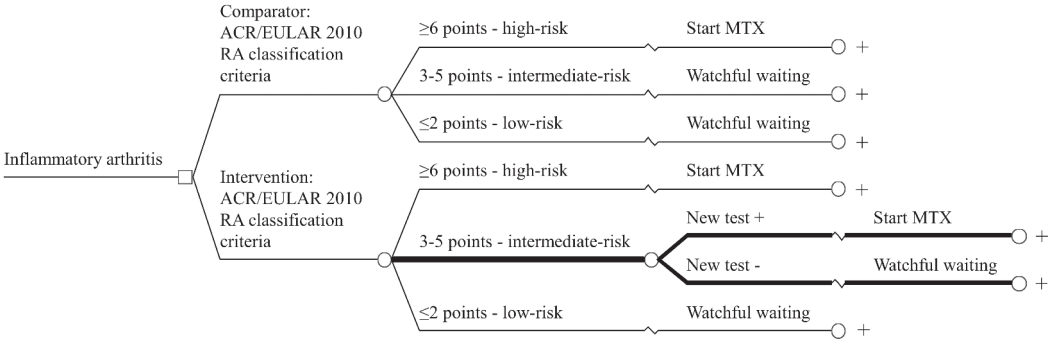
egy for intermediate risk patients is shown in Figure 4.1. Appendix 4.2 shows the follow-up branches (start MTX and watchful waiting) in the decision tree. As there is no gold standard for the diagnosis of RA, either Methotrexate (MTX) or other disease modifying anti-rheumatic drugs (DMARDs) if MTX was prescribed but discontinued due to side-effects, was employed as a surrogate. Table 4.1 shows the criteria used to assign true positive (TP), false positive (FP), true negative (TN), and false negative (FN) patients at 12 months.[11] Changes in diagnosis and treatment were modeled every 6 months as patients in the REACH cohort were seen every 6 months. We did not account for mortality during the 1-year time horizon of the model.

Table 4.1: Criteria used in classifying patients as true positive (TP), false positive (FP), true negative (TN) and false negative (FN). Each of the 4 criteria needs to be fulfilled as described

	≥6 points	Methotrexate – initiation at baseline	Methotrexate – 12 months[$]	Other diagnosis
True positive (TP)	+	+	+	-[#]
False positive (FP)	+	+	-	-
True negative (TN)	-	-	-	-
False negative (FN)	-	-	+	-[~]

+ feature is present; - feature is absent; [$] if a patient refrained from MTX due to side effects this item was also scored positively; [#] if at 12 months another diagnosis was made by the rheumatologist a patients was classified as false positive; ~ if at 12 months another diagnosis was made a patients was classified as true negative.

Figure 4.1: Decision tree comparing the current diagnostic strategy with the combination of the current strategy plus a new test for intermediate risk patients



°+ The branches for 6 and 12 months re-evaluation of the diagnosis were collapsed and can be found in Appendix 4.2.

Current diagnostic test strategy (comparator) – test characteristics

The set of RA 2010 criteria was used as the comparator.[11] The criteria were developed as a risk prediction model that runs from 0-10 points, with higher values indicating higher risk of having RA. A patient was classified as having RA if 6 or more points were scored. However, over time another classified diagnosis may be underlying, which we dealt with as described in Table 4.1. At 12 months a TP rate of 0.62 and a TN rate of 0.77 was observed in REACH.

New diagnostic test strategies (interventions) – test characteristics

The diagnostic tests were: MRI of hands and feet (4 separate MRIs), IL6 serum level test,[13] B-cell related gene expression in whole blood,[14] a genetic assay with susceptibility SNPs (Single Nucleotide Polymorphisms) for RA,[15] and an almost perfect diagnostic test with a sensitivity of 0.99 and specificity of 0.99. The tests were evaluated employing assumed test characteristics and costs supplied by the Dutch Healthcare Authority[16] and, where not available, by utilizing expert opinion. The latter involved discussions with test developers on the costs of equipment and disposal materials, and the use of test costs if a test was already available for another disease. REACH was used as a background cohort in which the new diagnostic tests were simulated.

MRI – test characteristics

MRI may reclassify patients to different joint domains of the RA 2010 criteria if more joints show swelling compared to physical examination.[17-20] MRI also adds information on bone marrow edema[21] and teno-synovitis.[21-23] We set the test characteristics at 0.90 sensitivity and 0.60 specificity. The costs of MRI were set equal to the price reimbursed by the Dutch Healthcare Authority at €189 per MRI (€756 for 2 hands and 2 feet)[16].

Cytomic test – IL6 serum level test - test characteristics

IL6 is a cytokine which is present in inflammation. A recent evaluation of IL6 serum test performance in an early arthritis ESPOIR cohort showed a sensitivity of 0.70 and a specificity of 0.53 for the detection of RA among the IA patients.[13] We used these test properties and assumed €50 per test based on expert opinion.

Genetic assay testing – test characteristics

Heritability for RA is estimated at around 50%-60%. It is regarded as a complex disease in which several genes play a role.[15] It has been suggested that using genetic risk factors with current knowledge would result in a sensitivity of 0.40 with a specificity between 0.80 and 0.90 to identify patients with RA.[15] We used these estimations and chose to apply 0.85 specificity with an estimated €750 per test based on expert opinion.

Genomic test – test characteristics of B-cell profiling

B-cell RNA expression decreases over time in the development of arthritis.[14] Although the exact mechanism is poorly understood, the marker is useful to predict early arthritis occurrence in patients with seropositive arthralgia.[14] Currently, no data is available from patients with early IA. Hence we used data from the seropositive positive arthralgia cohort and discussed with the test developers utilizing a sensitivity of 0.60, specificity of 0.90, and costs of €150.

Input parameters

Table 4.2 presents the input parameters for the model. Additional information on the values of the parameters of the distribution is given in Appendix 4.3. REACH was used to estimate the empirical distributions of the input parameters. As no negative utility values were observed, the Beta distribution was used.

Prior probability

The prior probability of RA varies depending on which early arthritis patients are considered eligible for additional testing.[2] We varied the prior probability (^{P}P) by selecting 3 samples of patients from REACH who would undergo the test:

1. All patients (n=552; ^{P}P: 54%).
2. Patients with intermediate risk (3-5 points on the RA 2010 criteria using MTX at 12 months and who had no other classified disease; n=263; ^{P}P: 37%)
3. Patients who tested negative for anti-CCP and rheumatoid factor using MTX at 12 months and who had no other classified disease (n=329; ^{P}P: 44%)

Treatment

MTX is the cornerstone of treatment in RA.[18,19] On average, 50-70% of patients respond well to this drug.[24-26] We modeled treatment in the first 12 months by starting MTX in patients who scored 6 or more points as suggested by the RA 2010 criteria, and by employing a wait-and-see policy for the patients scoring less than 6 points. Distribution of treatment, side-effects, use of other DMARDs and other classifiable diseases were based on the REACH data (see Table 4.2).

Quality of life

Scores on the EQ-5D utility scale for quality of life at time of diagnosis and follow-up were estimated based on the literature and the REACH cohort using the Dutch EQ-5D tariff,[26] (see Table 4.2). TP patients had a 0.1 improvement in utility in the first year after diagnosis with a baseline level of

0.6 based on REACH data and the literature.[27-31] For the other patients the literature is less clear. We assumed that FP patients had a 0.05 improvement in utility in the first year with a baseline value of 0.65 (based on REACH data). The TN patients were given a 0.1 increase with a baseline of 0.65, as observed in the REACH data, while the FN patients' value decreased by 0.05 with a baseline value of 0.6 using data from the placebo group in the STIVEA trial, assuming FN patients would receive little therapy.[31]

Costs

We used a societal perspective to estimate costs in 2014 Euros. Current diagnostic test strategies, treatment (DMARDs), and treatment monitoring, were elicited from interviews with rheumatologists. These direct cost components were transformed to costs based on the expenditures for personnel, equipment and materials using Dutch reference prices as recommended for economic evaluations.[27] The costs of the diagnostic tests were based on tariffs from the Dutch Healthcare Authority.[26] Where these were not available from the Authority, the test list price of one academic hospital (Erasmus MC) was used. In addition, diagnostic test type and frequency data were elicited from this hospital for patients at risk of having RA.[32] Indirect costs included productivity loss due to absence from work in the previous 6 months, as obtained from REACH during the 1 year time horizon, using the friction cost method.[33]

Table 4.2: Input parameters used in the model

Parameter	Value	SE	Distribution	Source
PREVALENCE OF RA BASED ON RA 2010				
All patients	54%	2.1%	Beta	REACH+
Intermediate risk patients	37%	3.0%	Beta	REACH+
Seronegative patients	44%	2.7%	Beta	REACH+
METHOTREXATE START AND CONTINUATION				
Patients with >=6 points at baseline				
Start at baseline	100%			Guideline[11,24]
Continuation at 6 months	84%	2.4%	Dirichlet	REACH+
Continuation at 12 months†	88%	2.6%	Dirichlet	REACH+
< 6 points at baseline				
Start at baseline	0%			Guideline[11,24]
Start at 6 months	35%	2.7%	Dirichlet	REACH+
Continuation at 12 months†	84%	3.5%	Dirichlet	REACH+
Start at 12 months	14%	5.3%	Dirichlet	REACH+

Table 4.2: Input parameters used in the model (continued)

Parameter		Value	SE	Distribution	Source
OTHER DIAGNOSIS					
≥6 points at baseline					
6 months		14%	2.2%	Dirichlet	REACH[+]
12 months		12%	2.1%	Dirichlet	REACH[+]
< 6 points at baseline					
6 months		51%	2.8%	Dirichlet	REACH[+]
12 months		0%	0%	Dirichlet	REACH[+]
COSTS (€)					
Initial costs RA 2010 strategy		€2,966	€445	Gamma	Estimate*
Baseline costs		€917	€138	Gamma	Estimate*
Follow up during first year		€676	€101	Gamma	Estimate*
MTX costs including monitoring of side effects (1 year)		€102	€15	Gamma	Estimate*
Other DMARD costs (1 year)		€111	€32	Gamma	REACH[+]
Productivity costs (1 year)		€1,160	€360	Gamma	REACH[+]
MRI of both hands and feet		€756			Tariffs DHA[^]
B-cell		€150			Estimate[#]
IL6		€50			Estimate[#]
Genetic profile		€750			Estimate[#]
UTILITIES					
True positive patients	– baseline	0.60	0.0152	Beta	REACH[+]
	12 months	0.70	0.0104		Estimate[#]
False positive patients	– baseline	0.65	0.0164	Beta	REACH[+]
	12 months	0.70	0.0098		Estimate[#]
True negative patients	– baseline	0.65	0.0164	Beta	REACH[+]
	12 months	0.75	0.0111		Estimate[#]
False negative patients	– baseline	0.60	0.0152	Beta	REACH[+]
	12 months	0.55	0.0065		Estimate[#]

SE = standard error; [+] REACH data on file, details about data collection can be found in Alves et al.[12];
^ DHA: Dutch Health Authority; * frequency laboratory tests and radiographics based on expert opinion; medical consultation time based on the Dutch manual for costing in economic evaluations[27]; Costs of these parameters derived from the DHA tariffs; [#] Estimate based on expert opinion, for the se of the utilities data from REACH was used; [†] continuation of MTX at 12 month is the % of MTX users at 6 months that still use MTX at 12 months.

Scenario analyses

We evaluated each of the four tests as add-ons to the RA 2010 criteria in all patients (n=552 in REACH), intermediate risk patients only (n=263 in REACH), and seronegative patients only (n=329 in REACH). We also tested

the impact of all four tests when used to replace the RA 2010 criteria for all patients. These diagnostic test scenarios were compared to the RA 2010 criteria in terms of costs, effectiveness and the incremental cost-effectiveness ratio (ICER). The ICER was defined as the additional costs when using the new test in addition to the RA 2010 criteria, divided by the additional quality-adjusted life years (QALYs) gained. Diagnostic effectiveness was evaluated using the change in the TP rate minus the change in the FP rate, divided by the number of patients that underwent the test (unweighted diagnostic Net Benefit, udNB).[34] The headroom per test was calculated as the Net Monetary Benefit (NMB) at a threshold value of €20,000 per QALY minus the difference in costs (excluding the price of the new test) between the two approaches. Headroom is the maximum costs of the test required to ensure a positive NMB. It is important to note that in the add-on strategies people already classified as RA patients based on the RA 2010 criteria were not reclassified, as a result of a negative add-on test.

Sensitivity analyses

In a univariate sensitivity analysis, we investigated the effects of varying the values of utility, sensitivity, specificity, and test costs for the add-on test for intermediate risk patients as shown in Appendix 4.4. As this was an early HTA we tested extreme cases in which no and perfect diagnostic value, and no and 20% utility gain for the different groups were assumed. To take into account second-order parameter uncertainty, 10,000 independent samples were drawn for each of the input parameter distributions (i.e., probabilistic sensitivity analysis), thus dealing with the joint uncertainty in QALYs and costs for each scenario. Cost-effectiveness planes and cost-effectiveness acceptability curves (CEACs) were used to express the uncertainty of the costs and effects.[35]

Model validation

Our model was internally validated by an independent modeler who used a standard protocol to check the model structure, all input parameters with distributions, macros, and the appropriateness of the probabilistic sensitivity analysis.

4.3 Results

The RA 2010 criteria correctly identified 185 RA patients out of 552 early IA patients in the REACH cohort and 195 patients without RA, resulting in 69% correct diagnoses at a cost of €2,966 per patient in the first year after diagnosis. For each test scenario, Table 4.3 shows the short term effects and costs of the five diagnostic pathways. It is important to note that the sensitivity of the combined tests is higher than those of the single tests, e.g., the

combined sensitivity is 0.85 when an add-on with Se 0.60 in all patients is used in addition to the RA 2010 criteria with Se 0.62. This is due to the *add-on* strategy that would not reclassify patients who fulfill the RA 2010 criteria (i.e., if any of the tests was positive the patient was classified as having RA).

All patients

Implementing one of the new diagnostic pathways for all patients (n=552) as an *add-on* test would result in an udNB between -3.9% (IL6) and 16.4% (B-cell assay), see Table 4.3. The QALY gain would be modest (0.01/year) at higher costs, resulting in ICERs varying from €8,702 for the IL6 serum test to €78,000 for the gene assay.

Intermediate risk patients

Among those patients that had 3-5 points on the RA 2010 criteria in REACH (n=263), 32% were using MTX at 12 months without having another classifiable disease. Implementing one of the new diagnostic tests would result in an udNB between -10.2% for the IL6 serum test and 42.9% for the B-cell assay. Costs and utility gains fell compared to using the tests in all patients, but costs fell more than QALYs, resulting in smaller ICERs.

Seronegative patients

Based on REACH, 44% percent of the seronegative patients (n=329) were using MTX at 12 months follow-up without having another classifiable disease. Similar patterns emerged from modeling the 4 diagnostic pathways in this patient group compared with intermediate risk patients (see Table 4.3).

Replacement among all patients

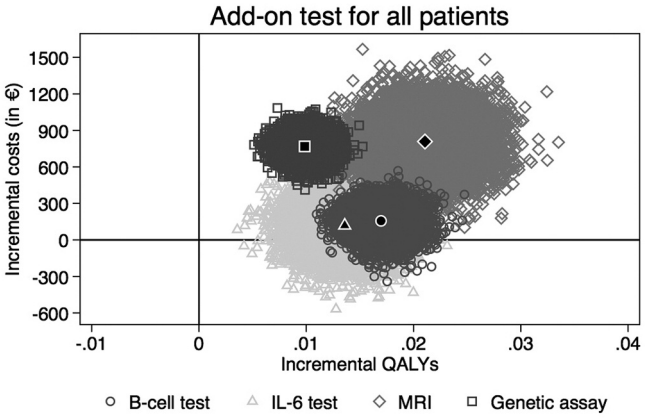
To show what would happen when the RA 2010 criteria were replaced by a single test, we also modeled a *replacement* scenario with each of the 4 diagnostic tests for all patients. Detailed results can be found in Appendix 4.5. More TP patients would be identified with the diagnostic alternatives compared to the RA 2010 criteria at an inflated risk of FPs except for the B-cell assay and gene assay. A limited QALY gain (0.001/year) was noted for the IL6 and B-cell assay, and a modest QALY gain (0.015/year) was observed for MRI and genetic profiling. All tests were cost-effective in the replacement strategy due to the lower costs compared to application of the RA 2010 criteria.

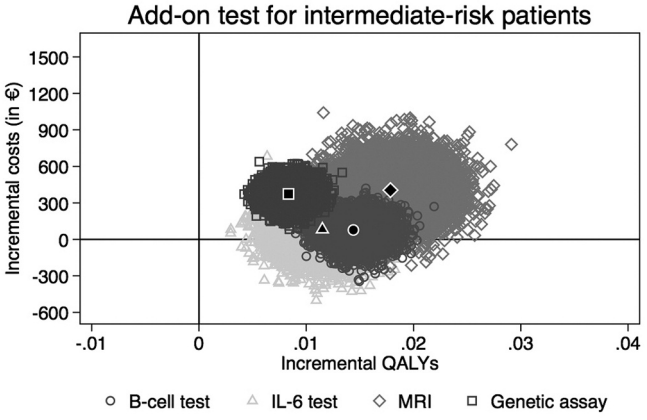
Table 4.3: Short term diagnostic results in REACH representing the 3 groups in which the tests were applied: all patients, intermediate risk patients only and seronegative patients only. Results are presented for the combination of the RA 2010 and the new test

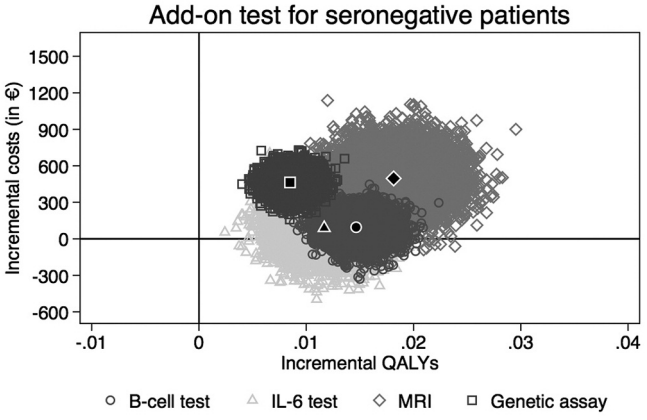
	RA2010	IL6 serum	Gene assay	B-cell assay	MRI
Se	0.62	0.70	0.40	0.60	0.90
Sp	0.77	0.53	0.85	0.90	0.60
Price	-	€50	€750	€150	€756
New tests used in all Patients		**RA2010 + IL6 serum**	**RA2010 + gene assay**	**RA2010 + B-cell assay**	**RA2010 + MRI**
Se		88.6%	77.1%	84.8%	96.2%
Sp		40.9%	65.6%	69.5%	46.3%
ΔTP		26.7%	15.2%	22.9%	34.3%
ΔFP		36.3%	11.6%	7.7%	30.9%
udNB		-3.9%	5.5%	16.4%	8.3%
QALY	0.6864	0.7000	0.6963	0.7034	0.7075
Costs	€2,966	€3,086	€3,736	€3,123	€3,776
ICER		€8,702	€78,000	€9,186	€38,370
Headroom		€204	€179	€334	€369
New tests used in intermediate risk patients		**RA2010 + IL6 serum**	**RA2010 + gene assay**	**RA2010 + B-cell assay**	**RA2010 + MRI**
Se		84.6%	74.9%	81.4%	91.1%
Sp		46.3%	67.3%	70.6%	50.9%
ΔTP		12.3%	7.0%	10.6%	15.8%
ΔFP		14.1%	4.5%	3.0%	12.0%
udNB		-10.2%	14.4%	42.9%	21.8%
QALY		0.6979	0.6948	0.7008	0.7043
Costs		€3,049	€3,340	€3,044	€3,372
ICER		€7,111	€44,658	€5,314	€22,661
Headroom		€172	€151	€283	€313
New tests used in seronegative patients		**RA2010 + IL6 serum**	**RA2010 + gene assay**	**RA2010 + B-cell assay**	**RA2010 + MRI**
Se		85.0%	75.1%	81.7%	91.6%
Sp		45.7%	67.2%	70.5%	50.4%
ΔTP		12.6%	7.2%	10.8%	16.1%
ΔFP		14.4%	4.6%	3.1%	12.2%
udNB		-11.4%	8.4%	28.4%	11.1%
QALY		0.6981	0.6949	0.7011	0.7046
Costs		€3,057	€3,430	€3,062	€3,464
ICER		€7,650	€54,536	€6,458	€27,345
Headroom		€174	€154	€288	€317

RA2010 = RA 2010 criteria; Se = sensitivity; Sp = specificity; TP = true positive; FP = false positive; udNB = unweighted diagnostic Net Benefit; QALY = quality-adjusted life year; ICER = incremental cost-effectiveness ratio.

Figure 4.2: Cost-effectiveness planes showing the scatter plot of 10,000 Monte Carlo trials of the probabilistic model for all patients, for intermediate risk patients and for seronegative patients







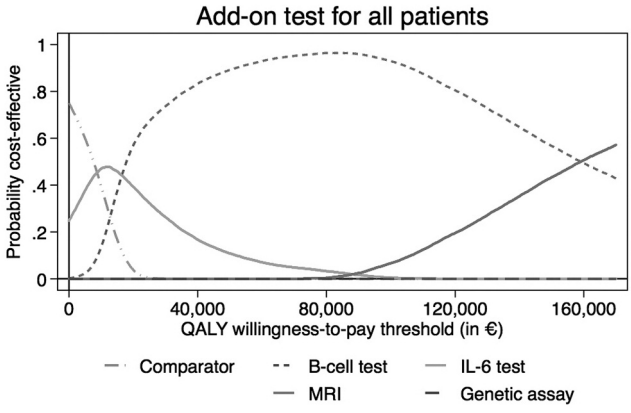
Sensitivity analysis

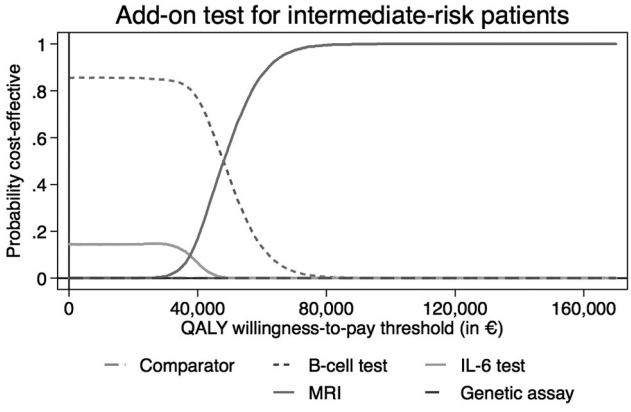
In the univariate sensitivity analysis among intermediate risk patients, we used the B-cell assay as the base case. Because effectiveness was the most important driver for the ICER, we show the impact of the change of test characteristics and utility on the overall effectiveness (see Appendix 4.4). The utility change assigned to TP and TN patients had the largest impact on overall effectiveness. No utility improvement for the TP patients would reduce overall effectiveness from 0.015 to 0.001, while improving the utility to 0.15 would increase overall effectiveness to 0.022. A similar picture was observed when assigning no utility change and 0.15 utility gain for TN patients. Adjusting sensitivity and specificity values had less impact, although reducing the sensitivity to 0.50 lowered the overall utility gain from 0.015 to 0.005.

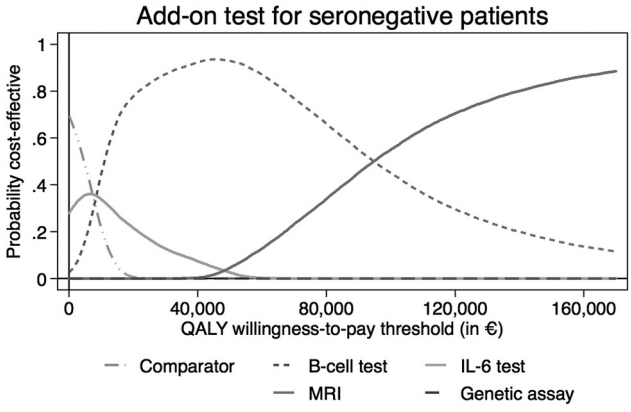
In the probabilistic sensitivity analysis, we evaluated the combined input parameter uncertainty on the differences in costs and effects. Differences in costs and effects varied as shown in the cost- effectiveness plane, with most data points lying in the north east quadrant, suggesting QALY gains at higher costs (see Figure 4.2). Uncertainty varied in a similar fashion for the 4 alternatives with greater cost uncertainty for MRI and the gene assay.

The CEACs in Figure 4.3 show the probability that the cost-effectiveness of a certain test would fall under the threshold value or maximum willingness-to-pay (WTP) for a QALY. If a threshold of €20,000 is used, there is a probability of cost-effectiveness for the add-on B-cell assay of 78% for intermediate risk patients, versus 57% for all patients, and 73% for seronegative patients.

Figure 4.3: Acceptability curves for the short term analysis illustrate the probability that a particular diagnostic strategy is cost-effective at different WTP per QALY threshold levels for the add-on strategies MRI, IL6 serum, B-cell assay and gene assay







4.4 Discussion

In this early HTA study we modeled the short term cost-effectiveness of 4 new diagnostic pathways with test characteristics and costs used from the literature and expert opinion. Each of the pathways was evaluated as an add-on in all patients at risk of having RA, in patients with an intermediate risk, and in patients that were negative for ACPA and rheumatoid factor. The highest udNB was found when using the B-cell assay in intermediate risk patients (42.9%; ICER €5,314), while the IL6 test in seronegative patients resulted in the lowest udNB (-11.4%; ICER €7,650). The negative net diagnostic benefit is a result of a bigger increase of FP compared to TP. If a threshold of €20,000 is used, the B-cell assay would be favored over the other alternatives with a probability of cost-effectiveness of 78% for intermediate risk patients, 57% for all patients and 73% for seronegative patients. For the 4 different tests, the available headroom, given a utility gain of 0.01 in the first year after diagnosis and a WTP threshold of €20,000/QALY, varied between €151 and €369 depending on characteristics and setting. We chose to evaluate values of sensitivity and specificity that seemed realistic in the current field of tests. If we had evaluated a test with 99% sensitivity and 99% specificity at a cost of €150 per test the probability of cost-effectiveness would have been 97%-99% in the 3 patient populations.

Ideally, the new diagnostic test as an add-on would have high sensitivity and high specificity. However, if test developers need to balance sensitivity against specificity, the recommendation would be to choose high specificity rather than high sensitivity. The RA 2010 criteria already perform well in ruling out RA, which would disappear if a subsequent test with lower specificity were to be added. This is illustrated in our results by the IL6 serum test and MRI, since both have a low specificity. For both tests the increase in TP cases was neutralized by the increase in FP cases. This was more distinct in the IL6 test, resulting in a negative diagnostic net benefit. The cost-effectiveness of the tests also depended on the patient sample evaluated. This is simply the result of the number of tests needed to achieve the maximum diagnostic information. In an add-on strategy among all patients, a new diagnostic test is redundant for those people who have already been classified as RA patients based on the RA 2010 criteria. These RA patients would not be reclassified and would thus only incur higher test costs. It would therefore be beneficial to consider the prior probability of reclassification when deciding the place of the new test in the diagnostic work-up of IA patients. Test developers may also consider replacing the RA 2010 criteria by their new test. In our example with respect to all patients, this would only be feasible if the test characteristics should equal those of the RA 2010 criteria at a lower price, which was not the case for any of the tested diagnostic strategies.

With new diagnostic biomarkers in the pipeline, our decision model is able to determine the maximum increase in diagnostic cost for which reclassification of patients is still likely to be cost-effective given a threshold value of €20,000 per QALY. A limitation of performing early HTA is that there is only a limited amount of clinical data available with which to populate the decision models. In particular the health gain in utility terms of adding a new diagnostic test is difficult to estimate without observed data. We chose to assign equal improvement for the TP and TN patients, while the FP patients had less gain and the FN patients' health deteriorated. Whether this reflects clinical practice needs to be proven. Adjustments in utilities had the largest impact on the ICERs in the univariate sensitivity analysis. This relates to the assumption we made about patients that were FNs on the RA 2010 criteria. We assumed that none of them start treatment at baseline, but in real-life this might not be the case, resulting in smaller utility gains from the new strategy.

4.5 Conclusion

Applying a new diagnostic test *as an add-on* in the work-up of patients at risk of having rheumatoid arthritis is probably cost-effective, with the largest diagnostic benefit arising in intermediate risk patients. Test developers should balance sensitivity and specificity in favor of specificity and take into account that under the current utility assumptions the headroom is between €151 and €369 per test.

4.6 Acknowledgments

This research was performed within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), and the Dutch Arthritis Foundation, project TRACER (grant O4I-202). We are grateful to Nasuh Büyükkaramikli for performing an internal validity assessment of the model.

4.7 References

1 Carbonell J, Cobo T, Balsa A, et al. The incidence of rheumatoid arthritis in Spain: results from a nationwide primary care registry. *Rheumatology (Oxford)* 2008;47(7):1088–92.

2 Hazes JM, Luime JJ. The epidemiology of early inflammatory arthritis. *Nat Rev Rheumatol* 2011;7(7):381–90.

3 Soderlin MK, Borjesson O, Kautiainen H, et al. Annual incidence of inflammatory joint diseases in a population based study in southern Sweden. *Ann Rheum Dis* 2002;61(10):911–5.

4 Savolainen E, Kaipiainen-Seppanen O, Kroger L, et al. Total incidence and distribution of inflammatory joint diseases in a defined population: results from the Kuopio 2000 arthritis survey. *J Rheumatol* 2003;30(11):2460–8.

5 Finckh A, Liang MH, Van Herckenrode CM, et al. Long-term impact of early treatment on radiographic progression in rheumatoid arthritis: A meta-analysis. *Arthritis Rheum* 2006;55(6):864–72.

6 Rezaei H, Saevarsdottir S, Forslind K, et al. In early rheumatoid arthritis, patients with a good initial response to methotrexate have excellent 2-year clinical outcomes, but radiological progression is not fully prevented: data from the methotrexate responders population in the SWEFOT trial. *Ann Rheum Dis* 2012;71(2):186–91.

7 Ter Wee MM, Lems WF, Usan H, et al. The effect of biological agents on work participation in rheumatoid arthritis patients: a systematic review. *Ann Rheum Dis* 2012;71(2):161–71.

8 Villeneuve E, Nam JL, Bell MJ, et al. A systematic literature review of strategies promoting early referral and reducing delays in the diagnosis and management of inflammatory arthritis. *Ann Rheum Dis* 2013;72(1):13–22.

9 Ten Cate DF, Luime JJ, Swen N, et al. Role of ultrasonography in diagnosing early rheumatoid arthritis and remission of rheumatoid arthritis - a systematic review of the literature. *Arthritis Res Ther* 2013;15(1):R4.

10 Peterfy C, Ostergaard M, Conaghan PG. MRI comes of age in RA clinical trials. *Ann Rheum Dis* 2013;72(6):794–6.

11 Aletaha D, Neogi T, Silman AJ, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* 2010;62(9):2569–81.

12 Alves C, Luime JJ, Van Zeben D, et al. Diagnostic performance of the ACR/EULAR 2010 criteria for rheumatoid arthritis and two diagnostic algorithms in an early arthritis clinic (REACH). *Ann Rheum Dis* 2011;70(9):1645–7.

13 Gottenberg JE, Dayer JM, Lukas C, et al. Serum IL-6 and IL-21 are associated with markers of B cell activation and structural progression in early rheumatoid arthritis: results from the ESPOIR cohort. *Ann Rheum Dis* 2012;71(7):1243–8.

14 Van Baarsen LG, Bos WH, Rustenburg F, et al. Gene expression profiling in autoantibody-positive patients with arthralgia predicts development of arthritis. *Arthritis Rheum* 2010;62(3):694–704.

15 Yarwood A, Huizinga TW, Worthington J. The genetics of rheumatoid arthritis: risk and protection in different stages of the evolution of RA. *Rheumatology (Oxford)* 2016;55(2):199–209.
16 Dutch Healthcare Authority (in Dutch: Nederlandse Zorgauthoriteit, NZa). NZa Zorgproducten Tariefapplicatie. Available from: http://dbc-zorgproducten-tarieven.nza.nl/nzaZpTarief/ZoekfunctieDbc.aspx. Accessed May 1, 2013.
17 Colebatch AN, Edwards CJ, Ostergaard M, et al. EULAR recommendations for the use of imaging of the joints in the clinical management of rheumatoid arthritis. *Ann Rheum Dis* 2013;72(6):804–14.
18 Sugimoto H, Takeda A, Hyodoh K. Early-stage rheumatoid arthritis: prospective study of the effectiveness of MR imaging for diagnosis. *Radiology* 2000;216(2):569–75.
19 Machado PM, Koevoets R, Bombardier C, et al. The value of magnetic resonance imaging and ultrasound in undifferentiated arthritis: a systematic review. *J Rheumatol Suppl* 2011;87:31–7.
20 Sugimoto H, Takeda A, Masuyama J, et al. Early-stage rheumatoid arthritis: diagnostic accuracy of MR imaging. *Radiology* 1996;198(1):185–92.
21 Duer-Jensen A, Horslev-Petersen K, Hetland ML, et al. Bone edema on magnetic resonance imaging is an independent predictor of rheumatoid arthritis development in patients with early undifferentiated arthritis. *Arthritis Rheum* 2011;63(8):2192–202.
22 Bird P, Conaghan P, Ejbjerg B, et al. The development of the EULAR-OMERACT rheumatoid arthritis MRI reference image atlas. *Ann Rheum Dis* 2005;64 Suppl 1:i8–10.
23 Nieuwenhuis WP, Krabben A, Stomp W, et al. Evaluation of magnetic resonance imaging-detected tenosynovitis in the hand and wrist in early arthritis. *Arthritis Rheum* 2015;67(4):869–76.
24 Smolen JS, Landewe R, Breedveld FC, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;73(3):492–509
25 Visser K, Katchamart W, Loza E, et al. Multinational evidence-based recommendations for the use of methotrexate in rheumatic disorders with a focus on rheumatoid arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3E Initiative. *Ann Rheum Dis* 2009;68(7):1086–93.
26 Lamers LM, McDonnell J, Stalmeier PF, et al. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. *Health Econ* 2006;15(10):1121–32.
27 Buitinga L, Braakman-Jansen LM, Taal E, et al. Comparative responsiveness of the EuroQol-5D and Short Form 6D to improvement in patients with rheumatoid arthritis treated with tumor necrosis factor blockers: results of the Dutch Rheumatoid Arthritis Monitoring registry. *Arthritis Care Res (Hoboken)* 2012;64(6):826–32.

28 Lillegraven S, Kristiansen IS, Kvien TK. Comparison of utility measures and their relationship with other health status measures in 1041 patients with rheumatoid arthritis. *Ann Rheum Dis* 2010;69(10):1762–7.
29 Linde L, Sorensen J, Ostergaard M, et al. Gain in quality-adjusted life-years in patients with rheumatoid arthritis during 1 year of biological therapy: a prospective study in clinical practice. *J Rheumatol* 2013;40(9):1479–86.
30 Sorensen J, Linde L, Ostergaard M, et al. Quality-adjusted life expectancies in patients with rheumatoid arthritis—comparison of index scores from EQ-5D, 15D, and SF-6D. *Value Health* 2012;15(2):334–9.
31 Verstappen SM, McCoy MJ, Roberts C, et al. Beneficial effects of a 3-week course of intramuscular glucocorticoid injections in patients with very early inflammatory polyarthritis: results of the STIVEA trial. *Ann Rheum Dis* 2010;69(3):503–9.
32 Benner BJM, Buisman LR, Vis M, et al. Costs of diagnostic tests in the workup of early arthritis patients. (Submitted for publication).
33 Brouwer WB, Koopmanschap MA. The friction-cost method : replacement for nothing and leisure for free? *Pharmacoeconomics* 2005;23(2):105–11.
34 Pepe MS, Kerr KF, Longton G, et al. Testing for improvement in prediction model performance. *Stat Med* 2013;32(9):1467–82.
35 Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves—facts, fallacies and frequently asked questions. *Health Econ* 2004;13(5):405–15.

4.8 Appendices

Appendix 4.1: Characteristics of the unclassified inflammatory arthritis patients in REACH

ALL PATIENTS (n=552)		
Age (in years; SD)		52.8 (15.0)
Sex (% female)		67.7%
DOMAINS OF ACR/EULAR 2010 RA CLASSIFICATION		
Swollen Joints	> 10	19.0%
	4-10 small	34.0%
	1-3 small	32.7%
	2-10 large	3.4%
	1 large	6.0%
Serology	high	34.1%
	low	9.0%
Disease duration: >6 weeks		85.0%
Elevated acute phase reactants		53.3%
SCORE ON ACR/EULAR 2010 RA CLASSIFICATION		
6 or more points		41.9%
3-5 points		47.5%
0-2 points		10.6%

SD = standard deviation.

Design of the Rotterdam Early Arthritis CoHort

The Rotterdam Early Arthritis CoHort (REACH) is an inception cohort study with 2 years of follow up. Assessments took place at baseline, 6 months, 12 months and 24 months. REACH aims to study the etiopathogenesis, diagnostic strategies, and outcome of patients with inflammatory joint conditions for <12 months. Both general practitioners and rheumatologists invited patients to participate in REACH from July 2004 to April 2011. For general practitioners, short educational courses on the importance of early treatment of RA and early referral were organized. Physicians that agreed to participate in REACH received written information and verbal instructions on the general aims of the study and on how to send patients for inclusion in the study. Data collection includes a large array of detailed medical examinations and questionnaires.

General practitioners selected patients with arthritis in ≥1 joint or patients experiencing conditions in ≥2 joints without synovitis. The general practitioners determined that conditions existed for <12 months and were

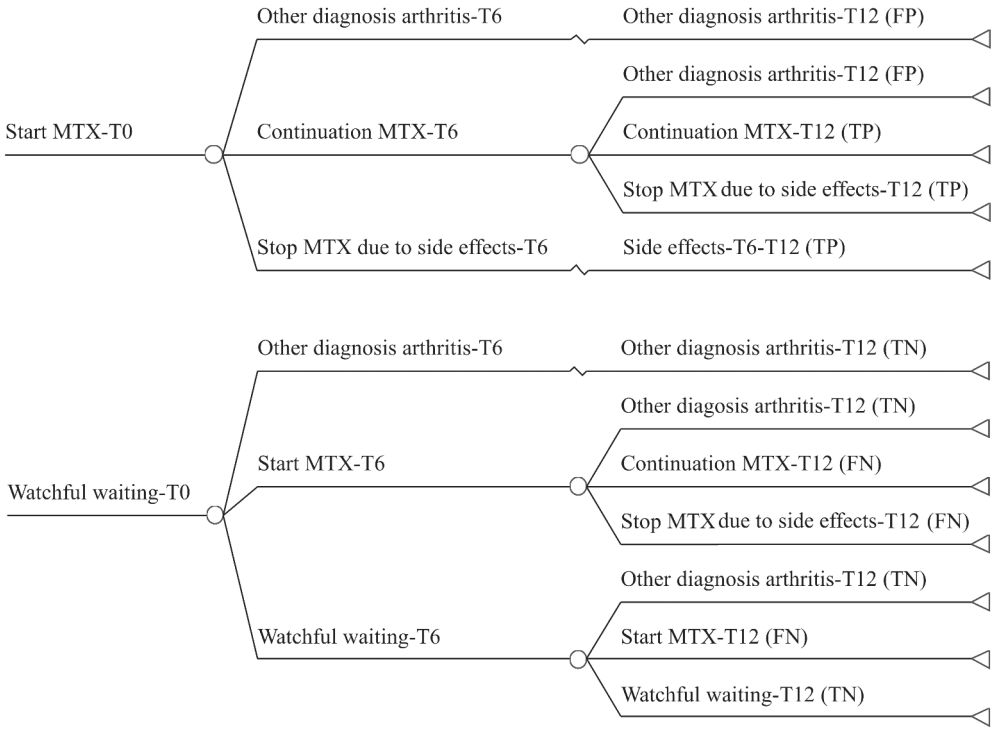
not due to trauma/mechanical problems. In addition, patients had to be age >16 years. During an interview by telephone and a subsequent medical examination by a rheumatologist, the inclusion criteria were verified. Patients were included if 1) joint conditions existed for <12 months with no requirement of a minimum duration; 2) they had arthritis in ≥1 joint or conditions in ≥2 joints in combination with at least 2 of the following criteria ascertained during medical examination by a rheumatologist: morning stiffness for >1 hour, bilateral compression pain in the metacarpophalangeal or metatarsophalangeal joints, symmetric presentation, positive family history, non-fitting shoes, non-fitting rings, pins and needles in fingers, or unexplained fatigue for <1 year; and 3) conditions were predominantly present in the morning and at night, and improved with movement. Patients were excluded if 1) conditions were due to trauma/mechanical problems, 2) they were age <16 years, 3) no written communication was possible in Dutch, or 4) a prior diagnosis of RA, ankylosing spondylitis, Sjögren's syndrome, systemic lupus erythematosus, or juvenile arthritis had been made by a rheumatologist before inclusion in this study.

For patients directly visiting rheumatologists, a similar verification procedure was applied. The final diagnosis was made by a rheumatologist in one of the 5 participating out-patient rheumatology clinics.

The study was approved by the Erasmus MC Medical Ethical Committee. All patients gave written informed consent.

In the current study we use 552 consecutive patients with unclassified inflammatory arthritis.

Appendix 4.2: Follow-up branches of decision tree model at 6 months and 12 months re-evaluation of the diagnosis



Appendix 4.3: Distribution values of the parameter inputs

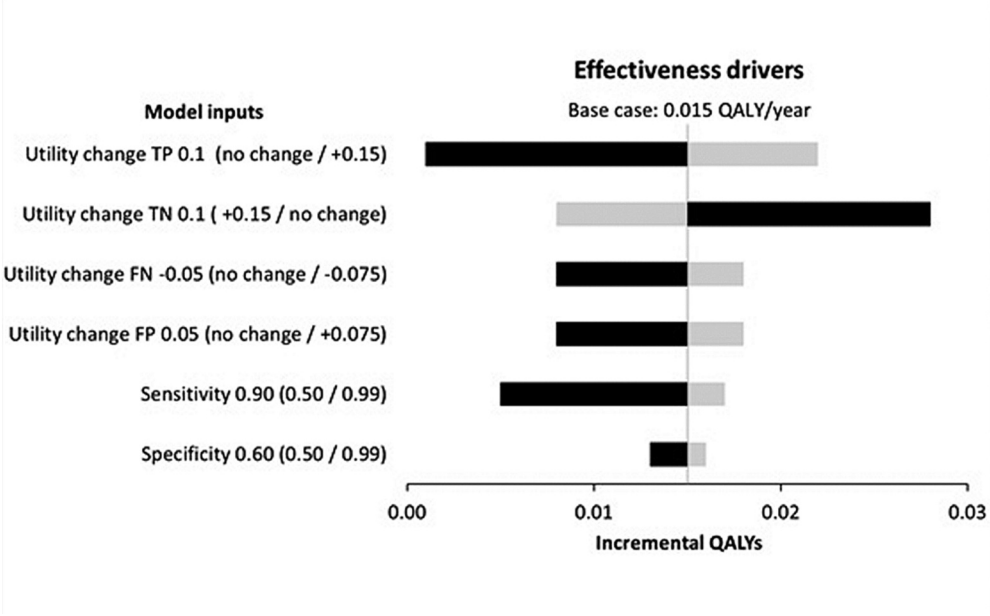
Parameter	Alpha	Beta	Distribution	Source
PREVALENCE OF RA BASED ON ACR/EULAR 2010 RA CLASSIFICATION CRITERIA				
All patients	300	252	Beta	REACH
Intermediate risk patients	97	166	Beta	REACH
Seronegative patients	146	183	Beta	REACH
METHOTREXATE START AND CONTINUATION				
In patients >=6 points at baseline				
Start at baseline				Guideline[11,19]
Continuation at 6 months	1,282	0.0007	Dirichlet	REACH
Continuation at 12 months	1,125	0.0008	Dirichlet	REACH
In patients < 6 points at baseline				
Start at baseline				Guideline[11,19]
Start at 6 months	167	0.0021	Dirichlet	REACH
Continuation at 12 months	592	0.0014	Dirichlet	REACH
Start at 12 months	7	0.0199	Dirichlet	REACH
OTHER DIAGNOSIS				
>=6 points at baseline				
6 months	35	0.0038	Dirichlet	REACH
12 months	201	0.0045	Dirichlet	REACH
< 6 points at baseline				
6 months	322	0.0016	Dirichlet	REACH
12 months	-	-	Dirichlet	REACH
COSTS (€)				
Costs ACR/EULAR 2010 RA classification criteria strategy	€44	€67	Gamma	Estimate*
Baseline costs	€44	€21	Gamma	Estimate*
Follow up during first year	€44	€15	Gamma	Estimate*
MTX costs including monitoring of side effects (1 year)	€44	€2	Gamma	Estimate*
Other DMARD costs (1 year)	€152	€1	Gamma	REACH
Productivity costs (1 year)	€12	€118	Gamma	REACH

Appendix 4.3: Distribution values of the parameter inputs (continued)

Parameter		Alpha	Beta	Distribution	Source
UTILITIES					
True positive patients	– baseline	626	417	Beta	REACH
	12 months	2,354	1,569		Estimate[#]
False positive patients	– baseline	548	295	Beta	REACH
	12 months	1,471	490		Estimate[#]
True negative patients	– baseline	548	295	Beta	REACH
	12 months	882	156		Estimate[#]
False negative patients	– baseline	626	417	Beta	REACH
	12 months	2,942	2,942		Estimate[#]

[+] REACH data on file, details about data collection can be found in Alves et al.[12]; [#] Estimate based on expert opinion and REACH data, for the se of the utilities data from REACH was used.

Appendix 4.4: Tornado diagram of the univariate sensitivity analysis on the effectiveness (QALY) in intermediate risk patients using the B-cell assay as base case (Se: 0.60; Sp: 0.90; €150)



Utility change is change in utility during the first year; Se = sensitivity; Sp = specificity; TP = true positive; TN = true negative; FN = false negative; FP = false positive.

Appendix 4.5: Results of replacing the ACR/EULAR 2010 RA classification criteria by one of the 4 hypothetical tests

When new test is applied in all patients	ACR/EULAR 2010 RA classification criteria	B-cell assay	Gene assay	IL6	MRI
Price		€150	€750	€50	€756
Se		0.60	40.0%	70.0%	90.0%
Sp		0.90	85.0%	53.0%	60.0%
TP		-1.9%	-21.9%	8.1%	28.1%
FP		-12.8%	-7.8%	24.2%	17.2%
udNB		8.9%	-15.3%	-12.3%	13.6%
QALY	0.6866	0.6880	0.6706	0.6877	0.7056
Costs	€2,967	€2,460	€3,028	€2,576	€3,304
ICER		-€319,608	-€3,832	-€304,241	€17,555
Headroom		€685	€368	€461	€797

Se = sensitivity; Sp = specificity; TP = true positive; FP = false positive; udNB = unweighted diagnostic Net Benefit; QALY = quality-adjusted life year; ICER = incremental cost-effectiveness ratio (a negative ICER for the B-cell assay was the result of a lower costs at higher levels of effectiveness (savings), while the negative ICER for the Gene assay is the result of higher costs and lower levels of effectiveness).

CHAPTER 5

A FIVE-YEAR MODEL TO ASSESS THE EARLY COST-EFFECTIVENESS OF NEW DIAGNOSTIC TESTS IN THE EARLY DIAGNOSIS OF RHEUMATOID ARTHRITIS

Leander R. Buisman[1,2], Jolanda J. Luime[3], Mark Oppe,[1,2,4], Johanna M.W. Hazes[3], Maureen P.M.H. Rutten-van Mölken[1,2]

[1] Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands
[2] Institute for Medical Technology Assessment, Erasmus University Rotterdam, Rotterdam, The Netherlands
[3] Department of Rheumatology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
[4] EuroQol Research Foundation, Rotterdam, The Netherlands

Forthcoming in *Arthritis Research & Therapy* 2016. doi:10.1186/s13075-016-1020-3.

ABSTRACT

Introduction: There is a lack of information what sensitivity, specificity and costs new diagnostic tests should have to improve early diagnosis of rheumatoid arthritis (RA). Our objective was to explore the early cost-effectiveness of various new diagnostic test strategies in the workup of patients with inflammatory arthritis (IA) at risk of having RA.

Methods: A decision tree followed by a patient-level state transition model, using data from published literature, cohorts and trials, was used to evaluate diagnostic test strategies. Alternative tests were assessed as add-on to the ACR/EULAR 2010 RA classification criteria for all patients and for intermediate-risk patients, and as replacement. Tests included B-cell gene expression (sensitivity: 0.60, specificity: 0.90, costs: €150), MRI (sensitivity: 0.90, specificity: 0.60, costs: €756), IL-6 serum level (sensitivity: 0.70, specificity: 0.53, costs: €50), and genetic assay (sensitivity: 0.40, specificity: 0.85, costs: €750). Patients with IA at risk of having RA were followed for 5 years using a societal perspective. Guideline treatment was assumed using tight controlled treatment based on DAS28; if patients had a DAS>3.2 at 12 months or later patients could be eligible for biological start. The outcome was expressed in incremental cost-effectiveness ratios (2014 Euros per quality-adjusted life year [QALY] gained) and headroom.

Results: The B-cell test was the least expensive strategy when used as an add-on for intermediate-risk patients and as replacement, making it the dominant strategy since it has better health outcomes and lower costs. As add-on for all patients, the B-cell test was also the most cost-effective test strategy. When using a willingness-to-pay threshold of €20,000 per QALY gained, the IL-6 and MRI strategies were not cost-effective, except as replacement. A genetic assay was not cost-effective in any strategy. Probabilistic sensitivity analysis revealed that the B-cell test was consistently superior in all strategies. Univariate sensitivity analysis for intermediate-risk patients showed the largest impact of specificity and DAS28 in the B-cell add-on strategy and DAS28 and sensitivity in the MRI add-on strategy.

Conclusions: This early-CEA indicated that new tests to diagnose RA are most likely to be cost-effective when the tests are used as an add-on for intermediate-risk patients, have a high specificity, and the test costs should not be higher than €200-€300.

Keywords: rheumatoid arthritis, diagnosis, treatment, tests, early cost-effectiveness analysis.

5.1 Introduction

Rheumatoid arthritis (RA) is a chronic inflammatory joint disease characterized by structural irreversible joint damage, leading to severe disability, serious loss of quality of life and premature death if left untreated.[1-6] Disease progression can be slowed down by synthetic and biologic disease modifying antirheumatic drugs (DMARDs), especially if started early in the disease.[7-11] This requires early detection of RA. However, early diagnosis is complex, since RA-related symptoms early in the disease course resemble those of other musculoskeletal disorders.[12] Early detection would result in early start of DMARDs and improved prognosis. Therefore, several diagnostic tests are currently being developed to improve the early diagnosis of RA in inflammatory arthritis (IA) patients, e.g., B-cell related gene expression, IL-6 serum level test, magnetic resonance imaging (MRI) of hands and feet, and genetic assay with susceptibility single nucleotide polymorphisms (SNPs) for RA. However, little is known about their potential cost-effectiveness.

To guide implementation of new diagnostic tests in the workup of patients at risk of having RA, a decision model could be used to evaluate the test performance (i.e., sensitivity and specificity), test costs, and positioning of the test on the clinical outcomes and societal costs. Evaluating the cost-effectiveness of new drugs before entering the market is well implemented since reimbursement authorities request this. However, it is less common to evaluate the cost-effectiveness of new medical tests that enter the market while this also affects healthcare spending directly. Given the constraints on healthcare budgets, it is likely that clinicians evaluate the impact of new tests on their departmental budget and consider diagnostic uncertainty. To inform this clinical decision problem, we conducted an early-cost-effectiveness analysis (early-CEA). In this analysis, the incremental costs of the new potential and current diagnostic test strategies are weighed against the gain in quality-adjusted life years (QALYs) and potential improved labor force participation.[13] The main assumption is that early diagnosis results in timely start of effective treatment that reduces disease activity and consequently postpones or prevents treatment with a biologic DMARD.

The aim of this study was twofold. The first objective was to develop an early-CEA model to evaluate the costs and health effects (in terms of QALYs) of new and current diagnostic test strategies for IA patients who were suspected of having RA, from a societal perspective. The second objective was to analyze the costs and health effects of new test strategies compared to the ACR/EULAR 2010 RA classification criteria (referred to as RA-2010 criteria).

5.2 Methods

We applied the framework with general steps of early-CEAs of medical tests as developed by Buisman et al.[14] This framework is a useful guidance for researchers performing early-CEAs of medical tests.

Current diagnostic test strategy

IA patients at risk of having RA undergo a diagnostic workup to establish the presence of RA. This entails joint counts, blood testing, and radiographs ordered or established by the rheumatologist. If no other explanation for the symptoms is found (e.g., SLE, gout, or psoriatic arthritis) the patient can classify as RA if at least 6 out of 10 points on the RA-2010 criteria are scored. Patients that score less than 6 but more than 2 points are regarded intermediate-risk patients that do not fulfil the RA-2010 criteria. The RA-2010 criteria were the comparator in our early-CEA.[6]

New diagnostic test strategies

The cost-effectiveness was assessed of four diagnostic tests that are currently being developed as part of the TRACER project:[15] B-cell related gene expression,[16] IL-6 serum level test,[17] MRI of hands and feet,[18-23] and genetic assay with susceptibility SNPs for RA.[24] For each of the tests (described below), three different test strategies were modelled: add-on to the RA-2010 criteria for all IA patients, add-on for intermediate-risk patients only, and replacement of all bloodwork and radiographs used to classify patients according to the RA-2010 criteria. For the add-on test strategies, the performance of the new test strategy was estimated by combining the sensitivity and specificity of the RA-2010 criteria and the new tests.

B-cell related gene expression

During the development of arthritis, B-cell RNA expression decreases over time.[15] Although, the exact mechanism is poorly understood, the marker is useful to predict early arthritis in patients with seropositive arthralgia.[16] Currently, no data is available for IA patients. Therefore, we used the data from the seropositive arthralgia cohort from Baarsen et al.[16] After discussion with the developers of this test, a sensitivity of 0.60 and specificity of 0.90 was used. The cost of the test was set at €150.

IL-6 serum level test

IL-6 is a cytokine that is present in inflammation. A recent evaluation of IL-6 serum level test performance in detecting RA in IA patients showed a sensitivity of 0.70 and specificity of 0.53.[17] We used these test sensitivity and specificity and assumed a cost of €50 per test.

MRI of hands and feet

MRI may reclassify patients to different joint domains of the RA-2010 criteria if more joints show swelling compared to physical examination. MRI also provides additional information on bone marrow edema[18] and tenosynovitis.[19,20] Based on literature[21-23] and discussions with test developers, we set MRI to be 0.90 sensitive and 0.60 specific. The costs of MRI were assumed to be equal to the unit costs currently used by the Dutch Healthcare Authority of €189 per MRI (€756 for 4 MRIs: 2 hands and 2 feet).[25]

Genetic assay with susceptibility SNPs for RA

RA is a complex disease in which several genes are involved. Heritability for RA is estimated to be around 50%-60%.[24] Literature discussed by experts suggests that using genetic risk factors with the current knowledge would result in a sensitivity of 0.40 with a specificity between 0.80 and 0.90 to identify patients with RA.[24] We used these estimations and applied a sensitivity of 0.40 and specificity of 0.85 with an estimated cost of €750 per test based on expert opinion from test developers.

Treatment

In the current and new diagnostic test strategies, test positive patients received MTX (25 mg/week orally).[26] Due to side effects of MTX, patients could switch to other synthetic DMARDs (e.g., Sulfasalazine, Leflunomide). After failure of two synthetic DMARDs, patients could switch to biologic DMARDs (i.e., TNF-inhibitors, IL6-inhibitors, B-cell depletion, or T-cell inhibition).[27]

RA patients that were additionally detected by the new diagnostic test strategies as compared to the RA-2010 criteria were assumed to get early treatment. As a result, we modelled that they had an improvement of 0.2 in DAS28 at 12 months as compared to patients in the current test strategy. The DAS28 improvement of 0.2 was based on a sensitivity analysis in which we evaluated the effect of changing this value on the model results (see univariate sensitivity analysis below).

Model structure

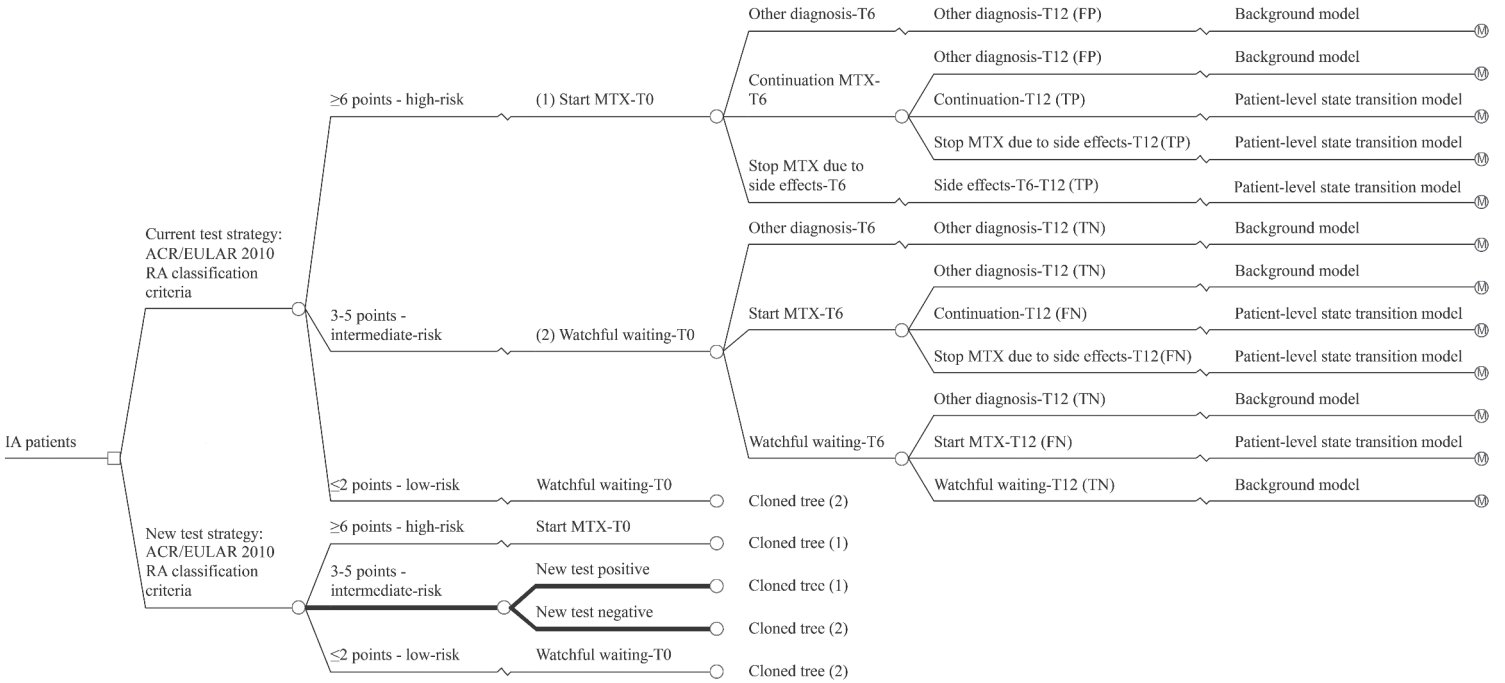
In RA, the diagnosis and subsequent prognosis are complex processes in which various tests and measures of disease activity influence treatment decisions and subsequently outcomes in terms of both costs and effects. As the diagnosis is often reconsidered in the first year of disease, especially in those initially not classified as RA, we decided to model the first year as a decision tree with chance nodes at 6 and 12 months to classify patients as true positive (TP), false positive (FP), true negative (TN), and false negative (FN) during the first year. Patients were classified as TP if they had a positive test result (≥6 points on the RA-2010 criteria or positive on the new test) at

baseline and at 12 months, used methotrexate (MTX) or stopped MTX due to side effects. Moreover, the symptoms should not be explained by another classified diagnosis.[6] Patients were considered as TN if they had a negative test result (<6 points on the RA-2010 criteria or negative on the new test) at baseline, did not use MTX at 12 months, and had symptoms explained by another classified diagnosis. FPs scored ≥6 points on the RA-2010 criteria or positive on the new test at baseline but did not use MTX at 12 months and had symptoms explained by another classified diagnosis. Patients were classified as FN if they scored <6 points on the RA-2010 criteria or negative on the new test at baseline but used MTX or stopped MTX due to side effects at 12 months, and had no symptoms explained by another classified diagnosis. Using a combination of initial RA-2010 criteria scores and the use of DMARD not explained by any other disease is a common way of dealing with a disease for which no hall-mark sign is available.[6,28]

The first year is followed by a four-year individual-level Markov model (i.e., patient-level state transition model) that simulates the change in disease activity (DAS28) over time in 3-month cycles. The cycle time is 3 months because patients are commonly seen every 3 months by the rheumatologists. This five-year model was used to simulate what would happen if a proportion of FNs in the current strategy were earlier diagnosed at lower levels of disease activity in the new test strategy. The time horizon was five years because the long term effects of biological use are unknown. Patients were categorised in three disease states: remission: DAS28≤2.6; low disease activity: DAS28>2.6–≤3.2; and moderate and severe disease activity: DAS28>3.2*. This categorization in DAS28 states is common in the field of RA.[29-31] Resource use, costs and utilities were linked to these three states. The patients who were classified as TP or FN at 12 months entered the patient-level state transition model. A proportion of patients with DAS28>3.2 was modelled to start a biologic DMARD in addition to MTX. They were assumed to stay on a biologic DMARD and could switch to another biologic DMARD for the remainder of the four years. The patients who were classified as TN or FP at 12 months entered a background model in which they stayed for the remaining four years, assuming no change in utilities, biologic DMARD costs for 10% of FPs in the first year after diagnosis, and otherwise no RA-related costs. Figure 5.1a and 5.1b show the decision model comparing the current diagnostic test strategy with the new add-on diagnostic test strategy for intermediate-risk patients (described below). This five-year cost-effectiveness model is an extension of our one-year model published elsewhere.[32]

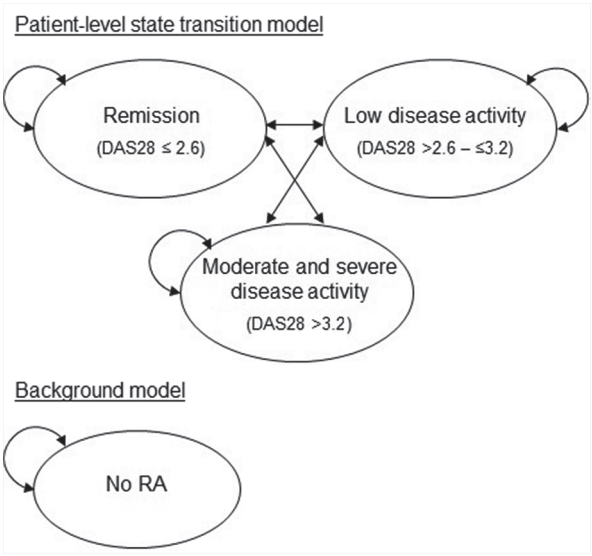
* Moderate and severe DAS28 were categorised in one disease state. Since treatment choice is based on DAS>3.2, our categorisation does not affect the model results.

Figure 5.1a: Decision tree of first year comparing the current diagnostic test strategy with the new diagnostic add-on test strategy for intermediate-risk patients



Current test strategy = ACR/EULAR 2010 RA classification criteria; New test strategy=add-on test for intermediate-risk patients (3-5 points) according to the ACR/EULAR 2010 RA classification criteria; IA = inflammatory arthritis; MTX = methotrexate; TP = true positive; FP = false positive; TN = true negative; FN = false negative.

Figure 5.1b: Patient-level transition state model and background model of second to fifth year for all test strategies



Data sources

To populate the model, we mainly used data from three different sources. Appendix 5.1 shows the characteristics of the three sources. First, data from the REACH (usual care) cohort was used to populate the one-year decision tree with 552 IA patients who were suspected of having RA (details about the cohort can be found in Appendix 5.2). Patients had to have at least one clinical synovitis that could not be classified as another inflammatory joint disease. The prevalence of RA was 54% at 12 months based on the RA-2010 criteria and MTX use.

Second, data from the tREACH trial was used in our model for RA patients after one-year follow-up.[33] The tREACH trial includes patients aged 18 year-old or older with arthritis in at least one joint, and symptom duration less than one year. Patients were randomized into three initial treatment strategies: triple DMARD therapy (MTX, sulfasalazine and hydroxychloroquine) with glucocorticoids intramuscular; triple DMARD therapy with an oral glucocorticoids tapering scheme; and MTX monotherapy with an oral glucocorticoids tapering scheme. For a detailed description of the tREACH trial, see Claessen et al.[34]

Third, summary data from the DREAM registry as published in Vermeer et al.[35] was used to inform biological start. This publication describes data

of two cohorts: a treat-to-target and usual care cohort of clinical RA. The treat-to-target strategy used a standardized treatment step-up protocol.[35] In contrast, the treatment switches were not protocolized in the usual care cohort.

Model inputs

Appendix 5.3 gives an overview of all model input parameters, their estimates and distributions for probabilistic sensitivity analysis, and data sources.

Estimation of transition probabilities

During the first 12 months of our model, the probabilities to be TP, FN, TN and FP were elicited from the REACH cohort in which patients were classified according to the RA-2010 criteria, use of MTX and use of other synthetic DMARDs at baseline, 6 and 12 months.

At start of the patient-level state transition model (i.e., at 12 months), the DAS28 of TPs and FNs at 12 months in REACH resulted in entering one of the three disease states. Patients that entered DAS>3.2 at start or later on in time could be eligible for biological start. To model this, summary data on biological start from DREAM was used, in which the observed use of biologic DMARDs in clinical practice was 15% at 24 months. We transformed this 15% rate into a 3-month transition probability of 2% to start biologic DMARDs in those patients with a DAS>3.2. This 2% was distributed over the three disease states in a 1:3:6 distribution (state1:state2:state3) based on flare rates (DAS28>3.2) in tREACH.

Estimates of resource use and costs

We distinguished two cost categories: direct medical and productivity costs. Direct medical costs include costs of visits to rheumatologists and other health professionals (e.g., physical therapist), laboratory tests including diagnostic tests and those to monitor side effects, and medication use. Productivity costs represent the number of days that a patient with a paid job was absent from work in the past 3 months. Resource use and productivity losses per disease state were obtained from REACH in the first year[28] and tREACH[33,34] in the second and third year. The latter was extrapolated to 5 years. In the background model, TNs were assumed to incur no RA-related costs, while 10% of FPs incurred biologic DMARD costs in the first year after diagnosis due to misdiagnosis for which the frequency was based on REACH.

The unit costs of visits and productivity losses were based on reference prices published in the Dutch Manual of Costing in economic evaluations.[36] Diagnostic test costs were based on tariffs published by the Dutch Healthcare Authority[25] and medication costs were obtained from the National

Health Care Institute.[37] All costs were adjusted to 2014 Euros using the general price index from the Dutch Central Bureau of Statistics.[38] All cost parameters can be found in Appendix 5.3.

Estimation of QALYs

When assessing the impact of a new test or treatment on quality of life over time, the health outcomes are usually measured in terms of quality-adjusted life years (QALYs). The QALY combines the number of life years with the level of health-related quality of life (i.e., utilities) in those years.[39] The EuroQol five dimension 3-level questionnaire (EQ-5D-3L) was used to estimate utilities. The baseline utilities of TP, FP, TN, and FN were obtained from REACH and were 0.60, 0.65, 0.65, and 0.60, respectively. Based on literature, we assigned an improvement of 0.10 over the first year to the TP.[40-44] Based on REACH, we assigned an improvement of 0.05 and 0.10 over the first year to the FP and TN, respectively. Based on the placebo group in the STIVEA trial, we assigned a 0.05 reduction over the first year for FN assuming that FN patients would receive little therapy.[44]

In the patient-level state transition model, patients were assigned EQ-5D values based on their DAS28 every 3 months stratified for biologic DMARD start. As observed in tREACH, the EQ-5D values for patients not using biologic DMARDs were higher. Furthermore, the EQ-5D values were not normally distributed. About 25% of patients in tREACH had at least once in 3 years a decrease in EQ-5D that led to a utility score lower than 0.50. Therefore, different distributions of utility values were estimated. One distribution for patients with at least one EQ-5D decrease below 0.50 over time and another distribution for patients who always had an EQ-5D higher than 0.50 over time. In the background model, patients were assumed to have an EQ-5D value of 0.75, that remained constant over time.

Analyses/Modelling

We performed a base case analysis with four diagnostic tests that were used in three diagnostic test strategies as described above. We calculated the incremental costs per QALY gained in each new test strategy compared with the current test strategy [i.e., incremental cost-effectiveness ratio (ICER)]. Probabilistic sensitivity analyses were performed in which incremental costs and QALYs were calculated as the mean of 1,000 Monte Carlo simulations, where each simulation samples simultaneously from the appropriate distributions of the input parameters (see Appendix 5.3 for the distributions). Cost-effectiveness planes and acceptability curves were constructed from the Monte Carlo simulation. In addition, we used the headroom (i.e., potential profit) method to assess the maximum additional cost for which each new diagnostic test was still likely to be cost-effective at a willingness-to-pay threshold of €20,000 per QALY gained.[45,46]

Furthermore, we explored the impact of our model parameters in univariate sensitivity analyses, varying the sensitivity, specificity, new test costs, DAS28 improvement of TPs in the new test strategy who were FN in the current test strategy, and costs of biologic DMARDs for FPs in the first year after diagnosis. We report these analyses for an add-on test for intermediate-risk patients. For each analysis, one model parameter was altered while the other parameters were held constant at the baseline value.

In our analyses, differential discounting was applied in accordance with the Dutch guidelines for economic evaluation research, with an annual discount rate of 4.0% for all costs and 1.5% for health effects.[47]

Model validation

The model structure and input parameters were checked on clinical correctness by rheumatologists. We also verified the model for coding and logical correctness by running extreme value scenarios. Furthermore, an independent modeler internally validated our model to check the model structure, all input parameters with distributions, and the visual basic code used to program the model in Excel.

Ethical approval and patient's consent

No ethical approval and consent from patients was needed for this study.

5.3 Results

Reclassification

In the add-on test strategy, only intermediate-risk patients could reclassify to high-risk or low-risk of having RA, and low-risk patients could reclassify to high-risk. Table 5.1 shows the reclassification for both add-on strategies. Due to a high specificity of B-cell test and genetic assay, more patients reclassified to low-risk, while due to a high sensitivity of IL-6 and MRI, more patients reclassified to high-risk.

Cost-effectiveness

Table 5.2 shows the results of the RA-2010 criteria and the four new tests used as add-on for all patients, add-on for intermediate-risk patients, and replacement of the RA-2010 criteria.

The B-cell test was the least expensive strategy when used as an add-on for intermediate-risk patients and as replacement test, making it the dominant strategy since it has better health outcomes and lower costs. As add-on for all patients, the B-cell test was also the most cost-effective test strategy. When using a willingness-to-pay threshold of €20,000 per QALY gained, the IL-6 and MRI test strategies were not cost-effective, except in the replacement strategy. A genetic assay was not cost-effective in any strategy.

Table 5.1: Reclassification table with the results for *intermediate-risk patients* according to the ACR/EULAR 2010 RA classification criteria and *all patients*

	Number of patients according to ACR/EULAR 2010 RA classification criteria	Number of patients reclassified to **high-risk**	Number of patients reclassified to **low-risk**	**Combined sensitivity**: RA-2010 criteria + new test	**Combined specificity**: RA-2010 criteria + new test
A: ADD-ON TEST FOR ALL PATIENTS					
B-CELL TEST				0.85	0.69
High-risk	243	243	0		
Intermediate-risk	**263**	**75 (29%)**	**188 (71%)**		
Low-risk	**46**	**13 (28%)**	**33 (72%)**		
IL-6 TEST				0.89	0.41
High-risk	243	243	0		
Intermediate-risk	**263**	**146 (56%)**	**117 (44%)**		
Low-risk	**46**	**26 (57%)**	**20 (43%)**		
MRI				0.96	0.46
High-risk	243	243	0		
Intermediate-risk	**263**	**154 (59%)**	**109 (41%)**		
Low-risk	**46**	**26 (59%)**	**20 (41%)**		
GENETIC ASSAY TEST				0.77	0.66
High-risk	243	243	0		
Intermediate-risk	**263**	**64 (24%)**	**199 (76%)**		
Low-risk	**46**	**11 (24%)**	**35 (76%)**		
B: ADD-ON TEST FOR INTERMEDIATE-RISK PATIENTS					
B-CELL TEST				0.81	0.71
High-risk	243	243	0		
Intermediate-risk	**263**	**75 (29%)**	**188 (71%)**		
Low-risk	46	0	46		
IL-6 TEST				0.85	0.46
High-risk	243	243	0		
Intermediate-risk	**263**	**146 (56%)**	**117 (44%)**		
Low-risk	46	0	46		
MRI				0.91	0.51
High-risk	243	243	0		
Intermediate-risk	**263**	**154 (59%)**	**109 (41%)**		
Low-risk	46	0	46		
GENETIC ASSAY TEST				0.75	0.67
High-risk	243	243	0		
Intermediate-risk	**263**	**64 (24%)**	**199 (76%)**		
Low-risk	46	0	46		

Combined sensitivity = sensitivity of ACR/EULAR 2010 RA classification criteria + sensitivity of new test; Combined specificity = specificity of ACR/EULAR 2010 RA classification criteria + specificity of new test.

When comparing the test strategies, a replacement of the RA-2010 criteria was the most cost-effective test strategy, followed by the add-on test for intermediate-risk patients, and the least favorable was the add-on test for all patients.

Headroom

Table 5.2 shows the maximum additional cost for which each new test was likely to be cost-effective at a willingness-to-pay threshold of €20,000 per QALY gained (i.e., the headroom). Given the sensitivity and specificity of the different tests, an IL-6 test will only be cost-effective with a unit cost below €59 in the replacement test strategy. The headroom of a genetic assay test (€210 as add-on for all patients, €173 as add-on for intermediate-risk patients, and €543 as replacement) also shows that the current unit costs of this test was too high, because the headroom was lower than the current unit costs of this test (€750).

Table 5.2: Five-year cost-effectiveness of new test strategies versus current test strategy

Test strategy	New test	Se*	Sp*	Test costs*	TP n(%)	FP n(%)	TN n(%)	FN n(%)	Costs	QALYs	ΔCosts	ΔQALYs	ICER	Headroom#
ACR/EULAR 2010 RA			**0.62**	**0.77**	**€1,593^**	**185(34)**	**58(11)**	**195(35)**	**114(21)**	**€16,784**	**3.430**			
Add-on all patients	B-cell	0.60	0.90	€150	254(46)	77(14)	175(32)	46(8)	€ 16,807	3.454	€ 23	0.024	€ 969	€602
	IL-6	0.70	0.53	€ 50	265(48)	149(27)	103(19)	34(6)	€ 17,387	3.451	€ 603	0.021	€ 28,171	-€125[&]
	MRI	0.90	0.60	€756	288(52)	135(24)	117(21)	11(2)	€ 17,848	3.461	€ 1,063	0.031	€ 34,318	€312
	Genetic	0.40	0.85	€750	231(42)	87(16)	166(30)	69(13)	€ 17,611	3.444	€ 827	0.014	€ 57,606	€210
Add-on intermediate-risk patients	B-cell	0.60	0.90	€150	244(44)	74(13)	178(32)	56(10)	€ 16,748	3.450	- € 37	0.020	Dominant[$]	€511
	IL-6	0.70	0.53	€ 50	254(46)	135(24)	117(21)	46(8)	€ 17,271	3.448	€ 487	0.018	€ 26,696	-€72[&]
	MRI	0.90	0.60	€756	273(49)	124(22)	128(23)	27(5)	€ 17,404	3.456	€ 620	0.026	€ 23,457	€269
	Genetic	0.40	0.85	€750	224(41)	82(15)	170(31)	75(14)	€ 17,211	3.442	€ 427	0.012	€ 35,233	€173
Example of what would happen if one would replace the RA classification criteria														
Replacement all	B-cell	0.60	0.90	€150	180(33)	25(5)	227(41)	120(22)	€15,983	3.432	- € 801	0.002	Dominant[$]	€984
	IL-6	0.70	0.53	€ 50	210(38)	119(22)	134(24)	90(16)	€16,849	3.434	€ 64	0.004	€17,526	€ 59
	MRI	0.90	0.60	€756	270(49)	101(18)	151(27)	30(5)	€17,139	3.458	€ 355	0.027	€12,906	€951
	Genetic	0.40	0.85	€750	120(22)	38(7)	214(39)	180(33)	€16,675	3.414	- € 110	-0.016	€ 6,914	€543

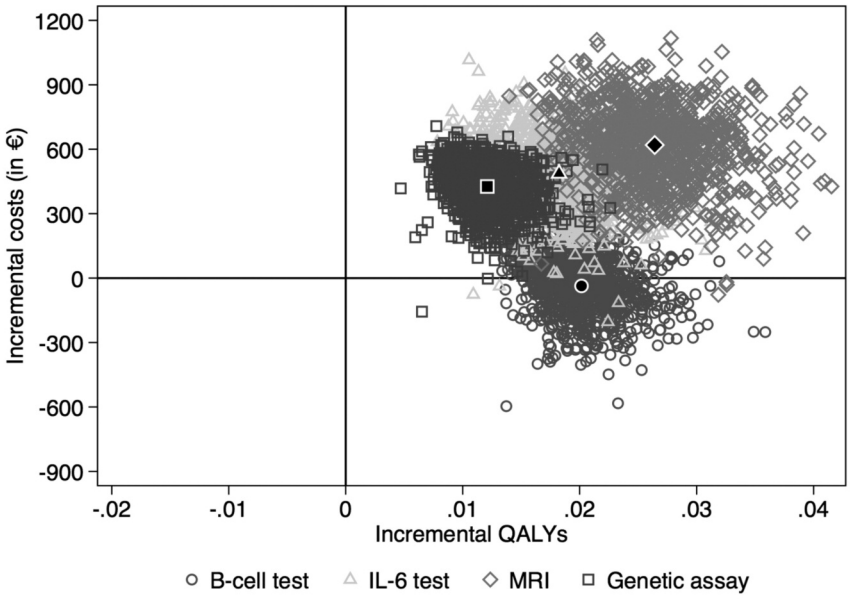
Se = sensitivity; Sp = specificity; TP = true positive; FP = false positive; TN = true negative; FN = false negative; QALY = quality-adjusted life year;
ICER = incremental cost-effectiveness ratio (€ per QALY gained); * These are the Se, Sp and costs of the new test; the Se and Sp of the combination of the new test plus the ACR/EULAR criteria are reported in the text; # Willingness to pay threshold is €20,000 per QALY gained; [$] Dominant=better health outcomes and lower costs;
^ Costs of visits and diagnostic tests during first year; [&] Mainly due to the low specificity, an IL6 test as add-on test can never be cost-effective compared to the current test strategy.

Probabilistic sensitivity analysis

Figure 5.2 shows the cost-effectiveness planes with the average costs and effects and the uncertainty around this average for the add-on test strategy in intermediate-risk patients. All estimates lie within the Northeast or Southeast quadrants, meaning improved health outcomes. In the Northeast quadrant, this is accompanied by higher costs as seen for IL-6, MRI, and gene assay. Nearly 100% of these estimates (99.6%, 99.7%, and 99.8%, respectively) lie within this quadrant. For B-cell, part of the estimates (57.2%) lie within the Southeast quadrant depicting lower costs and improved health outcomes. Uncertainty for the MRI and IL-6 test were greater than for B-cell and gene assay as shown by the width and height of the cloud, which is a consequence of the low specificity of 0.60 (MRI) and 0.53 (IL-6).

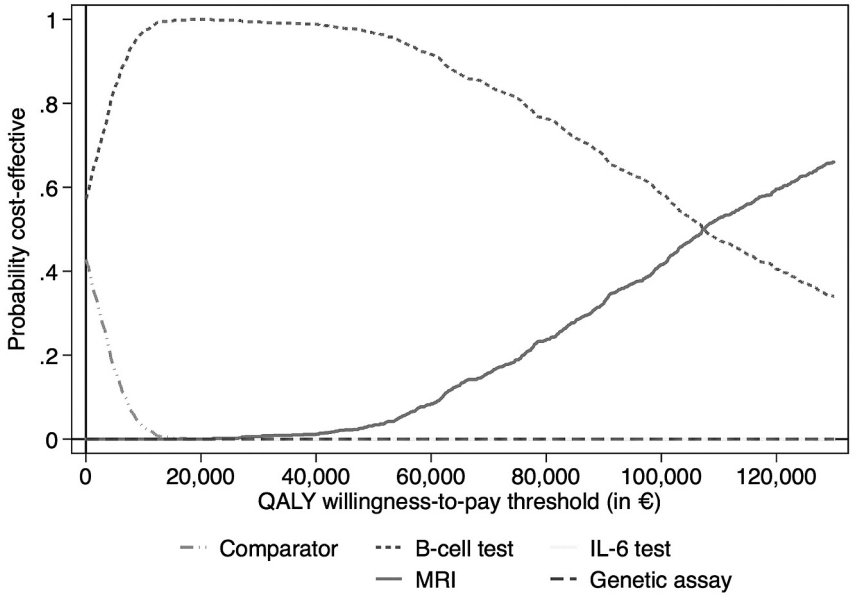
Figure 5.3 shows the cost-effectiveness acceptability curves for the add-on test strategy for intermediate-risk patients. If a willingness-to-pay threshold of €20,000 per QALY gained is used, there is a probability of cost-effectiveness for the B-cell test of 100%.

Figure 5.2: Cost-effectiveness planes of add-on for intermediate-risks test strategy versus the current diagnostic test strategy



Current test strategy = ACR/EULAR 2010 RA classification criteria; QALY = quality-adjusted life year.

Figure 5.3: Cost-effectiveness acceptability curves of add-on for intermediate-risks test strategy and the current diagnostic test strategy



Current test strategy = ACR/EULAR 2010 RA classification criteria; QALY = quality-adjusted life year.

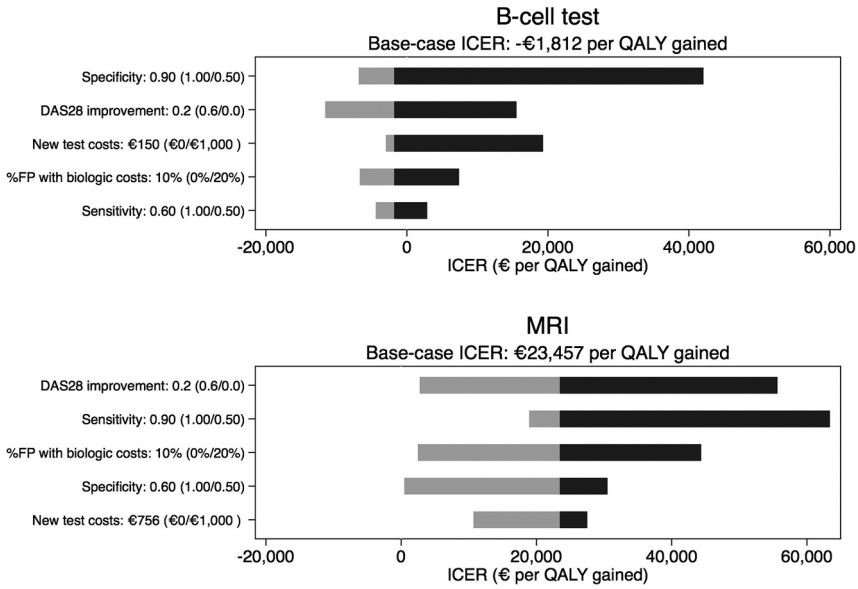
Univariate sensitivity analysis

The sensitivity of our results was evaluated by varying single model parameters and comparing the ICERs to the base-case ICERs of the most *cost-effective* test (B-cell) and the most *effective* test (MRI) when used as add-on in intermediate-risk patients. The base-case ICERs were estimated using a DAS28 improvement of 0.2 at 12 months in the FNs, 10% of FPs with biologic DMARD costs in the first year after diagnosis, and the test-specific sensitivity, specificity, and costs.

For a B-cell test, the change in specificity had the largest impact (see Figure 5.4), followed by DAS28 improvement of TPs in the new test strategy who were FN in the current test strategy. No DAS28 improvement resulted in an ICER of about €16,000, while a DAS28 improvement of 0.6 resulted in a dominant new test strategy (i.e., better health outcomes and lower costs). Varying the new test costs, the number of FPs having biological DMARD costs between 0% and 20%, and sensitivity had less impact on the ICER, but still caused variation.

For the MRI, the change in DAS28 improvement of TPs in the new test strategy who were FN in the current test strategy had the largest impact

Figure 5.4: Impact of varying model inputs on incremental cost-effectiveness ratio for an add-on B-cell test or MRI for intermediate-risk patients



QALY = quality-adjusted life year; ICER = incremental cost-effectiveness ratio.

(see Figure 5.4). No DAS28 improvement resulted in an ICER of about €56,000, while a DAS28 improvement of 0.6 resulted in an ICER of about €3,000. Changing sensitivity to 0.50 or 1.00 had about the same impact as varying the number between 0% and 20% of FPs having biological DMARD costs. Adjustments in specificity and costs had the least impact on the ICER.

5.4 Conclusions and discussion

Various new medical technologies enable the detection of RA in an increasingly earlier stage. RA is a disease in which early detection has high potential because there are treatments available that effectively reduce disease progression especially if introduced early in the disease. In our study, we compared four different tests and we found that a B-cell test was the most cost-effective test in all test strategies. This is mainly due to the high specificity (0.90) in combination with moderate sensitivity (0.60), and relative low test costs (€150). When the specificity of the test increases, the number of false positives who may get unnecessary expensive treatment is reduced, which largely increases the likelihood that the test becomes cost-effective.

A MRI was the second most cost-effective test as add-on for intermediate-risk patients. Mainly as a result of the higher costs of a MRI (€756) compared to a B-cell test (€150), the MRI was less cost-effective, even though it had a much higher sensitivity than the B-cell test (0.90 versus 0.60). However, the MRI had a lower specificity, which had more impact on the ICER than a lower sensitivity. The specificity is more important because an add-on test with a specificity lower than 100% results in more patients classified as FP while these patients were TN according to the current test strategy.

This study indicated where it is most likely to position a new diagnostic test given their sensitivity, specificity and costs in relation to the current RA-2010 criteria test strategy. B-cell gene expression as add-on test for *intermediate-risk patients* was the most likely cost-effective add-on strategy and could even replace the RA-2010 criteria given its moderate sensitivity (0.60), high specificity (0.90) and low costs (€150). However, it is unlikely that a replacement test would be discovered that provides the rheumatologists with the same richness of information as the current RA-2010 criteria. To use a new test in addition to the RA-2010 criteria in all patients is the least cost-effective, because it would not alter treatment for the high-risks (243 of 552 patients). These patients will have higher test costs without additional health gain.

The discussion about the potential benefits of early detection and treatment of RA is also relevant in the light of the current discussion about biosimilars. Currently, the price of biologic DMARDs is set at €14,000 per patient per year. Due to this high price (which is about the same for all biologic DMARDs), it is an important driver for the cost-effectiveness results. If the price of a biosimilar will be significantly lower, it will have a smaller impact on the results. On the one hand, this could result in earlier biological start resulting in more TPs using biologic DMARDs. On the other hand, a cheaper biosimilar may worsen the cost-effectiveness of new tests because the savings from postponing or preventing treatment with a biologic are less great. The net result is hard to predict as prescription behavior may change with lower costs of biologicals.

The results of our study are likely to be generalizable to other countries and different healthcare systems since the RA-2010 criteria are widely used internationally for classifying RA. The costs of diagnostic tests and treatment patterns obviously differ between countries and healthcare systems (e.g., higher costs in the US). However, this would not result in differences in the most cost-effective test strategy.

The results of our simulation should be interpreted taking into account that we used fixed values of sensitivity, specificity and costs for the alternative tests. This was done to show the differences in early cost-effectiveness between tests with different test characteristics. If we would have added distributions around these parameters, the differences in cost-effectiveness be-

tween the tests would likely decrease due to large uncertainties of these parameters in early CEAs. To overcome this, we performed univariate sensitivity analysis to explore the impact of changing these parameters. We found that multiple factors had impact on the cost-effectiveness of a new test strategy. The main drivers of the ICER include the sensitivity, specificity and costs of the new test, but also the DAS28 improvement for TPs in the new test strategy who were FN according to the RA-2010 criteria, and the percentage of FPs with biological DMARD costs in the first year after diagnosis.

In addition, since RA is a heterogeneous disease, a new test might provide additional diagnostic information in particular subgroups, such as obese or those with coexistent osteoarthritis. Furthermore, the choice for a new test might also be guided by additional diagnostic information that is expected in particular subgroups. Further research should investigate the impact of this additional diagnostic information on the cost-effectiveness of these new tests.

Like any early-CEA study, our study had limitations. One was that limited data was available on long-term disease progression from a cohort receiving usual care. Therefore, we had to synthesize data from different data sources and to make assumptions about the new test strategies based on expert opinion. Typically, the improvement in health outcomes of adding a new test without observed data is difficult to estimate. In the first year after diagnosis, we assigned equal improvements in utilities for TPs and TNs, less gain for FPs, and reduced utilities for FPs. Whether this reflects clinical practice needs to be proven. Another limitation could be that a death state was not included in our model. However, because of the low (1-2%) mortality risk of RA patients during a follow-up of five years,[48] and the equal mortality risk in the current and new diagnostic test strategy, the mortality risk would not influence the model results.

We have shown that the B-cell gene expression as add-on test for *intermediate-risk patients* was the most likely cost-effective add-on strategy. A new add-on test for intermediate-risk patients should have a high specificity and the costs should not be higher than €200-€300.

5.5 Acknowledgments

This research was performed within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), and the Dutch Arthritis Foundation, project TRACER (grant 04I-202). We are grateful to Nasuh Büyükkaramikli for performing an internal validity assessment of the model.

5.6 References

1 Pincus T, Callahan LF, Sale WG, et al. Severe functional declines, work disability, and increased mortality in seventy-five rheumatoid arthritis patients studied over nine years. *Arthritis Rheum* 1984;27(8):864–72.

2 Mitchell DM, Spitz PW, Young DY, et al. Survival, prognosis, and causes of death in rheumatoid arthritis. *Arthritis Rheum* 1986;29(6):706–14.

3 Scott DL, Symmons DP, Coulton BL, et al. Long-term outcome of treating rheumatoid arthritis: results after 20 years. *Lancet* 1987;1(8542):1108–11.

4 Isomäki H. Long-term outcome of rheumatoid arthritis. *Scand J Rheumatol Suppl* 1992;95:3–8.

5 Wolfe F. The natural history of rheumatoid arthritis. *J Rheumatol Suppl* 1996;44:13–22.

6 Aletaha D, Neogi T, Silman AJ, et al. Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;69(9):1580–8.

7 Bijlsma JW, Weinblatt ME. Optimal use of methotrexate: the advantages of tight control. *Ann Rheum Dis* 2007;66(11):1409-10.

8 Jacobs JW. Optimal use of non-biologic therapy in the treatment of rheumatoid arthritis. *Rheumatology (Oxford)* 2012;51(4):iv3–8.

9 Chen YF, Jobanputra P, Barton P, et al. A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness. *Health Technol Assess* 2006;10(42):iii-iv, xi-xiii, 1–229.

10 Donahue KE, Garthehner G, Jonas DE, et al. Comparative Effectiveness of Drug Therapy for Rheumatoid Arthritis and Psoriatic Arthritis in Adults. Rockville, MD: AHRQ Comparative Effectiveness Reviews; 2007.

11 Singh JA, Christensen R, Wells GA, et al. A network meta-analysis of randomized controlled trials of biologics for rheumatoid arthritis: a Cochrane overview. *CMAJ* 2009;181(11):787–96.

12 Hazes JM, Luime JJ. The epidemiology of early inflammatory arthritis. *Nat Rev Rheumatol* 2011;7(7):381-90.

13 Drummond MF, Sculpher MJ, Torrance GW, et al. Methods for the Economic Evaluation of Health Care Programmes. 3rd ed. Oxford: Oxford University Press; 2005.

14 Buisman LR, Rutten-van Mölken MPMH, Postmus D, et al. The early bird catches the worm: Early cost-effectiveness analysis of new medical tests. *Int J Technol Assess Health Care* 2016;32(1):1–8. Epub 2016 Mar 22.

15 The Center for Translational Molecular Medicine. Full article on project TRACER. Available from: http://www.ctmm.nl/en/downloads/full-article-downloads-pdf/full-article-on-project-tracer. Accessed January 27, 2015.

16 Van Baarsen LG, Bos WH, Rustenburg F, et al. Gene expression profiling in autoantibody-positive patients with arthralgia predicts development of arthritis. *Arthritis Rheum* 2010;62(3):694–704.

17 Gottenberg JE, Dayer JM, Lukas C, et al. Serum IL-6 and IL-21 are associated with markers of B cell activation and structural progression in early rheumatoid arthritis: results from the ESPOIR cohort. *Ann Rheum Dis* 2012;71(7):1243–8.
18 Duer-Jensen A, Hørslev-Petersen K, Hetland ML, et al. Bone edema on magnetic resonance imaging is an independent predictor of rheumatoid arthritis development in patients with early undifferentiated arthritis. *Arthritis Rheum* 2011;63(8):2192–202.
19 Bird P, Conaghan P, Ejbjerg B, et al. The development of the EULAR-OMERACT rheumatoid arthritis MRI reference image atlas. *Ann Rheum Dis* 2005;64(1):i8–10.
20 Nieuwenhuis WP, Krabben A, Stomp W, et al. Evaluation of magnetic resonance imaging-detected tenosynovitis in the hand and wrist in early arthritis. *Arthritis Rheumatol* 2015;67(4):869–76.
21 Sugimoto H, Takeda A, Hyodoh K. Early-stage rheumatoid arthritis: prospective study of the effectiveness of MR imaging for diagnosis. *Radiology* 2000;216(2):569–75.
22 Narváez J, Sirvent E, Narváez JA, et al. Usefulness of magnetic resonance imaging of the hand versus anticyclic citrullinated peptide antibody testing to confirm the diagnosis of clinically suspected early rheumatoid arthritis in the absence of rheumatoid factor and radiographic erosions. *Semin Arthritis Rheum* 2008;38(2):101–9.
23 Machado PM, Koevoets R, Bombardier C, Van der Heijde DM. The value of magnetic resonance imaging and ultrasound in undifferentiated arthritis: a systematic review. *J Rheumatol Suppl* 2011;87:31–7.
24 Yarwood A, Huizinga TW, Worthington J. The genetics of rheumatoid arthritis: risk and protection in different stages of the evolution of RA. *Rheumatology (Oxford)* 2016;55(2):199–209.
25 Dutch Healthcare Authority (in Dutch: Nederlandse Zorgauthoriteit, NZa). NZa Zorgproducten Tariefapplicatie. Available from: http://dbc-zorgproducten-tarieven.nza.nl/nzaZpTarief/ZoekfunctieDbc.aspx. Accessed May 1, 2013.
26 Smolen JS, Breedveld FC, Burmester GR, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis* 2016;75(1):3–15.
27 Dutch Institute for Healthcare Improvement CBO. Guideline diagnostics and treatment for rheumatoid arthritis (In Dutch: Richtlijn Diagnostiek en behandeling van reumatoïde artritis). Utrecht: Dutch Institute for Healthcare Improvement CBO; 2008.
28 Alves C, Luime JJ, Van Zeben D, et al. Diagnostic performance of the ACR/EULAR 2010 criteria for rheumatoid arthritis and two diagnostic algorithms in an early arthritis clinic (REACH). *Ann Rheum Dis* 2011;70(9):1645–7.
29 Van Gestel AM, Haagsma CJ, Van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998;41(10):1845–50.

30 Fransen J, Creemers MC, Van Riel PL. Remission in rheumatoid arthritis: agreement of the disease activity score (DAS28) with the ARA preliminary remission criteria. *Rheumatology (Oxford)* 2004;43(10):1252–5.
31 Prevoo ML, Van 't Hof MA, Kuper HH, et al. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38(1):44–8.
32 Luime JJ, Buisman LR, Oppe M, et al. A cost-effectiveness model for evaluating new diagnostic tests in the work-up of patients with inflammatory arthritis at risk of having Rheumatoid Arthritis. *Arthritis Care Res (Hoboken)* 2015. doi:10.1002/acr.22776. Epub 2015 Nov 10.
33 De Jong PH, Hazes JM, Han HK, et al. Randomised comparison of initial triple DMARD therapy with methotrexate monotherapy in combination with low-dose glucocorticoid bridging therapy; 1-year data of the tREACH trial. *Ann Rheum Dis* 2014;73(7):1331–9.
34 Claessen SJ, Hazes JM, Huisman MA, et al. Use of risk stratification to target therapies in patients with recent onset arthritis; design of a prospective randomized multicenter controlled trial. *BMC Musculoskelet Disord* 2009;10:71.
35 Vermeer M, Kievit W, Kuper HH, et al. Treating to the target of remission in early rheumatoid arthritis is cost-effective: results of the DREAM registry. *BMC Musculoskelet Disord* 2013;14:350.
36 Tan SS, Bouwmans CA, Rutten FF, et al. Update of the Dutch Manual for Costing in Economic Evaluations. *Int J Technol Assess Health Care* 2012;28(2):152–8.
37 National Health Care Institute. Available from: http://www.medicijnkosten.nl/default.asp. Accessed February 2, 2015.
38 Statistics Netherlands. Available from: http://statline.cbs.nl. Accessed February 2, 2015.
39 Pliskin JS, Shepard D, Weinstein MC. Utility functions for life years and health status. *Operations Research* 1980;28(1):206–24.
40 Buitinga L, Braakman-Jansen LM, Taal E, et al. Comparative responsiveness of the EuroQol-5D and Short Form 6D to improvement in patients with rheumatoid arthritis treated with tumor necrosis factor blockers: results of the Dutch Rheumatoid Arthritis Monitoring registry. *Arthritis Care Res* 2012;64(6):826–32.
41 Lillegraven S, Kristiansen IS, Kvien TK. Comparison of utility measures and their relationship with other health status measures in 1041 patients with rheumatoid arthritis. *Ann Rheum Dis* 2010;69(10):1762–7.
42 Linde L, Sørensen J, Østergaard M, et al. Gain in quality-adjusted life-years in patients with rheumatoid arthritis during 1 year of biological therapy: a prospective study in clinical practice. *J Rheumatol* 2013;40(9):1479–86.
43 Sørensen J, Linde L, Ostergaard M, et al. Quality-adjusted life expectancies in patients with rheumatoid arthritis—comparison of index scores from EQ-5D, 15D, and SF-6D. *Value Health* 2012;15(2):334–9.

44 Verstappen SM, McCoy MJ, Roberts C, et al. Beneficial effects of a 3-week course of intramuscular glucocorticoid injections in patients with very early inflammatory polyarthritis: results of the STIVEA trial. *Ann Rheum Dis* 2010;69(3):503–9.

45 Girling A, Young T, Brown C, et al. Early-stage valuation of medical devices: the role of developmental uncertainty. *Value Health* 2010;13(5):585-91.

46 Cosh E, Girling A, Lilford R, et al. Investing in new medical technologies: a decision framework. *J Commerc Biotechnol* 2007;13(4):263–271.

47 College voor zorgverzekeringen (CVZ). Guidelines for pharmacoeconomic research, updated version. Diemen: College voor zorgverzekeringen; 2006.

48 Toledano E, Candelas G, Rosales Z, et al. A meta-analysis of mortality in rheumatic diseases. *Reumatol Clin* 2012;8(6):334–41.

5.7 Appendices

Appendix 5.1: Cohort characteristics

Source	Patient population	Age (SD)	Sex (% female)	ACCP/RF positive	≥6 points on RA-2010 criteria	Intervention	Control	QALY data available	Work outcomes data available
REACH [1]	Newly detected early inflammatory arthritis patients	53 (15)	68%	43%	42%	Usual care	NA	Yes	Yes
tREACH [2]	Newly diagnosed RA patients	I: 53 (15) I: 54 (14) 7 C: 54 (14)	60% 2% 70%	81%/76% 72%/70% 77%/67%	97%	Methotrexate, Plaquenil, Sulphazalazine Prednisone as bridging therapy	Methotrexate, Prednisone as bridging therapy	Yes	Yes
DREAM registry [3]	Newly diagnosed RA-patients (clinical diagnosis)	I: 58 (14) C: 57 (13)	61.7%	68%*	NG	Treat- to-target step-up therapy	Usual care	No	No

I = Intervention; C = Comparator; SD = standard deviation; QALY = quality-adjusted life year; * RF only; NG = Not Given; NA = Not Applicable.

References Appendix 5.1

1 REACH data on file.

2 De Jong PH, Hazes JM, Han HK, et al. Randomised comparison of initial triple DMARD therapy with methotrexate monotherapy in combination with low-dose glucocorticoid bridging therapy; 1-year data of the tREACH trial. *Ann Rheum Dis* 2014;73(7):1331–9.

3 Vermeer M, Kievit W, Kuper HH, et al. Treating to the target of remission in early rheumatoid arthritis is cost-effective: results of the DREAM registry. *BMC Musculoskelet Disord* 2013;14:350.

Appendix 5.2: Characteristics of the unclassified inflammatory arthritis patients in REACH

		All patients (n=552)
Age (in years; SD)		53 (15)
Sex (% female)		68%
DOMAINS OF ACR/EULAR 2010 RA CLASSIFICATION		
Swollen Joints	> 10	19%
	4-10 small	34%
	1-3 small	33%
	2-10 large	3%
	1 large	6%
Serology	high	34%
	low	9%
Disease duration: >6 weeks		85%
Elevated acute phase reactants		53%
SCORE ON ACR/EULAR 2010 RA CLASSIFICATION		
6 or more points		42%
3-5 points		48%
0-2 points		11%

SD = standard deviation.

Design of the Rotterdam Early Arthritis CoHort

The Rotterdam Early Arthritis CoHort (REACH) is an inception cohort study with 2 years of follow up. Assessments took place at baseline, 6 months, 12 months and 24 months. REACH aims to study the etiopathogenesis, diagnostic strategies, and outcome of patients with inflammatory joint conditions for <12 months. Both general practitioners and rheumatologists invited patients to participate in REACH from July 2004 to April 2011. For general practitioners, short educational courses on the importance of early treatment of RA and early referral were organized. Physicians that agreed to participate in REACH received written information and verbal instructions on the general aims of the study and on how to send patients for inclusion in the study. Data collection includes a large array of detailed medical examinations and questionnaires.

General practitioners selected patients with arthritis in ≥1 joint or patients experiencing conditions in ≥2 joints without synovitis. The general practitioners determined that conditions existed for <12 months and were not due to trauma/mechanical problems. In addition, patients had to be age

>16 years. During an interview by telephone and a subsequent medical examination by a rheumatologist, the inclusion criteria were verified. Patients were included if 1) joint conditions existed for <12 months with no requirement of a minimum duration; 2) they had arthritis in ≥1 joint or conditions in ≥2 joints in combination with at least 2 of the following criteria ascertained during medical examination by a rheumatologist: morning stiffness for >1 hour, bilateral compression pain in the metacarpophalangeal or metatarsophalangeal joints, symmetric presentation, positive family history, non-fitting shoes, non-fitting rings, pins and needles in fingers, or unexplained fatigue for <1 year; and 3) conditions were predominantly present in the morning and at night, and improved with movement. Patients were excluded if 1) conditions were due to trauma/mechanical problems, 2) they were age <16 years, 3) no written communication was possible in Dutch, or 4) a prior diagnosis of RA, ankylosing spondylitis, Sjögren's syndrome, systemic lupus erythematosus, or juvenile arthritis had been made by a rheumatologist before inclusion in this study.

For patients directly visiting rheumatologists, a similar verification procedure was applied. The final diagnosis was made by a rheumatologist in one of the 5 participating out-patient rheumatology clinics.

The study was approved by the Erasmus MC Medical Ethical Committee. All patients gave written informed consent.

In the current study we use 552 consecutive patients with unclassified inflammatory arthritis.

Appendix 5.3: Model input parameters

Parameter	Value	SE	Distribution (alpha;beta)	Source
PREVALENCE OF RA BASED ON ACR/EULAR 2010 RA CLASSIFICATION CRITERIA				
All patients	0.540	0.021	Beta (300;252)	REACH+
Intermediate risk patients	0.370	0.030	Beta (97;166)	REACH+
PROBABILITIES				
METHOTREXATE START AND CONTINUATION				
Patients with ≥6 points at baseline				
Start at baseline	1.000			[1]
Continuation at 6 months	0.840	0.024	Dirichlet (1,282;0.0007)	REACH+
Continuation at 12 months†	0.880	0.026	Dirichlet (1,125;0.0008)	REACH+
<6 points at baseline				
Start at baseline	0.000			[1]
Start at 6 months	0.350	0.027	Dirichlet (167;0.0021)	REACH+
Continuation at 12 months†	0.840	0.035	Dirichlet (592;0.0014)	REACH+
Start at 12 months	0.140	0.053	Dirichlet (7;0.0199)	REACH+
OTHER DIAGNOSIS				
≥6 points at baseline				
6 months	0.140	0.022	Dirichlet (35;0.0038)	REACH+
12 months	0.120	0.021	Dirichlet (201;0.0045)	REACH+
<6 points at baseline				
6 months	0.510	0.028	Dirichlet (322;0.0016)	REACH+
12 months	0.000	0.000	Dirichlet (0.000; 0.000)	REACH+
Biological use at 12 months	0.100	0.021	Beta (0.027;21)	REACH+, DREAM [2]
Biological start – per 3 months				
Total	0.020	0.010	Beta (4;48)	DREAM [2]
DAS28≤2.6	0.002			
DAS28>2.6–≤3.2	0.006			
DAS28>3.2	0.012			
EQ-5D always >0.5 over time				
Patients not using a biological	0.752	0.024	Beta (245;81)	tREACH++
Biological patients	0.736	0.030	Beta (155;56)	tREACH++
DAS28 INCREASE AT TIME OF BIOLOGICAL START				
Delta DAS28≤2.6	2.064	0.120	Normal	tREACH++
Delta DAS28>2.6–≤3.2	0.860	0.058	Normal	tREACH++

Appendix 5.3: Model input parameters (continued)

Parameter	Value	SE	Distribution (alpha;beta)	Source
OLS FUNCTIONS OF DAS28 OVER TIME SINCE START BIOLOGICAL				
0–3 months from biological start				
Time	-0.290	0.038	Normal	tREACH++
Constant	4.019	0.109	Normal	tREACH++
4–15 months from biological start				
Time	-0.044	0.014	Normal	tREACH++
Constant	3.564	0.172	Normal	tREACH++
16–48 months from biological start				
Time	-0.015	0.011	Normal	tREACH++
Constant	3.426	0.289	Normal	tREACH++
COSTS				
COSTS DURING 1ST YEAR – 0-12 MONTHS				
Baseline visit and diagnostic tests	€917	€138	Gamma (€44;€21)	Estimate***
Follow-up visits and diagnostic tests during first year	€676	€101	Gamma (€44;€15)	Estimate***
MTX costs plus monitoring during first year	€102	€15	Gamma (€44;€2)	Estimate***
Other sDMARD costs				
True positive patients	€165	€10	Gamma (€253;€0.65)	REACH+
False positive patients	€103	€14	Gamma (€56;€1.83)	REACH+
True negative patients****	€57	€5	Gamma (€112;€0.51)	REACH+
False negative patients	€117	€11	Gamma (€104;€1.12)	REACH+
Productivity costs				
True positive patients	€1,368	€205	Gamma (€44;€31)	REACH+
False positive patients	€961	€144	Gamma (€44;€22)	REACH+
True negative patients	€710	€107	Gamma (€44;€16)	REACH+
False negative patients	€1,691	€254	Gamma (€44;€38)	REACH+
COSTS PER 3 MONTHS – 12-60 MONTHS				
MTX	€38	€3	Gamma (€44;€0.41)	[3]
Biologic DMARD	€3,500	€525	Gamma (€44;€79)	[3]
Other sDMARD	€27	€2	Gamma (€192;€0.14)	tREACH++
Direct medical costs				
DAS28≤2.6	€105	€3	Gamma (€964;€0.11)	tREACH++
DAS28>2.6–≤3.2	€133	€8	Gamma (€287;€0.47)	tREACH++
DAS28>3.2	€190	€6	Gamma (€959;€0.20)	tREACH++

Appendix 5.3: Model input parameters (continued)

Parameter	Value	SE	Distribution (alpha;beta)	Source
Productivity costs				
DAS28≤2.6	€493	€74	Gamma (€44;€11)	tREACH++
DAS28>2.6–≤3.2	€908	€136	Gamma (€44;€20)	tREACH++
DAS28>3.2	€1,247	€187	Gamma (€44;€28)	tREACH++
COSTS NON-RA PATIENTS				
12-60 months (for 4 years)	€5,600	€840	Gamma (€44;€126)	Estimate#
COST NEW DIAGNOSTIC TESTS				
MRI of both hands and feet	€756			[4]
B-cell test	€150			Estimate#
IL-6 test	€50			Estimate#
Genetic assay test	€750			Estimate#
UTILITIES				
True positive patients – baseline	0.600	0.015	Beta (626;417)	REACH+
12 months	0.700	0.010	Beta (2,354;1,569)	Estimate#
False positive patients – baseline	0.650		Beta (548;295)	REACH+
	0.700	0.016	Beta (1,471;490)	Estimate#
12 months		0.010		
True negative patients – baseline	0.650		Beta (548;295)	REACH+
12 months	0.750	0.016	Beta (882;156)	Estimate#
		0.011		
False negative patients – baseline	0.600		Beta (626;417)	REACH+
	0.550	0.015	Beta (2,942;2,942)	Estimate#
12 months		0.007		
PATIENTS NOT USING A BIOLOGICAL – 12-60 MONTHS				
EQ-5D always >0.5 over time				
DAS28≤2.6	0.858	0.003	Beta (11,632;1,923)	tREACH++
DAS28>2.6–≤3.2	0.815	0.006	Beta (3,387;770)	tREACH++
DAS28>3.2	0.763	0.005	Beta (5,220;1,625)	tREACH++
EQ-5D at least 1 time period <0.5				
DAS28≤2.6	0.746	0.015	Beta* (634;216)	tREACH++
DAS28>2.6–≤3.2	0.674	0.027	Beta* (201;97)	tREACH++
DAS28>3.2	0.462	0.025	Beta* (188;219)	tREACH++

Appendix 5.3: Model input parameters (continued)

Parameter	Value	SE	Distribution (alpha;beta)	Source
BIOLOGICAL PATIENTS – 12-60 MONTHS				
EQ-5D always >0.5 over time				
DAS28≤2.6	0.825	0.008	Beta (1,990;421)	tREACH++
DAS28>2.6–≤3.2	0.770	0.009	Beta (1,548;461)	tREACH++
DAS28>3.2	0.751	0.006	Beta (3,692;1,226)	tREACH++
EQ-5D at least 1 time period <0.5				
DAS28≤2.6	0.694	0.034	Beta** (124;55)	tREACH++
DAS28>2.6–≤3.2	0.623	0.049	Beta** (61;37)	tREACH++
DAS28>3.2	0.558	0.026	Beta** (199;157)	tREACH++
NON-RA PATIENTS				
12-60 months	0.750	0.010	Beta (1471;490)	Estimate#
DIAGNOSTIC PERFORMANCE OF TESTS				
B-cell test				
Sensitivity	0.60			[5]
Specificity	0.90			[5]
IL-6 serum level test				
Sensitivity	0.70			[6]
Specificity	0.53			[6]
MRI of both hands and feet				
Sensitivity	0.90			[7-9]
Specificity	0.60			[7-9]
Genetic assay test				
Sensitivity	0.40			[10]
Specificity	0.85			[10]
DISCOUNT RATES				
Costs	4.0%			[11]
Health effects	1.5%			[11]

SE = standard error; DHA = Dutch Healthcare Authority; + REACH data on file, details about data collection can be found in Alves et al.[12]; ++ tREACH data on file, details about data collection can be found in De Jong et al.[13]; * Beta distribution adjusted to allow values from -0.2 based on observed tREACH data; ** Beta distribution adjusted to allow values from -0.1 based on observed tREACH data; #Estimate based on expert opinion, for the SE data from REACH was used; *** Costs include diagnostic test costs (frequency laboratory tests and radiographics based on expert opinion; costs derived from the DHA tariffs), medical consultation time based on the Dutch manual for costing in economical evaluations[11]; † continuation of MTX at 12 month is the % of MTX users at 6 months that still use MTX at 12 months; **** True negative patients could use other sDMARDS due to other classifiable disease.

References Appendix 5.3

1 Aletaha D, Neogi T, Silman AJ, et al. Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;69(9):1580–8.
2 Vermeer M, Kievit W, Kuper HH, et al. Treating to the target of remission in early rheumatoid arthritis is cost-effective: results of the DREAM registry. *BMC Musculoskelet Disord* 2013;14:350.
3 National Health Care Institute. Available from: http://www.medicijnkosten.nl/default.asp. Accessed February 2, 2015.
4 Dutch Healthcare Authority (in Dutch: Nederlandse Zorgauthoriteit, NZa). NZa Zorgproducten Tariefapplicatie. Available from: http://dbc-zorgproducten-tarieven.nza.nl/nzaZpTarief/ZoekfunctieDbc.aspx. Accessed May 1, 2013.
5 Van Baarsen LG, Bos WH, Rustenburg F, et al. Gene expression profiling in autoantibody-positive patients with arthralgia predicts development of arthritis. *Arthritis Rheum* 2010;62(3):694–704.
6 Gottenberg JE, Dayer JM, Lukas C, et al. Serum IL-6 and IL-21 are associated with markers of B cell activation and structural progression in early rheumatoid arthritis: results from the ESPOIR cohort. *Ann Rheum Dis* 2012;71(7):1243–8.
7 Duer-Jensen A, Hørslev-Petersen K, Hetland ML, et al. Bone edema on magnetic resonance imaging is an independent predictor of rheumatoid arthritis development in patients with early undifferentiated arthritis. *Arthritis Rheum* 2011;63(8):2192–202.
8 Bird P, Conaghan P, Ejbjerg B, et al. The development of the EULAR-OMERACT rheumatoid arthritis MRI reference image atlas. *Ann Rheum Dis* 2005;64(1):i8–10.
9 Nieuwenhuis WP, Krabben A, Stomp W, et al. Evaluation of magnetic resonance imaging-detected tenosynovitis in the hand and wrist in early arthritis. *Arthritis Rheumatol* 2015;67(4):869–76.
10 Yarwood A, Huizinga TW, Worthington J. The genetics of rheumatoid arthritis: risk and protection in different stages of the evolution of RA. *Rheumatology (Oxford)* 2016;55(2):199–209.
11 College voor zorgverzekeringen (CVZ). Guidelines for pharmacoeconomic research, updated version. Diemen: College voor zorgverzekeringen; 2006.
12 Alves C, Luime JJ, Van Zeben D, et al. Diagnostic performance of the ACR/EULAR 2010 criteria for rheumatoid arthritis and two diagnostic algorithms in an early arthritis clinic (REACH). *Ann Rheum Dis* 2011;70(9):1645–7.
13 De Jong PH, Hazes JM, Han HK, et al. Randomised comparison of initial triple DMARD therapy with methotrexate monotherapy in combination with low-dose glucocorticoid bridging therapy; 1-year data of the tREACH trial. *Ann Rheum Dis* 2014;73(7):1331–9.

CHAPTER 6

CLINICAL PRACTICE VARIATION NEEDS TO BE CONSIDERED IN COST-EFFECTIVENESS ANALYSES: A CASE STUDY OF PATIENTS WITH A RECENT TRANSIENT ISCHEMIC ATTACK OR MINOR ISCHEMIC STROKE

Leander R. Buisman[1,2], Adriana J. Rijnsburger[1,2], Heleen M. den Hertog, MD[3,4], Aad van der Lugt, MD[5], William K. Redekop[1,2]

[1] Institute of Health Policy and Management,
Erasmus University Rotterdam, Rotterdam, The Netherlands
[2] Institute for Medical Technology Assessment,
Erasmus University Rotterdam, Rotterdam, The Netherlands
[3] Department of Neurology, Medical Spectrum Twente, Enschede,
The Netherlands
[4] Department of Neurology, Erasmus MC,
University Medical Center Rotterdam, Rotterdam, The Netherlands
[5] Department of Radiology, Erasmus MC,
University Medical Center Rotterdam, Rotterdam, The Netherlands

Published in *Applied Health Economics and Health Policy* 2016:14(1):67–75.
doi:10.1007/s40258-015-0167-4.

ABSTRACT

Background and Objective: The cost-effectiveness of clinical interventions is often assessed using current care as the comparator, with national guidelines as a proxy. However, this comparison is inadequate when clinical practice differs from guidelines, or when clinical practice differs between hospitals. We examined the degree of variation in the way patients with a recent transient ischemic attack (TIA) or minor ischemic stroke are assessed and used the results to illustrate the importance of investigating possible clinical practice variation, and the need to perform hospital-level cost-effectiveness analyses (CEAs) when variation exists.

Methods: Semi-structured interviews were conducted with 16 vascular neurologists in hospitals throughout the Netherlands. Questions were asked about the use of initial and confirmatory diagnostic imaging tests to assess carotid stenosis in patients with a recent TIA or minor ischemic stroke, criteria to perform confirmatory tests, and criteria for treatment. We also performed hospital-level CEAs to illustrate the consequences of the observed diagnostic strategies in which the diagnostic test costs, sensitivity and specificity were varied according to the local hospital conditions.

Results: 56% (9/16) of the emergency units and 63% (10/16) of the outpatient clinics use the initial and confirmatory diagnostic tests to assess carotid stenosis in accordance with the national guidelines. Of the hospitals studied, only one uses the recommended criteria for use of a confirmatory test, 38% (6/16) follow the guidelines for treatment. The most cost-effective diagnostic test strategy differs between hospitals.

Conclusions: If important practice variation exists, hospital-level CEAs should be performed. These CEAs should include an assessment of the feasibility and costs of switching to a different strategy.

Keywords: clinical practice variation, cost-effectiveness analysis, comparator, hospital-level, hospital-specific, diagnostic test, TIA, ischemic stroke, The Netherlands.

6.1 Introduction

The cost-effectiveness of clinical interventions (e.g., diagnostic tests, therapies or medicines) is normally assessed using current clinical care as a comparator, with national guidelines as a proxy for current care.[1,2] However, this comparison with guidelines is inadequate when clinical practice differs significantly from guidelines and is particularly problematic when clinical practice differs between hospitals.

The National Institute for Health and Care Excellence (NICE) in the UK provides guidance through the 'Guide to the Methods of Technology Appraisal' in evaluating clinical care strategies in terms of their cost-effectiveness.[3] In addition, the NICE Diagnostic Assessment Programme Manual specifically provides guidance in the evaluation of cost-effectiveness of diagnostic tests and technologies.[4] However, both documents pay little attention to clinical practice variation and its consequences when performing relevant cost-effectiveness analyses (CEAs). In practice, most cost-effectiveness studies do not take into account possible causes and consequences of clinical practice variation.[1,2]

There are many reasons why clinical practice guidelines are not used in daily practice.[5] In some cases, hospitals may wilfully deviate from the guidelines if those guidelines are not be viewed by hospitals as valid. Other hospitals may have no choice but to deviate from guidelines if they are simply impossible to implement in their hospital. In other cases, the guidelines might not be followed due to solvable problems with logistics or financing. Ultimately, various local hospital conditions may cause variation in clinical practice between hospitals, which may result in varying costs and health effects (i.e., quality-adjusted life years [QALYs]) between hospitals and consequently important differences in estimated cost-effectiveness.

A few studies have used large databases to investigate practice variation and the impact it has on costs and effects.[1,2] However, their approach is different from ours since we aim to illustrate the importance of investigating possible clinical practice variation and deviation from national guidelines, and the need to perform hospital-level CEAs which incorporate local hospital conditions when important clinical practice variation exists. We specifically focused on diagnostic imaging tests for the assessment of carotid stenosis and criteria for treatment of patients with a recent transient ischemic attack (TIA) or minor ischemic stroke in the Netherlands.

6.2 Methods

National stroke guidelines

After diagnostic evaluation and treatment in the acute phase, patients with a recent TIA or minor ischemic stroke undergo an assessment of carotid stenosis and subsequent treatment as part of the secondary prevention (i.e., to prevent a future stroke). For the assessment of carotid stenosis, Dutch guidelines recommend duplex ultrasonography (DUS) as the initial diagnostic test and computed tomography angiography (CTA) or magnetic resonance angiography (MRA) as a confirmatory test.[6] The criterion for performing a confirmatory test is moderate (50%-69%) carotid stenosis for men (based on the initial diagnostic test) or severe (70%-99%) carotid stenosis for women.[6] No distinction is made in the guidelines between the hospital's emergency unit and outpatient clinic. Furthermore, the Dutch guidelines recommend surgery (i.e., carotid endarterectomy) for patients (men and women) with a severe (70%-99%) carotid stenosis and a TIA or minor ischemic stroke in the past 6 months. In addition, a carotid endarterectomy is advised for men with moderate (50%-69%) carotid stenosis and a TIA or minor ischemic stroke in the past 3 months.[6]

Interviews and questionnaire

Semi-structured interviews were conducted with 16 vascular neurologists in 6 academic and 10 non-academic hospitals throughout the Netherlands. Only one hospital refused to participate (reason: not interested) resulting in a response rate of 94%. We included the hospitals that are participating in the Plaque at Risk (PARISk) cohort study.[7] In addition, we do not claim generalizability of the sample of hospitals that we used in our study, despite including 18% (16/89) of all Dutch hospitals. The interviews were conducted either face-to-face or by telephone, by four different interviewers from universities and various university medical centers. In total, eleven out of sixteen (69%) interviews were conducted face-to-face. All interviews were conducted between May 2012 and January 2013.

During the interviews, we queried vascular neurologists about the type and sequence of diagnostic imaging tests for the assessment of carotid stenosis as used in their hospitals, and criteria for subsequent treatment. In particular, questions were asked about the use of initial and confirmatory diagnostic imaging tests, and criteria for performing confirmatory diagnostic tests. For each hospital, we examined diagnostic practice in both the emergency unit and outpatient clinic because we expected differences in the use of diagnostic tests between these units. In addition, we asked which criteria were used to perform confirmatory diagnostic tests, and which criteria were used to decide for either surgery (i.e., carotid endarterectomy) in combination with medicines and lifestyle modification, or medicines only

(e.g., platelet aggregation inhibitors) and lifestyle modification. We also queried the reasons to deviate from the Dutch guidelines.

A questionnaire, including both open- and multiple-choice questions, was used during the interviews (see Appendix 6.1 for the questionnaire). The Dutch stroke guidelines served as the basis for the questionnaire.[6] The questionnaire was designed in collaboration with a radiologist and vascular neurologist to guarantee the clinical relevance of the questions. Subsequently, the questionnaire was reviewed by two experts (two vascular neurologists from two different hospitals), and adjusted where necessary. Additional questions were included in the questionnaire and several options were added to the multiple-choice questions on the use of initial and confirmatory tests.

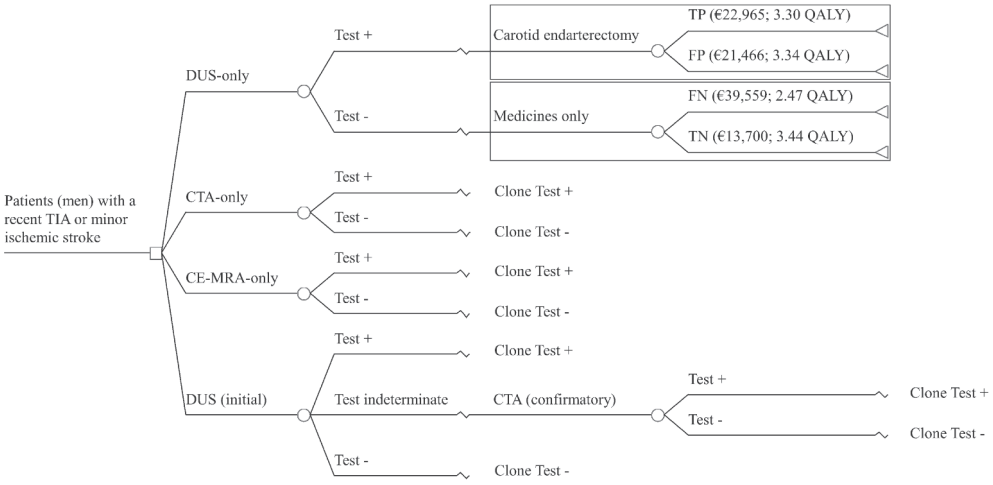
Analysis

We calculated the percentage of hospitals using the recommended test combination (both initial and confirmatory test), the percentage of hospitals complying with the guidelines regarding criteria for use of a confirmatory test, and the percentage of hospitals complying with recommended criteria for treatment. Hospitals were categorized as compliant with the recommended test combination when an initial DUS and confirmatory CTA or MRA (e.g., time-of-flight-MRA [TOF-MRA] or contrast-enhanced-MRA [CE-MRA]) is used. Hospitals were categorized as non-compliant to the guidelines if they use extra criteria (besides degree of carotid stenosis and gender) regarding the use of a confirmatory test or extra criteria for treatment (besides degree of carotid stenosis, gender, and time since TIA or minor ischemic stroke onset).

Case study of hospital-level CEAs

A case study was used to illustrate the value of performing hospital-level CEAs when important clinical practice variation exists or when clinical practice differs from guidelines. A five-year decision analytic model for men with a recent TIA or minor ischemic stroke was used, which incorporated four diagnostic strategies (see Figure 6.1). We assumed that patients who tested positive (i.e., patients with a high risk of a recurrent stroke) underwent a carotid endarterectomy while others received medicines only. Based on the performance (i.e., sensitivity and specificity) of the tests, patients were classified into four groups: true positive (TP), false positive (FP), false negative (FN), and true negative (TN). Final health outcomes were dependent on how patients were classified by the tests and the treatment that followed. The final health outcomes consisted of a minor, major, fatal (i.e., death), or no ischemic stroke event. Death from other causes was incorporated in the model by using the life expectancy from the Dutch population.[8]

Figure 6.1: Cost-effectiveness model



An indeterminate test for men means a moderate (50-69%) carotid stenosis found in men with an initial DUS and a TIA or minor ischemic stroke in the past 3 months. DUS = duplex ultrasonography, CTA = computed tomography angiography, CE-MRA = contrast-enhanced-magnetic resonance angiography, TP = true positive, FP = false positive, FN = false negative, TN = true negative.

First, a base case CEA was performed in which the unit costs of diagnostic tests were based on the national unit costs from the Dutch Healthcare Authority for 2012[9] and the average performance of the tests (see Table 6.1). The national unit costs represent a national average based on all Dutch hospitals. Second, two hospital-level CEAs were performed, which incorporated hospital-specific unit costs from 2012 and performance of diagnostic tests from two hospitals (see Table 6.1). Since the case study in this paper is an illustration of the association between practice variation and the cost-effectiveness of different test strategies, we chose to include two (out of the 16) hospitals that use two very different test strategies: one hospital that uses the test strategy recommended in the guidelines (initial DUS and confirmatory CTA) and one hospital that does not (initial CTA and no confirmatory test). The actual diagnostic strategies from each of the two hospitals, the guideline-based strategy and other strategies found in clinical practice were compared with each other in the hospital-level CEAs to investigate the most cost-effective strategy for each hospital.

The sensitivity and specificity of each test in the hospital-level CEAs were adjusted based on self-reported clinician expertise in performing certain tests using the limits of 95% confidence intervals as reported in the literature.[10,11] Since clinicians from hospital 1 reported having great expertise in performing a CTA and an average expertise in performing a DUS and

CE-MRA, we assumed that the sensitivity and specificity of a CTA performed in that hospital would be higher than average (and set their values at the upper limit of the 95% confidence interval) and an average performance of DUS and CE-MRA. In contrast, clinicians from hospital 2 reported having great expertise in performing DUS and CE-MRA, but low expertise in performing CTA. We therefore assumed higher values of sensitivity and specificity of DUS and CE-MRA (and used the upper limit of the 95% confidence intervals for DUS and CE-MRA) and a lower performance of CTA than average in hospital 2 (and used the lower limit of the 95% confidence interval for CTA).

Table 6.1: Input parameters of the model

Parameter		Base case	Hospital 1	Hospital 2	Source
PERFORMANCE OF TESTS					
DUS	- Sensitivity	89%	89%	92%	[10]
	- Specificity	84%	84%	89%	
CTA	- Sensitivity	91%	99%	71%	[10]
	- Specificity	99%	100%	98%	
CE-MRA	- Sensitivity	94%	94%	97%	[11]
	- Specificity	93%	93%	96%	
COSTS					
DUS		€63	€78	€60	Base case [9]; Hospital 1 and 2: internal unit costs
CTA		€209	€138	€167	Base case [9]; Hospital 1 and 2: internal unit costs
CE-MRA		€244	€244	€161	Base case [9]; Hospital 1 and 2: internal unit costs
Carotid endarterectomy		€6,836	€6,836	€6,836	[12]

DUS = duplex ultrasonography, CTA = computed tomography angiography, CE-MRA = contrast-enhanced-magnetic resonance angiography.

The total costs per patient consisted of the costs of diagnosis, treatment (i.e., carotid endarterectomy and medicines), and stroke-related societal costs (which comprise both healthcare and non-healthcare costs). The average diagnostic test costs per patient were calculated by combining the frequency of tests used in the assessment with their unit costs (see Figure 6.1). Figure 6.1 (upper right-hand corner) shows the five-year average treatment and stroke-related societal costs, and total QALYs for each category of patients (i.e., TP, FP, FN, and TN). The treatment costs of TP and FP patients

included the costs of a carotid endarterectomy, which were based on a recent cost analysis,[12] and the costs of medicines for each category of patients were based on expert opinion. The utility values and cost input parameters used in the model can be found in Appendix 6.2 along with their sources. The average treatment and stroke-related healthcare costs, and QALYs per patient were dependent on the sensitivity and specificity of the diagnostic tests. For example, a higher sensitivity results in a higher TP rate with more patients correctly identified for carotid endarterectomy. Likewise, a higher specificity results in a higher TN rate, which means that the correctly speified patients were prevented from unnecessary carotid endarterectomies resulting in lower costs and higher QALYs. All costs were calculated in 2012 Euros using a societal perspective. Differential discounting was applied in accordance with the Dutch guidelines, with an annual discount rate of 4.0% for all costs and 1.5% for health effects.[13]

6.3 Results

Tables 6.2 and 6.3 show the use of initial and confirmatory imaging tests at the emergency units and outpatient clinics, respectively. These tables also show the degree of compliance to the guidelines regarding use of initial and confirmatory tests and criteria for use of a confirmatory test.

Table 6.2 shows that 56% (9/16) of the hospitals' emergency units use the test combinations in accordance with the Dutch guidelines; the other seven hospitals use various other test combinations, with an initial CTA and confirmatory DUS as the most common combination.

Table 6.3 shows that 63% (10/16) of the hospitals' outpatient clinics use the test combinations as advised in the Dutch guidelines; the other six hospitals use various other test combinations. In addition, Tables 6.2 and 6.3 show that only one hospital uses the criteria for use of a confirmatory test according to the guidelines. In contrast, the other hospitals use broader criteria regarding the degree of carotid stenosis and/or other criteria (e.g., age and plaque characteristics). For example, some hospitals use a confirmatory test for men with a 50%-69% carotid stenosis or a confirmatory test for women with a >50% carotid stenosis.

Criteria for treatment

We found that 38% (6/16) of vascular neurologists strictly use the criteria for treatment as advised in the guidelines (i.e., degree of carotid stenosis, gender, and time since TIA or minor ischemic stroke onset) in their decision for either surgery or medicines only. The other vascular neurologists use additional criteria in their decision-making about surgery. Patient age and life expectancy play a role in 44% (7/16) of the hospitals. Other factors that influence decision-making are co-morbidity, risk of surgery, patient prefer-

Table 6.2: Clinical practice variation in use of diagnostic tests in the emergency unit

Number of hospitals (number of academic hospitals)	Initial test	Confirmatory test(s)	Compliance to guidelines regarding use of initial and confirmatory test?	Compliance to guidelines regarding criteria for use of a confirmatory test?
Dutch guidelines	*DUS*	*CTA or MRA*		*If carotid stenosis is >70% for women or 50-69% for men*
6 (1)	DUS	CTA	Yes	No
2 (0)	DUS	TOF-MRA	Yes	Yes (one hospital) No (one hospital)
1 (0)	DUS	CE-MRA	Yes	No
1 (1)	DUS or CTA	DUS or CTA*	No	No
1 (0)	DUS	DUS and CTA**	No	No
3 (3)	CTA	DUS	No	No
1 (1)	CTA	None	No	No
1 (0)	DUS or CTA	DUS***	No	No

DUS = duplex ultrasonography, CTA = computed tomography angiography, MRA = magnetic resonance angiography, CE-MRA = contrast-enhanced-MRA, TOF-MRA = time-of-flight-MRA.
* If a DUS is used as initial test, a CTA is used as confirmatory test. If a CTA is used as initial test, a DUS is used as confirmatory test.
** DUS is used as confirmatory test, even if an initial DUS is performed. CTA is used when the results of the initial DUS and confirmatory DUS differ.
*** DUS is used as confirmatory test, even if an initial DUS is performed. CTA is used when patients were included in a particular clinical study.

Table 6.3: Clinical practice variation in use of diagnostic tests in the outpatient clinic

Number of hospitals (number of academic hospitals)	Initial test	Confirmatory test(s)	Compliance to guidelines regarding use of initial and confirmatory test?	Compliance to guidelines regarding criteria for use of a confirmatory test?
Dutch guidelines	*DUS*	*CTA or MRA*		*If carotid stenosis is >70% for women or 50-69% for men*
7 (2)	DUS	CTA	Yes	No
2 (0)	DUS	TOF-MRA	Yes	Yes (one hospital) No (one hospital)
1 (0)	DUS	CE-MRA	Yes	No
1 (1)	DUS or CTA	DUS or CTA*	No	No
1 (0)	DUS	DUS and CTA**	No	No
1 (0)	DUS	None	No	No
1 (1)	DUS and CE-MRA or DUS and CTA***	None	No	No
1 (1)	CTA	None	No	No
1 (1)	CE-MRA	DUS	No	No

DUS = duplex ultrasonography, CTA = computed tomography angiography, MRA = magnetic resonance angiography, CE-MRA = contrast-enhanced-MRA, TOF-MRA = time-of-flight-MRA.
* If a DUS is used as initial test, a CTA is used as confirmatory test. If a CTA is used as initial test, a DUS is used as confirmatory test.
** DUS is used as confirmatory test, even if an initial DUS is performed. CTA is used when the results of the initial DUS and confirmatory DUS differ.
*** Choice of DUS and CE-MRA or DUS and CTA is based on logistical reasons.

ences, duration of complaints, severity of the stroke, and plaque characteristics.

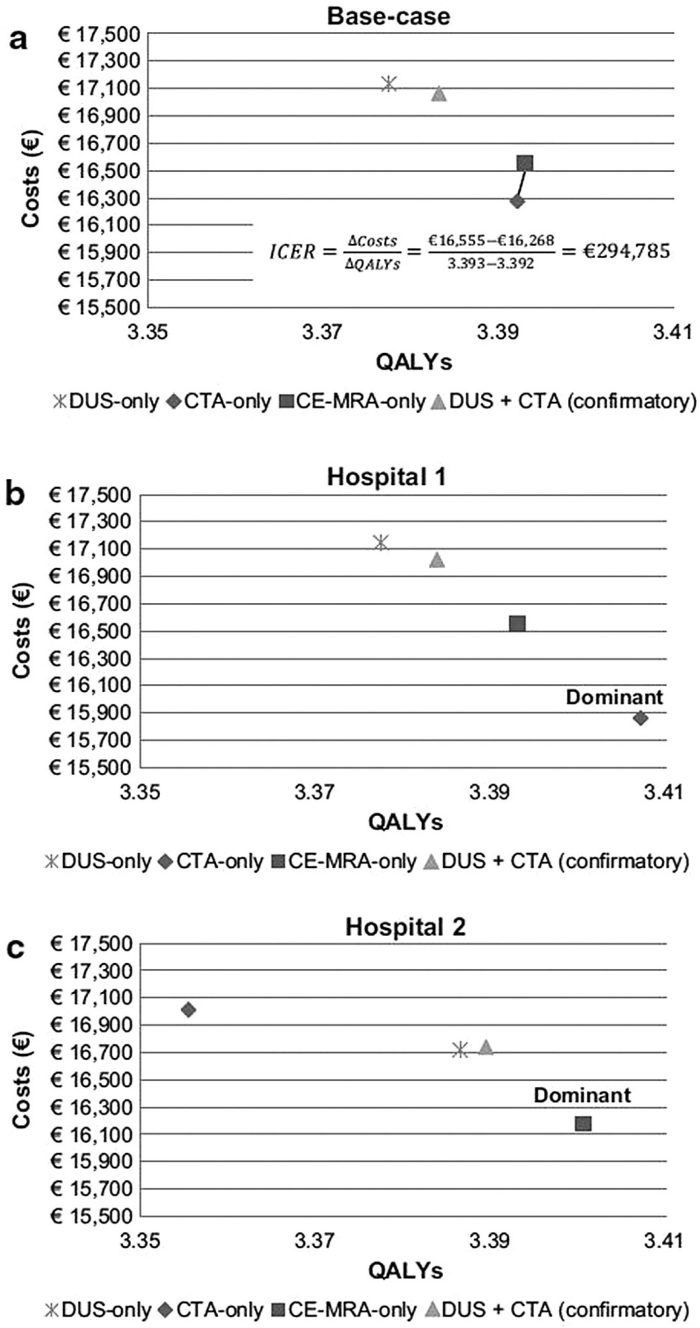
Reasons to deviate from guidelines regarding use of initial and confirmatory test

According to vascular neurologists, clinical practice variation in the choice of initial and confirmatory tests arises for different reasons, including varying degrees of expertise in performing diagnostic tests, patient case-mix, clinical reasons, financial incentives, logistics, availability of imaging technology, and preferences of radiologists, vascular surgeons and vascular neurologists. For example, one hospital uses a CTA as the initial test due to high expertise in CTA, even though the guidelines recommend CTA-only as a confirmatory test. Another hospital uses an initial CTA instead of the DUS recommended in the guidelines because CTA is more accurate than DUS. In addition, 25% (4/16) of hospitals use a more costly confirmatory test (i.e., CTA or MRA), even though the guidelines indicate that a DUS is sufficient. Examples of clinical reasons for deviating from the guidelines were that MRA could not be used with patients with a pacemaker and that CE-MRA could not be used for patients with kidney problems or allergy due to the contrast liquid.

Case study of hospital-level CEAs

Figure 6.2 shows the results of the base case analysis and two hospital-level CEAs. For each test strategy, this figure shows the average costs and health effects in QALYs per patient. It also shows that the most cost-effective strategy differs between the base case analysis and the two hospital-level CEAs. In the base case analysis, the CTA-only and CE-MRA-only strategies were the most cost-effective strategies (see Figure 6.2a). When comparing these two strategies, the CE-MRA-only strategy leads to slightly more QALYs (0.001) and higher costs (€287) versus the CTA-only strategy, resulting in an incremental cost-effectiveness ratio (ICER) of €294,785 per QALY gained. The high ICER means that the CE-MRA-only strategy is not cost-effective versus the CTA-only strategy; therefore, CTA-only is the preferred strategy in the base case analysis. However, the results changed when hospital-level values for unit costs and test performance were used. To start with, the CTA-only strategy was the dominant strategy in the first hospital, since it had the lowest costs and highest QALYs of all strategies (see Figure 6.2b). In contrast, the CE-MRA-only strategy was the dominant one in the second hospital (see Figure 6.2c).

Figure 6.2: Hospital-level cost-effectiveness results



Guideline-based strategy = DUS+CTA (confirmatory), Hospital 1 currently uses CTA-only strategy, Hospital 2 currently uses DUS+CTA (confirmatory) strategy. DUS = duplex ultrasonography, CTA = computed tomography angiography, CE-MRA = contrast-enhanced-magnetic resonance angiography, QALY = quality-adjusted life year, ICER = incremental cost-effectiveness ratio.

6.4 Conclusions and discussion

CEAs using guidelines as a comparator are inadequate if the guidelines are not used in clinical practice. The existence of important clinical practice variation and the resulting cost differences support the need to perform hospital-level CEAs which incorporate local hospital conditions (e.g., patient case-mix, costs, availability of facilities, and expertise). A hospital-level CEA could examine the cost-effectiveness of the hospital's current care strategy versus the strategies used in other hospitals as well as strategies incorporating new tests or treatments. This will result in multiple hospital-level ICERs that will help individual hospitals to explore the potential to improve effectiveness and cost-effectiveness by implementing a different strategy. One possible rule to decide if hospital-level CEAs should be performed is to compare which tests or treatments are performed, while another would be to see if the different strategies used in current care have different short term costs and effectiveness (i.e., when the costs and effectiveness of the different current care strategies are similar, then it would be irrelevant to use multiple comparators). One extreme solution would be to model the long term impact on costs and effectiveness before concluding whether the observed practice variation is actually important. Lastly, hospital-level CEAs may be of interest to other parties than just individual hospitals. For example, health insurers might want to use hospital-level CEAs to determine how costly and cost-effective the care currently provided in hospitals is compared to the most cost-effective strategy available.

The observed variation in the use of diagnostic tests for patients with a recent TIA or minor ischemic stroke means that the most cost-effective diagnostic strategy may differ between hospitals, as illustrated in our case study. In the first hospital-level CEA, the average five-year costs per patient range from €15,862 to €17,145 between strategies. While this range may seem small, its effect on budget may be important depending on the annual volume of patients.[14] For example, if this hospital were to assess 500 patients per year (i.e., the average number of patients with a TIA or ischemic stroke per hospital in the Netherlands in 2012),[8,15] the total five-year costs would range from €7,930,779 to €8,572,496, meaning a difference of €641,717.

There are several ways in which practice variation may result in differences in overall costs and health effects.[16,17] First of all, a cost difference arises from using different diagnostic tests. The use of different strategies may also result in short-term differences in health effects simply due to differences in complication risks or patient discomfort (e.g., a more invasive test leads to more discomfort for the patient). Moreover, the long term differences in costs and health effects between the diagnostic strategies are caused by differences in sensitivity and specificity of the diagnostic tests

used. For example, if patients are more often misclassified (i.e., FPs or FNs) by one test than by another test, this may lead to greater long term costs and less health. Even if hospitals use the very same diagnostic test, this too may result in different long term costs and health effects if they use the test results differently when making treatment decisions or when the test's diagnostic accuracy differs between hospitals (as was illustrated in our case study).

If hospitals perform their own hospital-level CEAs, they should consider the feasibility of the different strategies being considered as well as incorporate the costs of switching to a different strategy in the analysis. This holds true for strategies involving new interventions (such as new diagnostic tests, medicines or therapies) as well as existing ones. Switching to a different strategy is only worth considering when local hospital conditions can readily be modified (e.g., training clinicians and other healthcare personnel to use a more advanced imaging test).[18,19]

If clinicians lack the necessary expertise to perform the most cost-effective strategy, the hospital-level CEA must include the extra costs of training. In addition, capacity and logistical problems may arise because of limited availability of tests. For example, some small hospitals have DUS available, but not more expensive scanners like CTA or MRA. These hospitals currently refer patients to larger hospitals, resulting in higher costs (e.g., repetition of tests and travel costs) and a delay in decision-making. A CEA for such small hospitals should include the costs of purchasing an imaging test and possible training costs of personnel. Viewed in that way, the current strategy may be more cost-effective due to the relatively high implementation costs.

Hospitals should have a sufficient understanding of clinician attitudes when designing an implementation strategy, since attitudes influence behavior and the choice of strategy used in clinical practice.[18] Innovation managers or other professionals may assist in the implementation process, for example by showing clinicians convincing evidence about the improved effectiveness and/or cost-effectiveness of a new strategy versus the existing one.[19]

The aim of our paper was to present a major methodological issue that seems to be underestimated by many. We have illustrated the importance of investigating possible practice variation and deviation from the clinical guidelines, and have demonstrated the value of performing hospital-level CEAs based on local hospital conditions (e.g., unit costs) when important practice variation exists. The general principles of performing hospital-level CEAs are valid and should be applied when clinical practice differs between hospitals or when clinical practice differs significantly from guidelines, irrespective of the disease area or countries under study. Further research is recommended that applies the presented general principles of performing

hospital-level CEAs to other disease areas, types of care, and countries. We do not claim generalizability of the results from our case study based on two hospitals, to all Dutch hospitals. The aim of the case study was merely to illustrate the importance of performing hospital-level CEAs.

One limitation of performing hospital-level CEAs may be the feasibility. However, the additional data needed to perform hospital-level CEAs (when practice variation is found) are the hospital-specific costs of the tests and the hospital-specific sensitivity and specificity of tests. We recommend hospital-level CEAs if data for potentially influential variables (like the hospital-level costs and performance of tests in this study) can be retrieved or sufficiently estimated. Moreover, there is a good chance that hospitals unable to retrieve data on potentially influential parameters are not functioning very efficiently compared to those that are. The results of a CEA based on a hospital's own data can help a hospital to perform more efficiently; in the example described in this paper, this would involve comparing various test strategies and selecting the one that is most cost-effective for that particular hospital. This approach may demonstrate that one strategy is the most cost-effective in one hospital, while another strategy is the most cost-effective in another hospital. In this sense, the ultimate choice of a hospital may differ from national guidelines for justifiable reasons.

In conclusion, consideration of clinical practice variation and deviation from the clinical guidelines should be one of the first steps in any CEA. If important practice variation or deviation from the guidelines exists, hospital-level CEAs should be performed which compare the care that is actually provided in hospitals. Moreover, a hospital-level CEA should consider the causes of variation, since they will affect the feasibility and costs of implementing a new strategy.

6.5 Acknowledgments

This research was performed within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), project PARISk (grant 01C-202), and supported by the Dutch Heart Foundation. We also wish to thank Anouk van Dijk, Alexandra de Rotte, and Martine Truijman for their assistance in data collection, and Peter Koudstaal, Robert van Oostenbrugge, and Manuela Joore for their useful comments.

6.6 References

1 Karnon J, Caffrey O, Pham C, et al. Applying risk adjusted cost-effectiveness (RAC-E) analysis to hospitals: estimating the costs and consequences of variation in clinical practice. *Health Econ* 2013;22(6):631-42.

2 Pham C, Caffrey O, Ben-Tovim D, et al. Evaluating the effects of variation in clinical practice: a risk adjusted cost-effectiveness (RAC-E) analysis of acute stroke services. *BMC Health Serv Res* 2012;12:266.

3 National Institute for Health and Care Excellence (NICE). Guide to the Methods of Technology Appraisal 2013. London: National Institute for Health and Care Excellence; 2013.

4 National Institute for Health and Care Excellence (NICE). Diagnostic Assessment Programme Manual. London: National Institute for Health and Care Excellence; 2011.

5 Kennedy PJ, Leathley CM. Hughes CF. Clinical practice variation. *Med J Aust* 2010;193(8):S97-9.

6 Dutch Institute for Healthcare Improvement CBO. Guideline: Diagnostics, treatment and healthcare for patients with a stroke (In Dutch: Richtlijn: Diagnostiek, behandeling en zorg voor patiënten met een beroerte). Utrecht: Dutch Institute for Healthcare Improvement CBO; 2008.

7 Truijman MT, Kooi ME, Van Dijk AC, et al. Plaque At RISK (PARISK): prospective multicenter study to improve diagnosis of high-risk carotid plaques. *Int J Stroke* 2014;9(6):747–54.

8 Statistics Netherlands. Available from: http://statline.cbs.nl/Statweb/?LA=en. Accessed Feb 1, 2014.

9 Dutch Healthcare Authority (in Dutch: Nederlandse Zorgauthoriteit, NZa). NZa Zorgproducten Tariefapplicatie. Available from: http://dbc-zorgproducten-tarieven.nza.nl/nzaZpTarief/ZoekfunctieDot.aspx. Accessed Oct 10, 2014.

10 Tholen ATR, De Monyé C, Genders TSS, et al. Suspected Carotid Artery Stenosis: Cost-effectiveness of CT Angiography in Work-up of Patients with Recent TIA or Minor Ischemic Stroke. *Radiology* 2010;256(2):585-97.

11 Wardlaw JM, Chappell FM, Best JJ, et al. Non-invasive imaging compared with intra-arterial angiography in the diagnosis of symptomatic carotid stenosis: a meta-analysis. *Lancet* 2006;367(9521):1503-12.

12 Buisman LR, Tan SS, Nederkoorn PJ, et al. Hospital costs of ischemic stroke and TIA in the Netherlands. *Neurology* 2015;84(22):2208–15.

13 College voor zorgverzekeringen (CVZ). Guidelines for pharmacoeconomic research, updated version. Diemen: College voor zorgverzekeringen; 2006.

14 Trueman P, Drummond M, Hutton J. Developing guidance for budget impact analysis. *Pharmacoeconomics* 2001;19(6):609–21.

15 Vaartjes I, Bots ML, Poos MJJC. Hoe vaak komt een beroerte voor en hoeveel mensen sterven eraan? In: Volksgezondheid Toekomst Verkenning, Nationaal Kompas Volksgezondheid. Bilthoven: Rijksinstituut voor Volsgezondheid en Mileu (RIVM); 2014.
16 Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991;11(2):88-94.
17 Thornbury JR. Clinical efficacy of diagnostic imaging: love it or leave it. *AJR Am J Roentgenol* 1994;162:1-8.
18 Wensing M, Bosch M, Foy R, et al. Factors in theories on behaviour change to guide implementation and quality improvement in healthcare. Nijmegen: Centre for Quality of Care Research (WOK); 2005.
19 Bayley MT, Hurdowar A, Richards CL, et al. Barriers to implementation of stroke rehabilitation evidence: findings from a multi-site pilot project. *Disabil Rehabil* 2012;34(19):1633-8.

6.7 Appendices

Appendix 6.1: Questionaire

ctmm

ERASMUS UNIVERSITEIT ROTTERDAM
iMTA

QUESTIONNAIRE: PRACTICE VARIATION IN THE USE OF DIAGNOSTIC TESTS AND TREATMENT CRITERIA IN PATIENTS WITH A RECENT TIA / MINOR ISCHEMIC STROKE

Institute for Medical Technology Assessment, Erasmus University Rotterdam, The Netherlands

This research was performed within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), project PARISk (grant 01C-202), and supported by the Dutch Heart Foundation.

Q1: In which hospital are you currently working?……
…….....

Q2: Which initial test is used to assess the degree of carotid stenosis at the emergency unit/inpatient clinic of your hospital in patients with a TIA or minor ischemic stroke in the past 3 months? (more than 1 answer is possible)

- ○ Duplex ultrasonography (DUS)
- ○ Computed tomography angiography (CTA)
- ○ Magnetic resonance angiography - Time of flight (MRA-TOF)
- ○ Contrast-enhanced magnetic resonance angiography (CE-MRA)
- ○ Other test, namely …….

Q3: Which initial test is used to assess the degree of carotid stenosis at the outpatient clinic of your hospital in patients with a TIA or minor ischemic stroke in the past 3 months? (more than 1 answer is possible)

- ○ Duplex ultrasonography (DUS)
- ○ Computed tomography angiography (CTA)
- ○ Magnetic resonance angiography - Time of flight (MRA-TOF)
- ○ Contrast-enhanced magnetic resonance angiography (CE-MRA)
- ○ Other test, namely …….

If a different initial test is used at the emergency unit/inpatient clinic than the one at the outpatient clinic, as answered in question 2 and 3.
----> Go to question 4

If the same initial test is used at the emergency unit/inpatient clinic and outpatient clinic, as answered in question 2 and 3.
----> Go to question 5

Q4: What is the reason why different initial tests are used at the emergency unit/inpatient clinic and outpatient clinic?

……
……
……

Q5: Is a second (confirmatory) test used to assess the degree of carotid stenosis based on the test result of the initial test at the emergency unit/inpatient clinic of your hospital? If yes, which test(s)? (more than 1 test possible)

- ○ No;
- ○ Yes, namely: ○ DUS ○ CTA ○ TOF-MRA ○ CE-MRA ○ DSA

Q6: Is a second (confirmatory) test used to assess the degree of carotid stenosis based on the test result of the initial test at the outpatient clinic of your hospital? If yes, which test(s)? (more than 1 test possible)

- ○ No;
- ○ Yes, namely: ○ DUS ○ CTA ○ TOF-MRA ○ CE-MRA ○ DSA

Q7: Which criteria are used in your hospital to determine if a second (confirmatory) test is needed? (e.g., gender, degree of carotid stenosis found with the initial test, plaque characteristics, etc.)

Please specify the criteria in the table below. If more space is needed, please use the space below the table.

Criteria for confirmatory test					
DUS					
CTA					
TOF-MRA					
CE-MRA					
DSA					

……

……

……

……

……

Q8: The Dutch stroke guidelines recommend a DUS as initial test and a MRA or CTA as confirmatory test for women with a ≥ 70% carotid stenosis and men with a 50-69% carotid stenosis. Does your hospital deviate from the Dutch stroke guidelines?

○ Yes ----> Go to question 9 ○ No ----> Go to question 10

Q9: What is the main reason why your hospital deviates from the Dutch stroke guidelines in the use of diagnostic tests in patients with a recent TIA or minor ischemic stroke? (e.g., lack of expertise in certain test, capacity problems, lack of evidence, etc.)

……

……

……

Q10: Which criteria are used in your hospital to decide which patients should undergo a carotid endarterectomy (CEA)? (e.g., patient characteristics (age and gender), degree of stenosis, time interval between TIA/minor ischemic stroke and CEA, etc.)

……

……

……

Q11: Besides degree of carotid stenosis, are plaque characteristics used in decisions about which patients should undergo a carotid endarterectomy (CEA)? If yes, which plaque characteristics and which diagnostic test are used for this purpose? (e.g., plaque morphology (ulcerations), plaque compositions (calcifications, CT density, echo density).

……

……

……

Appendix 6.2: Utility values and cost input parameters of the model

Parameter	Value	Source
UTILITIES		
Baseline (recent TIA or minor ischemic stroke)	0.71	[1]
After minor ischemic stroke	0.66	[1,2]
After major ischemic stroke	0.31	[1]
COSTS		
Carotid endarterectomy	€6,836	[3]
Medicines (per year)	€120	Expert opinion
Minor ischemic stroke		[4]
first year	€8,433	
subsequent years (per year)	€1,443	
Major ischemic stroke		[4]
first year	€48,095	
subsequent years (per year)	€28,082	
Fatal ischemic stroke (in year of fatal ischemic stroke)		[5]
Men – aged <65	€6,436	
Men – aged 65-74	€10,505	
Men – aged 75-85	€8,438	
Men – aged >85	€10,278	
Women – aged <65	€7,103	
Women – aged 65-74	€9,644	
Women – aged 75-85	€11,632	
Women – aged >85	€13,992	

References Appendix 6.2

1 Hallan S, Asberg A, Indredavik B, Widerøe TE. Quality of life after cerebrovascular stroke: a systematic study of patients' preferences for different functional outcomes. *J Intern Med* 1999;246(3):309–16.

2 Sullivan PW, Ghushchyan V. Mapping the EQ-5D index from the SF-12: US general population preferences in a nationally representative sample. *Med Decis Making* 2006;26(4):401–9.

3 Buisman LR, Tan SS, Nederkoorn PJ, Koudstaal PJ, Redekop WK. Hospital costs of ischemic stroke and TIA in the Netherlands. *Neurology* 2015;84(22):2208–15.

4 Buskens E, Nederkoorn PJ, Buijs-Van Der Woude T, et al. Imaging of carotid arteries in symptomatic patients: cost-effectiveness of diagnostic strategies. *Radiology* 2004;233(1):101–12.

5 Struijs JN, Van Genugten ML, Evers SM, et al. Future costs of stroke in the Netherlands: the impact of stroke services. *Int J Technol Assess Health Care* 2006;22(4):518–24.

CHAPTER 7

HOSPITAL COSTS OF ISCHEMIC STROKE AND TIA IN THE NETHERLANDS

Leander R. Buisman[1,2], Siok Swan Tan[1,2], Paul J. Nederkoorn[3], Peter J. Koudstaal[4], William K. Redekop[1,2]

[1] Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands
[2] Institute for Medical Technology Assessment, Erasmus University Rotterdam, Rotterdam, The Netherlands
[3] Department of Neurology, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
[4] Department of Neurology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

Published in *Neurology* 2015;84(22):2208–15. doi:10.1212/WNL.0000000000001635.

ABSTRACT

Objectives: There have been no ischemic stroke costing studies since major improvements were implemented in stroke care. We therefore determined hospital resource use and costs of ischemic stroke and TIA in the Netherlands for 2012.

Methods: We conducted a retrospective cost analysis using individual patient data from a national Diagnosis Related Group registry. We analyzed four subgroups: inpatient ischemic stroke, inpatient TIA, outpatient ischemic stroke, and outpatient TIA. Costs of carotid endarterectomy and costs of an extra follow-up visit were also estimated. Unit costs were based on reference prices from the Dutch Healthcare Insurance Board and tariffs provided by the Dutch Healthcare Authority. Linear regression analysis was used to examine the association between hospital costs and various patient and hospital characteristics.

Results: A total of 35,903 ischemic stroke and 21,653 TIA patients were included. Inpatient costs of ischemic stroke were €5,328 ($6,845) and for TIA €2,470 ($3,173). Outpatient costs were €495 ($636) for ischemic stroke and €587 ($754) for TIA. Costs of carotid endarterectomy were €6,836 ($8,783). Costs of inpatient days were the largest contributor to hospital costs. Age, hospital type, and region were strongly associated with hospital costs.

Conclusions: Hospital costs are higher for inpatients and ischemic strokes compared with outpatients and TIAs, with length of stay (LOS) as most important contributor. LOS and hospital costs have substantially reduced over the past 10 years, possibly due to improved hospital stroke care and efficient integrated stroke services.

Keywords: ischemic stroke, TIA, cost analysis, diagnosis treatment combination, Diagnosis Related Group (DRG), hospital costs, length of stay, The Netherlands.

7.1 Introduction

Ischemic stroke is the second most common cause of death and a major cause of disability resulting in a huge burden on current healthcare and societal costs.[1] Advances in prevention and treatment of ischemic stroke and TIA during the past two decades have resulted in significant changes of medical practice, and therefore in improved clinical outcomes and potentially lower costs.[1] A significant improvement was the organization of stroke care into integrated stroke services which reduced both mortality and disability.[2] Furthermore, use of intravenous thrombolysis improved clinical outcome,[3] and advances in secondary prevention, such as carotid endarterectomy, antiplatelet agents, blood pressure lowering, and cholesterol reduction, have proven to be effective.[4-7]

In the past, several studies have determined the costs of ischemic stroke both in the Netherlands[8-10] and other developed countries.[11-14] However, these studies used resource use data from more than 10 years ago before the major improvements were implemented. At that time, treatment of stroke was far from optimal. On average, 10 of 28 inpatient days were explained by non-medical reasons: most frequently waiting to be discharged to a nursing home.[15] One would expect that the improvements in patient management reduced the length of stay (LOS) and therefore hospital costs of ischemic stroke. Therefore, the aim of this study was to determine hospital resource use and costs of ischemic stroke and TIA in the Netherlands for 2012. The secondary aim was to examine the association between hospital costs of ischemic stroke and TIA and various patient and hospital characteristics.

7.2 Methods

Patient population

The patient population included patients with a recent ischemic stroke or TIA. Ischemic stroke was defined as a focal neurological deficit of sudden onset of presumed vascular origin, lasting at least 24 hours, with brain imaging showing typical signs of brain infarction or no abnormalities. The diagnosis of TIA was defined as a focal neurologic deficit of sudden onset lasting less than 24 hours and with no signs of recent infarction on CT scan.

Four subgroups were analyzed: inpatient ischemic stroke, inpatient TIA, outpatient ischemic stroke, and outpatient TIA. Additionally, we estimated the costs of carotid endarterectomy, and costs of an extra follow-up visit necessary because of: (i) direct consequences of the ischemic event, such as cognitive dysfunction, inability to return to work, anxiety, headache, and epilepsy, and (ii) management of vascular risk factors, such as uncontrolled hypertension, hypercholesterolemia, and obesity. The costs of carotid en-

darterectomy and the costs of an extra follow-up visit are costs in addition to those of one of the four subgroups when a patient underwent carotid endarterectomy or had an extra follow-up visit, respectively.

Data

We performed a retrospective cost analysis using patient and resource use data from 2010 from the Diagnosis Treatment Combination (in Dutch: 'Diagnose Behandeling Combinatie', DBC) case mix system, a Diagnosis Related Group-like system which contains resource use data of all hospitalizations in the Netherlands. The length of follow-up is patient-specific for both inpatients and outpatients. It starts on the day of first hospital admission after a recent ischemic stroke or TIA to a maximum of 42 days for inpatients and 90 days for outpatients following the starting date. The length of follow-up included acute hospitalization, re-admission for the same ischemic event, hospital allied health services, and hospital outpatient visits if it occurred in the same hospital. A new follow-up was started and a new DBC was opened, if a patient was seen after the maximum length of follow-up, at another hospital, or for a new ischemic event.

Hospital costs were determined by multiplying the resource use by the corresponding unit costs for 2012. Resource use focused on the hospital only and included: inpatient days, intensive care days, outpatient and emergency room visits, ambulatory treatment, diagnostic and imaging tests, allied health services, and laboratory investigations (i.e., laboratory, microbiology, pathology, and blood products). The unit costs of inpatient days, intensive care days, outpatient visits, emergency room visits, ambulatory treatment, and allied health services were based on reference prices of the Dutch Manual of Costing.[16] The Dutch Manual of Costing provides a list of references prices for different types of care to facilitate in the standardization and comparability of economic evaluations. A reference price is an average unit cost estimated on the basis of large, diverse populations that can be directly used to value resource quantities. The unit costs of diagnostic and imaging tests, and laboratory investigations were based on tariffs provided by the Dutch Healthcare Authority.[17] The unit costs of surgery were estimated based on internal cost prices from a Dutch academic hospital. All unit costs reflected full hospitals costs, including material, labor, drug, overhead, and capital costs. Where necessary, unit costs were adjusted to 2012 Euros using the general price index from the Dutch Central Bureau of Statistics.[18] We have converted Euros to US Dollars and added that parenthetically. A mean exchange rate of €1=$1.2848 was used for 2012, because the daily exchange rates ranged from €1=$1.2089 to €1=$1.3454 in 2012.[19]

Regression analysis

Three ordinary least squares regression analyses were performed for three patient groups to examine the association between hospital costs and various patient and hospital characteristics: inpatient ischemic stroke and TIA (Analysis 1), outpatient ischemic stroke and TIA (Analysis 2), and patients who underwent carotid endarterectomy (Analysis 3). In each analysis, hospital costs were the dependent variable. The independent variables were various patient and hospital characteristics. The patient variables were age, gender, and the interaction between age and gender. Age was estimated using the midpoint of each 5-year category (i.e., 63 year was used if the 5-year category was 61-65 year). The hospital characteristics included hospital type (academic and different type of non-academic hospitals) and region in the Netherlands. The geographical regions used were the twelve Dutch provinces. In demographic and economic terms, these provinces can be classified into two areas: a metropolitan area including the four largest Dutch cities (in the provinces North Holland, South Holland, Utrecht, and Flevoland), and the rest of the Netherlands. The metropolitan area has developed in an advanced economy in many leading sectors, has the lowest unemployment rates in the OECD, and is one of the most densely populated areas in the OECD.[20] Statistical inference of associations was calculated using the two-tailed t-test.

7.3 Results

Descriptive statistics

Table 7.1 shows the patient and hospital characteristics of the four patient subgroups: inpatient ischemic stroke, inpatient TIA, outpatient ischemic stroke, and outpatient TIA. 24,671 inpatient ischemic stroke and 4,596 inpatient TIA patients were admitted to any Dutch hospital in 2010 (see Table 7.1). In the outpatient setting, 11,232 ischemic stroke and 17,057 TIA patients were admitted to any Dutch hospital in 2010.

Moreover, Table 7.1 shows that ischemic stroke patients are on average older than TIA patients (inpatient: 71±14 (mean±standard deviation) versus 69±13, P<0.0001; outpatient: 69±13 versus 68±13, P<0.0001) and that inpatient TIA and outpatient ischemic stroke had a slightly higher proportion of men. We observed no large differences between the four patient subgroups in the proportion of patients admitted per hospital type and region. Table 7.1 shows that most patients were admitted to small non-academic hospitals, followed by teaching and large non-academic hospitals. Most patients were admitted in the most densely populated provinces (South Holland, North Holland and North Brabant).

In addition, 3.3% (1,891/57,556) of the ischemic stroke and TIA patients admitted to a hospital in 2010 underwent carotid endarterectomy. Of

the inpatients alone, 6.5% (1,891/29,267) underwent carotid endarterectomy, the mean age was 70±10 and about two-thirds were men. Furthermore, 65% (1,231/1,891) of carotid endarterectomies were performed in academic, teaching non-academic, or large non-academic hospitals.

In total, 12% (6,868/57,556) of the patients who visited the hospital in 2010 (either as an inpatient or outpatient) had an extra follow-up visit. Of these patients, 92% (6,303/6,828) visited the outpatient clinic. The mean age of outpatients with an extra follow-up visit was 64±14 versus 67±14 years for inpatients (P<0.0001) and in both settings 54% were men.

Table 7.1: Patient and hospital characteristics of four patient subgroups

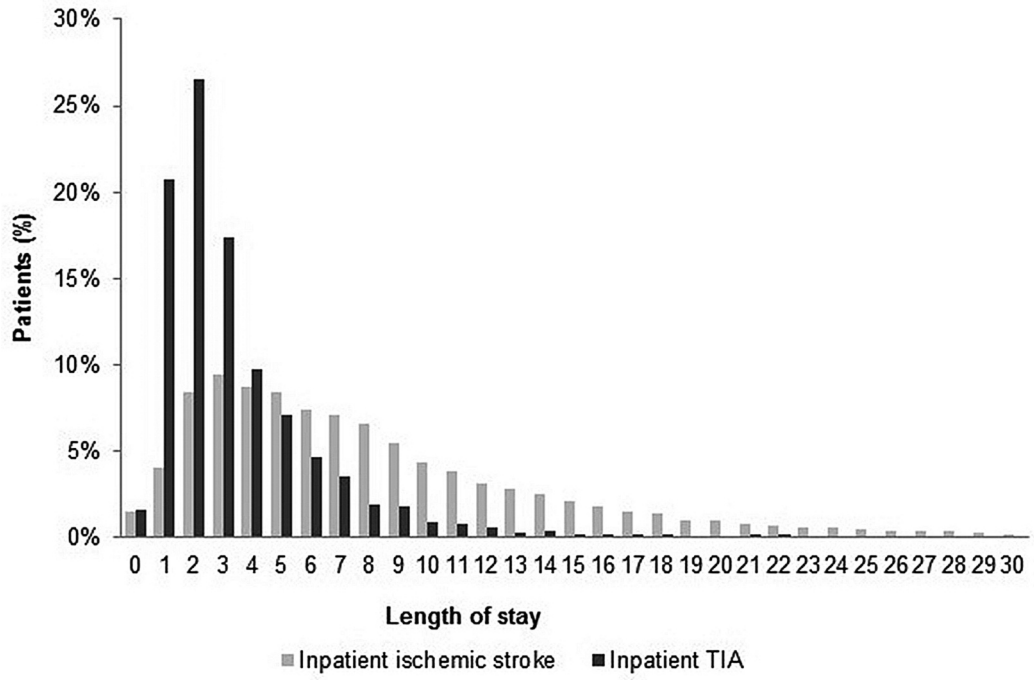
	Inpatient ischemic stroke	Inpatient TIA	Outpatient ischemic stroke	Outpatient TIA
Number of patients	24,671	4,596	11,232	17,057
AGE (mean±SD)	71±14	69±13	69±13	68±13
Men (%)	50%	53%	53%	50%
Hospital type (%)				
Academic	9%	7%	9%	8%
Non-academic (teaching)	28%	26%	25%	26%
Non-academic (large)	20%	25%	20%	21%
Non-academic (midsized)	14%	11%	12%	11%
Non-academic (small)	25%	28%	29%	28%
Non-academic (type unknown)	4%	3%	5%	6%
REGION (%)				
Drenthe	3%	3%	4%	4%
Flevoland	2%	2%	2%	2%
Friesland	3%	4%	4%	3%
Gelderland	13%	12%	11%	13%
Groningen	4%	3%	4%	3%
Limburg	8%	7%	10%	10%
North Brabant	15%	14%	19%	19%
North Holland	15%	15%	16%	13%
Overijssel	6%	5%	4%	6%
South Holland	22%	21%	18%	21%
Utrecht	6%	12%	6%	6%
Zeeland	3%	4%	2%	2%

SD = standard deviation.

Length of stay

Figure 7.1 shows the distribution of LOS for inpatient ischemic stroke and inpatient TIA patients up to 30 days; 2.1% of inpatient ischemic strokes and 0.2% of inpatient TIAs have more than 30 inpatient days. The average LOS of inpatient TIA patients was shorter than that of inpatient ischemic stroke patients (3.6 versus 8.8 days, P<0.0001), which is mainly caused by the higher proportion of TIA patients with 1 to 4 inpatient days and lower proportion of longer stays than ischemic stroke patients. About 1.5% of inpatient ischemic stroke and inpatient TIA patients had zero inpatient days, probably because they died within 24 hours after hospital admission.

Figure 7.1: Distribution of length of stay



Inpatient ischemic stroke n = 24,671; Inpatient TIA n = 4,596.

Costs of ischemic stroke and TIA

Table 7.2 presents the unit costs of the different cost items, breakdown of average resource use, and average hospital costs per subgroup. Average hospital costs were €5,328±€4,346 ($6,845±$5,584) for inpatient ischemic stroke, €2,470±€1,939 ($3,173±$2,491) for inpatient TIA, €495±€333 ($636±$428) for outpatient ischemic stroke, and €587±€303 ($754±$389) for outpatient TIA.

Costs of inpatient days were the largest contributor to costs of both inpatient ischemic stroke (80%) and inpatient TIA (71%). Diagnostic and imaging tests were the second largest contributor to costs of both inpatient groups. However, the contribution was substantially larger for TIA (17%) than for ischemic stroke (9%). The share of visits (i.e., outpatient, emergency room, and ambulatory treatment) was the third largest contributor for both inpatient groups.

A different pattern was observed in the contribution of outpatient costs. Costs of diagnostic and imaging tests were the largest contributor for both ischemic stroke (61%) and TIA (57%). Visits were the second largest contributor to outpatient costs (ischemic stroke: 35%, TIA: 41%).

Differences in costs between inpatients and outpatients were mainly caused by this difference in LOS. In addition to costs related to LOS, other contributors to higher costs for inpatients included more frequent use of laboratory and allied health services.

Table 7.2: Unit costs, average resource use, and average costs of inpatient and outpatient ischemic stroke and TIA

	Unit costs	Average resource use per patient				Average costs per patient			
		Inpatient		Outpatient		Inpatient		Outpatient	
COST ITEM		IS	TIA	IS	TIA	IS	TIA	IS	TIA
LOS									
Inpatient days	€ 485	8.794	3.602	0.006	0.002	€4,269	€1,748	€ 3	€ 1
Intensive care days	€2,319	0.011	0.002	0.000	0.000	€ 25	€ 5	€ 1	€ 0
VISITS									
Outpatient visits	€ 76	0.799	0.810	0.890	0.843	€ 61	€ 62	€ 68	€ 64
Emergency room visits	€ 160	0.889	0.868	0.256	0.242	€ 143	€ 139	€ 41	€ 39
Ambulatory treatment	€ 267	0.013	0.035	0.236	0.514	€ 3	€ 9	€ 63	€ 137
DIAGNOSTIC AND IMAGING TESTS									
CT (brain)	€ 197	1.177	1.003	0.641	0.772	€ 232	€ 198	€ 126	€ 152
Duplex (carotid)	€ 85	0.770	0.900	0.561	0.866	€ 66	€ 77	€ 48	€ 74
ECG	€ 20	0.775	0.701	0.383	0.581	€ 15	€ 14	€ 8	€ 12
X-ray (thorax)	€ 51	0.429	0.243	0.099	0.129	€ 22	€ 12	€ 5	€ 7
MRI (brain)	€ 253	0.249	0.178	0.281	0.162	€ 63	€ 45	€ 71	€ 41
EEG	€ 73	0.120	0.194	0.079	0.093	€ 9	€ 14	€ 6	€ 7
Diagnostic services (other)	€ 129	0.604	0.451	0.315	0.322	€ 78	€ 58	€ 41	€ 41
ALLIED HEALTH SERVICES									
Physical therapy	€ 38	2.975	0.532	0.010	0.002	€ 114	€ 20	€ 0	€ 0
Occupational therapy	€ 24	1.775	0.271	0.015	0.000	€ 42	€ 6	€ 0	€ 0
Speech therapy	€ 35	1.505	0.239	0.020	0.007	€ 53	€ 8	€ 1	€ 0
Psychotherapy	€ 82	0.045	0.010	0.008	0.001	€ 4	€ 1	€ 1	€ 0
LABORATORY INVESTIGATIONS									
Laboratory services	€ 13	7.101	3.235	0.706	0.830	€ 95	€ 43	€ 9	€ 11
Microbiological services	€ 15	2.044	0.484	0.156	0.078	€ 32	€ 7	€ 2	€ 1
Pathological services	€ 91	0.014	0.004	0.003	0.001	€ 1	€ 0	€ 0	€ 0
Blood (derived) products	€ 219	0.014	0.005	0.003	0.001	€ 3	€ 1	€ 1	€ 0
TOTAL						**€5,328**	**€ 2,470**	**€ 495**	**€ 587**

LOS = length of stay; IS = ischemic stroke; inpatient IS n=24,671; outpatient IS n=11,232; inpatient TIA n = 4,596; outpatient TIA n = 17,057.

Table 7.3: Unit costs, average resource use, and average costs of carotid endarterectomy

Cost item	Unit costs	Average resource use per patient	Average costs per patient
LOS			
Inpatient days	€ 485	4.828	€ 2,343
Intensive care days	€ 2,319	0.052	€ 120
VISITS			
Outpatient visits	€ 76	0.750	€ 57
Emergency room visits	€ 160	0.077	€ 12
Ambulatory treatment	€ 267	0.008	€ 2
DIAGNOSTIC AND IMAGING TESTS			
CT (brain)	€ 197	0.055	€ 11
Duplex (carotid)	€ 85	1.161	€ 99
ECG	€ 20	0.265	€ 5
X-ray (thorax)	€ 51	0.166	€ 8
MRI (brain)	€ 253	0.007	€ 2
EEG	€ 73	0.457	€ 33
Imaging services (other)	€ 129	0.951	€ 122
SURGERY			
Carotid endarterectomy	€ 3,486	1.000	€ 3,486
Surgery (other)	€ 3,486	0.129	€ 450
ALLIED HEALTH SERVICES			
Physical therapy	€ 38	0.292	€ 11
Occupational therapy	€ 24	0.035	€ 1
Speech therapy	€ 35	0.053	€ 2
Psycho therapy	€ 82	0.000	€ 0
LABORATORY INVESTIGATIONS			
Laboratory services	€ 13	3.369	€ 45
Microbiological services	€ 15	0.542	€ 8
Pathological services	€ 91	0.118	€ 11
Blood (derived) products	€ 219	0.031	€ 7
TOTAL			**€ 6,836**

LOS = length of stay; carotid endarterectomy n=1,891.

Costs of carotid endarterectomy

Table 7.3 shows the unit costs of the different cost items, breakdown of average resource use, and average hospital costs of patients who underwent carotid endarterectomy. Average costs of these patients were €6,836 ± €2,862 ($8,783 ± $3,677), with surgery (51%) and inpatient days (34%) as main components.

Costs of an extra follow-up visit

Substantial differences were observed in average hospital costs per patient between inpatients and outpatients which were mainly caused by the average LOS of inpatient ischemic stroke (7.6 days) and inpatient TIA (4.8 days) (P<0.01). Average costs of inpatients who had an extra follow-up visit after ischemic stroke were €4,567±€5,817 ($5,868±$7,474) and after TIA €2,983±€3,368 ($3,833±$4,327). The average costs of outpatients with an extra follow-up visit were €223±€282 ($287±$362) for ischemic stroke and €229±€316 ($294±$406) for TIA. The largest component of outpatient costs were outpatient visits.

Regression analysis

The regression analyses indicate that the overall regression models are statistically significant. However, the explained variance is low (2.1% to 5.6%). Table 7.4 shows the associations between hospital costs and the various patient and hospital characteristics of inpatient care for ischemic stroke and TIA (Analysis 1), outpatient care for ischemic stroke and TIA (Analysis 2), and carotid endarterectomy (Analysis 3). Inpatient costs were significantly related to age [i.e., one additional year of age was associated with a cost increase of €25 ($32)]. The impact of age on inpatient costs was high [e.g., a 20-year age difference results in additional costs of €500 ($640)]. Moreover, Analysis 1 shows that costs at teaching, small and unknown type of non-academic hospitals were significantly higher than at academic hospitals, and that inpatient costs at hospitals located in Drenthe, Flevoland, Groningen, Limburg, North Brabant, and Overijssel were significantly higher than in South Holland.

Outpatient costs of ischemic stroke and TIA were significantly associated with age. Outpatient costs decrease by €2 ($3) for each additional year of age and no difference was observed between genders. Significant differences in outpatient costs were found between academic and all types of non-academic hospitals, except midsized non-academic hospitals. Moreover, outpatient costs at hospitals located in most provinces were significantly lower than in South Holland.

Hospital costs of patients who underwent carotid endarterectomy were significantly related to age and region. Costs of carotid endarterectomy were significantly lower in all provinces compared to South Holland (except for

Groningen), ranging from €570 to €1,513 ($732 to $1,944) lower costs. In addition, significant differences were found between academic and mid-sized and small non-academic hospitals.

Table 7.4: Association between hospital costs and patient and hospital characteristics

	Analysis 1: Inpatient care Coefficient (SE)	**Analysis 2: Outpatient care** Coefficient (SE)	**Analysis 3: Carotid endarterectomy** Coefficient (SE)
PATIENT CHARACTERISTICS			
Age	25 (3)***	-2 (0)***	31 (11)***
Male	-417 (265)	30 (20)	457 (961)
Age*Male	5 (4)	0 (0)	-13 (14)
HOSPITAL CHARACTERISTICS			
Hospital type			
Academic	ref.	ref.	ref.
Non-academic (teaching)	613 (102)***	71 (8)***	272 (200)
Non-academic (large)	-114 (113)	34 (9)***	-21 (235)
Non-academic (midsized)	73 (102)	-3 (8)	-852 (206)***
Non-academic (small)	362 (108)***	55 (8)***	665 (323)**
Non-academic (unknown)	855 (155)***	22 (11)**	106 (332)
Region			
Drenthe	958 (150)***	14 (11)	-1,119 (397)***
Flevoland	482 (205)**	-70 (15)***	-950 (562)*
Friesland	179 (143)	-143 (12)***	-1,450 (394)***
Gelderland	-21 (86)	-16 (7)**	-872 (245)***
Groningen	245 (146)*	-77 (12)***	-183 (413)
Limburg	645 (103)***	-28 (8)***	-980 (264)***
North Brabant	492 (83)***	-41 (6)***	-570 (225)**
North Holland	99 (83)	-20 (7)***	-739 (240)***
Overijssel	1,320 (118)***	7 (9)	-1,116 (327)***
South Holland	ref.	ref.	ref.
Utrecht	50 (111)	19 (10)*	-1,307 (303)***
Zeeland	209 (164)	-118 (15)***	-1,513 (696)**
Number of observations	28,655	27,740	1,861
R^2	2.1%	2.6%	5.6%

* P<0.1; ** P<0.05; *** P<0.01; SE = standard error; ref. = reference category.

Subsequent analysis

We performed a subsequent analysis to examine whether hospital costs and LOS were associated with admission day of the week. We found that both LOS and costs were significantly lower at Mondays, Tuesdays, Wednesdays, and Thursdays compared to the other days of the week ($P<0.1$).

7.4 Discussion

This study is the first and most extensive cost analysis of inpatient and outpatient hospital costs of ischemic stroke and TIA patients since 10 years. Costs between inpatient and outpatient care for ischemic stroke and TIA differed extensively and was largely attributable to inpatient hospital stay. The cost difference between inpatient ischemic stroke (€5,328/$6,845) and inpatient TIA (€2,470/$3,173) was caused by a shorter LOS for TIA patients (3.6 days versus 8.8 for ischemic stroke), which could have been expected given the lower severity and shorter symptom duration of TIA.[21] Costs of outpatient care were higher for TIA than for ischemic stroke due to more frequent use of ambulatory treatment, and diagnostic and imaging tests.

During the past two decades, LOS and costs of ischemic stroke patients have substantially reduced, possibly due to major improvements in patient management (e.g., the organization of stroke care into integrated stroke services and the use of intravenous thrombolysis). Two studies from the early 90s estimated an average LOS of 27 days[8] and 28 days.[15] A study performed in the late 90s examined the ability of three experimental stroke services to reduce LOS. Only one hospital succeeded (13 days versus 32 and 33 in the two other hospitals).[10] We found that the average LOS for ischemic stroke has considerably been reduced to 8.8 days in 2010. In terms of hospital costs per patient, the difference of about 20 inpatient days alone has resulted in a cost reduction of €9,700 ($12,460) in 2012, with a unit cost of €485 ($623) per inpatient day. On the other hand, intravenous thrombolysis and other novel therapeutic options have increased the unit costs of an inpatient day.

Our finding that age, hospital type, and region were associated with hospital costs partially correspond with the results of an earlier study which found hospital costs to be dependent on region as well as stroke severity at discharge, waiting lists for nursing home care, and death from stroke in the acute phase. However, in that study, age was not found to be associated with hospital costs and hospital type was not considered.[10] Our regression analyses show significant associations between hospital costs and both hospital type and region. Both inpatient and outpatient costs were significantly higher at teaching, small and unknown type of non-academic hospitals compared to academic hospitals. The analyses also show that inpatient costs are significantly higher and outpatient costs significantly lower in most of

the provinces outside the Dutch metropolitan area. A possible explanation is that the patient case mix differs between hospital types and regions. Moreover, we found that LOS was associated with admission day of the week. A possible explanation is that LOS is shorter when patients are admitted during weekdays due to the availability of medical personnel. This finding and the determinants of hospital costs, such as hospital type and region merit further research.

In addition, we found that more than 90% of patients with an extra follow-up visit were outpatients. We assume that inpatients were discharged to long term care facilities where they received the necessary follow-up care. In contrast, the patients who were originally seen at the hospital's outpatient clinic were directly discharged home and therefore sometimes needed an extra follow-up visit.

The strength of our study is that we examined the hospital costs and LOS using data from all patients in the Netherlands who were seen at any hospital for an ischemic stroke or TIA in 2010. Our findings are highly relevant for neurologists since they reflect the costs of current patient care. The methods we used in our study can also be used by clinicians as an example to perform similar analyses for other diseases. In addition, our results may be used as input for economic evaluations to support decision making about reimbursement, investment, and pricing of healthcare interventions.

Our study also has some limitations. The DBC database does not contain information on stroke severity, such as the National Institutes of Health Stroke Scale or the Modified Rankin Scale, and the most probable cause, such as the TOAST classification. Potential proxy measures of stroke severity, such as clinical parameters and discharge location were also not included in the DBC database. We were therefore unable to estimate the costs and LOS for different types of ischemic events. Furthermore, we were unable to assess potential risk factors and co-morbidity that determine hospital costs, such as diabetes mellitus, hypertension and hyperlipidemia.[22] Moreover, the database includes only hospital care. The costs of stroke care outside hospitals (e.g., GPs, rehabilitation centers, long term care facilities, and nursing home care) could not be estimated with the database. Further research should estimate these costs to obtain a total estimate of costs of ischemic stroke and TIA patients.

Some of our results are not necessarily generalizable to different healthcare systems. For example, the absolute costs seen in our study do not necessarily reflect the costs expected elsewhere because, for example, the unit costs of particular components of stroke care (cost per inpatient day) differ between countries and healthcare systems. However, we expect that the trends and patterns in stroke care found in our study (e.g., reduction in length of stay and differences between hospital types and regions) are likely to be generalizable.

We have shown that hospital costs are higher for inpatients and ischemic strokes compared with outpatients and TIAs, with LOS as most important contributor. LOS and hospital costs have substantially reduced over the past 10 years, possibly due to improved hospital stroke care and efficient integrated stroke services.

7.5 Acknowledgments

This research was performed within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), project PARISk (grant 01C-202), and supported by the Dutch Heart Foundation and Dutch Organisation for Health Research and Healthcare Innovation (ZON-MW, grant no. 152002043). We also wish to thank Joke Hendriks, Frank van den Boogert, and Tamara van Kleef for their assistance in estimating specific unit costs.

7.6 References

1 Donnan GA, Fisher M, Macleod M, et al. Stroke. *Lancet* 2008;371:1612–23.

2 Minkman MM, Schouten LM, Huijsman R, et al. Integrated care for patients with a stroke in the Netherlands: results and experiences from a national Breakthrough Collaborative Improvement project. *Int J Integr Care* 2005;5:e14.

3 Saver JL, Fonarow GC, Smith EE, et al. Time to treatment with intravenous tissue plasminogen activator and outcome from acute ischemic stroke. *JAMA* 2013;309:2480–8.

4 Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis. North American Symptomatic Carotid Endarterectomy Trial Collaborators. *N Engl J Med* 1991;325:445–53.

5 MRC European Carotid Surgery Trial: interim results for symptomatic patients with severe (70–99%) or with mild (0–29%) carotid stenosis. European Carotid Surgery Trialists' Collaborative Group. *Lancet* 1991;337:1235–43.

6 Randomised trial of a perindopril-based blood pressure lowering regimen among 6,105 patients with prior stroke or transient ischaemic attack. PROGRESS Collaborative Group. *Lancet* 2001;358:1033–41.

7 Amarenco P, Bogousslavsky J, Callahan A, et al. High-dose atorvastatin after stroke or transient ischemic attack. *N Engl J Med* 2006;355:549–59.

8 Bergman L, Van der Meulen J, Limburg M, et al. Cost of medical care after first-ever stroke in the Netherlands. *Stroke* 1995;26:1830–6.

9 Evers SM, Engel GL, Ament AJ. Cost of stroke in the Netherlands from a societal perspective. *Stroke* 1997;28:1375–81.

10 Van Exel J, Koopmanschap MA, Van Wijngaarden JD, et al. Costs of stroke and stroke services: Determinants of patient costs and a comparison of costs of regular care and care organised in stroke services. *Cost Eff Resour Alloc* 2003;1(1):2.

11 Kolominsky-Rabas PL, Heuschmann PU, Marschall D, et al. Lifetime cost of ischemic stroke in Germany: results and national projections from a population-based stroke registry: the Erlangen Stroke Project. *Stroke* 2006;37(5):1179–83.

12 Gioldasis G, Talelli P, Chroni E, et al. In-hospital direct cost of acute ischemic and hemorrhagic stroke in Greece. *Acta Neurol Scand* 2008;118(4):268–74.

13 Cadilhac DA, Carter R, Thrift AG, et al. Estimating the long-term costs of ischemic and hemorrhagic stroke for Australia: new evidence derived from the North East Melbourne Stroke Incidence Study (NEMESIS). *Stroke* 2009;40(3):915–21.

14 Yoneda Y, Okuda S, Hamada R, et al. Hospital cost of ischemic stroke and intracerebral hemorrhage in Japanese stroke centers. *Health Policy* 2005;73(2):202–11.

15 Van Straten A, Van der Meulen JH, Van den Bos GA, et al. Length of hospital stay and discharge delays in stroke patients. *Stroke* 1997;28(1):137–140.

16 Tan SS, Bouwmans CA, Rutten FF, et al. Update of the Dutch Manual for Costing in Economic Evaluations. *Int J Technol Assess Health Care* 2012;28:152–8.

17 Dutch Healthcare Authority (in Dutch: Nederlandse Zorgauthoriteit, NZa). NZa Zorgproducten Tariefapplicatie. Available from: http://dbc-zorgproducten-tarieven.nza.nl/nzaZpTarief/ZoekfunctieDbc.aspx. Accessed May 1, 2013.
18 Statistics Netherlands (in Dutch: Central Bureau voor de Statistiek, CBS). CBS Statline. Available from: http://statline.cbs.nl. Accessed May 1, 2013.
19 European Central Bank. ECB: Euro foreign exchange reference rates. Available from: https://www.ecb.europa.eu/stats/exchange/eurofxref/html/index.en.html. Accessed October 24, 2014.
20 Organisation for Economic Co-operation and Development (OECD). 2010. OECD Territorial Reviews: Randstad Holland, Netherlands. Paris: OECD.
21 Chandratheva A, Geraghty OC, Luengo-Fernandez R, et al. ABCD² score predicts severity rather than risk of early recurrent events after transient ischemic attack. *Stroke* 2010;41:851–6.
22 O'Donnell MJ, Xavier D, Liu L, et al. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet* 2010;376:112–23.

CHAPTER 8

COST-EFFECTIVENESS OF NOVEL IMAGING TESTS TO SELECT ISCHEMIC STROKE PATIENTS FOR CAROTID ENDARTERECTOMY

Leander R. Buisman[1,2], Adriana J. Rijnsburger[1,2], Aad van der Lugt[3], Paul J. Nederkoorn[4], Peter J. Koudstaal[5], William K. Redekop[1,2]

[1] Institute of Health Policy and Management,
Erasmus University Rotterdam, Rotterdam, The Netherlands
[2] Institute for Medical Technology Assessment,
Erasmus University Rotterdam, Rotterdam, The Netherlands
[3] Department of Radiology, Erasmus MC,
University Medical Center Rotterdam, Rotterdam, The Netherlands
[4] Department of Neurology, Academic Medical Center,
University of Amsterdam, Amsterdam, The Netherlands
[5] Department of Neurology, Erasmus MC,
University Medical Center Rotterdam, Rotterdam, The Netherlands

Submitted for publication.

ABSTRACT

Background and Objective: Carotid endarterectomy (CEA) reduces the risk of recurrent stroke in patients with significant carotid stenosis, but carries the risk of treatment-related death and disability. Improved risk prediction by new imaging tests could help to select patients who benefit most from CEA. However, prognostic test performance and cost-effectiveness of these tests is unknown.

Methods: The guidelines recommend initial duplex ultrasonography (DUS) followed by a confirmatory test if DUS shows 30-69% stenosis; a positive CTA test is an indication for CEA. In an alternative strategy, we replaced CTA with a new imaging test and estimated the minimum prognostic performance that the new test must have in order to be cost-effective versus the guideline-based strategy. We assessed the potential cost-effectiveness in four age- and sex-specific subpopulations.

Results: For 60-year-old men, a perfect confirmatory test (100% sensitivity and specificity) improves health (0.066 quality-adjusted life years) and reduces costs (€110/$146) versus the guideline-based strategy. Potential health gain is smaller for 80-year-old men, while no health gain is expected for women. Assuming 100% sensitivity, a test must have a specificity of at least 66% for 60-year-old men and 87% for 80-year-old men to be cost-effective. Similarly, assuming 100% specificity, a test must have a sensitivity of at least 66% for 60-year-old men and 58% for 80-year-old men.

Conclusions: Information from new imaging technologies may improve stroke risk prediction and thereby improve decisions about which patients should undergo CEA. However, their cost-effectiveness strongly depends on the current test strategy and choice of patient subpopulation.

Keywords: ischemic stroke, TIA, secondary prevention, diagnostic imaging, prognostic test performance, sensitivity and specificity, cost-effectiveness analysis.

8.1 Introduction

Carotid endarterectomy (CEA) reduces the risk of recurrent stroke in selected populations: i.e., symptomatic patients with 70-99% stenosis, or patient subpopulations with 50-69% stenosis. The decision to select patients for CEA in daily clinical practice is mainly based on degree of carotid artery stenosis.[1-3] Several published studies have shown that the features of a vulnerable plaque (e.g., intraplaque hemorrhage, thin or ruptured fibrous cap, and large lipid-rich necrotic core) are related to risk of recurrent stroke.[4-10]

Imaging of the atherosclerotic plaque with MRI, CTA, duplex ultrasonography (DUS),[11] or biomechanical analysis based on imaging information,[12] can yield valuable information. This information can help to improve stroke risk prediction and thereby improve decisions about which patients should undergo CEA. This improvement in patient stratification is part of the promise of precision medicine.[13-15] However, the prognostic performance (i.e., sensitivity and specificity) of these tests is unknown. Several prospective cohort studies are currently ongoing determining the prognostic performance of various candidate tests (e.g., PARISk[16] and CAPIAS[17]). In this study, we estimated the potential cost-effectiveness of novel imaging tests using a computer model and explored the maximum health gain that these tests could achieve if they were perfect, in comparison with a guideline-based strategy. We also estimated the minimum prognostic performance that a new confirmatory test must have in order to be cost-effective versus the guideline-based strategy.

8.2 Methods

We applied the framework with general steps of early-CEAs of medical tests as developed by Buisman et al.[18] This framework is a useful guidance for researchers performing early-CEAs of medical tests.

Patient population

We defined the patient population as follows: patients with a recent TIA or minor ischemic stroke (mRS≤3). TIA was defined as a focal neurologic deficit of sudden onset lasting less than 24 hours and with no signs of recent infarction on CT or MRI. Ischemic stroke was defined as a focal neurological deficit of sudden onset of presumed vascular origin, lasting at least 24 hours, with brain imaging showing typical signs of brain infarction or no abnormalities.

Current imaging test strategies (including a guideline-based strategy)

As part of secondary prevention, patients are selected for CEA or optimal medical treatment (OMT) alone. The Dutch national stroke guidelines[19]

were examined and interviews were conducted with vascular neurologists from different hospitals to ascertain whether the guidelines were used in clinical practice (see Buisman et al.[20]).

In general, after diagnostic evaluation and treatment in the acute phase, patients with a recent TIA or minor ischemic stroke undergo an assessment of carotid artery stenosis and subsequent treatment as part of secondary prevention (i.e., preventing a recurrent stroke). In the assessment of carotid stenosis, Dutch guidelines recommend DUS as the initial test and CTA or MRA as confirmatory test.[19] According to the initial DUS, the criterion for performing a confirmatory test is moderate (50-69%) carotid stenosis for men or severe (70-99%) carotid stenosis for women.[19] CEA is recommended for both men and women with severe (70-99%) carotid stenosis and a TIA or minor ischemic stroke in the past 6 months. In addition, a CEA is recommended for men with moderate (50-69%) carotid stenosis and a TIA or minor ischemic stroke in the past 3 months.[19]

Since multiple imaging test strategies are used in clinical practice,[20] we included various strategies in our analysis as comparators, including a guideline-based strategy (initial DUS and confirmatory CTA) and three other strategies: DUS-only, CTA-only, and CE-MRA-only.

New imaging strategy

Since there are many imaging tests currently being developed to improve risk prediction with imaging features of a vulnerable plaque, leading to a large number of possible combinations of tests, vascular neurologists were queried about the optimal combination of tests. We determined that the most likely application for a new imaging test to improve risk prediction in patients with a recent TIA or minor ischemic stroke would be a confirmatory test for patients with 30-69% stenosis based on an initial DUS. DUS is the preferred initial test since it is often used in current care, relatively cheap and simple to use.

Model structure

A lifetime cost-effectiveness model was developed to perform the cost-effectiveness analyses. Figure 8.1 shows the first part of the model in which use of tests and subsequent treatment were modelled. Patients who test positive, i.e., patients with a high-risk of a recurrent stroke, undergo CEA, whereas others receive OMT alone. Based on the prognostic test performance, patients are classified into four groups: true positive (TP), false positive (FP), false negative (FN), and true negative (TN). Final health outcomes depend on these classifications and subsequent treatment. Figure 8.2 shows the health outcomes after CEA and OMT alone. If the test's sensitivity and specificity are less than 100%, the test misclassifies patients, resulting in inappropriate treatments and increased risk of ischemic stroke events. The

final health outcomes include minor, major, fatal or no ischemic stroke. Death from other causes is incorporated by using the life expectancy from the Dutch population.[21]

Figure 8.1: Decision model – tests and subsequent treatment

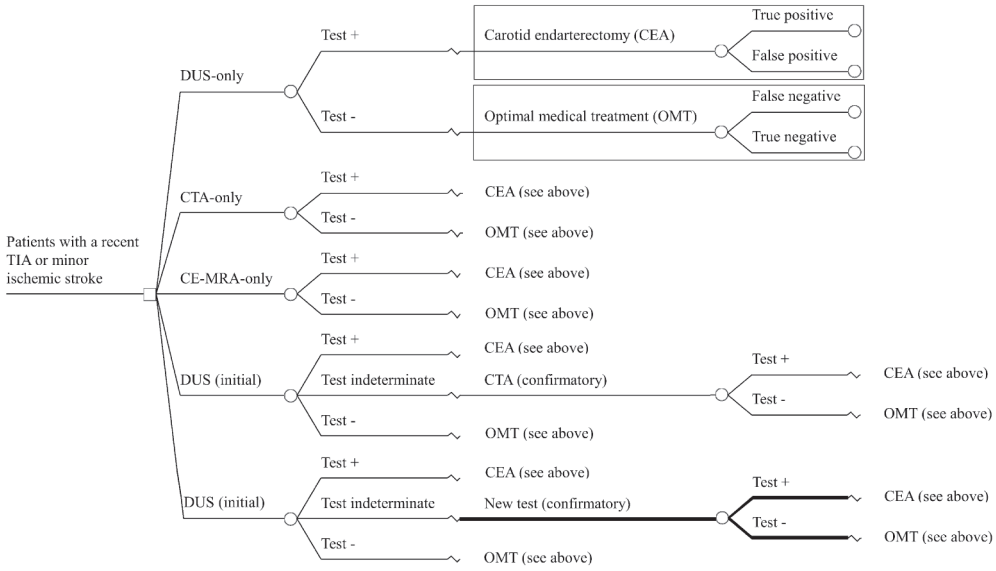
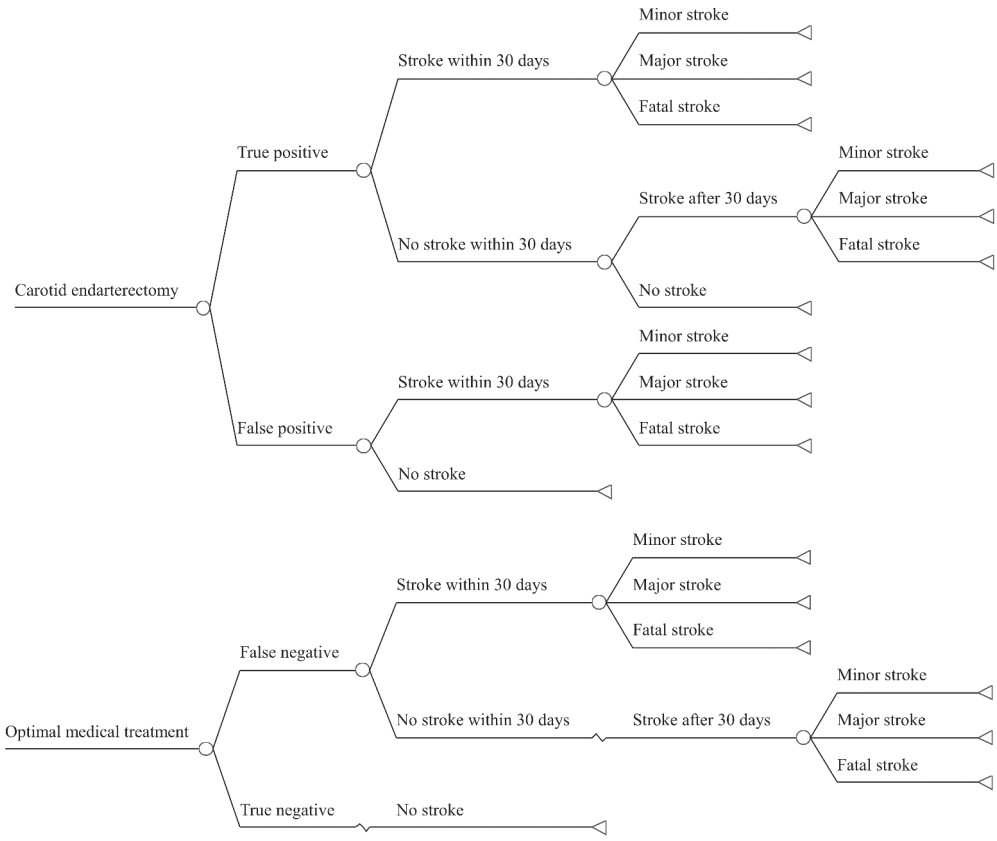


Figure 8.2: Decision model – health outcomes after treatment



Model inputs and assumptions

Test performance

The performance of the imaging tests used in *current care* is based on the sensitivity and specificity to diagnose 70-99% carotid stenosis (see Table 8.1). The sensitivity and specificity of DUS and CE-MRA were based on a meta-analysis of 41 studies.[22] The performance of CTA was derived from a prospective cohort study of 351 TIA/minor ischemic stroke patients.[23]

The performance of the *new* prognostic test was defined in terms of the ability to predict an ischemic stroke based on imaging features of vulnerable carotid plaque. TPs were defined as patients with a *positive test result* and a 100% lifetime risk of an ischemic stroke if they receive OMT alone. Since these patients subsequently undergo CEA, their risk of ischemic stroke is reduced. FNs were defined as patients with a *negative test result* and a 100% lifetime risk of an ischemic stroke if they receive OMT alone. Since these patients do not undergo CEA and instead receive OMT alone, their lifetime risk of ischemic stroke remains 100%. TNs were defined as patients with a *negative test result* and a 0% lifetime risk of an ischemic stroke if they receive OMT alone. Since these patients are correctly identified, they will receive OMT and continue to have a 0% risk of ischemic stroke. FPs were defined as patients with a *positive test result* and a 0% lifetime risk of an ischemic stroke if they receive OMT alone. Although these patients should receive OMT, they are misclassified, undergo CEA and therefore have a short-term risk of ischemic stroke due to surgical complications.

When these definitions are applied, a test with a higher sensitivity increases the chance that patients who will benefit from a CEA are correctly identified and treated. Similarly, a test with a higher specificity increases the chance that patients who will not benefit from a CEA are correctly identified and treated.

Health-related quality of life (utilities) and life expectancy

We included quality-adjusted life years (QALYs) as a measure of health outcomes. Table 8.1 shows the utility weights by ischemic stroke severity used in our model (at baseline after a recent TIA or minor ischemic stroke, after minor ischemic stroke, and after major ischemic stroke). At baseline, all patients have had a recent TIA or minor ischemic stroke and a utility of 0.71 was assigned.[24] We assigned a utility of 0.66 when a second minor ischemic stroke occurred, which was calculated by combining the utility loss due to a minor ischemic stroke (0.0524) with the baseline utility.[25] We used a utility of 0.31 after major ischemic stroke.[24]

Life expectancies of patients after a TIA, minor, or major ischemic stroke were estimated by combining survival data of ischemic stroke patients[26] with the life expectancy from the Dutch population in 2014.[21]

Table 8.1: Model input parameters

Parameter		Value	SE	Distribution (alpha;beta)	Source
PERFORMANCE OF TESTS					
DUS	– sensitivity	0.890	0.015	Beta (371;46)	[21]
	– specificity	0.840	0.026	Beta (173;33)	[21]
CTA	– sensitivity	0.910	0.041	Beta (44;4)	[22]
	– specificity	0.990	0.005	Beta (376;4)	[22]
CE-MRA	– sensitivity	0.940	0.015	Beta (225;14)	[21]
	– specificity	0.930	0.015	Beta (257;19)	[21]
HEALTH-RELATED QUALITY OF LIFE					
Baseline (recent TIA or minor ischemic stroke)		0.710	0.020	Beta (350;143)	[23]
After minor ischemic stroke		0.658	0.015	Beta (631;329)	[23,24]
After major ischemic stroke		0.310	0.015	Beta (283;629)	[23]
COSTS (in 2014 €)					
DUS		€125	€19	Gamma (€44;€3)	[27]
CTA		€189	€28	Gamma (€44;€4)	[27]
CE-MRA		€243	€36	Gamma (€44;€5)	[27]
CE-MRI (new test)		€362	€54	Gamma (€44;€8)	[27]
Carotid endarterectomy		€7,077	€68	Gamma (€10,788;€1)	[28]
Optimal medical treatment (per year)		€120	€18	Gamma (€44;€3)	Expert opinion
Minor ischemic stroke					[29]
first year		€8,731	€1,113	Gamma (€61;€142)	
subsequent years (per year)		€1,494	€190	Gamma (€62;€24)	
Major ischemic stroke					
first year		€49,790	€6,350	Gamma (€61;€810)	[29]
subsequent years (per year)		€29,072	€3,708	Gamma (€61;€473)	
Fatal ischemic stroke (in year of fatal ischemic stroke)					[30]
Men – aged <65		€6,663	€999	Gamma (€44;€150)	
Men – aged 65-74		€10,875	€1,631	Gamma (€44;€245)	
Men – aged 75-85		€8,735	€1,310	Gamma (€44;€197)	
Men – aged >85		€10,640	€1,596	Gamma (€44;€239)	
Women – aged <65		€7,353	€1,103	Gamma (€44;€165)	
Women – aged 65-74		€9,984	€1,498	Gamma (€44;€225)	
Women – aged 75-84		€12,043	€1,806	Gamma (€44;€271)	
Women – aged ≥85		€14,484	€2,173	Gamma (€44;€326)	

SE = standard error; DUS = duplex ultrasonography; CE-MRA = contrast enhanced MRA; CE-MRI = contrast enhanced MRI.

Costs

All costs were calculated in 2014 Euros and converted to US Dollars using a societal perspective (see Table 8.1). A mean exchange rate of €1=$1.3285 was used for 2014, because daily exchange rates ranged from €1=$1.2141 to €1=$1.3953 in 2014.[27] The total costs per patient consisted of costs of tests, treatment (i.e., CEA and medicines), and stroke-related societal costs. Test costs were based on tariffs provided by the Dutch Healthcare Authority.[28] Costs of CEA were based on a recent cost-analysis,[29] and the costs of OMT were based on expert opinion. Stroke-related costs in the first and subsequent years after a minor or major ischemic stroke, and costs related to a fatal ischemic stroke were based on literature.[30,31]

The cost of the new confirmatory test was arbitrarily set at €362/$481 which was based on the unit cost of a MRI of the carotid plaque including contrast,[28] because recent studies have shown an association between vulnerable plaque features assessed with MRI and the recurrence of ischemic stroke events in TIA/stroke patients.[7-9]

Risk of ischemic stroke

The risk of ischemic stroke after CEA or OMT alone was based on published estimates of the ECST-study[32,33] and reanalyzed according to the NASCET-method[34,35]. These estimates were combined with a risk reduction of 33% to reflect that medical treatment has improved (e.g., widespread use of statins, better antiplatelet therapy, and lower targets for blood pressure control).[36-38]

Short-term risks referred to the risk of events occurred within 30 days after start of OMT or CEA. Long-term risks referred to the risk of events occurred after 30 days until death. We assumed that the risk after 10 years, in each subsequent year, would be equal to risk in the 10th year (see Appendix 8.1).

Analysis

The cost-effectiveness of the new test strategy was assessed versus the different comparators. Incremental cost-effectiveness ratios (ICERs) were calculated using the guideline-based strategy as comparator. ICERs were calculated as the difference in costs divided by the difference in QALYs. Probabilistic sensitivity analysis with 1,000 simulations was performed. For each simulation, the values of the input parameters were randomly sampled from the appropriate distributions (see Appendix 8.1).

The cost-effectiveness of the new test strategy was assessed in four subpopulations (60-year-old men, 80-year-old men, 60-year-old women, and 80-year-old women). Cost-effectiveness planes were created to show the incremental costs and QALYs of the new imaging test strategies versus the guideline-based strategy. As illustration, we assumed that a perfect confir-

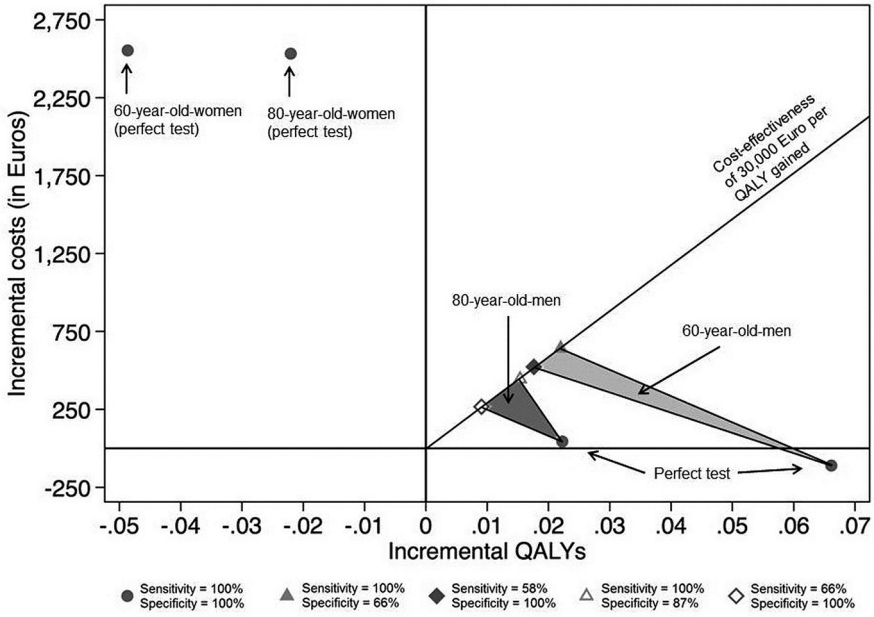
matory test (100% sensitivity and specificity) was used. However, a new confirmatory test is unlikely to be perfect. Therefore, we estimated the minimum prognostic performance that a new confirmatory test must have in order to be cost-effective versus the guideline-based strategy using a QALY threshold of €30,000/$39,855. We created four scenarios by varying the sensitivity and specificity. First, we estimated the minimum required specificity given a test with 100% sensitivity. Second, we estimated the minimum required sensitivity given 100% specificity. In the third and fourth scenarios, we estimated the minimum required specificity given a 90% sensitive test, and the minimum sensitivity required given a 90% specific test. Annual discount rates were 4.0% for costs and 1.5% for health outcomes in accordance with the Dutch guidelines.[39]

8.3 Results

A cost-effectiveness plane represents the difference between the new test strategy and guideline-based test strategy in health outcomes on the horizontal axis and in costs on the vertical axis. Figure 8.3 shows a cost-effectiveness plane that summarizes the results for the four subpopulations. For 60-year-old and 80-year-old women, a perfect confirmatory test strategy leads to poorer health outcomes and higher costs than the guideline-based strategy. For 60-year-old and 80-year-old men, the triangles represent all combinations of sensitivity and specificity that a test can have in order to be cost-effective compared to the guideline-based strategy. These triangles show that a cost-effective test (i.e., a test with cost-effective combinations of sensitivity and specificity) usually results in better health outcomes but higher costs. The exception to this rule is the use of a perfect or near-perfect confirmatory test with 60-year-old men, where the result would be better health outcomes and lower costs).

Table 8.2 shows the cost-effectiveness results of the new test scenarios compared to the guideline-based strategy. The costs and health outcomes are dependent on the sensitivity and specificity of the tests, age, and sex.

Figure 8.3: Cost-effectiveness plane with different subpopulations



Se = sensitivity; Sp = specificity; QALY = quality-adjusted life year; Cost-effectiveness of €30,000/$39,855 per QALY gained = willingness-to pay threshold line of €30,000/$39,855 per QALY gained.

Table 8.2: Cost-effectiveness of new test strategy versus current test strategies in different subpopulations

Sex	Age	Sensitivity	Specificity	Test strategy	QALYs	Δ QALYs (vs. GL)	Costs	Δ Costs (vs. GL)	ICER (vs. GL)
Men	60			GL	11.589	*Reference*	€29,134/$38,705	*Reference*	*Reference*
				DUS-only	11.556	-0.033	€29,112/$38,675	-€22/-$29	€667/$886
				CTA-only	11.606	0.017	€28,590/$37,982	-€544/-$723	Dominant*
				CE-MRA-only	11.625	0.036	€28,768/$38,218	-€366/-$486	Dominant*
		100%	100%	DUS + new test	11.655	0.066	€29,024/$38,558	-€110/-$146	Dominant*
		100%	**66%**	DUS + new test	11.611	0.022	€29,775/$39,556	€641/$852	€29,136/$38,707
		58%	100%	DUS + new test	11.607	0.018	€29,657/$39,399	€523/$695	€29,056/$38,601
		90%	**74%**	DUS + new test	11.610	0.021	€29,749/$39,522	€615/$817	€29,286/$38,906
		70%	90%	DUS + new test	11.608	0.019	€29,697/$39,452	€563/$748	€29,632/$39,366
Men	80			GL	4.705	*Reference*	€19,047/$25,304	*Reference*	*Reference*
				DUS-only	4.694	-0.011	€19,016/$25,263	-€31/-$41	€2,818/$3,744
				CTA-only	4.712	0.007	€18,360/$24,391	-€687/-$913	Dominant*
				CE-MRA-only	4.718	0.013	€18,654/$24,782	-€393/-$552	Dominant*
		100%	100%	DUS + new test	4.727	0.022	€19,091/$25,362	€44/$58	€2,000/$2,657
		100%	**87%**	DUS + new test	4.720	0.015	€19,491/$25,894	€444/$590	€29,600/$39,324
		66%	100%	DUS + new test	4.714	0.009	€19,314/$25,659	€267/$355	€29,667/$39,413
		90%	**91%**	DUS + new test	4.718	0.013	€19,434/$25,818	€387/$514	€29,769/$39,548
		91%	90%	DUS + new test	4.718	0.013	€19,458/$25,850	€411/$546	€31,615/$42,001
Women	60			GL	13.171	*Reference*	€32,957/$43,782	*Reference*	*Reference*
				DUS-only	13.041	-0.130	€37,638/$50,002	€4,681/$6,219	Dominated**
				CTA-only	13.078	-0.093	€35,897/$47,689	€2,940/$3,906	Dominated**
				CE-MRA-only	13.096	-0.075	€35,773/$47,524	€2,816/$3,741	Dominated**
		100%	100%	DUS + new test	13.122	-0.049	€35,526/$47,196	€2,569/$3,413	Dominated**
Women	80			GL	5.658	*Reference*	€20,540/$27,287	*Reference*	*Reference*
				DUS-only	5.607	-0.051	€24,083/$31,994	€3,543/$4,707	Dominated**
				CTA-only	5.622	-0.036	€22,799/$30,288	€2,259/$3,001	Dominated**
				CE-MRA-only	5.628	-0.030	€22,931/$30,464	€2,391/$3,176	Dominated**
		100%	100%	DUS + new test	5.636	-0.022	€23,097/$30,684	€2,558/$3,398	Dominated**

QALY = quality-adjusted life year; ICER = incremental cost-effectiveness ratio (€/$ per QALY gained); DUS = duplex ultrasonography; CE-MRA = contrast enhanced MRA; GL = guideline-based strategy (initial DUS + confirmatory CTA); new test = new confirmatory test; * Dominant = better health outcomes and lower costs; ** Dominated = poorer health outcomes and higher costs.

For 60-year-old men, the five new test scenarios and current CTA-only and CE-MRA-only strategies resulted in better health outcomes than the guideline-based strategy, while DUS-only resulted in poorer health outcomes. Furthermore, DUS-only, CTA-only, CE-MRA-only, and a perfect test resulted in lower costs than the guideline-based strategy, whereas the other four new test scenarios led to higher costs. Assuming 100% sensitivity, a test must have a specificity of at least 66% to be cost-effective. Similarly, a test with a specificity of 100% must have a sensitivity of at least 58%. Assuming 90% sensitivity, a test must have a specificity of at least 74% and a test with a specificity of 90% must have a sensitivity of at least 70%. Notice that the new imaging test must have a sufficiently high sensitivity and specificity to be cost-effective, and that the test specificity is more important than the test sensitivity.

For 80-year-old men, the minimum required test performance must be higher than for 60-year-old men. A test with a sensitivity of 100% must have a specificity of at least 87% (versus 66%) and a test with a specificity of 100% must have a sensitivity of at least 66% (versus 58%).

For 60-year-old and 80-year-old women, even a perfect confirmatory test is not cost-effective versus the guideline-based strategy since it leads to poorer health outcomes and higher costs. The poorer health outcomes were a result of the use of the confirmatory test in women with 30-69% stenosis. In the guideline-based strategy, women with 70-99% stenosis based on DUS undergo a more accurate confirmatory test. In the new test strategy, women with 70-99% stenosis only undergo DUS, resulting in a selection of women receiving inappropriate treatment.

We found that the maximum health gain achievable by using a perfect confirmatory test after a DUS is relatively small (i.e., 0.066 in 60-year-old men, 0.022 in 80-year-old men, and no gain in women).

8.4 Discussion

Information from new imaging technologies can help to improve stroke risk prediction and thereby improve decisions about which patients should undergo CEA. We examined how much these technologies can potentially improve the value of decision-making given the treatment options that are currently available.

We showed that multiple factors determine the cost-effectiveness of a new test strategy. First, the cost-effectiveness of the new confirmatory imaging test depends on which subpopulation is tested. In the two subpopulations of men that we examined a new test may be cost-effective if the test achieves minimum required prognostic performance, whereas even a perfect confirmatory test is not cost-effective among women with 30-69% stenosis. Amongst women, a new confirmatory test is likely to be cost-effective for

women with 70-99% stenosis since they have a higher risk of ischemic stroke.

Second, the positioning of new tests (e.g., as initial or confirmatory test) has an impact on the cost-effectiveness. Since new tests are likely to be costly, we examined a test strategy in which patients first undergo DUS and then undergo a new imaging test as confirmatory test only if they have 30-69% stenosis according to DUS. We found that the maximum health gain achievable by using a perfect confirmatory test after a DUS is relatively small per individual (0.066 and 0.022 QALY gain for 60-year-old and 80-year-old men, respectively). While these are small gains per individual, they could be important at a population level and cost-effective to achieve. Obviously, a new imaging test could be used in other ways, including as the initial test; an exploration of other ways to use new tests merits further research.

Third, the choice of current test strategy is important. When DUS-only, CTA-only or CE-MRA-only is used as the current test strategy for 60-year-old women instead of the guideline-based strategy, a perfect confirmatory test strategy leads to better health outcomes and lower costs.

Fourth, clinical effectiveness of existing treatments is important, because a test forms just one part of patient management. The available treatments should be able to reduce the risk of an ischemic stroke; a new test with excellent prognostic performance is useless if the available treatment options are not very effective. On the other hand, if available treatments are both effective and relatively cheap, a test will have no value since all patients could receive treatment without the test. While improvements of medical treatment have reduced the risk of ischemic stroke,[36-38] tests are still needed to identify the best candidates for treatment since existing treatments have both advantages and disadvantages.

Our study has some limitations. Only a single recurrent ischemic stroke event was allowed in our model even though patients can have more than one recurrent stroke. Furthermore, our model structure and inputs are based on literature and expert opinion. Ongoing studies such as PARISk[16] and CAPIAS[17] may yield better input for our model. Another limitation is the use of national tariffs instead of unit costs of tests. Hence, some of our results are not generalizable to other countries. The current test strategies used in our study do not necessarily reflect the test strategies seen elsewhere and the model inputs might be different (e.g., higher costs in the US). The US AHA/ASA guidelines recommend the use of at least one of the following tests to select suitable candidates for CEA: DUS, CTA or MRA[40]. Recommendations for the sequence of tests are not specified, which may result in a variety of test combinations. Furthermore, we expect that the life expectancy of ischemic stroke patients is generalizable to most developed countries because of similar overall life expectancies. However, differences in the quality of stroke care (e.g., risk of complications after CEA) might

result in differences in QALYs. Therefore, similar analyses should be performed in other countries.

8.5 Conclusions

While new imaging tests can help to select the most appropriate treatments for patients with a recent TIA or ischemic stroke, their impact on costs and health outcomes depends on the current test strategy and the choice of patient subpopulation. Our analyses show that the maximum health gain of a perfect confirmatory test strategy versus the guideline-based test strategy may be limited.

8.6 Acknowledgments

This research was performed within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), project PARISk (grant 01C-202), and supported by the Dutch Heart Foundation. We also wish to thank Sander Arons for his useful comments.

8.7 References

1 Rothwell PM, Eliasziw M, Gutnikov SA, et al. Analysis of pooled data from the randomised controlled trials of endarterectomy for symptomatic carotid stenosis. *Lancet* 2003;361:107–16.

2 European Carotid Surgery Trialists' Collaborative Group. Randomised trial of endarterectomy for recently symptomatic carotid stenosis: final results of the MRC European Carotid Surgery Trial (ECST). *Lancet* 1998;351:1379–87.

3 Mayberg MR, Wilson SE, Yatsu F, et al. Carotid endarterectomy and prevention of cerebral ischemia in symptomatic carotid stenosis. Veterans Affairs Cooperative Studies Program 309 Trialist Group. *JAMA* 1991;266:3289–94.

4 Sitzer M, Müller W, Siebler M, et al. Plaque ulceration and lumen thrombus are the main sources of cerebral microemboli in high-grade internal carotid artery stenosis. *Stroke* 1995;26:1231–3.

5 Spagnoli LG, Mauriello A, Sangiorgi G, et al. Extracranial thrombotically active carotid plaque as a risk factor for ischemic stroke. *JAMA* 2004;292:1845–52.

6 Redgrave JN, Lovett JK, Gallagher PJ, et al. Histological assessment of 526 symptomatic carotid plaques in relation to the nature and timing of ischemic symptoms: the Oxford plaque study. *Circulation* 2006;113:2320–8.

7 Takaya N, Yuan C, Chu B, et al. Association between carotid plaque characteristics and subsequent ischemic cerebrovascular events: a prospective assessment with MRI - initial results. *Stroke* 2006;37:818–23.

8 Altaf N, Daniels L, Morgan PS, et al. Detection of intraplaque hemorrhage by magnetic resonance imaging in symptomatic patients with mild to moderate carotid stenosis predicts recurrent neurological events. *J Vasc Surg* 2008;47:337–42.

9 Kwee RM, Van Oostenbrugge RJ, Mess WH, et al. MRI of carotid atherosclerosis to identify TIA and stroke patients who are at risk of a recurrence. *J Magn Reson Imaging* 2013;37:1189–94.

10 Wintermark M, Arora S, Tong E, et al. Carotid plaque computed tomography imaging in stroke and nonstroke patients. *Ann Neurol* 2008;64:149–57.

11 Van Engelen A, Wannarong T, Parraga G, et al. Three-dimensional carotid ultrasound plaque texture predicts vascular events. *Stroke* 2014;45:2695–701.

12 Chai CK, Akyildiz AC, Speelman L, et al. Local axial compressive mechanical properties of human carotid atherosclerotic plaques-characterisation by indentation test and inverse finite element analysis. *J Biomech* 2013;46:1759–66.

13 National Research Council. Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease. Washington, DC: National Academies Press; 2011.

14 Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015;372:793–5.

15 Obama B. Remarks by the President on Precision Medicine. Washington, DC: The White House. State of the Union speech, January 30, 2015. Available from: https://www.whitehouse.gov/the-press-office/2015/01/30/remarks-president-precision-medicine. Accessed August 30, 2015.
16 Truijman MT, Kooi ME, Van Dijk AC, et al. Plaque At RISK (PARISK): prospective multicenter study to improve diagnosis of high-risk carotid plaques. *Int J Stroke* 2014;9:747–54.
17 Bayer-Karpinska A, Schwarz F, Wollenweber FA, et al. The carotid plaque imaging in acute stroke (CAPIAS) study: protocol and initial baseline data. *BMC Neurol* 2013;13:201.
18 Buisman LR, Rutten-van Mölken MPMH, Postmus D, et al. The early bird catches the worm: Early cost-effectiveness analysis of new medical tests. *Int J Technol Assess Health Care* 2016;32(1):1–8. Epub 2016 Mar 22.
19 Dutch Institute for Healthcare Improvement CBO. Guideline: diagnostics, treatment and healthcare for patients with a stroke (in Dutch: Richtlijn: Diagnostiek, behandeling en zorg voor patiënten met een beroerte). Utrecht: Dutch Institute for Healthcare Improvement CBO; 2008.
20 Buisman LR, Rijnsburger AJ, Den Hertog HM, et al. Clinical Practice Variation Needs to be Considered in Cost-Effectiveness Analyses: A Case Study of Patients with a Recent Transient Ischemic Attack or Minor Ischemic Stroke. *Appl Health Econ Health Policy* 2016:14(1):67–75.
21 Statistics Netherlands (in Dutch: Centraal Bureau voor de Statistiek, CBS). CBS Statline. Available from: http://statline.cbs.nl. Accessed February 1, 2014.
22 Wardlaw JM, Chappell FM, Best JJ, et al. Non-invasive imaging compared with intra-arterial angiography in the diagnosis of symptomatic carotid stenosis: a meta-analysis. *Lancet* 2006;367:1503–12.
23 Tholen ATR, De Monyé C, Genders TSS, et al. Suspected Carotid Artery Stenosis: Cost-effectiveness of CT Angiography in Work-up of Patients with Recent TIA or Minor Ischemic Stroke. *Radiology* 2010;256:585–97.
24 Hallan S, Asberg A, Indredavik B, et al. Quality of life after cerebrovascular stroke: a systematic study of patients' preferences for different functional outcomes. *J Intern Med* 1999;246:309–16.
25 Sullivan PW, Ghushchyan V. Mapping the EQ-5D index from the SF-12: US general population preferences in a nationally representative sample. *Med Decis Making* 2006;26:401–9.
26 Gattellari M, Goumas C, Garden F, et al. Relative survival after transient ischaemic attack: results from the Program of Research Informing Stroke Management (PRISM) study. *Stroke* 2012;43:79–85.
27 European Central Bank. ECB: Euro foreign exchange reference rates. Available from: https://www.ecb.europa.eu/stats/exchange/eurofxref/html/index.en.html. Accessed September 4, 2015.

28 Dutch Healthcare Authority (in Dutch: Nederlandse Zorgauthoriteit, NZa). NZa Zorgproducten Tariefapplicatie. Available from: http://dbc-zorgproducten-tarieven.nza.nl/nzaZpTarief/ZoekfunctieDot.aspx. Accessed March 3, 2015.
29 Buisman LR, Tan SS, Nederkoorn PJ, et al. Hospital costs of ischemic stroke and TIA in the Netherlands. *Neurology* 2015;84:2208–15.
30 Buskens E, Nederkoorn PJ, Buijs-Van Der Woude T, et al. Imaging of carotid arteries in symptomatic patients: cost-effectiveness of diagnostic strategies. *Radiology* 2004;233:101–12.
31 Struijs JN, Van Genugten ML, Evers SM, et al. Future costs of stroke in the Netherlands: the impact of stroke services. *Int J Technol Assess Health Care* 2006;22:518–24.
32 Rothwell PM, Gutnikov SA, Warlow CP. Reanalysis of the final results of the European Carotid Surgery Trial. *Stroke* 2003;34:514–23.
33 U-King-Im JM, Hollingworth W, Trivedi RA, et al. Cost-effectiveness of diagnostic strategies prior to carotid endarterectomy. *Ann Neurol* 2005;58:506–15.
34 North American Symptomatic Carotid Endarterectomy Trial Collaborators. Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis. *N Engl J Med* 1991;325:445–53.
35 Barnett HJ, Taylor DW, Eliasziw M, et al. Benefit of carotid endarterectomy in patients with symptomatic moderate or severe stenosis. North American Symptomatic Carotid Endarterectomy Trial Collaborators. *N Engl J Med* 1998;339:1415–25.
36 Sillesen H, Amarenco P, Hennerici MG, et al. Atorvastatin reduces the risk of cardiovascular events in patients with carotid atherosclerosis: a secondary analysis of the Stroke Prevention by Aggressive Reduction in Cholesterol Levels (SPARCL) trial. *Stroke* 2008;39:3297–302.
37 Milionis HJ, Giannopoulos S, Kosmidou M, et al. Statin therapy after first stroke reduces 10-year stroke recurrence and improves survival. *Neurology* 2009;72:1816–22.
38 Bonati LH, Dobson J, Featherstone RL, et al. Long-term outcomes after stenting versus endarterectomy for treatment of symptomatic carotid stenosis: the International Carotid Stenting Study (ICSS) randomised trial. *Lancet* 2015;385:529–38.
39 College voor zorgverzekeringen (CVZ). Guidelines for pharmacoeconomic research, updated version. Diemen: College voor zorgverzekeringen; 2006.
40 Jauch EC, Saver JL, Adams HP, et al. Guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 2013;44:870–947.

8.8 Appendix

Appendix 8.1: Model input parameters

Parameter		Value	SE	Distribution (alpha;beta)	Source
PRIOR PROBABILITY OF HAVING AN ISCHEMIC STROKE IF PATIENTS RECEIVE OMT ALONE					[1,2]
70-99% carotid stenosis		0.216	0.028	Beta (47;171)	
RISK OF ISCHEMIC STROKE					[1,2]
70-99% STENOSIS – OMT ALONE					
Short-term outcome	- minor stroke	0.005	0.005	Beta (1;217)	
	- major stroke	0.000	0.000	Beta (0;217)	
	- fatal stroke	0.005	0.005	Beta (1;217)	
Long-term outcome	- minor stroke	0.169	0.044	Beta (12;61)	
	- major stroke	0.100	0.035	Beta (7;66)	
	- fatal stroke	0.055	0.027	Beta (4;69)	
50-69% STENOSIS – OMT ALONE					
Short-term outcome	- minor stroke	0.008	0.005	Beta (2;263)	
	- major stroke	0.004	0.004	Beta (1;264)	
	- fatal stroke	0.000	0.000	Beta (0;0)	
Long-term outcome	- minor stroke	0.093	0.024	Beta (14;136)	
	- major stroke	0.098	0.024	Beta (15;135)	
	- fatal stroke	0.056	0.019	Beta (8;142)	
0-49% STENOSIS – OMT ALONE					
Short-term outcome	- minor stroke	0.006	0.003	Beta (4;717)	
	- major stroke	0.001	0.001	Beta (1;720)	
	- fatal stroke	0.001	0.001	Beta (1;720)	
Long-term outcome	- minor stroke	0.078	0.011	Beta (42;501)	
	- major stroke	0.043	0.009	Beta (23;520)	
	- fatal stroke	0.025	0.007	Beta (14;529)	
70-99% STENOSIS – CEA					
Short-term outcome	- minor stroke	0.012	0.006	Beta (4;330)	
	- major stroke	0.030	0.009	Beta (10;324)	
	- fatal stroke	0.003	0.003	Beta (1;333)	
Long-term outcome	- minor stroke	0.081	0.019	Beta (17;192)	
	- major stroke	0.071	0.018	Beta (15;194)	
	- fatal stroke	0.030	0.012	Beta (6;203)	

Appendix 8.1: Model input parameters (continued)

Parameter		Value	SE	Distribution (alpha;beta)	Source
50-69% stenosis – CEA					
Short-term outcome	- minor stroke	0.037	0.010	Beta (14;363)	
	- major stroke	0.021	0.007	Beta (8;369)	
	- fatal stroke	0.011	0.005	Beta (4;373)	
Long-term outcome	- minor stroke	0.087	0.020	Beta (18;189)	
	- major stroke	0.058	0.016	Beta (12;193)	
	- fatal stroke	0.026	0.011	Beta (5;200)	
0-49% stenosis – CEA					
Short-term outcome	- minor stroke	0.023	0.005	Beta (25;1,055)	
	- major stroke	0.014	0.004	Beta (15;1,065)	
	- fatal stroke	0.006	0.002	Beta (6;1,074)	
Long-term outcome	- minor stroke	0.102	0.012	Beta (63;555)	
	- major stroke	0.068	0.010	Beta (42;576)	
	- fatal stroke	0.035	0.007	Beta (22;596)	
Risk reduction of ischemic stroke from the ECST-1 study to reflect the use of statins					[3-5]
Patients who underwent CEA		0.33	0.01	Beta (335;671)	
Patients who receive OMT alone		0.33	0.01	Beta (335;671)	

SE = standard error; OMT = Optimal medical treatment, CEA = carotid endarterectomy.

References Appendix 8.1

1 Rothwell PM, Gutnikov SA, Warlow CP. Reanalysis of the final results of the European Carotid Surgery Trial. *Stroke* 2003;34(2):514–23.

2 U-King-Im JM, Hollingworth W, Trivedi RA, et al. Cost-effectiveness of diagnostic strategies prior to carotid endarterectomy. *Ann Neurol* 2005;58(4):506–15.

3 Sillesen H, Amarenco P, Hennerici MG, et al. Atorvastatin reduces the risk of cardiovascular events in patients with carotid atherosclerosis: a secondary analysis of the Stroke Prevention by Aggressive Reduction in Cholesterol Levels (SPARCL) trial. *Stroke* 2008;39(12):3297–302.

4 Milionis HJ, Giannopoulos S, Kosmidou M, et al. Statin therapy after first stroke reduces 10-year stroke recurrence and improves survival. *Neurology* 2009;72(21):1816–22.

5 Bonati LH, Dobson J, Featherstone RL, et al. Long-term outcomes after stenting versus endarterectomy for treatment of symptomatic carotid stenosis: the International Carotid Stenting Study (ICSS) randomised trial. *Lancet* 2015;385:529–38.

CHAPTER 9

GENERAL DISCUSSION

9.1 Introduction

Novel medical tests can yield valuable information for use in various phases of disease prevention and treatment. This information can help to individualize prevention and treatment strategies based on people's unique characteristics, including genetics, microbiome composition, lifestyle, health history, and exposure to environment factors. This enhancement in patient stratification reflects part of the promise of precision medicine.[1-3] The increasing development of medical tests with various applications and target populations in different disease areas calls for early cost-effectiveness analyses (early-CEAs) to assess whether these tests have the potential to improve patient care, patient health outcomes, and reduce overall costs. The aim of early-CEAs is to improve the ability to develop effective and cost-effective medical tests in an efficient manner while considering the interests of different stakeholders (i.e., medical test manufacturers/developers (referred to as test developers), clinicians (as end-users of the tests), patients (as the main beneficiaries), hospital managers, and payers). This chapter provides an overview of the main findings of this thesis, the implications of the results of early-CEAs of medical tests for the different stakeholders, the methodological challenges of early-CEAs of medical tests, and recommendations for further research.

9.2 Main findings

Since little specific guidance on early-CEAs of medical tests was available, we developed a framework with the general steps of conducting early-CEAs of new medical tests. This framework consists of five general steps which can be applied to any new test that is being developed. A main characteristic of early-CEAs of medical tests is that the general steps are iterative, meaning that new insights and evidence that arise during the test development are continuously integrated and may convince test developers to return to ear-

lier development steps. The framework with general steps of early-CEAs of new medical tests was applied to two case studies in this thesis. The first case study focused on the evaluation of current and new diagnostic test strategies in the early diagnosis of rheumatoid arthritis (RA) to select patients for early start of optimal treatment. The second case study focused on the assessment of current and new prognostic imaging test strategies to improve ischemic stroke risk prediction and thereby improve decisions about which patients should undergo surgery (i.e., carotid endarterectomy). The main findings from both case studies will be discussed in the following sections.

Diagnostic tests for RA

We provided insight in the costs of diagnostic tests, which tests were requested, what differential diagnoses were considered, and what strategies were used by rheumatologists in patients with early inflammatory arthritis at risk for RA. A median of 25 [IQR 20-35] tests were requested during first visit at the rheumatologist in 2010 with mean costs of €422 (SD: €168) per patient. This information was used in the cost-effectiveness model.

We assessed the potential cost-effectiveness of new diagnostic tests in the work-up of patients with inflammatory arthritis at risk of having RA. The cost-effectiveness of four new diagnostic tests (B-cell gene expression, MRI, IL-6 serum level test, and genetic assay) was assessed, when these tests were either added to or replaced the current test strategy (i.e., ACR/EULAR 2010 RA classification criteria) in different target populations. The target populations were either: 1) all inflammatory arthritis patients, 2) intermediate-risk patients (3-5 points on the ACR/EULAR 2010 RA classification criteria), or 3) seronegative patients. We decided to present the short term cost-effectiveness of the new tests separately from the medium term cost-effectiveness because the treatment and health outcomes as a consequence of the test result have a larger impact on the cost-effectiveness of a test strategy than the test itself. Therefore, we used a one-year time horizon to assess the short term impact of the new test strategy, and a five-year time horizon incorporating the treatment effects and health outcomes to assess the medium term impact.

When using a one-year time horizon, diagnostic effectiveness [i.e., unweighted diagnostic Net Benefit (udNB)*], incremental cost-effectiveness ratios (ICERs), and the headroom of a new test were estimated for each new test strategy versus the ACR/EULAR 2010 RA classification criteria. The highest udNB was found when using the B-cell gene expression test in intermediate-risk patients (42.9%; ICER €5,314), while the IL-6 serum

* udNB = the change in the true positive (TP) rate minus the change in the false positive (FP) rate, divided by the number of patients that underwent the test.

level test in seronegative patients resulted in the lowest udNB (-11.4%; ICER €7,650). The negative udNB is a result of a larger increase of FPs compared to TPs. If a willingness-to-pay threshold of €20,000 per QALY gained is used, the B-cell gene expression test would be favored over the other alternatives with a probability of being cost-effective of 78% for intermediate-risk patients, 57% for all patients and 73% for seronegative patients. For the four different tests, the available headroom varied between €151 and €369 depending on the test characteristics and population to test given a willingness-to-pay threshold of €20,000 per QALY gained.

When using a five-year time horizon, we modelled that patients with DAS28>3.2 could switch to more expensive biologic DMARDs. The differences in ICERs between the new test strategies resulted in a clearer answer to the question which test strategy is potentially most cost-effective than in the one-year analysis. Of all add-on test strategies, we found that the B-cell gene expression test as add-on for intermediate-risk patients was the most cost-effective add-on strategy (it was the dominant strategy with improved health outcomes and lower costs than the current test strategy). MRI, IL-6 serum level test, and genetic assay were not cost-effective for all add-on test strategies at a willingness-to-pay threshold of €20,000 per QALY gained. The probability of cost-effectiveness for a B-cell gene expression test versus the alternative tests is 100% if a willingness-to-pay threshold of €20,000 per QALY gained is used. Moreover, a new add-on test for intermediate-risk patients should have sufficiently high specificity because the current ACR/EULAR 2010 RA classification criteria are sufficiently capable of ruling out RA and the test costs should not be higher than €200-€300.

Prognostic tests for recurrent ischemic stroke

We investigated the use of current test strategies in the secondary prevention of patients with a recent transient ischemic stroke (TIA) or minor ischemic stroke and deviation from the Dutch national guidelines. Of the hospitals studied, about 60% were found to use initial and confirmatory diagnostic tests to assess carotid stenosis in accordance with the national guidelines; the other hospitals used various other test combinations. According to vascular neurologists, the clinical practice variation in test strategies arises due to varying degrees of expertise in performing tests, patient case-mix, clinical reasons, financial incentives, logistics, availability of imaging technology, and preferences of radiologists, vascular surgeons and vascular neurologists. By means of a case study, we illustrated the importance of performing hospital-level CEAs by examining the cost-effectiveness of the hospital's current test strategy versus the strategies used in other hospitals while varying the test costs, sensitivity and specificity according to local hospital conditions. When hospital-level values were used, we found that the most cost-effective strategy differed between hospitals. In the base case analysis, the CTA-only

and CE-MRA-only strategies were the most cost-effective strategies. However, the results changed when hospital-level values for unit costs and test performance were used. The CTA-only strategy was the dominant strategy in the first hospital, since it had the lowest costs and highest QALYs of all strategies, while the CE-MRA-only strategy was the dominant one in the second hospital.

We performed the first and most extensive cost analysis of inpatient and outpatient hospital costs of ischemic stroke and TIA patients since 10 years and used resource use data of all hospitalizations in the Netherlands in 2010. The study showed that hospital costs were higher for inpatients and ischemic strokes compared with outpatients and TIAs, with length of stay (LOS) as the most important contributor. The cost difference between inpatient ischemic stroke (€5,328) and inpatient TIA (€2,470) was caused by a shorter LOS for TIA patients (3.6 days for TIA versus 8.8 for ischemic stroke). Costs of outpatient care were higher for TIA than for ischemic stroke due to more frequent use of ambulatory treatment, and diagnostic and imaging tests. The regression models show that age, hospital type, and region were strongly associated with hospital costs.

We evaluated the potential cost-effectiveness of a new prognostic imaging test to select patients with a recent TIA or minor ischemic stroke that benefit most from surgery (i.e., carotid endarterectomy). The new test was used as a confirmatory test for patients with a 30-69% carotid stenosis based on an initial duplex ultrasonography (DUS). Although the sensitivity, specificity and costs of the confirmatory test were unknown, an early-CEA can even be helpful at this stage since the results can help to guide further test development. In four different patient subpopulations, we estimated the maximum health gain that a confirmatory test could achieve if it was perfect (100% sensitivity and specificity) as well as the minimum prognostic sensitivity and specificity that a new confirmatory test must have in order to be cost-effective versus the guideline-based strategy. Our analyses showed that the maximum health gain for a confirmatory test may be limited. The new test strategy with new confirmatory test yields no health gain for 60-year-old and 80-year-old women and the QALY gain for 60-year-old and 80-year-old men is 0.066 and 0.022 QALY, respectively. In addition, the study showed that the prognostic sensitivity and specificity for 80-year-old men must be higher than for 60-year-old men and that even a perfect new confirmatory test is not cost-effective in 60-year-old and 80-year-old women because it leads to poorer health outcomes and higher costs. The negative outcomes for women arise because the confirmatory test is only used in women with a 30-69% carotid stenosis based on DUS. Our analysis showed that a confirmatory test for women with a 30-69% carotid stenosis based on DUS is not cost-effective. A more accurate confirmatory test for women with a 70-99% stenosis may be more likely to be cost-effective.

9.3 Implications of the results of early-CEAs of medical tests for different stakeholders

Early-CEAs performed during the development of medical tests have important benefits for different stakeholders: test developers, clinicians (as end-users of the tests), patients (as the main beneficiaries), hospital managers, and payers including healthcare insurers, governments, and managed care organizations. Table 9.1 summarizes the interests of the different stakeholders which are further described below.

Table 9.1: Interests of the different stakeholders of early-CEAs of medical tests

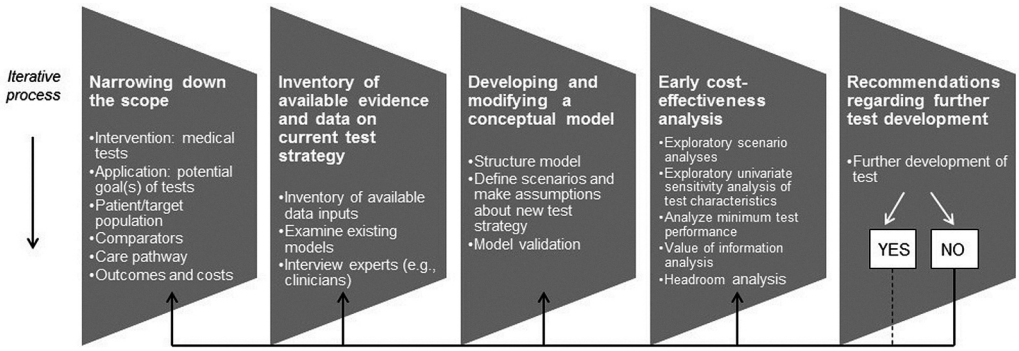
STAKEHOLDER	INTERESTS
Medical test manufacturers/ developers	Guidance on further development of tests
	Implementation of test in clinical practice
	Return on investment
	Design and manage a reimbursement strategy
Clinicians	Faster implementation of new tests
	Signal potential (sub)populations to test
	Impact on development of early-CEA model and tests
Patients	Including patients in development process to guarantee patients full benefits of new tests
	Faster access to medical innovations
	Reduction of uncertainty about diagnosis/prognosis by new tests
	Reduction of time before starting treatment
	Reduction of invasiveness of testing
Hospital managers	Budgetary possibilities to implement test
	Hospital efficiency
Payers	Active horizon scanning: *involved* in development of tests
	Passive horizon scanning: *informed* about the development of tests

Medical test developers

Early-CEAs provide important information to **test developers** to guide the further development of tests.[4-8] At the very early stages of test development, the potential application of the test and use in patient care might not exactly be known. At this stage, early-CEAs can help test developers to explore where in the healthcare system (i.e., in which target population, and in which setting including preventive, acute, or secondary care) a potential new test can be used to improve the selection of people for individualized prevention and/or treatment strategies, aiming to improve patient health outcomes. Throughout the different development stages of a test, early-CEAs can guide test developers in their decision to further develop the test.

Figure 9.1 shows the five general steps of early-CEAs of medical steps and the iterative character as described in Chapter 2. New insights, data, and evidence arise during test development and should continuously be integrated in an early-CEA to provide more informed decisions about the potential value of the new test in patient care, which may lead test developers to return to earlier steps of the early-CEA.

Figure 9.1: Framework with general steps of early-CEAs of medical tests



Steps of early-CEA were defined by the authors.

In both case studies presented in this thesis, we collaborated with test developers to discuss the potential use of the tests in different patient (sub)populations, the positioning of the new tests in the entire test strategy (i.e., as replacement test, add-on test, or confirmatory test), and the potential test performance (i.e., sensitivity and specificity). We presented and discussed our preliminary cost-effectiveness results with the test developers during the development process. In both case studies, the process was less iterative than anticipated due to unawareness among test developers of early-CEAs (see below), time to develop the early-CEA models, time available for building the early-CEA models, and limited availability of clinical data from usual care cohorts. It would have been more beneficial to perform more iterative early-CEAs to guide test development at the different stages. In general, the greater the interaction between test developers and early-CEA researchers, the more the analyses can be targeted to the specific test applications that the test developer is considering.

We experienced unawareness among test developers of the importance of performing early-CEAs of medical tests and the importance of cost-effectiveness results for implementing new tests in clinical practice. The po-

tential cost-effectiveness might not be seen as one of the most important criteria to further develop a test by developers. We experienced that factors such as the technical ability of the test, the clinical effectiveness, and safety have a higher priority than the potential cost-effectiveness of the test. However, the importance of cost-effectiveness is illustrated by the example of the development of digital X-ray radiogrammetry (DXR). After the development of DXR, rheumatologists were unwilling to use this in clinical practice because evidence on the effectiveness and cost-effectiveness was missing and the test costs were about €400. From a test developer perspective, early-CEA results are important to receive a sufficient return on investment. Therefore, the interaction between test developers and early-CEA researchers and the awareness of early-CEAs among test developers are important during medical test development. The test developers may contribute to collecting data that should be used by early-CEA researchers.

Clinicians

Early-CEAs provide important information to **clinicians** about in which patient (sub)populations a test is potentially cost-effective.[4-8] Such evaluations are valuable for clinicians as the potential end-users to signal promising new tests in specific patient populations during test development. An early-CEA can convince clinicians about the potential improvements of patient care and health outcomes which may result in faster take-up of tests in clinical practice. Clinicians are the key to actual implementation of tests in clinical practice. New tests will not be implemented without approval from clinicians. However, in current clinical practice, there is a knowledge gap between clinicians who see patients and laboratory personnel who develop new tests. It is important that opinion leaders among clinicians who know both worlds bridge the knowledge gap between clinicians and laboratory personnel. The role of key opinion leaders among clinicians would also ensure that a new test is considered to be incorporated in the clinical guidelines when it is found to be cost-effective.

Moreover, during the development of an early-CEA model, input and feedback from clinicians is crucial to build and validate the model. Clinicians can have an impact on the development of an early-CEA model and test development by providing feedback on the shortcomings of current methods to select patients for specific treatments. Clinicians can also help designing a new test strategy by proposing a new test strategy, indicating whether a proposed test strategy will work in practice, or anticipating the potential implementation issues of a proposed test strategy. In addition, clinicians can also provide their expert opinion about the plausibility of assumptions.

In the case of RA, no single test is currently used to diagnose patients for RA. Instead, the ACR/EULAR RA 2010 classification criteria are used, which are a combination of criteria, including joint involvement (number

of swollen and painful joints), blood test results, and duration of symptoms, in combination with the use of DMARDs not explained by any other disease. This is a common way of dealing with a disease for which no hallmark sign is available.[9,10] Given the tests that are currently being developed, rheumatologists mentioned that it is unlikely that a replacement test would be discovered that provides the same richness of information as the current test strategy (i.e., ACR/EULAR RA 2010 classification criteria) in diagnosing inflammatory arthritis patients at risk of having RA. Therefore, rheumatologists recommended evaluating the cost-effectiveness of new tests as an add-on to the current diagnostic test strategy for specific target populations at risk of having RA.

Chapter 4 and 5 presented two early-CEAs in which the cost-effectiveness of the add-on test strategies in different populations at risk was evaluated. Moreover, the impact of the new test's sensitivity, specificity and costs on the cost-effectiveness was evaluated. These studies indicated the most likely position of a new diagnostic test given the test's sensitivity, specificity and costs in relation to the current ACR/EULAR RA 2010 classification criteria test strategy. B-cell gene expression test as add-on for *intermediate-risk patients* was the most cost-effective add-on strategy given its moderate sensitivity (0.60), high specificity (0.90) and low costs (€150). Since the specificity of the current ACR/EULAR RA 2010 classification criteria is already high (0.70-0.90), a new add-on test for intermediate-risk patients should have sufficiently high specificity because all intermediate-risk patients based on the ACR/EULAR RA 2010 classification criteria are classified as TN or FP.

In current clinical practice, the ACR/EULAR RA 2010 classification criteria are not used as risk stratification tool by rheumatologists. Patients with an intermediate-risk of having RA according to the criteria are treated equally as patients with a low risk. Based on the results from our early-CEAs which showed that additional testing of intermediate-risk patients is cost-effective, we recommend rheumatologists to use the criteria as a risk stratification tool.

In the case of new prognostic tests to assess the risk of a recurrent ischemic stroke, both vascular neurologists and radiologists provided useful input. Since the new tests are currently being developed, the radiologists were interested in knowing what the minimum prognostic sensitivity and specificity of a new test must be in order to be cost-effective. Therefore, these minimum values were estimated using our early-CEA model. The vascular neurologists provided useful input about optimal combinations of tests and the positioning in clinical care. In their opinion, the most likely application for a new test to predict the risk of recurrent ischemic stroke due to plaque rupture in patients with a recent TIA or minor ischemic stroke would be as a confirmatory test for patients with a 30-69% stenosis based on an

initial DUS. DUS is the preferred initial test since it is often used in current care, relatively cheap and simple to use. Chapter 8 showed that the minimum required prognostic sensitivity and specificity of a confirmatory test differs between patient subpopulations and the minimum is not achievable in all patient subpopulations using the guideline-based strategy as comparator. In addition, the vascular neurologists also proposed a new prognostic test with high sensitivity and specificity as initial test. However, this test strategy would only be cost-effective if the test costs are not too high because the test would be used in all patients with a recent TIA or minor ischemic stroke.

Patients

The involvement of **patients** (by means of patient representatives) during test development would also be important. The patients will be the main beneficiaries of new medical tests. With the involvement of patients in early-CEAs of medical tests, patients can help to identify the shortcomings and lack of information of current tests in the selection of patients for specific treatments. An advantage of involving patients in de development process is that they may help to speed-up the process to enable sooner market access of medical innovations. In the case of diagnosing RA, inflammatory arthritis patients normally cope with an extensive period of uncertainty about whether the joint complaints are caused by RA or any other differential diagnosis. During this period, these patients are tested, re-tested and could switch medication either due to side-effects or if an incorrect diagnosis was made. This period of uncertainty can be a huge burden to patients. Patients expect a quick resolution of complaints. A potential improvement would be diagnostic tests that enable a diagnosis during the first visit with the rheumatologist since this would speed up the diagnostic process. New tests or improvements to current tests (e.g., reduced time to get ACCP, RF, or Rontgen test result) can help to make a quick diagnosis and reduce the period of uncertainty about the actual diagnosis. An earlier diagnosis of RA would result in earlier start of appropriate treatment and improve patient's quality of life. In addition, a reduction of the diagnostic uncertainty with the use of improved tests will likely reduce direct medical and productivity costs among patients who are at intermediate-risk according to the ACR/EULAR 2010 RA classification criteria.

Another interest of patients would be to reduce the invasiveness of testing. One example from our cases is digital subtraction angiography (DSA). DSA is the gold standard in diagnosing carotid artery stenosis in patients with a recent TIA or minor ischemic stroke. However, this test can cause a stroke by the catheter that is used during the procedure, it is invasive (thereby leading to discomfort to patients), and it is expensive. Therefore, the development of new tests is focused on non-invasive tests that have the same richness of information as the invasive test.

When involving patients, time and money should be invested in training patients to a certain level that they can contribute to the process of test development. We expect that patients can provide a useful contribution to the development of a medical test by sharing their interests and concerns regarding the safety, costs, impact on quality of life, and waiting time before getting the test result. The expectation would also be that involving patients would increase the acceptance of using tests in clinical practice and that model results are understood and seen as reliable.

Hospital managers

The engagement of **hospital managers** in stakeholder meetings during new test development would help to identify the budgetary possibilities of hospitals to implement new tests and barriers to implementation. In Europe, medical tests are most often reimbursed as part of a Diagnosis Related Group payment system. A good working DRG system would result in efficient use of resources. For hospital managers, the incentive for hospital efficiency in a DRG system is that hospitals can keep a positive difference between what is reimbursed for a DRG and what are the actual costs. From the perspective of a hospital manager, the test with the lowest cost would be recommended unless a new test is shown to be superior over its alternatives in terms of patient outcomes. In many cases, the heads of the different hospital departments, who are most often clinicians, are the hospital managers because each department needs to manage their own department's budget. As a result, the clinicians decide about the use of tests. One example would be a new prognostic imaging test with higher sensitivity and specificity but also higher costs than the test that is currently used to select patients with a recent TIA or minor ischemic stroke for surgery. This new test may result in improved treatment decisions and consequently a shorter hospital length of stay. The decision to implement this new prognostic test is made by the heads of the neurology and radiology departments taking into consideration that the test improves health outcomes and reduces overall costs due to a shorter hospital length of stay.

In addition, the implementation of a new medical test is only possible if hospital conditions can readily be modified. The use of a new test requires that clinicians and other healthcare personnel are trained to use the test. Furthermore, a new medical test might not always be cost-effective in all hospitals. Smaller hospitals in remote areas might have a smaller number of patients requiring the test and the associated high implementation and maintaining costs will result in negative recommendations from managers in smaller hospitals to use the test (i.e., economies of scale). For example, in the Netherlands, it would not be cost-effective to purchase, use, and maintain a MRI or CT scan in small hospitals with few patients requiring the test. Instead, patients are referred to larger (most often academic) hos-

pitals in the region. Viewed from the perspective of a small hospital, the current test strategy in small hospitals may be more cost-effective than the use of a highly accurate and expensive new test strategy. In countries like the United States and Canada with large distances between hospitals, the conclusion about using a MRI or CT scan in a small hospital might be different when the traveling costs of patients are incorporated in the early-CEA model.

Payers

The results of early-CEAs are also important for **payers**, because they decide in collaboration with key opinion leaders among clinicians about the reimbursement of new tests. As mentioned above, medical tests are most often reimbursed as part of a Diagnosis Related Group payment system in Europe. Although decision-making about reimbursement is usually considered when the test is mature and possibly already implemented in clinical practice, published early-CEA results can help payers to signal promising new tests that are in the pipeline, resulting in timely and better informed decisions about reimbursement (i.e., passive horizon scanning). Payers can also actively be involved in the development of new tests by participating in stakeholder meetings and discussing the potential impact of new tests on the national or regional budget (i.e., active horizon scanning). A budget impact analysis (BIA) can help payers to assess what the potential impact of a new test will be on the national or regional budget.[11]

Nowadays, new tests can essentially enter the market without any decisions about reimbursement. Clinicians can decide to use a test when they believe it can improve effectiveness. The test will then automatically be reimbursed as part of the current DRG payment, unless a request is submitted to increase the price of the DRG. A more evidence-based reimbursement of particularly the expensive medical tests is recommended

The involvement of the different stakeholders in early-CEAs of medical tests and dissemination and discussion of the results are crucial and should continue during test development by means of regular meetings with stakeholders. For the successful development and implementation of new medical tests, input from the different stakeholders with both shared and conflicting interests is important. The results of early-CEAs should then be discussed with the stakeholders and additional feedback and interests should be taken into account to further develop the medical test. Early-CEAs are therefore helpful to prioritize the interests and needs of the different stakeholders.

9.4 Methodological challenges

Application of a new medical test

Both the test developers and clinicians as the end-users of the test should assess where and how (e.g., diagnostic or prognostic) the new test has the potential to improve patient care and patient health outcomes. However, the application of a new test may not always be clear from the start of the development process. Chapter 2 described a framework that we developed with general steps of conducting early-CEAs of medical tests. The assessment of the application of a new test is part of the first step in which the Patient Population, Intervention, Comparator, and Outcomes (PICO) method[12] was extended to the Application, Patient Population, Comparator, Outcomes, Intervention (APCOI) method.[8]

In both case studies presented in this thesis, the application of the new tests was already clear to some extent when the first early-CEAs were performed (i.e., diagnostic tests for the early diagnosis of RA and prognostic tests to predict the risk of recurrent ischemic stroke based on plaque rupture in the carotid artery). However, in many instances the application may not be clear until much later in the development of the test. For example, when engineers try to optimize the technical ability of a medical test, the new application of the test, its potential use, and usefulness in clinical practice may not be known at that time. Therefore, the interests of stakeholders in a new test should be determined when developing a new test, including the interests of clinicians, hospital managers, healthcare insurers, policymakers, and patients. Therefore, stakeholder meetings with discussions about the potential application and target population are essential during test development.

Choice of comparator

The choice of comparator in any CEA is important since it determines the estimated impact of the new strategy. Therefore, an extensive assessment of the comparators is one of the first crucial steps, as emphasized in our framework of conducting early-CEAs of medical tests presented in Chapter 2. The test strategy recommended in clinical guidelines is often used as the comparator in a CEA. However, the use of guidelines as a proxy of current care is inadequate when clinical practice differs from the guidelines or when clinical practice differs between hospitals. If *important* practice variation or deviation from the guidelines exists, hospital-level CEAs should be performed which compare the care that is actually provided in hospitals by using multiple comparators and incorporate local hospital conditions (e.g., hospital-specific costs of tests).[13] Practice variation or deviation from the guidelines becomes important when the different test strategies have varying costs and health effects.

By means of a case study, Chapter 6 illustrated the importance of performing hospital-level CEAs when practice variation exists. The most cost-effective test strategy for patients with a recent TIA or minor ischemic stroke differed between hospitals when hospital-level values of sensitivity, specificity and costs of tests were used. Hospital-level CEAs are important for at least two stakeholders: hospitals and payers. Such CEAs will help individual hospitals to explore the potential to improve effectiveness, cost-effectiveness, and budgetary situation by implementing a different strategy. Payers (i.e., healthcare insurers, governments, and managed care organizations) might want to use hospital-level CEAs to determine how costly and cost-effective the care currently provided in hospitals is compared to the most cost-effective strategy available, and assess what the impact is on the hospital budget (i.e., BIA). The case study of hospital-level CEAs in Chapter 6 showed that the average five-year costs per patient with a recent TIA or minor ischemic stroke range from €15,862 to €17,145 between test strategies in one hospital. While this range may seem small, its effect on the overall budget may be important depending on the annual volume of patients.[11] For example, if this hospital was to assess 500 patients per year (i.e., the average number of patients with a TIA or ischemic stroke per hospital in the Netherlands in 2012),[14,15] the total five-year costs would range from €7,930,779 to €8,572,496, meaning a difference of €641,717.

In Chapter 3, we have shown that no single test strategy was used by rheumatologists to diagnose patients for RA in current clinical practice. When patients with early inflammatory arthritis visit a rheumatologist for the first time, multiple diagnoses are considered simultaneously by the rheumatologist. We found that rheumatologists request a median of 25 diagnostic tests (excluding standard laboratory tests) at first consultation. The choice of diagnostic tests greatly depends on which other diagnoses are being considered. In other words, two rheumatologists that independently evaluate the same patient can consider different diagnoses. This results in variation between rheumatologists in the use of tests. Therefore, it was not possible to define multiple comparators and perform hospital-level CEAs. As a result, we decided to include a mean diagnostic test cost per patient in the model which is an average of all diagnostic tests used for all inflammatory arthritis patients at risk of having RA.

Iterative nature of early-CEAs, uncertainty, and limited availability of data

At the different steps of early-CEAs of medical tests, limited data are typically available for use in early-CEA models. For example, the sensitivity, specificity, and costs of new tests are rarely known during test development. Moreover, robust evidence for the clinical and economic impact of a new test strategy in terms of improved patient care, patient health outcomes and costs is often missing or based on a small sample of patients and expert

opinion. Therefore, various assumptions on the new test strategy were defined for the early-CEAs of medical tests in this thesis. The limited availability of data results in more uncertainty around the cost-effectiveness estimates than in late-CEAs. In the RA case, univariate sensitivity analyses were performed to determine the impact of various model input parameters on the potential cost-effectiveness of a new test strategy. In the base case analysis, inflammatory arthritis patients that were diagnosed earlier with RA with the new test were assigned a lower disease activity of 0.2. This value was varied in the univariate sensitivity analyses. In addition, the test's sensitivity, specificity, and costs, and the percentage of false positive patients using biologic DMARDs were varied to determine the impact on the potential cost-effectiveness of the new test strategy. We learned from the univariate sensitivity analyses that seemingly small changes in disease activity have a large impact on the cost-effectiveness of a new add-on test for intermediate-risk patients, because a lower disease activity leads to delayed start of more expensive biologic DMARDs. If a MRI (sensitivity: 0.90, specificity: 0.60, costs: €756) as add-on test for intermediate-risk patients is used, the impact of changing the DAS28 improvement from 0.6 to no improvement resulted in a change in ICER from €2,729 to €55,654 per QALY gained.

During the process of test development, new data, insights, and evidence emerge which should continuously be integrated in early-CEA models. In this regard, a simple model during the early stages of development is useful and should continuously be updated if needed with data on important parameters and changes to the model structure to generate more robust cost-effectiveness estimates during the development process. One example from the RA case is that we initially assumed an overoptimistic improvement of health status due to timely treatment after using a new test. After discussions with rheumatologists we adjusted this in the model. This change had an important impact on the ICERs and headroom of a new test, which decreased from €2,062 to €417. This change illustrates the role of uncertainty in early-CEAs. By updating the early-CEA model, the impact of new evidence on the uncertainty around the cost-effectiveness estimates can be assessed and the uncertainty can potentially decrease.

Examples from the case of prognostic tests for ischemic stroke patients are updates of the input parameters in our early-CEA model. We initially modelled the risks of recurrent ischemic stroke for patients who underwent carotid endarterectomy and patients who received medical treatment alone from the ECST-study.[16,17] However, medical treatment has improved since the ESCT-study (e.g., widespread use of statins, better antiplatelet therapy, and lower targets for blood pressure control)[18-20] and has led to a relative reduction in the risk of ischemic stroke of approximately 33%. Therefore, we adjusted the estimates from the ECST-study with a risk reduction of 33% to reflect improved medical treatment. Another example is the update

of the unit costs of carotid endarterectomy based on our recent cost analysis[21] as presented in Chapter 7 of this thesis.

9.5 Recommendations for further research

What we have learned from both case studies is that the process of developing a cost-effectiveness model involves time to build the model, and that interaction with stakeholders (mainly clinicians and test developers) is crucial to develop an early-CEA model that includes a representation of current and future clinical care with the new test. The time that the process might take to perform an early-CEA is mostly underestimated. The processing time depends on multiple factors such as the availability of an appropriate cost-effectiveness model, the availability of clinical data from a usual care cohort, the level of expertise of the modeler, the time available by the modeler to perform the early-CEA, and the interaction with other stakeholders (e.g., clinicians) to provide input and validate the model and assumptions.

The advantages of early-CEAs are fully realized if they begin even sooner than when they started in the current case studies. The model used in this very early phase might be a simplified version given the limited data available, but can be updated and become more complex later. A disadvantage, therefore, might be that the uncertainty is high at the very early stages of test development, but the uncertainty will generally decrease by continuously updating the model. An advantage of having an early-CEA model at the very early stages of the development of tests is that it can be used to guide the further development of tests. Our framework with general steps of early-CEAs can be used for this purpose to support go/no go decisions about how to proceed further in new test development. The ability of this framework to guide the further development of tests should be investigated.

As mentioned above, we experienced unawareness among test developers of the importance of performing early-CEAs of medical tests and the importance of cost-effectiveness results for implementing new tests in clinical practice. It should be further investigated what techniques or methods may improve the involvement of stakeholders in early-CEAs.

In addition, a limitation of early-CEAs is the uncertainty around cost-effectiveness results due to a limited availability of data and assumptions on model input parameters and test scenarios. It would be useful to use our early-CEA models to perform value of information analyses to decide whether and which additional research is needed to inform the decision problem.[22-24] Given the high parameter uncertainty in early-CEAs of medical tests it is important to consider when during test development sufficient information is available to perform value of information analyses. In case of large parameter uncertainties and assumptions on key input parameters in the very early stages of test development (e.g., uncertainty about the sensi-

tivity and specificity of a new test and the extent of improved health outcomes due to optimal treatments), the uncertainty may be too high to perform a value of information analysis.

In the RA case study, we assessed the potential cost-effectiveness of alternative test strategies using one-year and five-year time horizons. The differences in results between these two models showed that the treatment of patients had a large impact on the cost-effectiveness of the new tests. Although the effect of treatment should always be incorporated in a CEA model of diagnostic tests, showing the immediate impact of the tests separately nevertheless generates useful information.

The importance of performing hospital-level CEAs when important practice variation exists was illustrated in Chapter 6 using a case study. A limitation of performing hospital-level CEAs might be the feasibility. However, the additional data that is needed to perform hospital-level CEAs (when practice variation exists) are the hospital-specific variables (e.g., hospital-specific test's costs, sensitivity, and specificity), which may be retrievable in at least some cases. We recommend hospital-level CEAs if data for potentially influential variables can be retrieved or sufficiently estimated. We also recommend the use of hospital-level CEAs in other diseases to consolidate the feasibility of performing hospital-level CEAs.

In conclusion, we highly recommend the use of our framework containing the general steps of early-CEAs in the very early stages of test development as well as the stages thereafter. Such early-CEAs should be performed in close collaboration with the different stakeholders. Well-designed early-CEAs are a useful tool to develop effective and cost-effective medical tests in an efficient manner.

9.6 References

1 National Research Council. Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease. Washington, DC: National Academies Press; 2011.

2 Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015;372(9):793–5.

3 Obama B. Remarks by the President on Precision Medicine. Washington, DC: The White House. State of the Union speech, January 30, 2015. Available from: https://www.whitehouse.gov/the-press-office/2015/01/30/remarks-president-precision-medicine. Accessed August 30, 2015.

4 Pietzsch JB, Paté-Cornell ME. Early technology assessment of new medical devices. *Int J Technol Assess Health Care* 2008;24(1):36-44.

5 IJzerman MJ, Steuten LM. Early assessment of medical technologies to inform product development and market access: a review of methods and applications. *Appl Health Econ Health Policy* 2011;9(5):331-47.

6 Hartz S, John J. Contribution of economic evaluation to decision making in early phases of product development: A methodological and empirical review. *Int J Technol Assess Health Care* 2008;24(4):465-72.

7 O'Prinsen AC, Gaultney J, Redekop WK. Universal Steps for Conducting Early-Stage Medical Techonolgy Assessment. *ISPOR Connections* 2009;November/December:4-6.

8 Buisman LR, Rutten-van Mölken MPMH, Postmus D, et al. The early bird catches the worm: early cost-effectiveness analysis of new medical tests. *Int J Technol Assess Health Care* 2016; 32(1):1–8. Epub 2016 Mar 22.

9 Aletaha D, Neogi T, Silman AJ, et al. Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;69(9):1580–8.

10 Alves C, Luime JJ, Van Zeben D, et al. Diagnostic performance of the ACR/EULAR 2010 criteria for rheumatoid arthritis and two diagnostic algorithms in an early arthritis clinic (REACH). *Ann Rheum Dis* 2011;70(9):1645–7.

11 Trueman P, Drummond M, Hutton J. Developing guidance for budget impact analysis. *Pharmacoeconomics* 2001;19(6):609–21.

12 O'Connor D, Green S, Higgins JPT. Defining the review question and developing criteria for including studies. In: Higgins, JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration. 2011. Available from: http://www.cochrane-handbook.org. Accessed May 2, 2014.

13 Buisman LR, Rijnsburger AJ, Den Hertog HM, et al. Clinical Practice Variation Needs to be Considered in Cost-Effectiveness Analyses: A Case Study of Patients with a Recent Transient Ischemic Attack or Minor Ischemic Stroke. *Appl Health Econ Health Policy* 2016:14(1):67–75.

14 Statistics Netherlands. Available from: http://statline.cbs.nl/Statweb/?LA=en. Accessed Feb 1, 2014.
15 Vaartjes I, Bots ML, Poos MJJC. Hoe vaak komt een beroerte voor en hoeveel mensen sterven eraan? In: Volksgezondheid Toekomst Verkenning, Nationaal Kompas Volksgezondheid. Bilthoven: Rijksinstituut voor Volsgezondheid en Mileu (RIVM); 2014.
16 Rothwell PM, Gutnikov SA, Warlow CP. Reanalysis of the final results of the European Carotid Surgery Trial. *Stroke* 2003;34(2):514–23.
17 U-King-Im JM, Hollingworth W, Trivedi RA, et al. Cost-effectiveness of diagnostic strategies prior to carotid endarterectomy. *Ann Neurol* 2005;58(4):506–15.
18 Sillesen H, Amarenco P, Hennerici MG, et al. Atorvastatin reduces the risk of cardiovascular events in patients with carotid atherosclerosis: a secondary analysis of the Stroke Prevention by Aggressive Reduction in Cholesterol Levels (SPARCL) trial. *Stroke* 2008;39(12):3297–302.
19 Milionis HJ, Giannopoulos S, Kosmidou M, et al. Statin therapy after first stroke reduces 10-year stroke recurrence and improves survival. *Neurology* 2009;72(21):1816–22.
20 Bonati LH, Dobson J, Featherstone RL, et al. Long-term outcomes after stenting versus endarterectomy for treatment of symptomatic carotid stenosis: the International Carotid Stenting Study (ICSS) randomised trial. *Lancet* 2015;385(9967):529–38.
21 Buisman LR, Tan SS, Nederkoorn PJ, et al. Hospital costs of ischemic stroke and TIA in the Netherlands. *Neurology* 2015;84(22):2208–15.
22 Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999;18(3):341-64.
23 Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty: when is there sufficient evidence? *Value Health* 2005;8(4):433-46.
24 Claxton KP, Sculpher MJ. Using value of information analysis to prioritise health research: some lessons from recent UK experience. *Pharmacoeconomics* 2006;24(11):1055-68.

CHAPTER 10

SUMMARY

Many new tests are currently being developed to improve the classification of individuals or patients into different subpopulations that differ in the risk or prognosis of a specific disease, or response to treatment. Preventive or therapeutic interventions can then be targeted to the subpopulations that will benefit most from it. This enhancement in patient stratification reflects part of the promise of precision medicine. The goal of precision medicine is to enable healthcare providers to identify individuals or patients and provide individualized treatment and prevention strategies that are targeted to people's unique characteristics, including genetics, microbiome composition, lifestyle, health history, and environment factors. The use of new medical tests with improved test characteristics may contribute to the identification of subjects and optimization of precision medicine. The effectiveness and cost-effectiveness of these new test strategies should be assessed before these strategies can be used in daily clinical practice.

Chapter 2 provided a framework with the general steps of an early cost-effectiveness analysis (early-CEA) of new medical tests. This framework was developed and applied to the two case studies of this thesis. The first case study focused on new diagnostic tests for the early diagnosis of rheumatoid arthritis (RA) to select patients for timely and optimal treatments. The second case study focused on new prognostic tests to assess the risk of a recurrent ischemic stroke in patients with a recent transient ischemic attack (TIA) or minor ischemic stroke, and thereby improve decisions about which patients should undergo surgery (i.e., carotid endarterectomy).

Chapter 3 provided insight in the costs of diagnostic tests, which tests are requested, what differential diagnoses are considered, and what strategies are used by rheumatologists in patients with early inflammatory arthritis at risk of having RA. A median of 25 [IQR 20-35] tests were requested during the first rheumatologist visit in 2010 with mean costs of €422 (SD: €168) per patient. This information was used in the cost-effectiveness model.

In Chapters 4 and 5, two early-CEAs were presented using models that were developed for evaluating new diagnostic tests in the work-up of patients with inflammatory arthritis at risk of having RA. The early-CEA model in Chapter 4 assessed the one-year cost-effectiveness of four diagnostic tests (B-cell gene expression, MRI, IL-6 serum level, and genetic assay). These tests were evaluated as add-on test for three different patient populations (all inflammatory arthritis patients, intermediate-risk patients (3-5 points on the ACR/EULAR 2010 RA classification criteria), and seronegative patients). It also assessed the cost-effectiveness of these tests when used to replace the currently used ACR/EULAR 2010 RA classification criteria. An add-on test in the work-up of patients at risk of having RA is potentially cost-effective, with the largest diagnostic benefit in intermediate-risk patients. The early-CEA model presented in Chapter 5 is an extension of the one-year model presented in Chapter 4 in which disease progression and treatment changes were modelled by changes in disease activity after the first year. This model was used to analyze the cost-effectiveness of the four tests evaluated in Chapter 4 in three test strategies (add-on for all inflammatory arthritis patients, add-on test for intermediate-risk patients [3-5 points on the ACR/EULAR 2010 RA classification criteria]), and as replacement test compared to the ACR/EULAR 2010 RA classification criteria over a five-year time horizon. This study showed that B-cell gene expression as add-on test for intermediate-risk patients was the most cost-effective add-on strategy and could even replace the ACR/EULAR 2010 RA classification criteria given its high specificity (0.90), moderate sensitivity (0.60), and low costs (€150). However, this test does not provide rheumatologists the same richness of information as the individual components of the currently used ACR/EULAR 2010 RA classification criteria do.

Chapter 6 used a case study to illustrate the importance of performing hospital-level CEAs when important practice variation between hospitals or deviation from the clinical guidelines exists. The study found that about 60% of the hospitals use the initial and confirmatory diagnostic tests to assess carotid stenosis in accordance with the national guidelines; the other hospitals used various other test combinations. The hospital-level CEA showed that the use of different test strategies resulted in important differences in costs and health effects between hospitals. As a consequence, the most cost-effective test strategy differed between hospitals. In the base case analysis, the CTA-only and CE-MRA-only strategies were the most cost-effective strategies. However, the results changed when hospital-level values for unit costs and test performance were used. The CTA-only strategy was the dominant strategy in the first hospital, since it had the lowest costs and highest QALYs of all strategies, while the CE-MRA-only strategy was the dominant one in the second hospital.

Chapter 7 determined the hospital resource use and costs of ischemic stroke and TIA patients in the Netherlands for 2012. In addition, the association between hospital costs of ischemic stroke and TIA and various patient and hospital characteristics was examined. This study is the first and most extensive cost analysis of inpatient and outpatient hospital costs of ischemic stroke and TIA patients in the last 10 years. The study showed that hospital costs are higher for inpatients and ischemic strokes compared with outpatients and TIAs, with length of stay as the most important contributor. The cost difference between inpatient ischemic stroke (€5,328) and inpatient TIA (€2,470) was caused by a shorter length of stay for TIA patients (3.6 days for TIA versus 8.8 for ischemic stroke). Hospital costs of outpatient care were higher for TIA than for ischemic stroke due to more frequent use of ambulatory treatment, and diagnostic and imaging tests. The regression models show that age, hospital type, and region were strongly associated with hospital costs.

Chapter 8 is an early-CEA in which the potential cost-effectiveness was evaluated of a new prognostic imaging test as confirmatory test to select patients with a recent TIA or minor ischemic stroke that benefit most from carotid endarterectomy. Since the sensitivity, specificity and costs of the confirmatory imaging test were unknown, the maximum health gain that a confirmatory imaging test could achieve was evaluated in four different patient subpopulations if it was perfect (100% sensitivity and specificity). The minimum required prognostic sensitivity and specificity were also investigated, for the new confirmatory imaging test to become cost-effective compared with the guideline-based strategy. The study showed that the maximum health gain for a confirmatory imaging test may be limited. Furthermore, the minimum prognostic sensitivity and specificity for 80-year-old men must be higher than for 60-year-old men. Even a perfect new confirmatory imaging test is not cost-effective in 60-year-old and 80-year-old women with 30-69% stenosis because it leads to poorer health outcomes and higher costs. The poorer health outcomes are due to women with 70-99% stenosis who have a higher risk of ischemic stroke but are not tested with the new confirmatory imaging test. Amongst women, a new confirmatory test is likely to be cost-effective for women with 70-99% stenosis.

Chapter 9 provides a discussion of the main findings of this thesis, the implications of the results of early-CEAs for medical test developers, clinicians as end-users of the medical tests, patients as main beneficiaries, hospital managers, and payers. The discussion then continues with a detailed discussion of several methodological challenges of early-CEAs of medical tests. The thesis concludes with recommendations for further research.

CHAPTER 11

SAMENVATTING

Veel nieuwe medische testen worden op dit moment ontwikkeld om de classificatie te verbeteren van individuen of patiënten in verschillende subpopulaties die verschillen in het risico of de prognose van een bepaalde ziekte of de respons op behandeling. Preventieve of therapeutische interventies kunnen dan worden gericht op de subpopulaties die er het meest baat bij hebben. Deze verbetering in patiënt stratificatie reflecteert een deel van de belofte van precision medicine. Het doel van precision medicine is om zorgverleners in staat te stellen individuen of patiënten te identificeren en te zorgen voor geïndividualiseerde behandeling en preventieve strategieën die gericht zijn op de individuele persoonskarakteristieken, inclusief genetica, microbiota, leefstijl, gezondheidsgeschiedenis en omgevingsfactoren. Het gebruik van nieuwe medische testen met verbeterde testkarakteristieken kan bijdragen aan de identificatie van personen en optimalisatie van precision medicine. De effectiviteit en kosteneffectiviteit van deze nieuwe test strategieën moeten worden onderzocht voordat deze strategieën kunnen worden gebruikt in de dagelijkse klinische praktijk.

In hoofdstuk 2 is een stappenplan gepresenteerd met de algemene stappen van een vroege kosteneffectiviteitsanalyse (vroege KEA) van nieuwe medische testen. Dit stappenplan is ontwikkeld en toegepast op de twee casestudies van dit proefschrift. De eerste casestudie richtte zich op nieuwe diagnostische testen in de vroege diagnose van reumatoïde artritis (RA) om patiënten tijdig te selecteren voor optimale behandelingen. De tweede casestudie richtte zich op nieuwe prognostische testen om het risico van een terugkerende ischemische beroerte te bepalen bij patiënten met een recente TIA of kleine ischemische beroerte en daarbij de beslissingen te verbeteren over welke patiënten een operatie (carotis endarterectomie) moeten ondergaan.

Hoofdstuk 3 heeft inzicht gegeven in de kosten van diagnostische testen, welke testen aangevraagd zijn, welke differentiaaldiagnoses zijn overwogen en welke strategieën door reumatologen zijn gebruikt bij patiënten met

vroege inflammatoire artritis die het risico hebben om RA te hebben. Een mediaan van 25 [IQR 20-35] testen zijn aangevraagd tijdens het eerste bezoek bij de reumatoloog in 2010 met gemiddelde kosten van €422 (SD: €168) per patiënt. Deze informatie is gebruikt in het kosteneffectiviteitsmodel.

Hoofdstukken 4 en 5 beschrijven twee vroege KEAs waarin modellen zijn gebruikt die zijn ontwikkeld om nieuwe diagnostische testen te evalueren voor patiënten met vroege inflammatoire artritis die het risico hebben om RA te hebben. Het vroege KEA model in hoofdstuk 4 schatte de kosteneffectiviteit van vier diagnostische testen (B-cell gene expression, MRI, IL-6 serum level, en genetic assay) bij een tijdshorizon van één jaar. Deze testen zijn geëvalueerd als een aanvullende test voor drie verschillende patiëntpopulaties [alle inflammatoire artritis patiënten, inflammatoire artritis patiënten met een middelmatig risico om RA te hebben (3-5 punten op de ACR/EULAR 2010 RA classificatie criteria) en seronegatieve patiënten]. Dit model werd ook gebruikt om de kosteneffectiviteit van deze testen te evalueren als vervanging van de huidige ACR/EULAR 2010 RA classificatie criteria. Een aanvullende test voor patiënten met het risico om RA te hebben is mogelijk kosteneffectief met het grootste diagnostische voordeel voor patiënten met een middelmatig risico. Het vroege KEA model dat gepresenteerd is in hoofdstuk 5 is een uitbreiding van het één-jaar model dat gepresenteerd is in hoofdstuk 4. Na het eerste jaar zijn de ziekteprogressie en het veranderen van behandeling gemodelleerd door veranderingen in ziekteactiviteit. Dit model is gebruikt om de kosteneffectiviteit van de vier testen die in hoofdstuk 4 zijn geëvalueerd te analyseren in drie test strategieën. Deze test strategieën zijn een aanvullende test voor alle inflammatoire artritis patiënten, een aanvullende test voor inflammatoire artritis patiënten met een middelmatig risico om RA te hebben (3-5 punten op de ACR/EULAR 2010 RA classificatie criteria) en een vervangende test in vergelijking met de ACR/EULAR 2010 RA classificatie criteria over een vijfjaar tijdshorizon. Deze studie heeft aangetoond dat B-cell gene expression als een aanvullende test voor inflammatoire artritis patiënten met een middelmatig risico de meest kosteneffectieve aanvullende test strategie was en het zelfs de ACR/EULAR 2010 RA classificatie criteria kan vervangen door de hoge specificiteit (0,90), middelmatige sensitiviteit (0,60) en de lage kosten (€150). Echter, deze test geeft reumatologen niet dezelfde schat aan informatie als de afzonderlijke componenten van de ACR/EULAR 2010 RA classificatie criteria die momenteel worden gebruikt.

In hoofdstuk 6 is aan de hand van een casestudie het belang van het uitvoeren van ziekenhuis-specifieke KEAs aangetoond wanneer belangrijke praktijkvariatie tussen ziekenhuizen of afwijking van de klinische richtlijnen bestaat. Deze studie toonde aan dat 60% van de ziekenhuizen de initiële en bevestigende diagnostische testen gebruiken die in de nationale richtlijnen

aanbevolen worden om de carotis stenose te beoordelen; de andere ziekenhuizen gebruiken verscheidene andere test combinaties. De ziekenhuis-specifieke KEA toonde aan dat het gebruik van verschillende test strategieën resulteerde in belangrijke verschillen in de kosten en gezondheidseffecten tussen ziekenhuizen. Als gevolg daarvan verschilde de meest kosteneffectieve strategie tussen ziekenhuizen. In de base case analyse waren de CTA-only en CE-MRA-only strategieën de meest kosteneffectieve strategieën. Echter, de resultaten veranderden wanneer ziekenhuis-specifieke kostprijzen en test karakteristieken werden gebruikt. In dit geval was de CTA-only strategie de dominante strategie in het eerste ziekenhuis omdat het de laagste kosten en hoogste gezondheidseffecten, uitgedrukt in voor kwaliteit gecorrigeerde levensjaren (QALYs), had van alle strategieën, terwijl de CE-MRA-only strategie de dominante strategie was in het tweede ziekenhuis.

Hoofdstuk 7 beschrijft het zorggebruik en de kosten in het ziekenhuis van ischemische beroerte en TIA patiënten in Nederland in 2012. Verder is de associatie tussen ziekenhuiskosten van ischemische beroerte en TIA en verscheidene patiënt- en ziekenhuiskarakteristieken onderzocht. Deze studie is de eerste en grootste kostenanalyse van klinische en poliklinische ziekenhuiskosten van ischemische beroerte en TIA patiënten in de laatste 10 jaar. Deze studie heeft aangetoond dat de ziekenhuiskosten hoger zijn van klinische en ischemische beroerte patiënten in vergelijking met poliklinische en TIA patiënten met ligduur als belangrijkste factor. Het kostenverschil tussen klinische ischemische beroerte patiënten (€5.328) en klinische TIA (€2.470) is veroorzaakt door de kortere ligduur voor TIA patiënten (3,6 dagen voor TIA vergeleken met 8,8 voor ischemische beroerte). Ziekenhuiskosten van poliklinische zorg waren hoger voor TIA dan voor ischemische beroerte patiënten door een hoger gebruik van dagbehandelingen en diagnostische en beeldvormende tests. De regressiemodellen laten zien dat leeftijd, ziekenhuistype en regio sterk geassocieerd waren met ziekenhuiskosten.

Hoofdstuk 8 beschrijft een vroege KEA waarin de potentiële kosteneffectiviteit van een nieuwe prognostische beeldvormende test is geëvalueerd als een bevestigende test om patiënten met een recente TIA of kleine ischemische beroerte te selecteren die het meest baat hebben om een carotis endarterectomie te ondergaan. Omdat de sensitiviteit, specificiteit en kosten van de bevestigende beeldvormende test niet bekend waren, is de maximale gezondheidswinst geëvalueerd dat een bevestigende beeldvormende test zou kunnen behalen wanneer deze perfect zou zijn (100% sensitiviteit en specificiteit). Ook is onderzocht wat de minimale benodigde prognostische sensitiviteit en specificiteit van de nieuwe bevestigende test zouden moeten zijn om kosteneffectief te zijn in vergelijking met de op de richtlijn gebaseerde strategie. Deze studie heeft aangetoond dat de maximale gezondheidswinst van een bevestigende beeldvormende test beperkt kan zijn. Ook moeten de minimale prognostische sensitiviteit en specificiteit voor 80-jarige mannen

hoger zijn dan voor 60-jarige mannen. Zelfs een perfecte nieuwe bevestigende beeldvormende test is niet kosteneffectief voor 60-jarige en 80-jarige vrouwen met een 30-69% stenose omdat het leidt tot lagere gezondheidswinsten en hogere kosten. De lagere gezondheidswinsten worden veroorzaakt doordat vrouwen met 70-99% stenose een hoger risico hebben op een ischemische beroerte maar niet getest zijn met de nieuwe bevestigende beeldvormende test. Bij vrouwen zal een nieuwe bevestigende test waarschijnlijk kosteneffectief zijn bij vrouwen met 70-99% stenose.

In hoofdstuk 9 zijn de hoofdbevindingen van dit proefschrift en de implicaties van de resultaten van vroege KEAs bediscussieerd voor medische test ontwikkelaars, clinici als eindgebruikers van de medische testen, patiënten als voornaamste begunstigden, ziekenhuismanagers en besluitvormers. Vervolgens zijn meerdere methodologische uitdagingen van vroege KEAs van medische testen gedetailleerd bediscussieerd. Het proefschrift sluit af met aanbevelingen voor vervolgonderzoek.

CHAPTER 12

ACKNOWLEDGMENTS

Het doen van promotieonderzoek bewijst maar weer dat het belangrijk is om de juiste mensen om je heen te hebben. De bijdrage van anderen is belangrijk geweest bij het tot stand komen van mijn proefschrift.

Allereerst wil ik mijn promotor Maureen Rutten-van Mölken en copromotor Ken Redekop bedanken. Maureen, bedankt voor al je inhoudelijke adviezen en de prettige samenwerking. Ik bewonder je verscheidenheid aan kennis, grote discipline en je motivatie en drive om mensen te motiveren. Ook bedankt dat je mij telkens in staat stelde om kritisch naar mijn eigen werk te blijven kijken. Ken, bedankt voor de kans die je mij hebt gegeven om te gaan promoveren en alles wat je mij hebt bijgeleerd. Je was altijd bereid om tijd vrij te maken om van gedachte te wisselen en adviezen te geven. Ik bewonder je visuele manier van uitleggen en het gebruik van metaforen die je altijd uit je mouw lijkt te schudden.

Verder wil ik alle co-auteurs bedanken die hebben bijgedragen aan de verschillende hoofdstukken. Hierbij wil ik een aantal personen in het bijzonder noemen.

Jolanda, bedankt voor je geduld om mijn eindeloze vragen over de diagnose en behandeling van reumatische patiënten te beantwoorden. Ook heb ik van je geleerd om altijd in perspectief te plaatsen wat de bijdrage van ons onderzoek zal betekenen voor het verbeteren van de klinische praktijk. Bovenal vond ik het altijd erg prettig om met je samen te werken.

Mark, bedankt voor het meedenken en je waardevolle adviezen bij het ontwikkelen van het TRACER model. Jouw benadering, inzicht en kennis van modelleren hebben mij enorm geholpen om de uitdagingen van het vroeg modelleren van de kosteneffectiviteit van nieuwe medische testen aan te pakken.

Rian, bedankt voor je hulp aan het begin van mijn promotieonderzoek en de prettige samenwerking bij de verschillende PARISk artikelen van mijn proefschrift.

Siok Swan, jouw kennis over DBC's en de DBC database zijn van onschatbare waarde geweest bij het onderzoek naar de ziekenhuiskosten van beroertezorg.

Voor de klinische input van de verschillende artikelen wil ik Peter, Paul, Aad, Bernadette en Mieke bedanken voor hun bijdrage.

Verder wil ik alle collega's en vrienden bij het iBMG bedanken voor jullie interesse, betrokkenheid en gezelligheid. Ik prijs mij gelukkig om mijn promotieonderzoek op een afdeling met zo'n goede werksfeer te hebben verricht. Ook heb ik genoten van de activiteiten naast het werk, zoals de deelname aan de Roparun, reisjes voor of na congressen en de vele borrels. In het bijzonder wil ik Melinde en Arthur bedanken en ben ik erg blij dat jullie mijn paranimfen zijn tijdens de verdediging van mijn proefschrift.

Daarnaast wil ik alle partners binnen de verschillende workpackages van de twee CTMM projecten bedanken (TRACER en PARISk) waar het werk van mijn promotieonderzoek deel van uitmaakt. Jullie input en feedback tijdens de presentaties en vergaderingen zijn belangrijk geweest. Ook wil ik de neurologen bedanken die zijn geïnterviewd bij het onderzoek naar de praktijkvariatie in het gebruik van beeldvormende testen.

Ik wil de leden van de beoordelingscommissie bedanken voor het beoordelen van mijn proefschrift en de leden van de promotiecommissie voor het opponeren tijdens de verdediging van mijn proefschrift.

Marc, bedankt voor de grafische vormgeving van mijn proefschrift. Het is erg fijn dat ik dit aan jou heb kunnen toevertrouwen.

I also wish to thank Adam, Caryn, Patrick, Paolo, Tania and Aparnaa at the World Bank, and Davidson at the Results for Development Institute for your trust in me and the opportunity you gave me to work on global health economic topics. Although, the published and forthcoming papers from this work are not part of this thesis, they are of important value to me.

Ellen, Eddy en Owen, bedankt voor de mogelijkheid om mijn wetenschappelijke carrière als postdoctoraal onderzoeker voort te zetten.

Mijn ouders wil ik bedanken voor alle steun die jullie mij altijd hebben gegeven en alles wat jullie mij hebben geleerd. Ook wil ik Florian, Sebastiaan, Anne Vera, familie en vrienden bedanken voor jullie interesse en betrokkenheid tijdens mijn promotieonderzoek.

CHAPTER 13

LIST OF PUBLICATIONS

Publications in peer-reviewed journals

Buisman LR, Luime JJ, Oppe M, Hazes JMW, Rutten-van Mölken MPMH. A five-year model to assess the cost-effectiveness of new diagnostic tests in the early diagnosis of rheumatoid arthritis. Forthcoming in *Arthritis Research & Therapy* 2016. doi:10.1186/513075-016-1020-3.

Bredenkamp C, **Buisman LR**. Providing financial protection from out-of-pocket spending in the Philippines: policies and progress. Published in *Health Policy and Planning* 2016. Epub 2016 Apr 11. doi:10.1093/heapol/czw011.

Buisman LR, Rutten-van Mölken MPMH, Postmus D, Luime JJ, Uyl-de Groot CA, Redekop WK. The early bird catches the worm: early cost-effectiveness analysis of new medical tests. Published in *International Journal of Technology Assessment in Health Care* 2016;32(1):1–8. Epub 2016 Mar 22. doi:10.1017/S0266462316000064.

Wagstaff A, Cotlear D, Hoang-Vu Eozenou P, **Buisman LR**. Measuring progress towards universal health coverage: with an application to 24 developing countries. *Oxford Review of Economic Policy* 2016;32(1):147–89. doi:10.1093/oxrep/grv019.

Buisman LR, Rijnsburger AJ, Den Hertog HM, Van der Lugt A, Redekop WK. Clinical Practice Variation Needs to be Considered in Cost-Effectiveness Analyses: A Case Study of Patients with a Recent Transient Ischemic Attack or Minor Ischemic Stroke. *Applied Health Economics and Health Policy* 2016:14(1):67–75. doi:10.1007/s40258-015-0167-4.

Luime JJ, **Buisman LR**, Oppe M, Hazes JMW, Rutten-van Mölken MPMH. A cost-effectiveness model for evaluating new diagnostic tests in the work-up of patients with inflammatory arthritis at risk of having Rheumatoid Arthritis. *Arthritis Care & Research* 2015. Epub 2015 Nov 10. doi:10.1002/acr.22776.

Wagstaff A, Dmytraczenko T, Almeida G, **Buisman L**, Hoang-Vu Eozenou P, Bredenkamp C, Cercone JA, Diaz Y, Maceira D, Molina S, Paraje G, Ruiz F, Sarti F, Scott J, Valdivia M, Werneck H. Assessing Latin America's Progress Toward Achieving Universal Health Coverage. *Health Affairs* 2015;34(10):1704–12. doi:10.1377/hlthaff.2014.1453.

Buisman LR, Tan SS, Nederkoorn PJ, Koudstaal PJ, Redekop WK. Hospital costs of ischemic stroke and TIA in the Netherlands. *Neurology* 2015;84(22):2208–15. doi:10.1212/WNL.000000 0000001635.

Buisman LR, García-Gómez P. Inequity in inpatient healthcare utilisation 10 years after Apartheid. *Development Southern Africa* 2015:32(2):193–208. doi:10.1080/0376835X.2014.984374.

Wagstaff A, Bredenkamp C, **Buisman LR**. Progress on Global Health Goals: Are the Poor Being Left Behind? *World Bank Research Observer* 2014;29(2):137–62. doi:10.1093/wbro/lku008.

Bredenkamp C, **Buisman LR**, Van de Poel E. Persistent inequalities in child undernutrition: evidence from 81 countries, from 1990 to today. *International Journal of Epidemiology* 2014;43(4):1328–35. doi:10.1093/ije/dyu075.

Submitted papers for publication in peer-reviewed journals

Buisman LR, Rijnsburger AJ, Van der Lugt A, Nederkoorn PJ, Koudstaal PJ, Redekop WK. Cost-effectiveness of novel imaging tests to select ischemic stroke patients for carotid endarterectomy. Submitted for publication.

Benner BJM, **Buisman LR**, Vis M, Rutten-van Mölken MPMH, Hazes JMW, Luime JJ. Costs of diagnostic tests in the work-up of early inflammatory arthritis patients: a mixed-methods approach. Submitted for publication.

Gwatkin DR, **Buisman LR**. Towards a Fuller Assessment of Need for Financial Protection in Health: By Including People Who Forgo Care Because of Its Cost. Submitted for publication.

Working papers

Wagstaff A, Cotlear D, Hoang-Vu Eozenou P, **Buisman LR**. 2015. Measuring Progress Towards Universal Health Coverage: With an Application to 24 Developing Countries. World Bank Policy Research Working Paper No. 7470. Washington, DC: World Bank.

Bredenkamp C, **Buisman LR**. 2015. Universal Health Coverage in the Philippines: Progress on Financial Protection Goals. World Bank Policy Research Working Paper No. 7258. Washington, DC: World Bank.

Wagstaff A, Bilger M, **Buisman LR**, Bredenkamp C. 2014. Who Benefits from Government Health Spending and Why? A Global Assessment. World Bank Policy Research Working Paper No. 7044. Washington, DC: World Bank.

Wagstaff A, Bredenkamp C, **Buisman LR**. 2014. Progress Toward the Health MDGs: Are the Poor Being Left Behind? World Bank Policy Research Working Paper No. 6894. Washington, DC: World Bank.

CHAPTER 14

PHD PORTFOLIO

PhD candidate:	Leander R. Buisman
Promotor:	Prof.dr. M.P.M.H. Rutten-van Mölken
Copromotor:	Dr. W.K. Redekop
PhD period:	2011-2016

PhD training

MTA and making decisions for new medical devices, Center for Translational Molecular Medicine (CTMM), Amsterdam, The Netherlands, 2013.

Mentoring training, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2013.

Supervision of theses, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2012.

Academic writing in English, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2012.

Diagnostic Research; Prognostic Research, NIHES Erasmus Winter Program, short courses, Rotterdam, The Netherlands, 2012.

Topics in Meta-analysis; Clinical Decision Analysis; Markers in Prognostic Research; Survival Analysis, NIHES Erasmus Summer Program, Rotterdam, The Netherlands, 2011.

Ready in 4 years, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2011.

Advanced economic modelling methods for health economic evaluation, University of Glasgow, Scotland, 2011.

Teaching

Supervision and co-supervision of several theses, master program Health Economics, Policy and Law, Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2011–2016.

Supervision and co-supervision of several theses, bachelor program Health Sciences, Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2011–2016.

Health Technology Assessment, master program Health Economics, Policy and Law, Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2015.

Equity in Health Care: Using Household Survey data for Policy Analysis and ADePT software, Summer School, Lviv, Ukraine, 2015.

Internship, bachelor program Health Sciences, Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2015.

Private Household Health Expenditure in Moldova: Using Household Survey data for Policy Analysis and ADePT software, Ministry of Health, Chisinau, Moldova, 2014.

Writing and research skills, pre-master program Health Economics, Policy and Law. Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2014.

Scientific skills, bachelor program Health Sciences, Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2011–2014.

Mentor, bachelor program Health Sciences, Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2011–2014.

Internship, bachelor program Health Sciences, Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2011–2013.

Inequities in health and health care, post-graduate program International Health and Policy Evaluation, Institute of Health Policy and Management and International Institute of Social Studies, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2011.

Podium presentations

Early HTA in TRACER: What have we learned? TRACER final symposium, Utrecht, The Netherlands, 21 November 2015.

Development of novel imaging tests to select patients for individualized therapies: are they worth further investment? The 6th lowlands Health Economists' Study Group (lolaHESG), Oostvoorne, The Netherlands, 22-23 May 2014.

Development of novel imaging tests to select patients for individualized therapies: are they worth further investment? PARISk semi-annual meeting, AMC, Amsterdam, The Netherlands, 15 November 2013.

PARISk: Highlights WP6 – MTA. PARISk semi-annual meeting, Erasmus MC, Rotterdam, The Netherlands, 14 June 2013.

Practice variation needs to be considered in cost-effectiveness analyses of diagnostic tests. The 5th lowlands Health Economists' Study Group (lolaHESG), Nunspeet, The Netherlands, 23-24 May 2013.

Cost-effectiveness analyses of diagnostic tests at the early stages of development. Health Economics Seminar, Institute of Health Policy and Management, Rotterdam, The Netherlands, 9 April 2013.

Development of novel and optimized imaging tests to stratify patients to individualized therapies: is it worth further investing? PARISk semi-annual meeting, UMC Utrecht, Utrecht, The Netherlands, 23 November 2012.

TRACER WP 11: Medical technology assessment (MTA). TRACER symposium, Utrecht, The Netherlands, 6 September 2012.

Topic: Trends in inequalities in child malnutrition. Health, Nutrition & Population meeting, World Bank, Washington, DC, USA, 31 August 2012.

PARISk: Highlights WP6 – MTA. PARISk semi-annual meeting, Maastricht UMC+, Maastricht, The Netherlands, 15 June 2012.

RA – hitting a moving target: Methodological challenges in evaluating new diagnostic tests in rheumatoid arthritis. CQM Working Group Methods for the Evaluation of Diagnostic Tests meeting, Erasmus MC, Rotterdam, The Netherlands, 22 November 2011.

Equity and Financial Protection in Health – Measurement and explanation. The 8th World Congress on Health Economics of the International Health Economics Association (iHEA), Toronto, Canada. Involved in the pre-congress workshop, 10-13 July 2011.

PARISk - Medical Technology Assessment (WP6). PARISk semi-annual meeting, Erasmus MC, Rotterdam, The Netherlands, 27 May 2011.

Poster presentations

A five-year model to assess the cost-effectiveness of new diagnostic tests in the early diagnosis of rheumatoid arthritis. International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 18th Annual European Congress, Milan, Italy. Research Poster Presentation, 7-11 November 2015.

Novel Imaging Technology to Select Patients with a Recent Transient Ischemic Attack or Minor Ischemic Stroke for Carotid Endarterectomy: The Relationship between Test Performance and Cost-Effectiveness. International Stroke Conference, Nashville, USA. Research Poster Presentation, 11-13 February 2015.

Novel imaging technology to select patients for individualized therapies: test performance and cost-effectiveness. International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 17th Annual European Congress, Amsterdam, The Netherlands. Research Poster Presentation, 8-12 November 2014.

Hospital costs of ischemic stroke and transient ischemic attack in the Netherlands. International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 17th Annual European Congress, Amsterdam, The Netherlands. Research Poster Presentation, 8-12 November 2014.

Cost-effectiveness analysis of diagnostic tests in the work-up of patients with intermediate risk of developing rheumatoid arthritis. International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 16th Annual European Congress, Dublin, Ireland. Research Poster Presentation, 2-6 November 2013.

Development of novel imaging tests to select patients for individualized therapies: are they worth further investment? International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 16th Annual European Congress, Dublin, Ireland. Research Poster Presentation, 2-6 November 2013.

Cost-effectiveness analysis of diagnostic tests in the work-up of patients with intermediate risk of developing rheumatoid arthritis. ACR/ARHP Annual Meeting, San Diego, USA. Research Poster Presentation, 26-30 October 2013.

Practice variation in diagnostic imaging workup and treatment criteria following a recent TIA or minor ischemic stroke: consequences for early economic evaluations. International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 15th Annual European Congress, Berlin, Germany. Research Poster Presentation, 4-7 November 2012.

An evaluation of the non-invasive imaging tests used in current care of TIA and minor ischemic stroke in the Netherlands: how much practice variation is there? International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 14th Annual European Congress, Madrid, Spain. Research Poster Presentation, 5-8 November 2011.

Research visits

World Bank, Washington, DC, USA, 2014.

World Bank, Manila, Philippines, 2014.

World Bank, Washington, DC, USA, 2012.

Grants and awards

Poster Finalist Award, Best Poster Presentation, International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 18th Annual European Congress, Milan, Italy, 2015.

The Best New Research from across the Bank, Winning paper: Wagstaff A, Bredenkamp C, Buisman LR. 2014. Progress Toward the Health MDGs: Are the Poor Being Left Behind? World Bank Policy Research Working Paper No. 6894. Washington, DC: World Bank.

Travel grant for international congress, Erasmus Trustfonds, Rotterdam, The Netherlands, 2013.

VPU Team Award, Training in Measurement of Health Equity and Financial Protection using ADePT, Human Development Network, World Bank, Washington, DC, USA, 2012.

CHAPTER 15

CURRICULUM VITAE

Leander Robert Buisman was born in Schiedam, The Netherlands on November 3th, 1987. He obtained a bachelor's degree in Health Sciences in 2009 and a master's degree in Health Economics, Policy and Law in 2010 at the Institute of Health Policy and Management from the Erasmus University Rotterdam. During his master's, he specialized in Health Economics.

After graduation, Leander started as a PhD candidate at the section Health Technology Assessment from the Institute of Health Policy and Management, Erasmus University Rotterdam. His PhD research was mainly focused on early cost-effectiveness of medical tests in the early diagnosis of rheumatoid arthritis and the prediction of recurrent ischemic stroke events.

Alongside his PhD research, Leander was working on global health economic topics such as international comparisons of equity in health outcomes, healthcare utilization and financial protection. He was involved in multiple projects from the World Bank and the Results for Development Institute.

After completing his PhD research, Leander continued his career as a postdoctoral researcher at the section Health Economics from the Institute of Health Policy and Management, Erasmus University Rotterdam.


9 789070 116743