



De mysterieuze massa

Oplage: 2000

Omslagfoto: Bianca Das

Vormgeving: Anna Bosselaar | Zoiets Communicatie, Capelle aan den IJssel

Druk: Optima Grafische Communicatie, Rotterdam

ISBN 978-90-8559-135-1

© H.J. de Koning, oratie Erasmus MC

26 juni 2009

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvuldigd zonder voorafgaande toestemming van de auteur.

Voorzover het maken van kopieën uit deze uitgave is toegestaan op grond van art. 16h t/m 16m Auteurswet 1912 j°. Besluit van 27 november 2002, Stb. 575, dient men de daarvoor wettelijk verschuldigde vergoeding te voldoen aan de Stichting Reprorecht te Hoofddorp (Postbus 3060, 2130 KB).

De mysterieuze massa*

Inaugurele rede
Bijzondere leerstoel
Evaluatie van vroegopsporing van ziekten

Prof.dr. H.J. de Koning, arts-epidemioloog
Erasmus MC
Afdeling Maatschappelijke Gezondheidszorg

Vrijdag 26 juni 2009



* Ook naar: W. Vandersteen: De malle mergpijp (143) en zijn eerste alliteratie De kleppende klipper (15).



Originele cartoon 'Agent 327', aangeboden door Monique Roobol, afd. Urologie, Erasmus MC.

VOOR
HARRY DE KONING
VAN MONIQUE ROOBOL
GROEFEN EN FELICITATIES
VAN MARTIN LOBENYK

*Mijnheer de Rector Magnificus,
Leden van het College van Bestuur van de Erasmus Universiteit Rotterdam,
Leden van de Raad van Bestuur van het Erasmus MC,
Leden van het bestuur van de Vereniging Trustfonds,*

*Lieve familie en vrienden, zeer gewaardeerde collega's en overige toehoorders,
dear friends,*

Het leven is een mysterie.

Er geldt dit niet des te meer voor zijn tegenhanger, de dood?

Er zijn talrijke theorieën omtrent de oorsprong van het leven, dan wel het einde van het leven. In dit Darwinjaar is de evolutietheorie, en daarmee de eindigheid van ons bestaan, populairder dan ooit. Er is het scheppingsverhaal met een indicatie van het oneindige van ons bestaan; er is de theorie van een oneindig bewustzijn, geformuleerd op basis van bijna-dood ervaringen van patiënten¹.

Er bestaan Japanse verhalen over shinigami's, engelen des doods, die met het schrijven van je naam in een zogenaamde 'death note' je tijdstip van overlijden bepalen. Deze shinigami's zouden tegenwoordig, uit verveling of om niet als streber te worden aangemerkt, steeds minder vaak namen noteren².

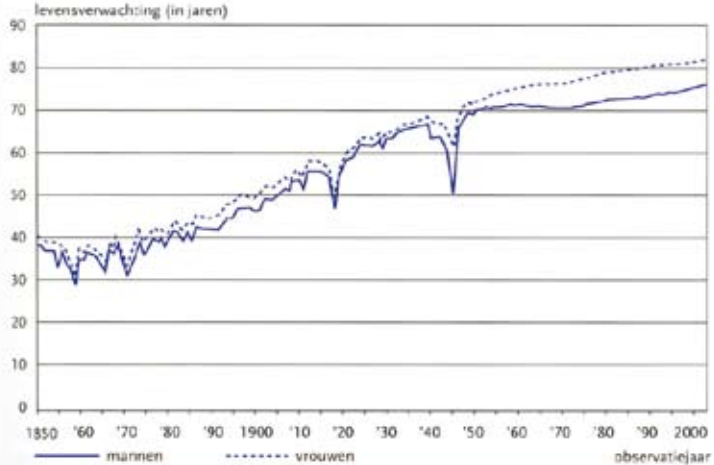
Het is een feit dat onze levensverwachting in de laatste 150 jaar spectaculair is toegenomen. Rond 1850 bedroeg de levensverwachting in Nederland gemiddeld 39 jaar; op dit moment bedraagt deze voor pasgeborenen bijna 80 jaar³.

Het is niet eenvoudig om aan te geven welke specifieke ontwikkelingen, bijvoorbeeld in de gezondheidszorg aan deze gezondheidswinst hebben bijgedragen. In een soort IDOLS-verkiezing heeft het British Medical Journal in 2007 lezers gevraagd te laten stemmen op de 15 belangrijkste 'medische' ontwikkelingen in deze 150 jaar. Met trots kan ik vermelden dat de ontwikkeling die ons afdelingshoofd, professor Johan Mackenbach, kort had beschreven – na een spannende eindstrijd – bovenaan eindigde: simpelweg de aanleg van

Het leven en de dood



DEATH NOTE © 2003 by
Tsugumi Ohba, Takeshi Obata /
Shueisha Inc.



Figuur 2.9 Ontwikkeling van de levensverwachting bij de geboorte in Nederland, 1850-2003
Bron: Centraal Bureau voor de Statistiek, Bevolkingstrends, 2005.

riolen en waterleiding, geruime tijd nadat duidelijk werd dat niet-hygiënisch watermanagement de oorzaak was van de verspreiding van ziektekiemen, werd als belangrijkste ontwikkeling gekozen⁴.

Het is een mooi voorbeeld van ons vakgebied, dat van de 'public health', ook wel omschreven als "the science and art of preventing disease, prolonging life and promoting health through the organized efforts of society". Het heeft een voorkeur voor preventieve maatregelen, is gericht op de volksgezondheid, in feite een optelling van de gezondheid van alle individuen, en gaat veelal over collectief aangestuurde programma's.

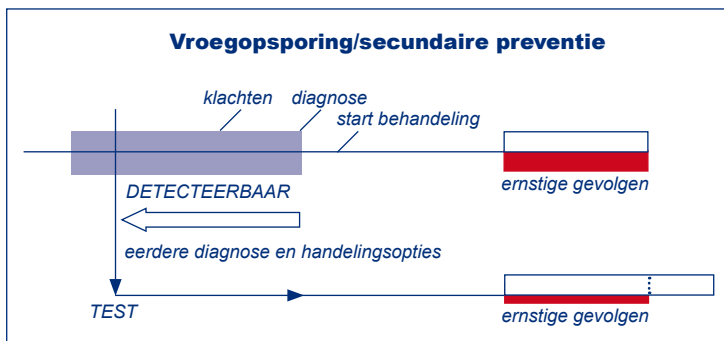
Er zijn prachtige voorbeelden in de 'public health' waarbij het aantonen van de determinant, en vervolgens actie ondernemen effectief is geweest in de bevordering van de gezondheid: de schatting is dat in de Verenigde Staten anti-rookbeleid vanaf de jaren vijftig heeft geleid tot 800.000 voorkomen

sterfgevallen aan longkanker. In Nederland heeft het verbieden van asbest ertoe geleid dat we dit jaar vermoedelijk 250 minder sterfgevallen aan longvlieskanker hebben⁵. Sinds de invoering van adviezen tot zijligging op het consultatiebureau is het aantal gevallen van wiegendood met meer dan 50% afgenomen⁶.

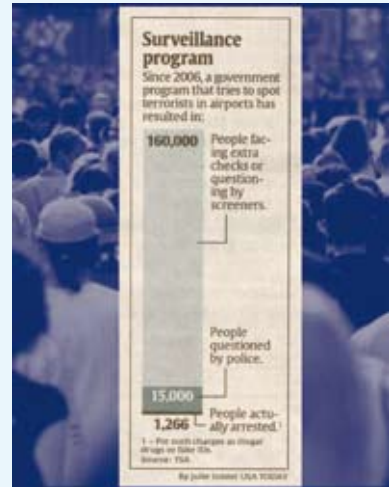
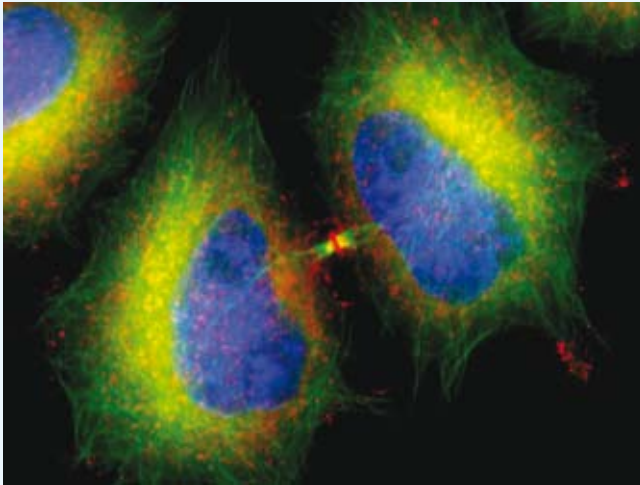
Voorkomen is beter dan genezen, zult u denken.

Maar voorkomen is vaak niet mogelijk; veelal omdat we te weinig specifieke kennis hebben over de mogelijke factoren die een rol spelen. Als voorkomen niet mogelijk is⁷, beschikken we eventueel ook nog over secundaire preventie of vroegopsporing. Hierbij bedoelen we het opsporen en behandelen van een ziekte, aandoening, afwijking of achterstand in een stadium dat er nog geen sprake is van geuite symptomen of klachten. Het doel van deze vroegopsporing is de voortgang van het ziekteproces of de ongunstige ontwikkeling tot stilstand te brengen of te vertragen, en daarmee gezondheidswinst te bereiken, of om in een eerder stadium dan gebruikelijk handelingsopties aan te kunnen bieden (zoals bij bepaalde zwangerschaps-screenings het geval is). Het principe is dat als gevolg van de toepassing van een relatief eenvoudige screeningstest de diagnose eerder wordt gesteld, de vroegtijdige behandeling effectiever is dan late behandeling, en daardoor de persoon langer leeft of minder nadelige gevolgen ondervindt, dan in de situatie zonder screening (zie figuur).

Feitelijk zijn preventie en therapie de twee belangrijkste methoden waarmee we in de geneeskunde en de gezondheidszorg de gezondheid van mensen kunnen verbeteren. Maar dit gebeurt vaak op een nogal verschillende manier.



Individuele therapie vs populatie-preventie



De afbeelding links, met dank aan onderzoekster Megan Chircop uit Sydney die hiermee de New Scientist Eureka Prize won, toont een delende kankercel, waarbij – in rood – vermoedelijk een eiwitplitsend enzym betrokken is om eiwit-filamenten – dit laatste eiwit-filament – door te knippen. Het onderzoek beoogt therapie te ontwikkelen die dit specifieke enzym blokkeert, zodat cellen zich niet of niet afdoende kunnen delen. Men probeert op individueel niveau de beste op-maat therapie voor deze patiënt of patiënte vast te stellen. Bij vroegopsporing en preventie is op individueel niveau vrijwel geen adequate voorspelling te doen, maar voor de populatie als geheel verwachten we gezondheidswinst. Het is niet van tevoren vast te stellen bij wie de aandoening reeds aanwezig is, bij wie deze door screening ontdekt kan worden en bij wie vroegtijdige behandeling effect zal hebben.

Ter illustratie, ernaast een overzicht van de resultaten van de, voor vliegreizigers zoals ik, bekende gedragscreeningen en screenpoortjes op vliegvelden,

bedoeld om voor terrorisme verdachte mensen en materialen te ontdekken: 160 duizend mensen ondergaan deze procedure, 15.000 mensen worden doorgestuurd naar de politie en uiteindelijk werden 1266 personen (minder dan 1%) gearresteerd – voor geheel andere zaken dan terrorisme nota bene. 159 duizend mensen worden achteraf onnodig ondervraagd of gecheckt (USA Today, 18 November 2008).

Er moeten dus hoge eisen worden gesteld aan de afweging welke screening of screeningen er aan de bevolking zouden moeten of kunnen worden aangeboden.

De eerste doelstelling van ons onderzoek, evaluatie van vroegopsporing van ziekten is dan ook om de (eventuele) positieve gezondheidseffecten van screening te kwantificeren.

Ik illustreer dit graag eerst aan de hand van borstkankerscreening.

In 1927 wordt voor het eerst de röntgenfoto diagnostisch gebruikt bij een vrouw met onduidelijke klinische bevindingen. Het duurt tot 1960 voordat de voorspelling wordt gemaakt dat het op deze manier mogelijk zou moeten zijn om 95% van alle kwaadaardige borsttumoren voor operatie vast te stellen en dat de introductie van routine mammografie, een mooi woord voor het maken van röntgenfoto's van de borsten, hopelijk tot een sterftedaling zou kunnen leiden⁸.

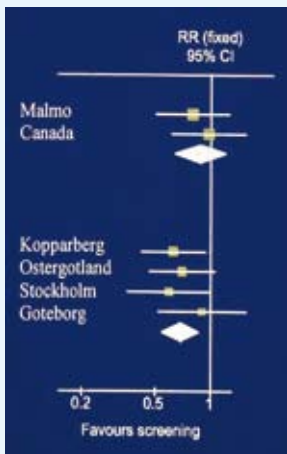
Om dit aan te tonen werd in 1963 het eerste gerandomiseerde onderzoek in New York bij 63.000 verzekerde vrouwen opgezet⁹. Op willekeurige wijze werden vrouwen in twee groepen verdeeld: de helft kreeg het aanbod van vroegopsporing door middel van een klinisch palpatie-onderzoek van de borst en mammografie, terwijl de andere groep dit aanbod niet kreeg. Er was tenslotte geen bewijs dat een dergelijke vroegopsporing gunstig was voor de vrouw. In 1971 komen de eerste gunstige resultaten van deze studie naar buiten: 32% minder sterfte aan borstkanker in de groep die vroegopsporing kreeg aangeboden¹⁰. Vanaf 1976 starten voornamelijk in Europa 6 onafhankelijke trials met het doel of vernieuwde beeldvormende techniek alleen (en zonder

klinisch onderzoek) ook – via de weg van vroegontdekking en vroegtijdige behandeling – de borstkankersterfte kan doen dalen.

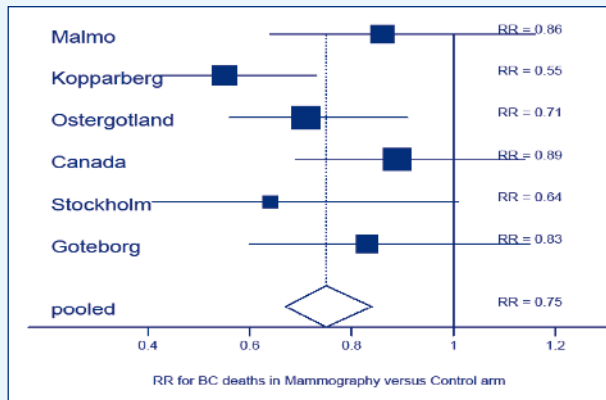
Na de eerste gunstige resultaten gepubliceerd in *The Lancet*¹¹, verrichten wij in 1987 een kosten-effectiviteitsanalyse die gunstige effecten voorspelt indien vrouwen in de leeftijdsgroep 50-69 jaar elke 2 jaar zouden worden uitgenodigd (in Nederland)¹². 28 jaar na de voorspelling kan het beleid definitief worden bepaald.

In 2000, inmiddels weer jaren na de publicatie van een gunstig effect uit de laatst gestarte trial, koppen de kranten met de vraag of borstkankerscreening wel gerechtvaardigd is?

Links in de figuur ziet u de puntschattingen van effect (en hun betrouwbaarheidsintervallen) van de 6 moderne hoog-kwalitatieve trials voor vrouwen van 50 jaar en ouder. Een (Deense) Cochrane review vat samen dat het niet gerechtvaardigd is: als je de onderste 4 Zweedse studies buiten beschouwing



Meta-analyse borstkankerscreening



De Koning *Lancet* 2000/Cochrane 2009/Rijnsburger et al. *IntJG* 2004.
Met dank aan Caspar Looman.

laat, zou er helemaal geen overtuigend bewijs zijn, getuige het diamantje dat deze rechte lijn $RR=1$ (geen effect) ook omvat¹³.

Philip Freriks opent het 8-uur journaal met de mededeling dat De Koning de vloer aanveegt met de Denen, en dat klopte eigenlijk wel zo'n beetje¹⁴. Er was geen reden om alleen de bovenste 2 studies als superieur te beschouwen, zoals de Denen deden.

Het verschil tussen de groepen personen wordt uiteindelijk namelijk door een groot aantal verschillende factoren bepaald: eenvoudige designkwesties als het screeningsinterval, de uitgenodigde leeftijdsgroepen, de opkomst, en opportunistische screening, op eigen initiatief in de controlegroep een foto laten maken, beïnvloeden de uiteindelijke grootte van het verschil tussen de 2 studiegroepen. Door modelleren toonden we aan dat deze relatief simpele designverschillen alleen al ervoor konden zorgen dat de ene trial 25% meer effect zou kunnen laten zien dan de andere¹⁵.

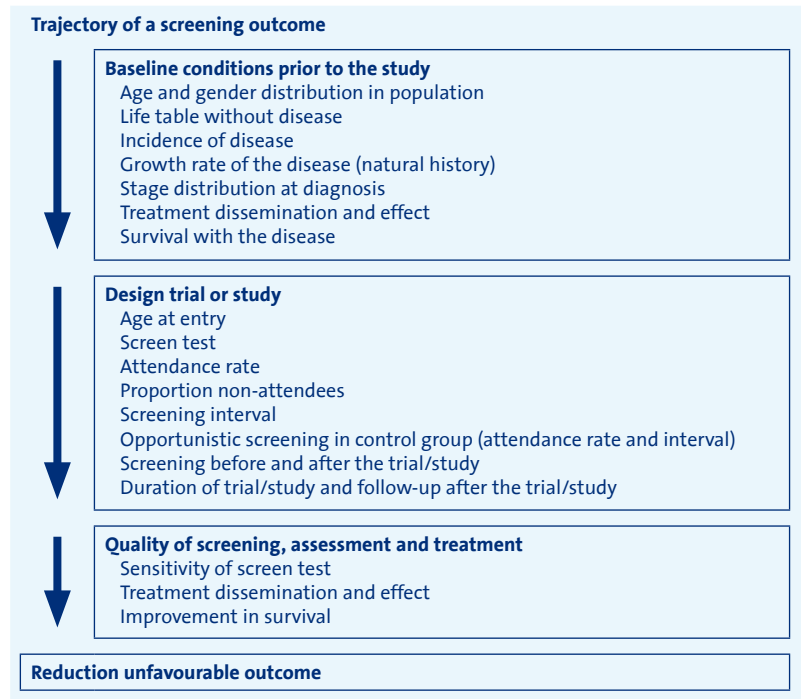
Maar nog veel belangrijker dan designkwesties is de mate van diagnosevervroeging en het effect hiervan. De belangrijke vraag gaat dus om de kwaliteit van de screening, hoeveel vrouwen worden er verwezen voor verder onderzoek en hoeveel tumoren worden hierdoor in een vroeg stadium ontdekt. En om de uitgangssituatie bij start van de studie. Indien in een regio vrouwen gemiddeld al vroegtijdiger bij een hulpverlener komen, kan de stadiumverschuiving die door screening wordt veroorzaakt in de ene studie minder zijn dan bij de andere, al is de kwaliteit van screening en therapie hetzelfde (zie stappenschema).

Zo werd juist in de Canadese studie bij de controle groep vrouwen de borsten door verpleegkundigen intensief onderzocht. Hierdoor kon ook in de controle-arm feitelijk als gevolg van eerdere ontdekking en behandeling de sterfte gedaald zijn, en het verschil tussen de 2 groepen vrouwen feitelijk verkleind hebben.

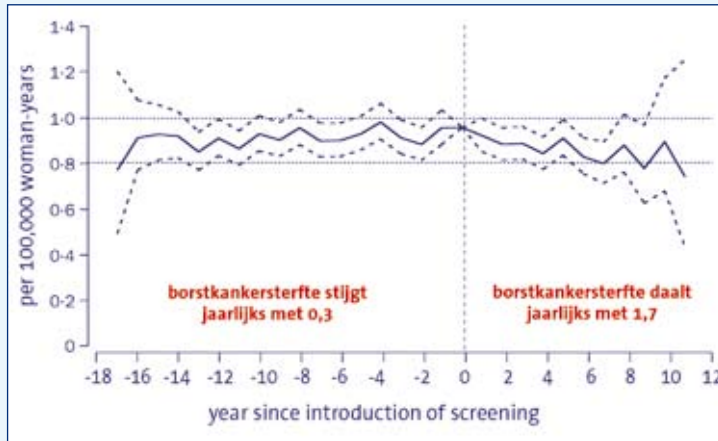
Met de individuele gegevens van deze Canadese trial hebben we met ons computermodel, MISCAN genaamd, aan kunnen tonen dat de kwaliteit van

de borstfoto's in deze Canadese studie geringer was dan in de andere Zweedse studies, maar dat tegelijkertijd de intensieve screening door verpleegkundigen in de zogenaamde controlegroep effect zou moeten hebben gehad¹⁶.

Rechts in de figuur op pagina 10 hebben we nu de modelmatige schatting voor de Canadese studie toegevoegd: als de studie opgezet zou zijn geweest zonder screening in de controlegroep, zoals alle andere studies dan zouden we minimaal een 11% sterfteverschil hebben verwacht, en is deze trial dus consistent met alle andere uitgevoerde borstkankerscreeningstrials.



Effect borstkankerscreening Nederland



Otto et al. Lancet 2003

Wereldwijd hebben vrouwen nu profijt gehad van deze lange evaluatiecyclus. In 2003 waren wij de eersten die in The Lancet aantoonde dat in Nederland borstkankerscreening (gevolgd door therapie) een stijging van de borstkankersterfte van 0,3 per 100.000 vrouwen per jaar wist om te draaien in een 1.7 (per 100.000 vrouwelijke populatie) daling per jaar¹⁷. Het knikpunt was bij start van screening in de gemeenten in Nederland (punt nul in de figuur).

Het is goed om te beseffen dat aan het begin van de lange cyclus een aantal zeer belangrijke beslissingen liggen, die vaak niet meer terug te draaien zijn. Twee grootschalige gerandomiseerde onderzoeken, een in Europa en een in Amerika werden geïnitieerd om een antwoord te geven of vroegopsporing met de PSA-test tot een sterftedaling kan leiden. De PSA-test – dat staat voor Prostaat Specifiek Antigeen – is een uitzonderlijke marker die een verhoogde waarde laat zien bij een goedaardige en een kwaadaardige vergroting van de prostaat, lang voordat de ziekte tot symptomen leidt. Maar bij het ontwerpen van de trial spelen de volgende zaken: Wat is de te prefereren doelgroep om

uit te nodigen, wat is de beste manier van recrutereren, hoeveel personen zijn er nodig, welke screeningstesten passen we toe en met welk screeningsinterval?

Cruciaal is dat het effect, in dit geval prostaatankersterfte, eenduidig en zonder vertekening wordt vastgesteld. In de screeningsgroep worden als gevolg van de test meer mannen met prostaatanker gediagnostiseerd. De kans is daarmee groter dat bij overlijden prostaatanker op de overlijdensakte komt te staan, terwijl dit misschien niet de doodsoorzaak was. In de Europese prostaatankertrial zijn door commissies van arts-deskundigen van alle patiënten met een prostaatankerdiagnose de doodsoorzaak bepaald aan de hand van volledige medische dossiers, zonder te weten tot welke groep de patiënt behoorde, screening of niet. In het Nederlandse deel van de prostaatstudie bleek uiteindelijk de echte doodsoorzaak in 15% te verschillen van de officiële doodsoorzaak^{18,19}.

Naast deze meer epidemiologische, public health criteria, is dan de juiste 'test-performance' (hoe goed is mijn test eigenlijk) van belang.

De juiste evaluatiebeslissingen vooraf

De juiste trialkeuze:

- uit te nodigen risicogroep (geslacht, leeftijd, etniciteit, ...)
- wijze van rekrutering
- willekeurige verdeling personen tussen S en C groep
- grootte en lengte van de studie
- screeningstest(en)
- screeningsinterval

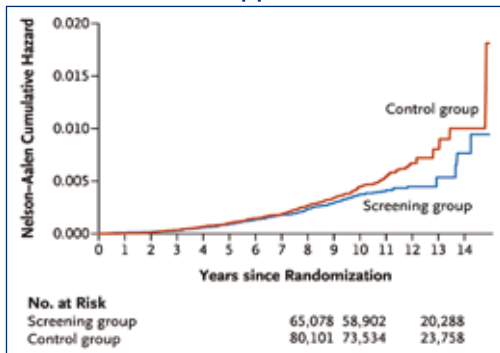
De juiste uitkomstmaat:

- blinde beoordeling doodsoorzaak (of andere uitkomstmaat)
- frequentie van interim-analyses
- complete follow up van data (via koppelingen)

De juiste 'testperformance'

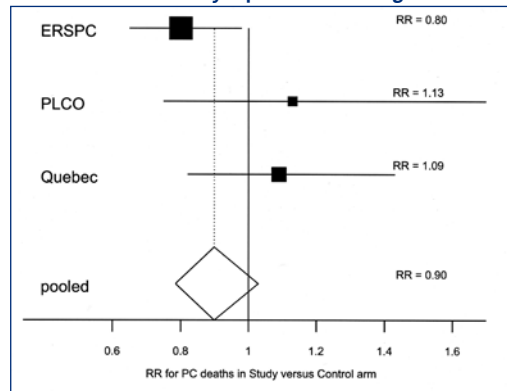
Het juiste geduld

Cumulatieve risico op prostaatankersterfte



Schröder F et al. N Engl J Med 2009

Meta-analyse prostaatscreening



Enkele maanden geleden kwamen na 9 jaar follow up de resultaten beschikbaar van de Europese prostaatstudie onder 162.000 mannen van 55-69 jaar: in de groep die uitgenodigd was voor PSA-testen was de kans om aan prostaatanker te overlijden met 20% gedaald²⁰. De bovenste lijn in de linker figuur geeft de totale sterfte aan prostaatanker in de controlegroep aan, en de onderste lijn die voor de voor screening uitgenodigde groep: 20% lager.

Inmiddels zijn er 3 gerandomiseerde prostaatstudies gepubliceerd, en het zal niet lang duren* voordat de volgende meta-analyse gepubliceerd zal worden (rechter figuur). Er zijn 2 kleinere studies die geen reductie laten zien: samen tonen de studies een 10% prostaatankersterfte reductie aan.

U zult inmiddels begrijpen dat een dergelijke meta-analyse van geen enkele waarde is, indien niet zorgvuldig gekeken is naar de kenmerken van de geïncludeerde personen, van de omvang en de kwaliteit van screening in de screeningsgroep, van opportunistische screening in de controlegroep, en van de uitgangssituatie van de populatie bij start van elke studie.

The New York Times verwoordde het met “The European trial is the most important study for men’s health in history”²¹.

* Inmiddels verschenen sept. 2010 BMJ.

Toch is dit niet voldoende om prostaatscreening te kunnen gaan aanbieden. Het is cruciaal om ook de nadelen te kwantificeren, en dat is daarmee de 2e belangrijke doelstelling in de evaluatie van vroegopsporing.

Elke screening op populatieniveau kent ook nadelen voor een grote groep personen, zoals met de luchthavencontroles al vluchtig was aangegeven. U ziet hieronder de belangrijkste ongunstige neveneffecten van vroegopsporing (in het algemeen) opgesomd. Er zijn fout-positieven: de persoon wordt doorgestuurd vanwege een ongunstige screeningsuitslag, maar hij/zij blijkt geen ziekte te hebben. De diagnose wordt eerder gesteld; dat lijkt gunstig, maar tegelijkertijd weten personen dan ook eerder en dus over een langere periode dat ze de ziekte hebben en zij zullen mogelijk ook over een langere periode de bijwerkingen van behandeling ondervinden. Voor een gedeelte van de groep staat er helemaal geen winst tegenover deze vroegontdekking, en er zullen zelfs personen zijn bij wie de diagnose anders nooit gesteld zou zijn, omdat ze in die vervroegingsfase aan een andere ziekte overlijden. Dergelijke personen worden dan zelfs overbehandeld.

Aan de screening en vervolgdagnostiek zelf kunnen risico's zitten, en de screening kan leiden tot onbedoelde detectie van andere, misschien niet behandelbare, aandoeningen. Ten slotte kan er onterechte geruststelling ontstaan, hetgeen betekent dat personen na een gunstige screeningsuitslag

Tabel 1

Belangrijkste neveneffecten vroegopsporing

- Foutpositieven
- Eerder (de wetenschap van) de diagnose
- Eerder (en over een langere periode) de bijwerkingen van behandeling
- Vroegontdekking, maar geen gezondheidswinst
- Extra ontdekken van ziekte, en overbehandeling
- Screeningsrisico's, en onbedoelde detectie van andere ziektes
- Mogelijk onterechte geruststelling: toch geconfronteerd worden met de diagnose, of de aandoening nu later ontdekken dan gebruikelijk
- Vroegopsporing leidt tot vroeg opsporen

onterecht een bezoek aan de arts uitstellen bij klachten omdat ze menen zekerheid uit de screeningsuitslag te kunnen putten.

In vroegopsporing is de mate van vroeg opsporen tegelijkertijd ook een indicatie van de ongunstige effecten, “als u begrijpt wat ik bedoel”.*

Een van de belangrijkste redenen om bij het aanbieden van vroegopsporing vaak ook een bovenste leeftijdsgrens te kiezen heeft te maken met deze balans van gunstige en ongunstige effecten²². Bij screening op borstkanker is daarom een bovengrens van 76 jaar bepaald voor vrouwen. Twee jaar geleden werd de Staat in kort geding gedaagd deze grens af te schaffen.

De rechter formuleerde het als volgt:

“Met het stellen van een leeftijdsgrens van 76 jaar voor bevolkingsonderzoek naar borstkanker maakt de Staat onderscheid naar leeftijd. De kern van dit kort geding kan worden samengevat in de vraag of de afweging van nut en risico's voor de specifieke leeftijdsgroep van 76-jarigen en ouder anders uitvalt dan voor de groep met een leeftijd tussen 50 en 76 jaar (de huidige doelgroep voor het bevolkingsonderzoek). De hier bedoelde afweging vindt op collectief niveau plaats. Bevolkingsonderzoeken betreffen immers naar hun aard categorieën van personen.

De redenen waarom de Staat de hier bedoelde bovengrens toepast kunnen niet als onvoldoende solide worden beschouwd. Met het stijgen van de leeftijd van de betrokken personen moet ernstig rekening worden gehouden met het gegeven dat een vroegtijdige ontdekking en een daarop gevolgde behandeling, waaraan ook altijd nadelen zijn verbonden, in steeds mindere mate de kwaliteit van leven verhoogt of levensverlengend werkt.

Het gehandhaafde onderscheid is niet discriminerend en dus niet onrechtmatig.”²³ Inmiddels is dit in hoger beroep nogmaals bevestigd²⁴.

* Naar Olivier B. Bommel

Vonnis in kort geding van 30 oktober 2007:

“Met het stellen van een leeftijdsgrens van 76 jaar voor bevolkingsonderzoek naar borstkanker maakt de Staat onderscheid naar leeftijd. De kern van dit kort geding kan worden samengevat in de vraag of de afweging van nut en risico's voor de specifieke leeftijdsgroep van 76-jarigen en ouder anders uitvalt dan voor de groep met een leeftijd tussen 50 en 76 jaar. De hier bedoelde afweging vindt op *collectief* niveau plaats. Bevolkingsonderzoeken betreffen immers naar hun aard categorieën van personen. De redenen waarom de Staat de hier bedoelde bovengrens toepast kunnen niet als onvoldoende solide worden beschouwd. Met het stijgen van de leeftijd van de betrokken personen moet ernstig rekening worden gehouden met het gegeven dat een vroegtijdige ontdekking en een daarop gevolgde behandeling, waaraan ook altijd nadelen zijn verbonden, in steeds mindere mate de kwaliteit van leven verhoogt of levensverlengend werkt.

Het gehandhaafde onderscheid is niet discriminerend en dus niet onrechtmatig”.

Vonnis 9 februari 2010 na hoger beroep:

“Uit de onderzoeken kan worden afgeleid dat het advies van de Gezondheidsraad op goede leest is geschoeid. Uit de stukken kan hooguit worden afgeleid dat er discussie gaande is binnen de medische wetenschap over de vraag of de leeftijdsgrens voor deelname in het BOB gehandhaafd moet blijven. En dat binnen de medische wetenschap daarover geen eenduidig standpunt wordt ingenomen. Dit is onvoldoende om vast te stellen dat de Staat bij het vormen van zijn beleid het advies van de Gezondheidsraad niet moet opvolgen.

(Andere) aanwijzingen voor de onjuistheid van het beleid van de Staat zijn er niet. Ook in omringende landen is er een eindleeftijd voor het BOB (in de meeste landen bij 69 jaar, soms bij 64 of 65 jaar en soms bij 74 of, zoals in Nederland, 75 jaar).

Gelet op het voorgaande heeft de Staat voorshands (in dit kort geding) aannemelijk gemaakt dat er een rechtvaardiging is voor het stellen van de leeftijdsgrens op 75 jaar. Anders dan Clara Wichmann c.s. aanvoeren is dit niet in strijd met het zelfbeschikkingsrecht van vrouwen boven 75 jaar. Het gaat hier immers niet om de vraag of individuele vrouwen boven de 75 jaar zich mogen en kunnen laten onderzoeken op mogelijke borstkanker: dat recht hebben zij. Aan de orde is de vraag of de Staat de hele categorie van vrouwen van die leeftijd in het door de instellingen georganiseerde BOB moet laten opnemen.

Het principaal hoger beroep is ongegrond”.

Overdiagnose en overbehandeling zijn specifieke ernstige nadelen van vroegopsporing.

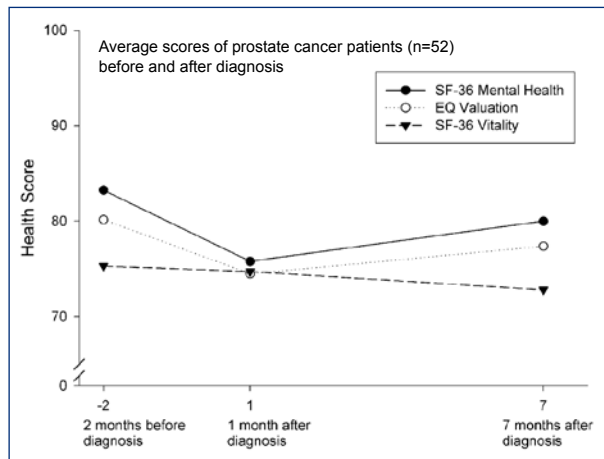
Wij hebben aangetoond dat bijvoorbeeld alleen de diagnosestelling van kanker al een tijdelijke 20%-dip in iemands kwaliteit van leven veroorzaakt²⁵. De figuur geeft 3 verschillende instrumenten door de tijd weer. Een persoon scoort ongeveer 84 op een schaal van 0-100 in de situatie zonder klachten, maar 75 na het horen van de diagnose kanker.

7 maanden nadat personen de diagnose hebben gehoord, zijn deze verschillen weer wat kleiner, personen passen zich mogelijk aan, want in feite hopen ze dat de vroege ontdekking niet voor niets is geweest.

Ik heb eerder al genoemd dat dat voor een gedeelte helemaal niet het geval is: sommige personen zouden anders nooit de diagnose en de bijwerkingen van diagnostiek en behandeling gehad hebben, en anderen hebben helemaal geen profijt van de eerdere behandeling.

Vroeg vaststellen van de ziekte als gevolg van vroegopsporing kan dus ook een belangrijk nadelig effect zijn.

Impact van de diagnose



Korfage et al. EJC 2006



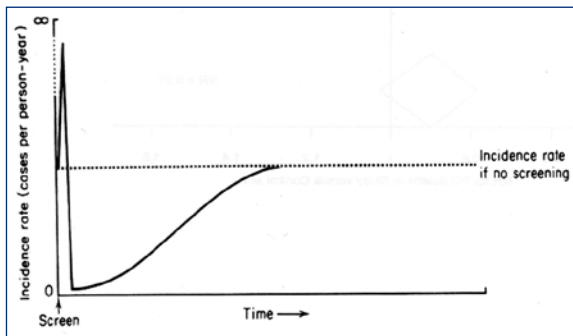
Maar wat voor het ene individu komt als een donderslag (bij heldere hemel), tegen hetzelfde fenomeen kan door een ander individu totaal anders aangekeken worden²⁶.

Een manier om met dit bliksem-ei fenomeen om te gaan, ligt in het concept 'geïnformeerde besluitvorming' tot deelname. Het is feitelijk aan het individu om af te wegen of de voorgespiegelde voordelen opwegen tegen de onvermijdelijke nadelen. Er is discussie of de huidige informatievoorziening voor grootschalig aangeboden programma's in die zin voldoen. In inmiddels 4 door ons verrichte onderzoeken tonen we aan dat de mate van geïnformeerde besluitvorming behoorlijk goed is in Nederland: 88% voor borstkankerscreening, 86% voor zwangere vrouwen die deelnemen aan het structureel echoscopisch onderzoek bij 20 weken en 76% bij prenatale screening naar Down syndroom. Als geen voorlichtingsinformatie wordt toegestuurd, ligt een dergelijk percentage zeer laag.

Ik heb uitgebreid stil gestaan bij het vaststellen van de gunstige en vaststellen van de ongunstige effecten van vroegopsporing. De 3e doelstelling van evaluatie-onderzoek op dit terrein is ten slotte een correcte schatting maken van al deze effecten en de kosten van verschillende screeningsscenario's en het meest optimale scenario aanbevelen²⁷⁻²⁹.

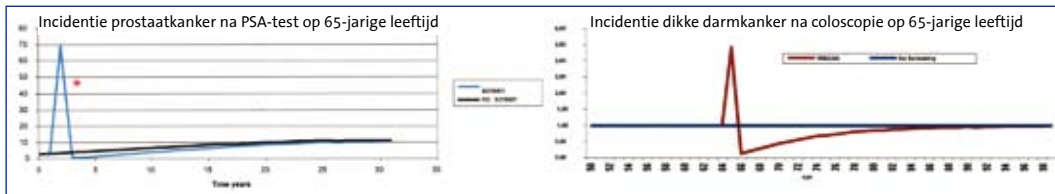
Bij toenemende investeringen, bijvoorbeeld heel frequent screenen kan wel meer effecten worden bereikt, maar relatief steeds minder: de gewonnen levensjaren nemen steeds minder toe voor de populatie.

Zoals eerder reeds aangehaald, een heel belangrijke beslissing is ook die van het te kiezen screeningsinterval³⁰. In de eerste figuur ziet u het theoretisch kader dat ik net schetste; screenen van tot dan toe nog asymptomatische personen zal aanvankelijk leiden tot een belangrijke stijging van het aantal personen bij wie de ziekte vroegtijdig wordt ontdekt. Afhankelijk van de gevoeligheid van de test, zal in de directe periode daarna weinig of geen ziekte klinisch worden vastgesteld, waarna sneller of minder snel weer ziektegevalen aan het licht komen en klinisch worden vastgesteld. De grootte van de veranderingen en het patroon bepalen welk screeningsinterval vermoedelijk verstandig is. Een te kort interval betekent veel testen, veel kosten, veel nadelen van screening, zonder veel ontdekken van nieuwe vroege afwijkingen. In onderstaande figuren kunt u de door ons berekende detectie van vroege tumoren zien, bij eenmalige prostaatscreening en eenmalige dikke darmkanker screening met coloscopie.



Indicaties screeningsinterval – theorie
Morrison 1992

Indicaties screeningsinterval – praktijk



U ziet dat het maar liefst meer dan 25 jaar na de screening duurt voordat we weer terug zijn naar het uitgangsniveau, de incidentie van voordat we begonnen met screenen. Een hele snelle screeningstest erna zal dus niet erg nuttig zijn. Het probleem bij de start van een trial, is dat er vaak slechts weinig informatie beschikbaar is om het juiste interval te kiezen. Bij de Europese prostaatrial hebben wij zo voor aanvang het 4-jaars interval bepaald met – het sterretje in de figuur – de schaarse data over slechts 2500 screeningen die we hadden in 1992³¹.

In het meest rechtse plaatje is dit uitgedrukt voor dikke darmkanker screening, maar nu staat op de X-as de leeftijd van de persoon: na een screening d.m.v. coloscopie, met een kijk slang de dikke darm van binnen bekijken, op 65-jarige leeftijd is de kans 5-10 jaar erna fors verkleind (van Balgeooijen, persoonlijke mededeling).

Deze afwegingen moeten bij elk aanbod van vroegopsporing worden gemaakt.

Als we het hebben over de ontwikkeling van kinderen, dan kunnen de nadelige gevolgen van een bepaalde achterstand bijvoorbeeld ook omvangrijk zijn. Alle tot nu toe beschreven afwegingen bij vroegopsporing gelden hier echter onverminderd³².

In een grootschalig gerandomiseerd onderzoek onder 11.440 kinderen hebben we aangetoond dat de toevoeging van een specifieke taalsignalering op 2-jarige leeftijd ertoe leidde dat op 3-jarige leeftijd 40% meer kinderen vroegtijdig voor taalproblemen werden behandeld dan in een vergelijkbare controlegroep, zonder die specifieke screening³³. Uiteindelijk blijkt consistent³⁴ dat als gevolg van deze vroegtijdige ontdekking en behandeling 30% minder kinderen in het speciaal onderwijs beland te zijn, op 8-jarige leeftijd (2,7% i.p.v. 3,7%)³⁵.

We gingen verder door een uniform screeningsscenario vast te stellen. Op jonge leeftijd zijn er verschillende signaleringsmomenten op het consultatiebureau (zie ook pagina 50). Met een eenduidige scoreverdeling op 2-jarige leeftijd kan vervolgleid worden bepaald. Een kleine 5% van kinderen



op 2-jarige leeftijd (VTO-Taal signaleringsinstrument) hebben dusdanig problemen met taal, dat zij profijt hebben van vroegtijdige interventies en op jonge leeftijd reeds gesignaleerd moeten worden, en direct verwijzing voor diagnostiek verdienen. Een middengroep van 20% dient eerder terug te worden gezien en met aanvullende screeningsinstrumenten tussen de 2e en 3e verjaardag te worden gescreend voor een definitief oordeel³⁶.

Toch hebben kinderen in de afgelopen 3 jaar nog niet geprofiteerd van deze nieuwe vroegdetectie, en kunnen we dus gerust zeggen dat er nu al onnodig 7500 kinderen in het speciaal onderwijs zullen belanden.

Waar heeft dat mee te maken? Spreken we niet voldoende elkaars taal? Een veel gehoorde mening is: "Maar, we kunnen toch niet een screening instellen, waarbij uitsluitend op basis van een testuitslag wordt vastgesteld of verwijzing noodzakelijk is?"³⁷

Dat is een misvatting. Let wel, een rechtvaardiging ontbreekt voor het stellen van ontwikkelingsvragen, zonder een goede mogelijkheid van kwantificeren of scoren, en zonder een duidelijk afkappunt voor verwijzen. "If screening is worth doing, it is only worth doing if done well" ³⁸.

"Maar, wat met mijn eigen expertise dan?"

Dat is een onterechte angst. Let wel, in vrijwel elke vorm van screening gaat het om mensenwerk: alleen een goed opgeleide professional weet aardig te schiften tussen plus en niet-plus. De screening is zijn/haar belangrijke hulpmiddel. "Do not rob the worker of his right to pride in workmanship" ³⁸.

Er moet juiste informatie zijn om publiek en beroepsgroepen in te lichten. Wilson en Jungner waren een van de eersten die in 1968 10 principes opstelden³⁹. Inmiddels lijken de 10 principes soms wel uitgegroeid te zijn tot dogma's. Ik heb reeds menigmaal gehoord dat als we het nu geen screening noemen, we dan tenminste niet aan de criteria van Wilson en Jungner hoeven te voldoen.

De nieuwe criteria anno 21e eeuw voor vroegopsporing zijn echter duidelijk (zie tabel).

Er moet omvangrijke gezondheidswinst zijn, in termen van gewonnen levens-

Tabel 3

De nieuwe vroegopsporingscriteria

Omvangrijke gezondheidswinst

- gewonnen levensjaren
- cognitieve, motorieke en/of sociaal-emotionele verbetering
- vergroting van (be)handelingsopties

Effecten eenduidig vastgesteld, liefst in gerandomiseerd onderzoek

Geringe mate van ongunstige neveneffecten

- mate van vervroeging, overdiagnose en neveneffecten geschat
- voor kwaliteit van leven gecorrigeerde levensjaren

Balans inzichtelijk voor deelname

Redelijke verhouding tussen kosten en effecten

Implementatie leidt niet tot omvangrijke onbedoelde effecten (ongelijkheid)

Ontwikkelingen op andere terreinen veranderen de afweging op korte termijn niet

jaren, verbetering in cognitieve, motorieke en/of sociaal-emotionele ontwikkeling, en/of een belangrijke vergroting van handelings- of behandelingsopties. Deze effecten moeten eenduidig zijn vastgesteld, liefst, indien mogelijk, in gerandomiseerd onderzoek.

Ten tweede moet er in geringe mate sprake zijn van ongunstige neveneffecten. De mate van vervroeging, de mate van overdiagnose en de neveneffecten dienen te zijn geschat, en in termen van voor kwaliteit van leven gecorrigeerde levensjaren te zijn uitgedrukt. De verwachte balans tussen gezondheidswinst en neveneffecten op groepsniveau dienen voor deelname inzichtelijk te zijn gemaakt.

Er moet een redelijke verhouding tussen kosten en effecten zijn (zodoende ontstaat er ruimte voor programma's waar voor dezelfde investering meer effecten mogen worden verwacht), en er dient voor gezorgd te zijn dat implementatie niet tot omvangrijke onbedoelde effecten leidt. Ten slotte dienen ontwikkelingen op andere terreinen, die de vroegopsporing zouden kunnen gaan beïnvloeden, de afweging niet substantieel te doen veranderen.

Nu u de criteria weet, hoe nu verder?

Op de foto ziet u dat mijn ouders hun tijd ver vooruit waren. Elke 26e juni deden ze mee aan de, inmiddels bekend staande, landelijke screeningsdag. Hier ziet u ze met 2 van hun 3 kleindochters, en komen ze net van het Centrum voor Jeugd en Gezin, waar ze allereerst uitleg hebben gekregen om de kleinkinderen niet op scoliose te laten onderzoeken, ondanks de herhaaldelijke flyers die ze hierover in de bus krijgen van 3 commerciële bedrijven. De kinderen zijn daarna ondervraagd met de 'hoe ga ik me lekker voelen in de toekomst'-lijst, en er werd vastgesteld of de eerdere risicotaxaties op 1- en 2-jarige leeftijd terecht zijn geweest.

Ze zijn nu op weg naar de preventieve vaatpoli, waar ze elkaar een onderzoek cadeau doen; mijn vader krijgt een



echo van zijn buikslagader, en mijn moeder een echo van de slagaderen naar de eierstokken, met continue meting van het CEA-gehalte.

Met de indrukwekkende technologische ontwikkelingen en mogelijkheden die uit fundamenteel onderzoek voortvloeien, neemt de vraag naar vroegopsporing spectaculair toe. Veel aandoeningen, afwijkingen of achterstanden zijn in een vroeg stadium herkenbaar, of zullen dat in de nabije toekomst worden. De verkalking van de belangrijkste slagaderen, de ontkalking van de botten en beginnende stadia van dementie kunnen mogelijk worden gevisualiseerd, maar zelfs bij de geboorte kunnen biomarkers in navelstrengbloed mogelijk de kans op obesitas voorspellen.

Tegelijkertijd wordt de individuele persoon veeleisender, kan zelfsteden aanvragen via internet, wordt op zijn/haar wenken bediend door commerciële bedrijven of welgemeende hulpverleners. Kan met alle risk-calculators op het internet het individu niet zelf goed bepalen of nu juist zij in aanmerking zouden moeten komen? Is er een betere inschatting van risicogroepen in de toekomst?

U moet zich echter goed realiseren dat het risico op de ziekte slechts één, hoewel belangrijk, facet van vroegopsporing is. Maar bevolkingsonderzoek is niets anders dan de juiste risicogroep vaststellen waarvoor – als groep – de voordelen in belangrijke mate opwegen tegen de nadelen. Individuen hebben veelal niet meer informatie beschikbaar om een voor hun specifiekere afweging te kunnen maken. Tegelijkertijd zullen er veel groepen, met minder geld, minder internet of afwegingsvermogen, de boot missen bij een dergelijke afweging.

Er zullen naast deze ontwikkelingen ook professionele veranderingen en economische ontwikkelingen zich voordoen. Huisartsen, EHBO's, bedrijfsartsen, paramedisch personeel zullen zich meer en meer willen toeleggen op screeningsmogelijkheden, als ze de middelen en de menskracht ervoor hebben.

Dit alles vergt des te meer een kwantitatieve onafhankelijke afweging van voor- en nadelen.

Op een groot deel van deze terreinen willen we ons dan ook in de toekomst richten, om enerzijds optimale gezondheidswinst te kunnen bereiken, en anderzijds gezondheidsschade te voorkomen of tegen te gaan.

En dan hebben we het niet alleen maar over high-tech(nology).

Simpele technologie verdient hetzelfde toetsingskader als hiervoor geschetst. Want feitelijk bestaat er in vroegopsporing geen simpele technologie, en is de term simpele technologie een understatement.

De zwangerschapsecho bij 20 weken, vaak eufemistisch de pretecho genoemd, heeft zich dusdanig ontwikkeld dat veel vroege afwijkingen kunnen worden gevisualiseerd. Vroege ontdekking kan eerdere erkenning en eerdere handelingsopties geven.

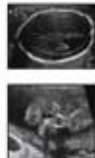
Links ziet u het beeld van een "open ruggetje". De middelste plaatjes tonen de 2 boezems en kamers van het hart, met daaronder een situatie waar duide-

Van low-tech naar hightech naar multi-tech

Structureel echoscopisch
onderzoek
(20 weken zwangerschap)



Open ruggetje



foutpositieven

eerdere diagnose

overdiagnose
(tijdelijk afwijkend)

risico's diagnostiek

onterechte
geruststelling



Hartafwijking



lijk nog maar 2 of 3 grote holttes zichtbaar zijn. Rechts nog meer details van andere lichaamsdelen.

Alle eerder genoemde ongunstige neveneffecten van vroegopsporing zijn echter ook van toepassing. Het ontdekken van vroege afwijkingen, waarvan het nut van vroeg ontdekken niet altijd vaststaat, waarvan sommige passagère zijn, en waarbij fout-positieve uitslagen, met alle mogelijke nadelige effecten van dien aanzienlijk kunnen zijn. Ook erg belangrijk hier zijn de risico's van aanvullende diagnostiek, mogelijk een niet-bedoelde zwangerschapsafbreking, maar ook onterechte geruststelling. Verschillende afwijkingen komen alsnog na de geboorte aan het licht.

In een grootschalige studie verricht in de regio Rijnmond waarbij we naar 79.000 mensen een simpele lintcentimeter stuurden bleek in de hoogrisico groep bij 6% de glucosewaarde verhoogd en bij nog eens 13% het cardiovasculaire risicoprofiel ongunstig te zijn, zonder dat deze personen of hun huisartsen hier van af wisten⁴². Er is dus een interessant low-tech instrument, waar menige internet riskcalculator nog een puntje aan kan zuigen, waarmee we afwijkingen die er echt toe doen vroegtijdig kunnen vaststellen. Mijns inziens wordt een interessante ontwikkeling om low-tech en high-tech te gaan combineren, zoals met een lintcentimeter en vragenlijst voor velen en een coronaire kalkscore voor sommigen. Maar, ook voor simpele technologie moet eenduidig bewijs zijn dat een grootschalige aanpak meer voordelen dan nadelen oplevert.

Dames en heren, ik heb u laten zien dat er duidelijke en universeel toepasbare nieuwe criteria voor (evaluatie van) vroegopsporing zijn.

Duidelijke afwegingen om een bepaalde vorm van vroegopsporing wel of niet aan te bieden, kunnen en moeten gemaakt worden, zowel om gezondheidswinst te behalen als gezondheidsschade te voorkomen. Er zijn geen andere criteria voor reeds bestaande programma's of nieuw in te voeren programma's. Er zijn geen andere criteria als we het risico-taxatie of signalering noemen. Er moet durf zijn om bestaande screeningen te verbeteren,

te veranderen of af te schaffen. Er moet wil zijn om belangen opzij te zetten; er moet regie zijn om in te kunnen grijpen op basis van correct uitgevoerd evaluatie-onderzoek.

Vroegopsporing vereist een aanpak van hoge kwaliteit met een uniform protocol en verwijsbeleid; dit staat niet haaks op respect voor autonomie van het individu en de professional.

Voorop staat de juiste hoogrisico groep screening aan te bieden; Individuele preventie bestaat niet; individuelere op-maat preventie misschien wel, maar dit vereist nog veel en onafhankelijk onderzoek. Er zal in toenemende mate behoefte zijn aan scholing van publiek, patiënten en professionals over de zin en onzin van vroegopsporing, en graag wil ik dat in NIHES-verband versterken.

Voedingsmiddelenfabrikanten geven tegenwoordig hun gunstige en minder gunstige nutriënten weer op de verpakkingen voor het publiek. Is dit de toekomstige screen bijsluiters?

Screen Facts/Valeurs de depistage	
per 1.000 servings (PSA test)	
Minder prostaatkankersterfte	8
Gewonnen levensjaren	66
Max. voorkomen prostaatkanker sterfte	32%
Minder behandelingen tegen uitzaaiingen	32%
Eerdere diagnose	75
Langer weten	11 jr.
Overbehandeling	31
Foutpositief	176
Gewonnen kwaliteitsjaren	20

Gebaseerd op een 162.000 grote populatiestudie. Uw dagelijkse waarden kunnen groter of lager zijn afhankelijk van uw drempel voor doorsturen, uw behandeling en uw voorgeschiedenis. We kennen momenteel geen andere interessante alternatieven.



Dames en heren, via het eerder getoonde scala aan onderzoeken tracht onze onderzoeksgroep, inmiddels met bijna 35 vrouwen en mannen, de gezondheid in Nederland en daarbuiten te verbeteren.

Zonder samenwerking met externe groepen bestaan we echter niet, of zoals mijn collega Lex Burdorf het uitdrukte, de lijst met personen in het dankwoord is soms langer dan de resultaten sectie zelf.

Velen kan ik, nu of straks, nog bedanken.

Een aantal mensen niet meer persoonlijk. Ik denk aan Kadri El Naggar, Marijke Vervoort, Jan Hendriks, Rinse Huizinga en aan mijn moeder.

Ik prijs me gelukkig dat ik mijn 90-jarige vader nog wel persoonlijk kan bedanken, voor deze toga en zijn gehele opvoeding.

Verder wil ik allereerst mijn zonen bedanken o.a. voor het feit dat ze weer te vaak genoeg moesten nemen met een voetbalshirt als ik weer van een werkreis terug kwam, en ik wil uiteraard Bianca bedanken voor haar, vrijwel onvoorwaardelijke liefde en steun voor mij in al die jaren.

Tabel 4

Evaluatie van vroegopsporing

Kwantificeren van de eventuele positieve gezondheidseffecten

- evalueren RCTs
- zelf uitvoeren RCTs, case-control of cohort studies
- evaluatie lopend bevolkingsonderzoek
- modelleren van ziekte en invloed van test en therapie

Kwantificeren van de ongunstige gezondheidseffecten

- evalueren van de impact op kwaliteit van leven
- modelleren van ziekte en invloed van test en therapie
- schatten van mate van geïnformeerde besluitvorming

Voorspellen van het meest optimale screeningsscenario

- modelleren van ziekte en invloed van test en therapie
- kosten en effecten schatten



© Ed de Jong

Werk gerelateerd gaat mijn dank in eerste instantie uit naar Johan Mackenbach en de benoemingscommissie voor het in mij gestelde vertrouwen. Mijn dank gaat ook uit naar Paul van der Maas als feitelijk mijn opleider in het doen en leiden van wetenschappelijk onderzoek op dit gebied. Veel dank ben ik verschuldigd aan alle huidige en voorgaande medewerkers van evaluatieonderzoeken, waarbij ik m.n. de seniormodellereurs in het zonnetje wil zetten: Gerrit van Oortmarsen, Rob Boer, Gerrit Draisma, Iris Lansdorp-Vogelaar en Eveline Heijnsdijk.

Ik wil Dik Habbema en Marjolein van Ballegooijen in het bijzonder bedanken voor de prettige samenwerking en afstemming, waarbij we altijd –zonder onderlinge problemen- de evaluatie van vroegopsporing als doelstelling hadden.

Speciale dank gaat natuurlijk uit naar Arry de Bruijn, mijn steun en toeverlaat op het werk, waarvan ik hoop dat ze tot aan mijn pensionering bij me blijft. Ik wil uiteraard heel de afdeling bedanken voor de jarenlange prettige sfeer, en i.h.b. wil ik dan Lex Burdorf, Ed van Beeck, Rikard Juttmann, Ewout Steyerberg en Marie-Louise Essink-Bot bedanken.

Ik wil Jan Klijn en Joop van Dongen bedanken als de eerste klinici na mijn coschappen, van wie ik leerde wetenschappelijk onderzoek in ziekenhuizen te doen. Ik wil uiteraard alle samenwerkingspartners bedanken, welke zo talrijk zijn dat ik hier slechts enkele kan noemen, i.h.b. de afdeling Urologie, Interne Oncologie, Longziekten, i.h.b. Rob van Klaveren, de afdeling Verloskunde en Vrouwenziekten, Kindergeneeskunde, Oogheelkunde, Cardiologie en Epidemiologie, Huisartsgeneeskunde, Gynaecologie en Gastro-enterologie van het Erasmus MC, het Landelijk Evaluatie Team voor bevolkingsonderzoek naar Borstkanker, en het Landelijk ReferentieCentrum voor bevolkingsonderzoek naar Borstkanker, en alle externe samenwerkingspartners in de NELSON-studie, i.h.b. Matthijs Oudkerk en Willem Mali. Ik wil de Nederlandse Stichting voor het Dove en Slechthorende Kind bedanken voor de lange samenwerking op een blijkbaar taalgevoelig onderwerp; ik wil van alle ziekenhuizen buiten de stad Rotterdam graag het Antoni van Leeuwenhoekziekenhuis bedanken, en alle financiers van ons onderzoek, i.h.b. het RIVM, ZonMw en KWF Kankerbestrijding bedanken voor hun vertrouwen. I would like to thank the NCI and the CISNET group for their trust in our work. I would especially like to thank the members of the Data Monitoring Committee of the European Randomised Screening for Prostate Cancer Study, Philip Smith, Jan Adolfs-son, Timo Hakulinen and Sue Moss, with whom I have worked together since 1994. The success of the European Prostate Trial is also their success.



Links: portret van Thomas More.

Rechts: portret van Desiderius Erasmus.

Beide door Hans Holbein de Jonge.

Dames en heren, het leven is een mysterie.

Van mijn 4e tot mijn 13e liep ik dagelijks 4-maal enkele kilometers over de Erasmusweg om van mijn huis naar de lagere school en vice versa te gaan.

Van mijn 13e tot mijn 18e ging ik naar het Thomas More College, in een geheel andere wijk van Den Haag.

Erasmus beschouwde Thomas More als een van zijn belangrijkste leermeesters, aan wie hij de Lof der Zotheid opdroeg.

Zojuist heb ik de Thomas Morelaan overgestoken om bij de Erasmus Universiteit mijn ambt te aanvaarden.

In het mysterie dat leven heet, hoop ik met onze onderzoeksgroep en externe partners een bescheiden, maar belangrijke bijdrage te leveren, in de verbetering van de gezondheid van mensen.

Ik heb gezegd.

Mijn mysterieuze massa



Opgedragen aan mijn vader die op 10.02.2010 is overleden.





Geni?



THE EUROPEAN
PROSTATE CANCER
COALITION

Chairman
Steering Committee
Tom HUDSON, Ireland
rthudson@eircom.net

Antwerpen, 5 juni 2009

Ann Dr. H.J. de Koning
Erasmus Universiteit Rotterdam
Complex Woudestein
Burgemeester Oudlaan 50
3062 PA Rotterdam
NEDERLAND

Beste Harry,

Joaquim DA CRUZ DOMINGOS
Treasurer, Portugal
cruzdomingos@nscabo.pt

Louis DENIS
Secretary, Belgium
louis.denis@skynet.be

Lars ELIASON
Member, Sweden
lars@eliason.se

Günter FEICK
Member, Germany
g.feick@web.de

Mike LOCKETT
Vice Chairman, UK
mike.lockett@tiscali.co.uk

Roland MUNTZ
Member, France
info@anamacap.fr

Hannu TAVIO
Vice Chairman, Finland
hannu.tavio@cancer.fi

Ex officio
Alberto COSTA, Italy
European School of Oncology
director@eso.net

Ex officio
Hein VAN POPPEL, Belgium
Chair Scientific Committee
hendrik.vanpoppeel@uzleuven.be

Het nieuws van uw aanstaande benoeming komt niet als een verrassing maar eerder als eindelijk. Onze meest hartelijke gelukwensen.

Met spijt, we zijn in het buitenland, moeten we ons verontschuldigen maar we hopen alsnog een exemplaar van uw rede te krijgen voor onze bibliotheek. De titel op zichzelf is al goed voor een Harry Potter succes. De mysterieuze massa. Veel vreugde nadien en tot binnenkort.

Met vriendelijke groeten,

Louis Denis

Secretariaat Europa Uomo
Oncology Centre Antwerp (OCA)
Lange Gasthuisstraat 35-37, 2000 Antwerp, Belgium
tel.: +32 3 223 53 51 - fax : +32 3 223 53 52 - e-mail: louis.denis@skynet.be
www.cancerworld.org/europauomo

Introduction

Life is a mystery.

And does the same not apply equally to its counterpart – death?

There are countless variations on theories surrounding both life's beginnings and its conclusion. In the commemorative 2009 Darwin year, the theory of evolution (and with it, the finite nature of our existence) is perhaps enjoying more popularity than ever. Examples of other variants include the Creation narrative and the infinite nature of our existence, or the theory of an infinite consciousness, based in part on patients' near-death experiences¹. There exist Japanese tales of shinigamis, angels of death, who determine the moment of a person's passing by writing their name in a 'death note'. It is supposed that shinigamis nowadays do not write down names as often, due either to boredom or to avoid seeming over-industrious².

It is a fact that our life expectancy has increased spectacularly over the last 150 years. At around 1850, average life expectancy in the Netherlands was 39 years; for newborns today it is almost 80 years, 70 of which are, on average, spent in relatively good health³.

It is no simple matter to give an indication of which specific developments (such as those in health care) have contributed to these improvements in health over the past few decades. In 2007, in an Idols-like voting event, the British Medical Journal asked readers to vote on 15 significant 'medical' milestones (medical in the broadest sense of the word) from the past 150 years. These included the discovery of antibiotics, of oral rehydration therapy, the development of imaging technology, of anaesthesia or of evidence-based medicine. The development described briefly by our head of department, Professor Johan Mackenbach, came in first: the simple installation of sewers and water pipes some time after it became clear that unhygienic water management was what was causing the spread of germs⁴.

It is also a fine example of public health, described as 'the science and art of preventing disease, prolonging life and promoting health through the organised efforts of society'. It embodies a strong preference for preventive measures, is focused on national health (in fact, the cumulative health of all individuals) and is often concerned with programmes run collectively. As part of all this, we conduct both research into the determinants of public health as well as evaluations of interventions and policy measures.

There are beautiful examples in public health whereby identification of the determinant and subsequent action have been effective in promoting health: demonstrating the harmful effects of smoking and adjusting anti-smoking campaigns and pricing policy accordingly, demonstrating the harmful effects of asbestos, as well as establishing the link between sleep positioning and sudden infant death syndrome (SIDS). In the Netherlands, prohibition of asbestos has not produced a quick reduction in the number of asbestos victims, but in 2009 it is likely that we will have had 250 fewer deaths from mesothelioma⁵. Since infant health centres started issuing recommendations for children to sleep on their sides, the number of deaths from SIDS has dropped by more than 50%⁶.

Screening

Prevention is better than a cure, one might think. But prevention is often not possible, often because we have too little specific knowledge of the possible determinants that can play a part. Even once we have demonstrated significant determinants, actually influencing them can still prove to be a thorny problem⁷. If primary prevention is not possible, there is the possibility of secondary prevention or screening. It refers to the detection and treatment of a disease, condition, abnormality or stunted development at a stage in which no symptoms have yet appeared. The purpose of early detection is to halt or delay the adverse progression of the disease and thereby

improve health, or to offer treatment options in an earlier stage than usual (as is the case with certain prenatal screening tests). The principle is that the application of a relatively straightforward screening test will result in earlier diagnosis, and that earlier treatment will be more effective, enabling the person to live longer or with less serious consequences than in the scenario with no screening.

Although in reality prevention and therapy are the two most-important methods available to us for the improvement of people's health in the health care system, they are carried out in different ways. In the curative sector, the aim is to determine the best course of therapy for the patient at an individual level. Early detection and prevention at the public level does not guarantee effective prediction at an individual level, however we do expect positive outcomes on health of the population as a whole. To achieve these outcomes we often need to implement programmes for large groups of people, only some of whom will benefit. It is not possible to determine in advance who will benefit from early therapy.

Quantifying possible favourable effects

Stringent standards must therefore be applied when considering which screening test (or tests) could or should be made available to the population. The primary objective of evaluation of early detection of diseases, is to quantify any positive effects of screening on public health. A large emphasis is therefore placed on the execution and analysis of empirical studies on the effects of screening – research of a long-term nature.

The example of breast-cancer screening can be used as illustration. In 1913, the first X-ray image was created of surgically removed breast tissue; in 1927 a mammogram was used for the first time to diagnose a woman with unclear clinical findings, and in 1932 it first became standard for all women with breast tumours in some hospitals to have a breast X-ray prior to surgery. It was not until 1960 that predictions were made that it should

be possible to use this method to diagnose 95% of all malignant tumours prior to surgery, and that the introduction of routine mammography might hopefully lead to a reduction in deaths⁸. To demonstrate this, in 1963 the first randomised study took place among 63,000 insured women in New York⁹. Women were randomly divided into two groups. Half were offered early detection via clinical palpation of the breasts and mammography; those in the other group were not. After all, there was no proof that women benefited from this type of early detection. In 1971 the first positive results of the study were released: a 32% reduction in breast-cancer deaths among the group that had access to a combination of palpation and imaging technology¹⁰. Around 1976, six independent studies were launched in Europe with the aim of determining whether new imaging technology alone (without clinical examination) might also reduce breast-cancer deaths by means of early detection and treatment. Following the initial positive results published in *The Lancet*¹¹, we conducted a cost-effectiveness analysis that predicted positive results if women aged 50-69 were to be invited for mammography every 2 years (in the Netherlands)¹². Finally, 28 years after the initial prediction, the policy was made official.

Factors influencing the final outcome in screening

The evaluation of randomised controlled screening trials is crucial and no simple matter. In the year 2000, years after the publication of positive outcomes from the last mammography trial that was conducted, newspaper headlines were asking whether breast cancer screening is justified. Based on the same previously published data, a (Danish) Cochrane review concluded that it is not: they suggested to separate the Malmö and Canadian trial from the others. "If the Swedish trials (apart from the Malmö trial) are judged to be biased, there is no reliable evidence that screening decreases breast cancer mortality"¹³.

In the Netherlands Philip Freriks opened the 8 o'clock news with the announcement that de Koning was mop-

ping the floor with the Danes, which was actually not too far from the truth¹⁴. There was no reason to regard these 2 studies as superior, as the Danish did. The difference between the groups of subjects is ultimately determined by a large number of varying factors: simple design questions such as the screening interval, demography, attendance, opportunistic screening and having the control group undergo mammograms on their own initiative all influence the ultimate extent of the effects. Through modelling, we showed that these relatively simple differences alone could make one trial exhibit a 25% greater effect than the other¹⁵.

However, it is much more important to consider how much earlier the diagnosis is made and the effect this has. The important question therefore relates to the quality of the screening, how many women are referred for further examination and how many tumours are detected earlier as a result. The baseline conditions prior to the study, or in this case those of the control group, are also significant. If women in one region on average receive health care at an earlier stage, this can mean that the difference between 'early' and 'late' is smaller in one region than in the other, although the quality of screening and therapy may be the same (figure 1).

Meta-analysis breast screening

Using computer models, we can simulate this quality and the initial conditions to predict the effects. We should keep in mind that therapy trials are often much more homogenous, and that these other facets are less relevant. One important aspect of the Danish interpretation was the almost complete lack of effect in the Canadian trial, described by the Danish as the 'Mecca' of proof and of very high quality. However, this Canadian study had no control group of women who were not offered screening: in the control group, the women's breasts were closely examined by nurses. Death rates in the control group may therefore also have decreased as a result of earlier detection and treatment in the "control" arm, in effect reducing the difference between the 2 groups of women.

Using our computer model called 'MISCAN', we have been able to use and reproduce the individual data from this Canadian trial and demonstrate that, although the quality of the mammograms in this trial was somewhat lower than in the other Swedish trials, at the same time the intensive screening by nurses in the so-called control group must have had an effect (given that tumours were being diagnosed considerably earlier)¹⁶.

Figure 2 shows the pooled estimate and (95% confidence intervals) of the 6 modern mammography trials for women aged 50 (or 55) and over at entry, including a model-based estimate for the Canadian trial: if the trial had been set up without any screening in the control group, we would then have expected a minimum 11% breast cancer mortality reduction, and the trial would therefore be consistent with all of the other breast cancer screening trials. Combined a 25% decrease in breast cancer mortality for women invited for mammography screening compared to not invited for screening, including the Canadian trial, would be the correct pooled estimate for women aged 50 and over.

We in the Netherlands were one of the first to show in *The Lancet* that breast-cancer screening (followed by earlier therapy) had succeeded in turning a 0.3 rise in breast-cancer deaths per year (per 100,000 woman-years) into a 1.7 drop at municipality level. The turning point was the commencement of screening in the municipality and had nothing to do with changes in therapy, which of course also took place¹⁷.

Meta-analysis prostate screening

Another very prominent example is the evaluation of prostate-cancer screening. Each year, 2,500 men in the Netherlands die as a result of this disease. The 1980s saw the discovery of the PSA test (Prostate-Specific Antigen), a unique marker that is present almost exclusively in the prostate, and which shows increased levels in the event of a benign or malignant enlargement of the prostate long before the appearance of any symptoms. Two large-scale randomised trials were initiated in Europe and

America to determine whether early detection using the PSA test would reduce prostate cancer deaths. It is important to realise that in order to do this, a number of extremely important decisions must be made at the beginning of the long evaluation cycle. These include matters such as the preferred target group, the best way to recruit subjects, how many participants are necessary, and which screening tests and screening interval to use. It is essential to reach clear agreement in advance regarding the treatment and diagnostics of (all) referred persons in the region, and to ensure that these are independent of the way in which people are admitted to the hospital, whether via screening or not.

It is crucial that the final measurement, in this case deaths due to prostate cancer, be established both clearly and without distortion. In the screening group, more men were diagnosed with prostate cancer as a result of the test. This increases the likelihood of prostate cancer appearing on the death certificate (a 'sticky diagnosis'), although this may not have been the cause of death. In the European prostate trial, committees of medical experts determined the cause of death of patients diagnosed with prostate cancer using complete medical histories, without knowing which group the patient belonged to, screened or not^{18,19}.

In addition to this type of epidemiological/public health criteria, the correct 'test performance' is also important.

Some months ago, after 9 years of follow-up, the results of the European prostate trial were released, which was conducted among 162,000 men aged 55-69. In the group that had been invited to take PSA tests, the likelihood of death due to prostate cancer had dropped by 20% – a crucial result, after 20 years of hard work²⁰.

Currently 3 randomised prostate PSA screening trials have been published, so the meta-analysis in figure 3 could be published, stating that together, the studies indicate a 10% reduction in prostate cancer deaths in men invited for PSA screening.

You will have realised by now that a meta-analysis of this nature is of little value unless there have been careful reviews of e.g.; the risk of the subjects in the various trials, the scope and quality of screening, diagnostics and therapy in the screening group, opportunistic screening in the control group and the baseline conditions at the start of the study. This requires a model-based (MISCAN) approach. Still, the New York Times summarised the European trial as follows: 'The most important study for men's health in history'²¹.

Quantifying unfavourable effects

A 20-30% reduction in deaths due to prostate cancer in the Netherlands would mean several hundred prevented prostate cancer deaths per year. Although this information is crucial, it is nevertheless insufficient to make the offer of prostate screening official.

It is also essential to quantify the disadvantages, which is the second important objective in the evaluation of screening. It involves establishing the scope and duration of the negative consequences using 'quality of life' research, as well as an estimation of the number of people who must undergo such consequences. In reality, it is a matter of creating an independent quantitative balance between positive and negative effects. This can therefore mean that, despite the demonstration of clear positive effects, the scope of the disadvantages for the population does not justify.

Each type of population-wide screening also imposes disadvantages for large groups of people, summarized in table 1. There are false positives – people who are referred because of an adverse screening result, but who turn out not to have any illness. The diagnosis is made earlier, which seems favourable, but this also means that people know that they are ill over a longer period and will possibly suffer side effects from treatment for a longer period. Some people in the group do not benefit at all from early detection, and there will also be people who die from other diseases and who would otherwise never

have been diagnosed. As a result, these people even are overtreated.

Risks can be associated with the screening and follow-up diagnosis, and the screening can also lead to the unintended detection of other, possibly untreatable conditions. Lastly, false peace of mind can be created, and people who receive a favourable scan result may be put off seeing a doctor because of symptoms as they believe they can draw reassurance from the screening results. In fact, in the field of early detection, the effects of detecting the disease earlier are likewise also the worst effects.

One of the most important reasons for setting an upper age limit when offering screening has to do with this balance of positive and adverse effects²².

Or as it was formulated in Dutch court:

‘By setting an age limit of 76 years for population breast-cancer screening, the State is making a distinction according to age. The key to this short case can be summarised by the question of whether the balance of benefits and risks tips differently for the specific 76+ age group than for those aged between 50 and 76. This balance is drawn at a *collective* level. After all, population studies pertain to categories of people by their very nature. The reasons why the State applies the aforementioned upper age limit cannot be viewed as insufficiently substantiated. As the age of the persons concerned increases, serious consideration must be taken of the fact that early detection and ensuing treatment, which must always entail adverse effects, improves quality of life and life expectancy by an increasingly smaller degree.

The distinction here maintained is not discriminatory and therefore not unjust.”^{23,24}

In screening, overdiagnosis and overtreatment are specific unfavourable effects.

In an elegant study design, we have demonstrated that the mere diagnosis of cancer causes a temporary 20% drop in the day-to-day quality of a person’s life²⁵. Most studies do not begin to measure the effect until long

after people have been diagnosed, have had the chance to adjust, and in fact are hopeful that early detection has not been for nothing. This is not the case for everybody: some people would otherwise never have had the diagnosis nor the side effects of diagnosis and treatment, and others are completely unable to benefit from early treatment.

So, what can come as a thunder lightning/bombshell (from out of the blue) for some people, can be viewed completely differently by others²⁶. One way to manage these individual differences lies in the concept of an ‘informed decision/choice’ to participate. In reality it is up to the individual to consider whether the imagined benefits outweigh the inevitable disadvantages. But be warned: we estimate these at population level, and do not know how they will manifest at individual level. Individuals do not know this either. It is essential that the individual possess or be given enough information about the condition that is being tested for and its repercussions, about the screening and its effects, and about the possible advantages and the most significant disadvantages. If this information is sufficient, participation (or non-participation) should be consistent with the test subject’s general attitude towards screening in the particular area. There is currently much debate on whether current information provision for large-scale programmes is sufficient in this regard. We have now conducted studies which actually show that the level of informed decision-making in the Netherlands is quite good for breast cancer screening and prenatal screening.

The optimal screening scenario

The 3rd objective of evaluation studies in this field is making an accurate cost/benefit assessment of a range of scenarios, and recommending one of the more cost-effective ones. Model estimates are required here, since the variation in scenarios is many times greater than the variation in the research conducted²⁷⁻²⁹. One very important decision is that of which screening interval to

choose. The problem at the beginning of a trial is that there is often little information available to help with choosing the right interval. After all, the trial is intended to determine the effects of screening and the qualities of the test. It is therefore important to estimate beforehand (and later to quantify) how much earlier the test is capable of detecting the disease. However every disease is unique, and the test (or the interpretation of the test) will also vary among those who perform the screening. Computer models can be used to simulate this natural variation as well as the sensitivity of the test, based on data observed with and without screening.

Figure 4a shows the theoretical framework for a chronic disorder like cancer³⁰. Screening people who until that time were asymptomatic will initially lead to a significant rise in the number of people diagnosed with the disease early. Depending on the sensitivity of the test, the period immediately following will contain few clinical diagnoses of the disease, after which (in this case, depending on the rate of growth of the type of tumour) cases will start appearing and being clinically diagnosed again at varying rates. The magnitude of the variations and the pattern will determine which screening interval will likely be most effective. If the interval is too long, it could mean that although a number of tumours will be detected by the initial screening, during the interval the clinical baseline conditions will quickly reassert themselves. In this case a new round of tests may again be effective, but only mildly so. After all, the initial round of tests detects tumours that have already been growing for so long that they are detected at a relatively late stage and many were already close to being clinically diagnosed. A screening interval that is too short means lots of tests, high costs, the many disadvantages of screening and the possibility of few additional positive health outcomes. Figures 4b and 4c show the calculated detection rates of cancer using one-off prostate screening and one-off colorectal cancer screening via colonoscopy in a cohort of men aged 65. As you can see, after the PSA screening it

takes a total of 25 years to return to the baseline level, or the rate of incidence before we started screening (X-axis time). Another screening test very soon afterwards will therefore not be very useful. For colorectal cancer screening, the x-axis shows the person's age: after a screening at the age of 65, the likelihood of detection remains significantly lower for more than 5 years afterwards. We also determined beforehand the 4-year interval for the European prostate trial in this manner. The figure 4b shows the expected asterix, based on the sparse data that was available to us around 1992³¹.

The new criteria for screening

Wilson and Jungner were the first to attempt to try to make sense of this ocean of possibilities and obscurities in 1968 because there was 'growing importance, as evidenced by the controversies over matters such as cytological testing for cancer of the uterine cervix, and regular medical check-ups of key executive personnel'³⁹. The authors wrote and called these 10 principles, but nowadays they sometimes seem to have grown into dogmas, despite the actual spirit of the authors in the foreword. Several times I have heard people say "if we just call it something other than screening, then at least we won't need to satisfy the Wilson and Jungner criteria".

The newest criteria for early detection, however, are unambiguous (Table 2).

There must be substantial positive health outcomes, in terms of additional years of life, improvements to cognitive, motor and/or socio-emotional development or a significant increase in management or treatment options. These effects must have been established with certainty, if possible preferably by means of a randomised trial. Secondly, any adverse side-effects must be limited in nature. The extent of early detection, the level of over-diagnosis and the side-effects need to have been estimated and expressed in terms of quality-adjusted life-years. The anticipated balance between positive health outcomes

and side-effects at group level must have been clarified prior to possible participation. There must be a reasonable balance between costs and benefits (this allows the creation of programmes that provide increased benefits for the same investment), and measures must be taken to ensure that implementation will not lead to substantial unintended effects (such as widespread inequality). Lastly, nearby other developments would not influence the weighing of this balance substantially.

Evaluation cycle screening for language delays, and implementation

I am convinced that if parents were presented with an accurate overview of the pros and cons of screening in youth healthcare, that youth healthcare would start to look very differently.

The consequences can be just as great when discussing the development of children. Nonetheless, all of the above-mentioned considerations apply here to the same extent. The point is not whether a questionnaire or test is capable of demonstrating a deviation (sensitivity), it is about whether children (and/or parents) benefit afterwards from early treatment, and whether this outweighs the inevitable disadvantages. So do not add another 'how am I going to feel good in the future?'-questionnaire at the Youth and Family Centre, but instruct people not to have the children examined for scoliosis – despite commercial interest suggesting otherwise – ³².

In a large-scale randomised trial among 11,440 children, we demonstrated that the inclusion of a specific language screening test at age 2 led to 40% more children at age 3 being treated earlier for language problems than in a comparable control group who did not receive the specific screening³³. This early detection and treatment ultimately resulted in 30% fewer children ending up in special education at age 8³⁴⁻³⁵.

We went further, by putting together a uniform package of a range of screening instruments that ought to be ap-

plied at various ages (figure 5). A clear cut-off for scores at age 2 determines the follow-up policy. If so, only 5% of children at age 2 have enough problems with language that they possibly benefit from early intervention, and need to be identified and referred directly to an audio-logical centre at a young age. However, a middle group of around 20% need to be seen again and screened between their 2nd and 3rd birthday using additional instruments and visits for a definitive assessment³⁶.

Over the last 3 years children have not yet been able to benefit from this new form of early detection. We can therefore safely say that in the Netherlands 7,500 children will needlessly end up in special education. Why is that?

The opinion that 'there is insufficient clarity regarding the implementation of a formal screening, which determines whether referral is required based exclusively on a test result³⁷' does injustice to the usefulness of every kind of detection. There is totally no justification to pose developmental questions to parents and children, without an effective means of quantifying or scoring and without a clear cut-off point for referral. "If screening is worth doing, it is ONLY worth doing if done well"³⁸.

"But, what about my own individual experience?"

Screening remains a human job: only a well-educated professional can tell the difference between what looks normal and what seems wrong. The technology is an important tool to them, but "one does not rob the worker of his right to pride in workmanship"³⁸. There needs to be an understanding that screening is not an isolated act performed by, say, a doctor, but that there is much to be gained from uniformity. Uniformity is quality, and clear guidelines for referral ensures quality. But the right information must be available to inform the public and professional groups.

To transform from evidence-based medicine towards actual implementation is a challenge, or as the famous Canadian ice hockey player Gretzky would put it: "Skate to where the puck will be".

It takes guts to improve, change or abolish existing screening methods. There needs to be a willingness to set interests aside and oversee an independent analysis of the beneficial and adverse effects, based on properly executed evaluation research. Verloove-Vanhorick summarised a successful transformation/change (read implementation of a new public health measure) can only be achieved if: there are clearly shared goals, there is a clear strategy, there is leadership and multidimensional thinking, authorities can be set aside, there is attention for detail and the individual professional, and finally there are financial or logistic constraints to be met, it is actually transferable and one really cooperates with each other⁴⁰.

Future developments

The impressive technological developments and opportunities emerging from fundamental research mean that the demand for early detection of diseases is growing spectacularly. Many conditions, abnormalities and cases of stunted development are recognisable at an early stage, or will be in the near future. In addition to producing images of calcification of the most important arteries, osteoporosis or the early stages of dementia, it may become possible to detect the levels of significant biomarkers as early as in the umbilical cord blood, markers that could increase the likelihood of obesity.

I believe that the testing framework for these developments should be the same as that described above, in order to achieve both optimum positive health outcomes for an evenly-distributed group and to prevent or counteract any negative health outcomes.

However, it may look as that the testing framework and technological developments cannot keep up with general developments in society, particularly the advent of individualisation. Individuals are becoming more demanding, can request self-tests online, and are waited on hand and foot by commercial companies or well-meaning care workers. Couples start treating each other to tests for

their birthday. The American Cancer Society has declared themselves official sponsor of birthdays. Does this mean that improved risk-group assessment is the way of the future? With all of the online risk calculators, can individuals not determine for themselves whether they ought to be eligible?

What you must realise, however, is that although it is important, the risk of disease is only one facet of early detection. Being at particular risk of a condition, abnormality or reduced development is no stringent condition or minimum requirement for being tested for it, just as a low risk could be a good reason to do so. Although better identification of the target group may help, the general considerations remain valid for the large group as a whole.

Population screening is nothing more than establishing the correct risk group for whom – as a group – the advantages considerably outweigh the disadvantages. There is often no more information available to individuals that would allow them to make a better or different quantitative assessment for themselves. At the same time, many groups that lack the funds, the Internet or the ability to weigh things up will be left out of such considerations. In addition to these developments, professional and economic developments will present themselves. GPs, first-aid officers, company doctors and paramedics will want to concentrate more and more on the possibilities offered by screening, if they have the money and resources available. In the future, insurers may also wish to distinguish themselves on the basis of screening and new target groups will see the scene. All of this constitutes an even better reason for a quantitative, independent consideration of the pros and cons of screening.

Simple technology or low-tech differently?

The same testing framework as described earlier also applies to simple technology. Because in actual fact there is no simple technology, and is the term 'simple technology' a redundancy, an understatement. The 20-week preg-

nancy ultrasound, once referred to using a euphemism meaning the 'ultrasound for fun', has undergone such development that on the images now taken many early abnormalities can be seen. Early abnormalities whose early detection is not necessarily always useful, some of which are transitory, and which can be particularly subject to false positives, including the entire range of possible negative effects.

Supposedly for years it has been possible to use a simple blood test to establish whether a person has high glucose levels. Such people run a very high risk of developing diabetes, and reducing glucose levels earlier through weight loss, increased exercise and/or medication may generate major positive health outcomes. So why is there no national screening programme on diabetes and metabolic syndrome?

Last year, a study demonstrated that taking the blood-sugar levels regularly of diabetes patients had no positive effect whatsoever on their health, but that it did negatively affect their quality of life and levels of depression⁴¹. By this I want to say that simple technology does not exist, at least in terms of 'it can't do any harm'. Even for simple technology, there must be clear proof that a large-scale approach creates more advantages than disadvantages. For diabetes screening this proof simply does not exist.

In a large-scale study conducted in the Rijnmond area, as part of which we sent tape measures to 79,000 people, 6% of those in the high-risk group turned out to have increased levels of glucose and another 13% had an adverse cardiovascular profile without them or their doctors being aware of it⁴². Therefore, there is an interesting low-tech instrument that can be used for the early detection of abnormalities that are genuinely relevant. Whether people are then able to benefit from early treatment can only be established through an RCT.

Another interesting development will be the combination of high-tech and low-tech solutions, e.g. using a tape measure for most cases and coronary calcium levels for

some cases. Cardiovascular mortality and morbidity are still top priorities.

In conclusion

The (new) criteria for early detection (and evaluation thereof) have been clearly developed and are universally applicable. Clear consideration regarding whether to offer a particular form of early detection or not, can and should be made both for reasons of increasing positive health outcomes as well as for the prevention of negative health outcomes. This requires national coordination and partnerships traversing a range of professions and fields. There are no other criteria for new or existing programmes. There are no other criteria if we call it 'risk assessment' or 'signalling'.

Early detection demands a high-quality approach with a uniform referral policy that does not compromise respect for the autonomy of individuals or professionals. The main priority is to provide screening to the correct high-risk group with the correct screening and possible interval; the system that does the providing is secondary. There is no such thing as individual prevention; more individual, customised prevention perhaps, however this demands large amounts of research. The need of educating the public, scholars, patients and professionals about the perils and promise of early detection is set to increase.

In this mystery called life, I hope that in evaluating screening our research team and external partners can make a modest but important contribution to the improvement of people's health.

Figure 1

Trajectory of a screening outcome



Baseline conditions prior to the study

Age and gender distribution in population
Life table without disease
Incidence of disease
Growth rate of the disease (natural history)
Stage distribution at diagnosis
Treatment dissemination and effect
Survival with the disease



Design trial or study

Age at entry
Screen test
Attendance rate
Proportion non-attendees
Screening interval
Screening in control group (attendance rate and interval)
Screening before and after the trial/study
Duration of trial/study and follow-up after the trial/study



Quality of screening, assessment and treatment

Sensitivity of screen test
Treatment dissemination and effect
Improvement in survival

Reduction unfavourable outcome

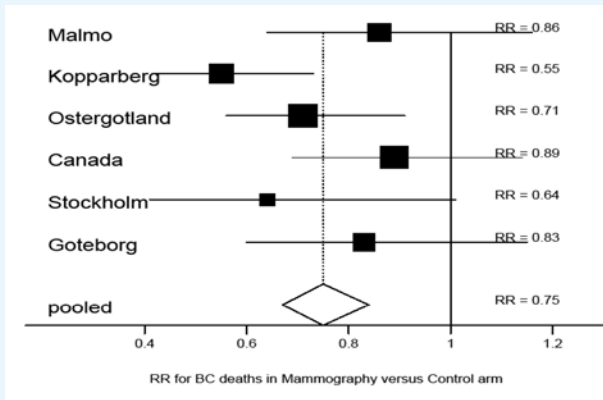


Figure 2

Point estimate, 95% confidence intervals and pooled estimate of the 6 modern randomised controlled breast screening trials for women aged 50 (55) and over, including an adjusted estimate for the Canadian trial assuming no protocolized screening in the control arm (as in the other trials)

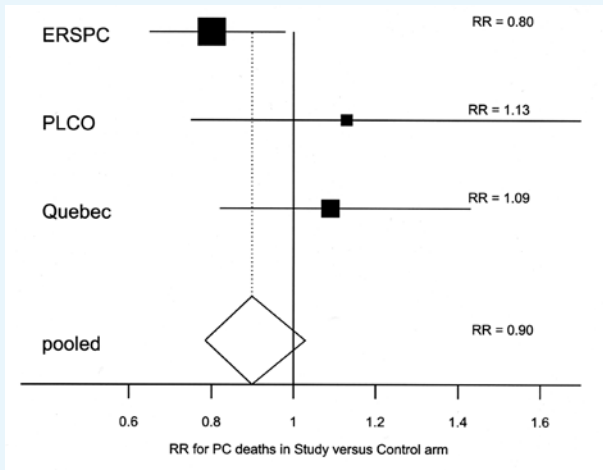


Figure 3

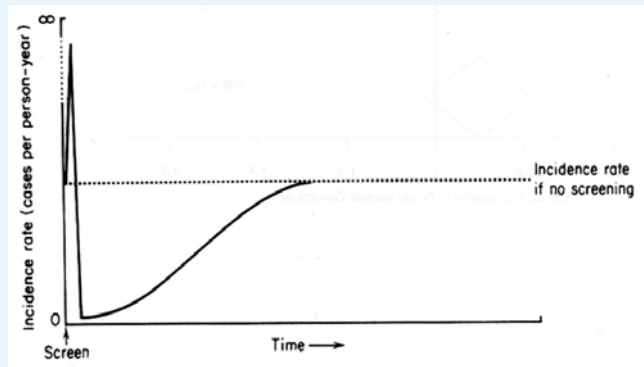
Point estimate, 95% confidence intervals and pooled estimate of the 3 randomised controlled prostate cancer PSA screening trials

Figures 4a, 4b, 4c

Incidence rates of disease in situation with and without screening

- a) Theoretical (by Morrison 1992)
- b) One PSA-test applied to a cohort of men 65-years of age (X-axis time)
- c) One colonoscopy applied to a cohort of men 65-years of age (X-axis age)

Figure 4a



Incidence prostate cancer after PSA test at age 65 years

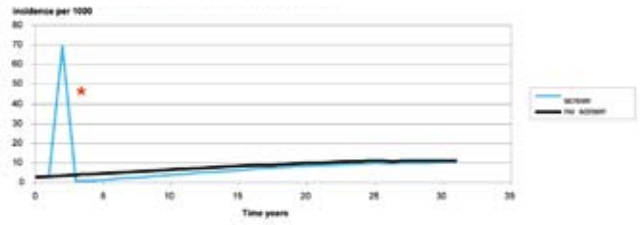


Figure 4b

Incidence colon cancer after colonoscopy at age 65 years

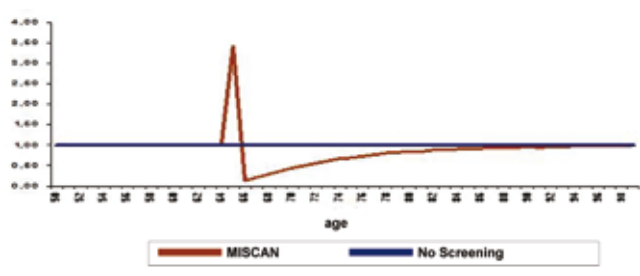


Figure 4c

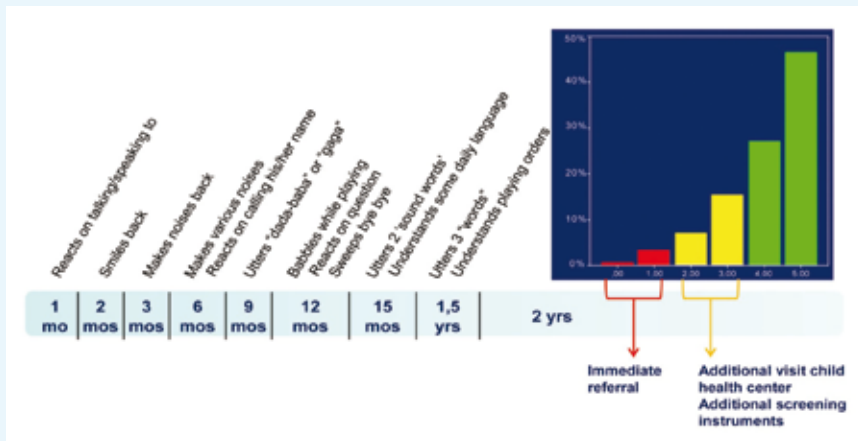


Figure 5
Proposed screening schedule for language delays 0-36 months

Table 1

Favourable and unfavourable effects of screening in general

Favourable effects

- Reduction in unfavourable outcome of disease (e.g. cancer deaths, developmental disturbance)
- Less treatment for advanced stages
- Less primary treatment
- Less intensive or mutilating treatment
- More efficient diagnostic work up (less clinically suspicious cases)

Unfavourable effects

- False positives
- Earlier (knowledge of) diagnosis
- Earlier (and during a longer time frame) side effects of treatment
- Early detection, but no health benefit
- Extra detection (overdiagnosis) and overtreatment
- Risks of screening and assessment, and unintended detection of other diseases
- Possible false reassurance (confrontation with the diagnosis, or detecting a disease, later as usual)
- Possible license to unhealthy behaviour

Table 2

The new criteria for implementation of screening

Substantial positive health outcomes

- life-years gained
- improvements to cognitive, motor and/or social-emotional development
- significant increase in management or treatment options

Effects established with certainty, preferably in RCTs

Limited adverse side-effects

- extent of early detection, overdiagnosis and side-effects estimated
- quality-adjusted life-years gained

Anticipated balance clarified prior to participation

Reasonable ratio between costs and benefits

Implementation will not lead to substantial unintended effects

Other developments do not change this ratio in the short run



My mysterious mass(es).



Dedicated to my dad who passed away February 10, 2010.

References

1. Lommel P van, van Wees R, Meyers V, Elfferich I. Near-death experience in survivors of cardiac arrest: a prospective study in the Netherlands. *Lancet* 2001 ;**358**:2039-45. Erratum in: *Lancet* 2002;**359**:1254.
2. Tsugumi O and Takeshi O. *Death Note* (manga). Part I. Publisher Kana, Antwerpen, 2003. ISBN 978-90-86770-37-3.
3. JP Mackenbach en PJ van der Maas (eds.). *Volksgezondheid en gezondheidszorg*. Elsevier Gezondheidszorg, Maarssen. ISBN 978 90 352 2952 5.
4. Mackenbach JP. Medical Milestones. Sanitation: pragmatism works. *Br Med J* 2007;**334**:s17.
5. Driece HA, Siesling S, Swuste PH, Burdorf A. Assessment of cancer risks due to environmental exposure to asbestos. *J Expo Sci Environ Epidemiol* 2010 in press (Epub ahead of print. PMID: 19865072).
6. Pickett KE, Luo Y, Lauderdale DS. Widening social inequalities in risk for Sudden Infant Death Syndrome. *Am J Public Health* 2005;**95**:1976–81.
7. Veral V, Bull D, Doll R, et al. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. *Lancet* 2002;**360**:187-95.
8. Ingleby H and Gershon-Cohen J (eds.). *Comparative anatomy, pathology and roentgenology of the breast*. Philadelphia: University of Pennsylvania, 1960.
9. Shapiro S and Venet L. Evaluation of periodic breast cancer screening with mammography. *J Amer Med Ass* 1966;**195**: I 11 - I 18.
10. Shapiro S, Strax P, Venet L. Periodic breast cancer screening in reducing mortality from breast cancer. *J Am Med Ass* 1971;**215**:1777-85.
11. Tabár L, Fagerberg CJ, Gad A, et al. Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. *Lancet* 1985;**1**:829-32.
12. Maas PJ van der, de Koning HJ, van Ineveld BM, et al. The cost-effectiveness of breast cancer screening. *Int J Cancer* 1989;**43**:1055-60.
13. Götzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;**355**(9198):129-34.
14. Koning HJ de. Assessment of nationwide cancer-screening programmes. *Lancet* 2000;**355**:80-1.
15. Koning HJ de, Boer R, Warmerdam PG, Beemsterboer PM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. *J Natl Cancer Inst* 1995;**87**:1217-23.
16. Rijnsburger AJ, van Oortmarssen GJ, Boer R, Draisma G, To T, Miller AB, de Koning HJ. Mammography benefit in the Canadian National Breast Screening Study-2: a model evaluation. *Int J Cancer* 2004;**110**:756-62.

17. Otto SJ, Fracheboud J, Looman CW, Broeders MJ, Boer R, Hendriks JH, Verbeek AL, de Koning HJ; National Evaluation Team for Breast Cancer Screening. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. *Lancet* 2003;**361**:1411-17.
18. Koning HJ de, Blom J, Merkelbach JW, *et al.* Determining the cause of death in randomized screening trial(s) for prostate cancer. *BJU Int* 2003;**92**(Suppl. 2):71-8.
19. Otto SJ, van Leeuwen PJ, Hoekstra JW, Merkelbach JW, Blom JHM, Schröder FH, Bangma CH, de Koning HJ. Effects of causes of death verification in the outcome of a trial: a major influence? *Eur J Cancer* 2010 (special issue), in press.
20. Schröder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis LJ, Recker F, Berenguer A, Määttänen L, Bangma CH, Aus G, Villers A, Rebillard X, van der Kwast T, Blijenberg BG, Moss SM, de Koning HJ, Auvinen A; ERSPC Investigators. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med* 2009;**360**:1320-28.
21. New York Times (Kolata G). *Prostate test found to save few lives.* New York Times, March 18, 2009.
22. Fracheboud J, Groenewoud JH, Boer R, Draisma G, de Bruijn AE, Verbeek AL, de Koning HJ. Seventy-five years is an appropriate upper age limit for population-based mammography screening. *Int J Cancer* 2006;**118**:2020-25.
23. Civil Court Den Haag, October 30, 2007, Hofhuis HFM (judge). Summary Procedure, Case nr KG 07/1099, www.rechtspraak.nl, LJN BB6695. *Judgement concerning the National breast cancer screening programme – age limit 75 years.*
24. Civil Court Den Haag, February 9, 2010. Judgment Higher Court, Case nr KG 07/1099 LJN BB6695. *Judgement concerning the National breast cancer screening programme – age limit 75 years. Final.*
25. Korfage IJ, de Koning HJ, Roobol M, Schröder FH, Essink-Bot M-L. Prostate cancer diagnosis: the impact on patients' mental health. *Eur J Cancer* 2006;**42**:165-70.
26. White S. *North Country. A graphic novel.* NBM, New York, 2005. ISBN 10:1-56163-435-2 and ISBN-13:978-1-56163-435-4.
27. Akker-van Marle ME van den, van Ballegooijen M, van Oortmarssen GJ, Boer R, Habbema, JDF. Cost-effectiveness of cervical cancer screening: comparison of screening policies. *J Natl Cancer Inst* 2002;**94**:193-204.
28. Okonkwo Lamberts Q, Draisma G, der Kinderen A, Brown M, de Koning HJ. Breast cancer screening policies in developing countries: a cost-effectiveness analysis for India. *J Natl Cancer Inst* 2008;**100**:1290-00.
29. Kok IMCM de, Ballegooijen M van, Habbema JDF. Cost-Effectiveness Analysis of Human Papillomavirus Vaccination in the Netherlands. *J Natl Cancer Inst* 2009;**101**:1083-92.
30. Morrison AS. *Screening in chronic disease.* 2nd edition. New York: University Press, 1992.

31. Gustafsson O, Norming U, Almgard L-E, *et al.* Diagnostic methods in the detection of prostate cancer: a study of a randomly selected population of 2,400 men. *J Urol* 1992;**148**:1827-31.
32. Bunge EM, Juttman RE, van Biezen FC, Creemers H, Hazebroek-Kampschreur AA, Luttmer BC, Wiegersma PA, de Koning HJ; Netherlands Evaluation Study on Screening for Scoliosis (NESCI) Group. Estimating the effectiveness of screening for scoliosis: a case control study. *Pediatrics* 2008;**121**:9-14.
33. Koning HJ de, de Ridder-Sluiters JG, van Agt HME, *et al.* A cluster-randomised trial of screening for language disorders in toddlers. *J Med Screening* 2004;**11**:109-16.
34. Ploeg CPB van der, Lanting CI, Verkerk PH. Should we screen for language delay in toddlers? [Letter to the editor]. *Pediatrics* 2008;**121**:1299-00.
35. Agt HME van, van der Stege HA, de Ridder-Sluiters H, Verhoeven LTW, de Koning HJ. A cluster-randomized trial of screening for language delay in toddlers: effects on school performance and language development at age 8. *Pediatrics* 2007;**120**:1317-25.
36. Agt H van, Blumenthal M, van den Brink G, van Denderen-Lubbers M, Dieleman L, Goorhuis-Brouwer S, Strik M, de Koning HJ. Een uniforme signalering van een vermoeden van taalachterstand en een uniform follow-up beleid. Draagvlak voor implementatie in de jeugdgezondheidszorg in Nederland. Erasmus MC, afd. Maatschappelijke Gezondheidszorg, 2007.
37. Ploeg CPB van der, Lanting CI, Galindo Garre F, Verkerk PH. *Screening op taalachterstanden en spraakstoornissen bij kinderen van 1 tot 6 jaar door de jeugdgezondheidszorg. Deelrapport 1: inventarisatie van instrumenten.* TNO Kwaliteit van Leven, TNO-rapport KvL/P&Z 2007.025, Leiden, 2007.
38. Raffle A and Gray M. *Screening. Evidence and practice.* Oxford University Press, Oxford, 2007.
39. Wilson JMG and Jungner G. *Principles and practice of screening for disease.* Public Health Papers nr 34. Geneva: WHO, 1968.
40. Verloove-Vanhorick SP. 'Het kind van nu, het kind van de toekomst'. Farewell speech on February 1 2008, LUMC/TNO, Leiden.
41. O'Kane MJ, Bunting B, Copeland M, Coates VE (ESMON study group). Efficacy of self monitoring of blood glucose in patients with newly diagnosed type 2 diabetes (ESMON study): randomised controlled trial. *Br Med J* 2008;**336**:1174-77.
42. Klijs B, Otto SJ, Willems JL, *et al.* Feasibility of a population-based randomized controlled trial of screening for type 2 diabetes and high CVD-risk using home measured waist circumference. 2010 (submitted).

Harry de Koning received his MD degree at State University Leiden in 1986 and a PhD degree in Medicine at Erasmus University Rotterdam in 1993 for his studies on the effects and costs of breast cancer screening. He had epidemiological training at the New England Epidemiology Institute and public health training at the department of Public Health (Erasmus MC) for which he is leading research on evaluation of screening since 1995. The research aims at quantifying the health benefits, unfavourable side-effects, impact on quality of life, and the cost consequences of introducing screening. This may lead to advice to introduce, or not to introduce, screening for a specific disease, or to introduce it in a specific way, e.g., for selective groups of the population only. His team has a.o. been partially responsible for the implementation and evaluation of the breast screen programme in the Netherlands with a consequent 29% reduction in BC mortality, for evidence on the effect of screening on language delays in toddlers and for the recognition of overdiagnosis in prostate cancer screening. He was appointed international member of the Breast Screen Australia Evaluation Committee, Deputy-chair of the Population and Behavioural Sciences Committee of Cancer Research UK and member of the “Signaleringscommissie KWF/NKB”.

Key publications

Wever E, Draisma G, Heijnsdijk EAM, Roobol MJ, Boer R, Otto SJ, **de Koning HJ**. Prostate-Specific Antigen screening in the United States vs in the European Randomized Study of screening for Prostate Cancer-Rotterdam. *J Natl Cancer Inst* 2010;102(5):352-5.

Draisma G, Etzioni R, Tsodikov A, Mariotto A, Wever E, Giulati R, Feuer E, **de Koning HJ**. Lead time and overdiagnosis in Prostate-Specific Antigen screening: importance of methods and context. *J Natl Cancer Inst* 2009;101:374-83.

Klaveren RJ van, Oudkerk M, Prokop M, Scholten ETH, Nackaerts K, Vernhout R, Iersel CA van, Bergh KAM van de, Westeinde SC van 't, van der Aalst C, Thunissen

sen E, Xu DM, Wang Y, Zhao Y, Gietema HA, de Hoop B-J, Groen HJM, Bock GH, van Ooijen P, Weenink C, Verschakelen J, Lammers JWJ, Timens W, Willebrand D, Vink A, Mali WPTM, **de Koning HJ**. Management of lung nodules detected by Volume CT Scanning. *N Eng J Med* 2009;361(123):2221-9.

Schröder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis LJ, Recker F, Berenguer A, Maattanen L, Bangma CH, Aus G, Villers A, Rebillard X, van der Kwast T, Blijenberg BG, Moss SM, **de Koning HJ**, Auvinen A. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med* 2009 Mar 26;360(13):1320-8.

Okonkwo LQ, Draisma G, Kinderen A der, Brown M, **de Koning HJ**. Breast cancer screening policies in developing countries: a cost-effectiveness analysis for India. *J Natl Cancer Inst* 2008;100(18):1290-300.

Agt HME van, van der Stege HA, de Ridder-Sluis H, Verhoeven LTW, **de Koning HJ**. A cluster-randomized trial of screening for language delay in toddlers: effects on school performance and language development at age 8. *Pediatrics* 2007;120:1317-25.

Kriege M, Brekelmans CTM, Boetes C, Besnard PE, Zonderland HM, Obdeijn IM, Manoliu RA, Kok Th, Peterse H, Tilanus-Linthorst MMA, Muller SH, Meijer S, Oosterwijk JC, Beex LVAM, Tollenaar RAEM, **de Koning HJ**, Rutgers EJT, Klijn JGM (for the M-RISC Study Group). Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition. *New Engl J Med* 2004;351(5):37-47.

Otto SJ, Fracheboud J, Looman CWN, Broeders MJM, Boer R, Hendriks JHCL, Verbeek ALM, **de Koning HJ** (NETB). Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systemic review. *Lancet* 2003;361:1411-7.

de Koning HJ, Boer R, Warmerdam PG, Beemsterboer PM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. *J Natl Cancer Inst*. 1995 Aug 16;87(16):1217-23.

