

## STAND VAN ZAKEN

# Onderscheidend vermogen, reclassificatie en netto-nut

## METHODOLOGIE VAN ONDERZOEK 7: DE VOORSPELLENDE WAARDE VAN BIOMARKERS

Ewout W. Steyerberg, Michael J Pencina en Ben Van Calster

- Er komen steeds meer biomarkers beschikbaar om de kans op ziekte te voorspellen. Om de voorspellende waarde van een biomarker uit te drukken worden allerlei maten gebruikt.
- In dit artikel gebruiken we als voorbeeld het Framingham-risicomodell met en zonder de hdl-cholesterolwaarde als risicofactor om het optreden van hart- en vaatziekten (HVZ) te voorspellen.
- We bespreken 3 maten voor diagnostische of prognostische kwaliteit: (a) de oppervlakte onder de 'receiver operating characteristic'(ROC)-curve (AUC), een traditionele maat voor onderscheidend vermogen; (b) 'net reclassification improvement' (NRI), een populaire maat die is gebaseerd op de verbetering van de indeling in hoog- en laag-risicopatiënten door de hdl-waarde als risicofactor toe te voegen; (c) 'net benefit' (netto-nut), een maat met een besliskundige basis.
- Het netto-nut is een betere maat voor de waarde van een biomarker, omdat deze meer rekening houdt met de klinische context en makkelijker te interpreteren is dan de AUC en NRI.

**In de rubriek Stand van zaken verschijnen de komende periode regelmatig bijdragen over methoden die gebruikt worden bij het opzetten van wetenschappelijk onderzoek. De artikelen in deze serie illustreren op begrijpelijke wijze wat een bepaalde methode behelst, zonder dat hier uitvoerige methodologische kennis voor nodig is. Zowel oude als nieuwe methodologische principes worden zo inzichtelijk gemaakt voor artsen die klinische onderzoeken goed willen interpreteren.**

Voor diagnose en prognose zijn steeds meer biomarkers beschikbaar. Zo wordt er intensief gezocht naar biomarkers voor het later optreden van hart- en vaatziekten (HVZ).<sup>1,2</sup> Maar hoe kunnen we de waarde vaststellen van dergelijke biomarkers voor het identificeren van personen met een hoog risico? Voor een hoog-risicogroep zijn de voordelen van behandelen groter dan de nadelen van overbehandelen. Een klassiek voorbeeld is de behandeling van acuut myocardinfarct, waarbij agressievere trombolysie alleen in de patiëntengroep met een hoog risico de nadelen van een groter bloedingsrisico overtreft.<sup>3,4</sup>

De voorspellende waarde van een biomarker blijkt uit de verschillen tussen personen mét en zonder ziekte, vaak uitgedrukt in de vorm van een relatief risico, bijvoorbeeld een 'oddsratio' of 'hazardratio'.<sup>5</sup> Na de eerste studies die aanwijzingen geven dat een biomarker een voorspellende waarde heeft, moeten grondige validatiestudies uitgevoerd worden met nieuwe groepen personen.<sup>5</sup> Hierbij gaat het erom te bepalen wat de toegevoegde waarde is

*Erasmus MC, afd. Maatschappelijke  
Gezondheidszorg, Rotterdam.*

*Prof.dr. E.W. Steyerberg, klinisch besliskundige;*

*Dr B. Van Calster, statisticus*

*(tevens: afd. Ontwikkeling en Regeneratie,*

*KU Leuven, Leuven, België).*

*Boston University, dept. of Biostatistics, and*

*Harvard Clinical Research Institute, Boston, VS.*

*Dr M.J. Pencina, statisticus.*

*Contactpersoon: prof.dr. E.W. Steyerberg*

*(e.steyerberg@erasmusmc.nl).*

## UITLEG

**Netto-nut**

**Netto-nut (NB) is gedefinieerd als het aantal terecht-positieven (TP) minus het gewogen (met gewicht  $w$ ) aantal fout-positieven (FP) in de steekproef van grootte  $n$ . In formulevorm:**

$$NB = (TP - wFP)/n.$$

**Het gewicht  $w$  wordt bepaald door de verhouding tussen de voordelen van terecht behandelen en de nadelen van onterecht behandelen.**

van biomarkers boven de voorspellende waarde van gegevens die al routinematig beschikbaar zijn.<sup>6</sup> Vaak wordt de toename in oppervlakte onder de 'receiver operating characteristic'(ROC)-curve ('area under the curve', AUC) gebruikt als maat voor de toegevoegde waarde.<sup>1</sup>

Tegenwoordig wordt ook vaak de 'net reclassification improvement' (NRI) gerapporteerd, een indicator die is gebaseerd op verschuivingen in de aantallen personen geclassificeerd als 'hoog risico' of 'laag risico'.<sup>7,8</sup> Verderop in dit artikel geven wij een nadere toelichting op deze methode. Een 3e methode maakt gebruik van een maat voor het zogenoemde netto-nut ('net benefit'). In dit artikel laten we zien dat nettonut een betere methode is dan de andere 2 om de diagnostische of prognostische kwaliteit van een biomarker te bepalen.

**VOORBEELD: BIOMARKERS VOOR HART- EN VAATZIEKTEN**

Voorspellende modellen spelen een belangrijke rol bij de preventie van hart- en vaatziekten (HVZ), bijvoorbeeld in de NHG-standaard 'Cardiovasculair risicomanagement'.<sup>9</sup> Het risico op HVZ wordt daarin geschat aan de hand van een beperkt aantal factoren, net als met het Framingham-model.<sup>10</sup> Behandeling met antihypertensiva of statines is alleen geïndiceerd bij een verhoogd risico op HVZ; de

NHG-standaard 'Cardiovasculair risicomanagement' adviseert deze medicatie bij een 10-jaarsrisico  $\geq 20\%$ .

Nieuwe biomarkers voor HVZ staan sterk in de belangstelling, bijvoorbeeld de mate van coronaire kalkafzetting,<sup>2</sup> of een verhoogde concentratie N-terminaal pro-breinnatriuretisch peptide (NT-proBNP) in het bloed.<sup>11</sup> In dit artikel gebruiken we een klassieke risicofactor, namelijk hdl-cholesterol, als voorbeeld om te laten zien hoe de voorspellende kwaliteit van deze biomarker berekend kan worden. Dit doen we aan de hand van gegevens van 3264 personen uit de Framingham-studie; bij 183 personen manifesteerden zich HVZ, met een 10-jaarsincidentie van 5,6% voor myocardinfarct, angina pectoris, hartfalen of overlijden aan HVZ.<sup>7,12</sup>

**MATEN VOOR VOORSPELLENDE KWALITEIT**

We bespreken 3 maten om de kwaliteit van een voorspellende biomarker aan te geven: onderscheidend vermogen, reclassificatie en netto-nut (tabel 1). Bij elk van de 3 methoden vergelijken we de voorspellingen van het Framingham-model mét en zonder de hdl-waarde als risicofactor.<sup>7,12</sup> De vraag is steeds in hoeverre het opnemen van de hdl-waarde in het risicomodel de voorspelling verbetert. Een andere vraag is welke afkappunt we moeten hanteren om de hoog-risicogroep te onderscheiden van de laag-risicogroep. Deze tweede vraag is essentieel als gebruik van het model klinische consequenties heeft, bijvoorbeeld preventieve maatregelen bij personen in de hoog-risicogroep.

**ONDSCHIEDEND VERMOGEN: DE ROC-CURVE**

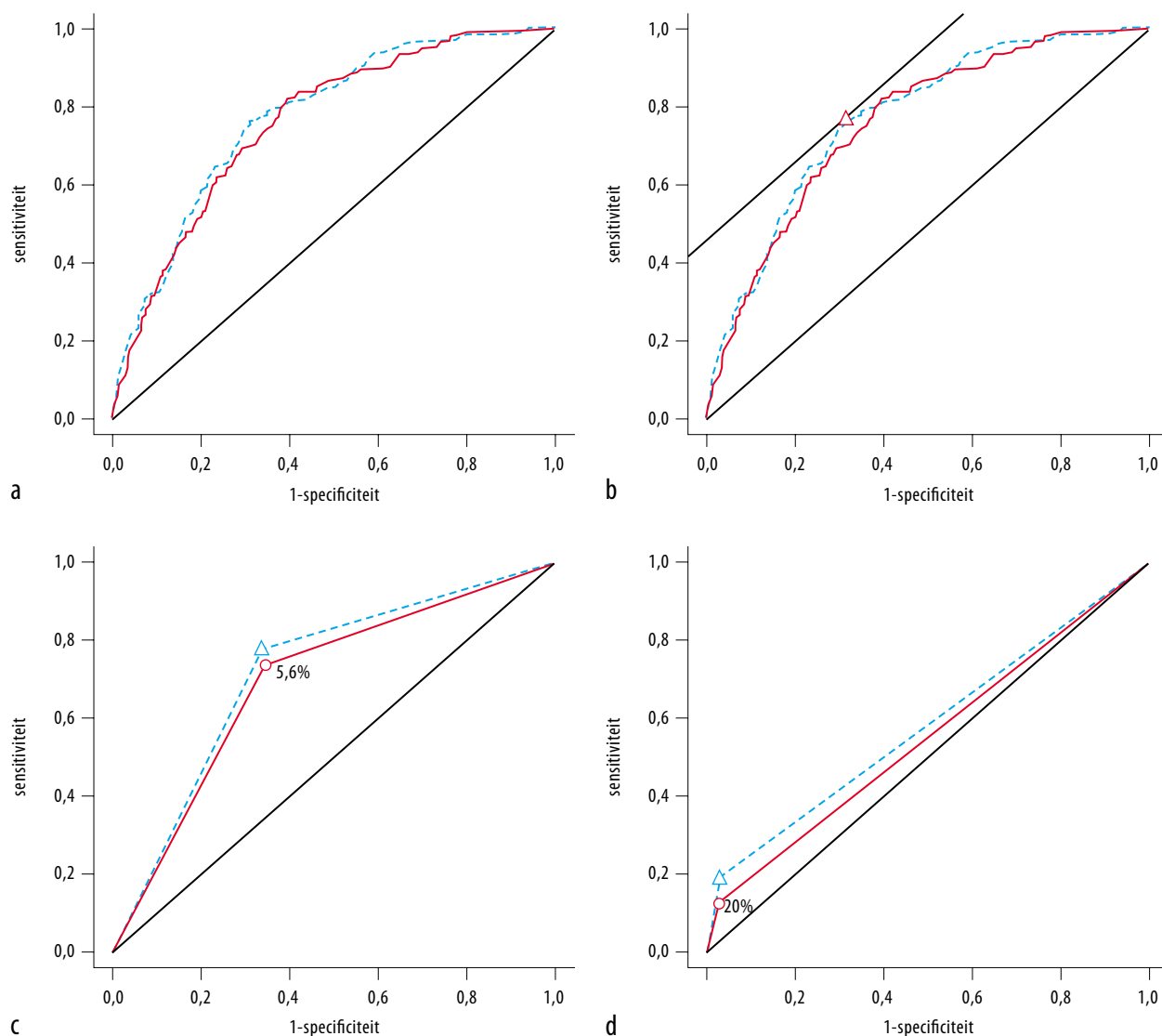
De toegevoegde waarde van hdl in het risicomodel is weer te geven in een ROC-curve. In deze curve worden de sensitiviteit en specificiteit van het risicomodel weergegeven bij alle denkbare afkappunten voor het onder-

**TABEL 1** 3 aspecten en maten om de kwaliteit van een voorspellende marker aan te geven

aspect	maat	opmerkingen
onderscheidend vermogen	toename in oppervlak onder de ROC-curve (AUC)	de AUC is de kans op correcte voorspelling van een ziekte of uitkomst bij een willekeurig paar personen, van wie 1 met en 1 zonder de ziekte of uitkomst die voorspeld moet worden; meestal neemt de AUC maar weinig toe door een marker toe te voegen aan het voorspellende model*
reclassificatie	'net reclassification improvement' (NRI)	de NRI is de nettoverbetering van het aantal correct geclassificeerde personen door de toevoeging van een marker aan een model voor het classificeren van personen in een risicogroep; de NRI vereist 1 of meer afkappunten voor classificatie en wordt uitgedrukt als een fractie
netto-nut	verschil in 'net benefit'	het 'net benefit' is de nettofractie terecht-positieven bij het vergelijken van classificaties op basis van een model met en zonder de marker; hierbij wordt meegewogen hoeveel fout-positieve classificaties nog acceptabel zijn om 1 terecht-positief persoon te identificeren

ROC = 'receiver operating characteristic'; AUC = 'area under the curve'.

\* In het voorbeeld van dit artikel, gebaseerd op de Framingham-studie, nam de AUC toe van 0,76 tot 0,77 door de hdl-waarde als voorspeller toe te voegen.



**FIGUUR 1** 'Receiver operating characteristic'(ROC)-curves voor predictiemodellen die het 10-jaarsrisico op hart- en vaatziekten (HVZ) voorspellen. De modellen zijn gebaseerd op gegevens uit de Framingham-studie en worden gebruikt om personen met een hoog risico op HVZ te identificeren. In elke figuur worden curven weergegeven voor een predictiemodel waarin de hdl-cholesterolwaarde wél (—) of niet (—) is meegeteld als risicofactor. De oppervlakte onder de curve (AUC) is steeds iets groter voor het model waarin hdl-cholesterol wordt meegeteld. (a) ROC-curve voor alle grenswaarden tussen hoog en laag risico; (b) raaklijn aan de curve om het punt met maximale sensitiviteit en specificiteit te vinden; (c) curve met een gemiddelde 10-jaarsrisico van 5,6% als grenswaarde; (d) 20% als grenswaarde (een klinisch relevante grenswaarde om personen met hoog risico op HVZ te identificeren). De zwarte lijn is de referentielijn (dat wil zeggen: een test die op deze lijn uitkomt heeft geen onderscheidend vermogen).

scheid tussen een hoog en laag risico op HVZ (figuur 1a). De AUC is een maat voor het onderscheidend vermogen van het risicomodel; de waarde 0,5 staat voor geen onderscheidend vermogen en 1,0 staat voor een optimaal onderscheidend vermogen.

Het oppervlak onder de curve (AUC) kan in ons voorbeeld geïnterpreteerd worden als de kans dat bij een wil-

lekeurig persoon die HVZ krijgt het risico hoger ingeschat wordt dan bij een willekeurig persoon zonder HVZ. De AUC is 0,762 als wij de hdl-waarde niet meetellen in het risicomodel, en 0,774 als de hdl wél wordt meegeteld. De meeste onderzoekers vinden deze toename van 0,012 teleurstellend, gezien het bereik van de AUC van 0,5 (waardeloos) tot 1 (perfect).

## RECLASSIFICATIE

Een meer recente ontwikkeling is om biomarkers te beoordelen naar hun vermogen om personen beter in te delen in risicocategorieën.<sup>7,8,13,14</sup> Bij het identificeren van een hoog-risicogroep wordt vaak gestreefd naar een maximale sensitiviteit en specificiteit.<sup>15,16</sup> Die vinden we in het meest linksboven gelegen punt op de ROC-curve.<sup>12</sup> In dit punt loopt de raaklijn aan de ROC-curve evenwijdig aan de referentielijn (figuur 1b); dit punt correspondeert met een diagnostische aannemelijkheidsverhouding van 1, ofwel een non-informatief testresultaat.<sup>17</sup> ‘Non-informatief’ betekent hier dat het afkappunt gelijk is aan de gemiddelde incidentie (5,6% 10-jaarsrisico in ons voorbeeld).<sup>18</sup>

‘Reclassificeren’ houdt in dat we bij dit afkappunt de 3264 personen uit de Framingham-studie 2 maal indelen in een hoog- en een laag-risicogroep: eerst zonder de hdl-waarde mee te nemen als risicofactor, de 2e keer mét de hdl-waarde. Bij de 2e keer blijken 321 van de 3264 personen (9,8%) in een andere risicogroep terecht te komen (tabel 2); met andere woorden: er zijn 321 reclassificaties.

Het gaat ons echter om terechte reclassificaties: personen die HVZ ontwikkelen moeten de classificatie ‘hoog risico’ krijgen, en personen zonder HVZ ‘laag risico.’ Bij de reclassificatie in ons voorbeeld komen 10 van de 183 personen met HVZ in de hogere en 3 in de lagere risicogroep. Netto zijn dit 7 terechte reclassificaties ( $7/183 = 3,8\%$ ). Omgekeerd worden er netto 24 personen zonder HVZ terecht gereclassificeerd ( $24/3081 = 0,8\%$ ). De som van deze 2 percentages (uitgedrukt in decimalen: 0,046) staat bekend als de NRI.

**TABEL 2** Reclassificatie: indeling van 3264 personen in een hoog-risicogroep (10-jaarsrisico op hart- en vaatziekten > 5,6%) en een laag-risicogroep (10-jaarsrisico ≤ 5,6%), eerst met een predictiemodel zonder hdl-cholesterol als risicofactor, daarna met hetzelfde model maar mét hdl-cholesterol; links zijn patiënten weergegeven bij wie gedurende 10 jaar follow-up geen hart- en vaatziekten optraden, rechts patiënten bij wie binnen 10 jaar follow-up hart- en vaatziekten optraden

aantallen patiënten zonder HVZ* (n = 3081)			aantallen patiënten met HVZ* (n = 183)		
risico zonder hdl	risico met hdl		risico zonder hdl	risico met hdl	
	laag	hoog		laag	hoog
laag	1872	142	laag	38	10
hoog	166	901	hoog	3	132

HVZ = hart- en vaatziekten; hdl = hdl-cholesterol als risicofactor voor HVZ.

\*Rood en cursief gedrukte aantallen zijn onterechte reclassificaties;

groen en vet gedrukte aantallen zijn terechte reclassificaties.

## ROC-CURVE MET 1 AFKAPPUNT

Het is eerlijker om de NRI te vergelijken met de AUC van een ROC-curve voor het afkappunt van 5,6% (zie figuur 1c) in plaats van de AUC voor alle mogelijke afkappunten. De oppervlaktes zijn dan 0,696 zonder de hdl-waarde mee te nemen als risicofactor en 0,719 mét de hdl-waarde, een verbetering van 0,023. De AUC met 1 afkappunt is per definitie gelijk aan het gemiddelde van sensitiviteit en specificiteit. Hieruit is mathematisch af te leiden dat de NRI gelijk is aan 2 maal het verschil in AUC's bij gebruik van 1 afkappunt.<sup>7,8,12</sup> Zo is de NRI 0,046 en de toename in AUC 0,023 bij een afkappunt van 5,6%. Wellicht maakt de NRI daardoor meer indruk dan de AUC. Overigens is 20% als afkappunt klinisch beter te beargumenteren, omdat iemand met een 10-jaarsrisico op HVZ groter dan 20% in aanmerking komt voor medicatie. Bij het afkappunt van 20% is de toename in AUC 0,029, en de toename in NRI 0,058 (inderdaad: 2 maal 0,029; zie figuur 1d).

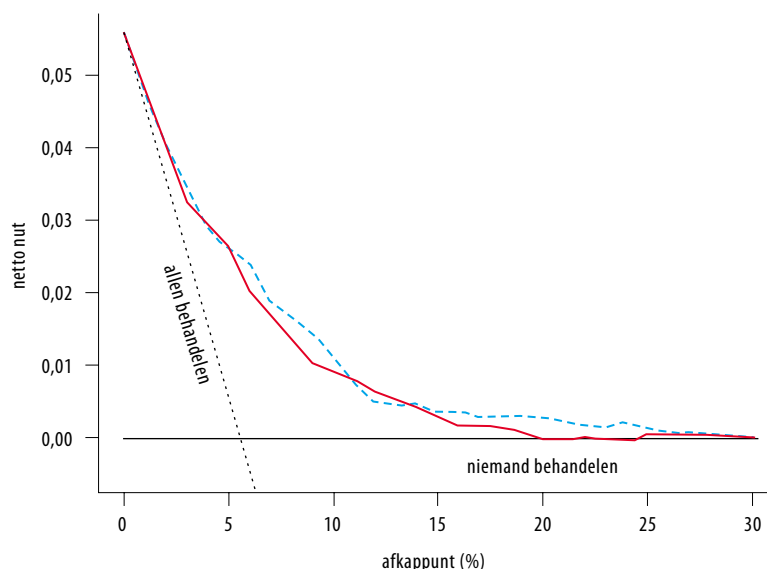
## NETTO-NUT

Reeds in 1884 beschreef Peirce ‘net benefit’ (netto-nut) als een maat voor de waarde van een risicoclassificatie. Deze maat gaat verder dan de bepaling van het onderscheidend vermogen of de reclassificatie.<sup>19</sup> Bij het netto-nut bepaalt de klinische context het relatieve gewicht van terechte en verkeerde classificaties (zie uitlegkader voor de berekening van het netto-nut).<sup>20</sup>

Wat is het idee achter ‘netto-nut’? Terecht-positieven – personen van wie correct voorspeld is dat zij HVZ krijgen – ontvangen een betere behandeling dankzij de goede risicoclassificatie (‘benefit’, winst), terwijl fout-positieven alleen de lasten van de behandeling ondergaan (‘harm’, nadelen). De verhouding tussen de nadelen van onterecht behandelen en de winst door terecht behandelen bepaalt het gewicht  $w$ .

Het gewicht  $w$  kunnen we gebruiken om het optimale afkappunt tussen hoog en laag risico vast te stellen. Een voorbeeld:  $w = 1:4$  als we de onterechte behandeling van 4 personen over hebben voor de winst door 1 terechte behandeling. Bij  $w = 1:4$  hebben personen netto-nut van behandeling als de kans op een terecht-positieve classificatie groter is dan 20%; de ‘odds’ is namelijk 1:4 en de kans wordt berekend als  $\text{odds}/(1+\text{odds})$ .<sup>16,21</sup> Als iemand nu volgens het predictiemodel een risico > 20% heeft, is de verwachte winst door behandelen groter dan het nadeel van mogelijk onterecht behandelen.

Omgekeerd kan het gewicht  $w$  berekend worden door uit te gaan van een bepaald afkappunt. Als we 5,6% als afkappunt nemen, is te berekenen dat we 17 fout-positieven moeten accepteren voor elke extra terecht-positieve (94,4:5,6 ≈ 17). Het maximaliseren van de sensitiviteit plus specificiteit – zoals we deden bij het reclassificeren – houdt dus in dat we veel meer fout-positieve classifica-



**FIGUUR 2** Besliscurve voor predictiemodellen met hdl-waarde (— —) en zonder hdl-waarde (—) om het 10-jaarsrisico op hart- en vaatziekten te voorspellen. Het net van behandelen op basis van het predictiemodel met hdl is voor de meeste afkappunten iets hoger dan voor behandelen op basis van het model zonder hdl. De referentielijnen geven het netto-nut weer bij 'allen behandelen' en 'niemand behandelen' (netto-nut = 0); bij een afkappunt boven de algemene incidentie (in dit voorbeeld een 10-jaarsrisico van 5,6%) zijn de nadelen van iedereen behandelen groter dan de voordelen (het netto-nut wordt negatief).

ties moeten accepteren dan bij het klinisch gebruikelijke afkappunt van 20% (80:20 = 4).

Aan de hand van tabel 2 kan men berekenen wat het netto-nut is van het toevoegen van de hdl-waarde aan het predictiemodel. De uitkomst van die berekening is 0,26%; dit betekent een netto toename van 2,6 terecht-positieven per 1000 personen. Met andere woorden: om 1 extra persoon met HVZ te identificeren moet hdl-cholesterol gemeten worden bij 385 personen (1:0,0026 of 1000:2,6 ≈ 385). Bij een afkappunt van 20% in plaats van 5,6% is dit verschil in netto-nut ongeveer even groot.

De balans tussen winst en nadeel van behandeling is vaak onzeker. Soms zijn te weinig onderzoeksgegevens beschikbaar om 1 afkappunt vast te stellen. Bovendien kan de afweging per persoon verschillen. Daarom is het verstandig om het netto-nut over een bereik van klinisch relevante afkappunten te bekijken.<sup>20,22</sup> Voor preventie van HVZ is een zinvol bereik 5-30%. Dit bereik is gehanteerd in figuur 2, een besliscurve die weergeeft wat het netto-nut is bij behandeling volgens de predictiemodellen zonder en mét hdl-waarde. Deze figuur laat zien dat het predictiemodel mét hdl een klein voordeel biedt ten opzichte van het model zonder hdl.

## PLAATSBEPALING VAN DE METHODEN

Voorspellingen met en zonder een extra biomarker kunnen we vanuit een methodologisch perspectief evalueren, zoals met het oppervlak onder de ROC-curve, het kiezen van een afkappunt op basis van de som van sensitiviteit en specificiteit, of het tellen van terechte reclassificaties (de NRI-componenten en hun som). Om de mogelijke waarde van zo'n biomarker voor de kliniek te bepalen moeten we echter rekening houden met de context waarin voorspellingen worden gebruikt: hoeveel fout-positieven hebben we over voor het identificeren van 1 terecht-positieve?<sup>19,22,23</sup> Een vergelijkbaar concept – het 'number willing to treat' – werd onlangs in dit tijdschrift uitgelegd.<sup>24</sup> Dit besliskundige concept ligt ook ten grondslag aan de maat 'netto-nut' en de besliscurves, die een stap verder gaan dan maten voor onderscheidend vermogen en reclassificatie.

In ons voorbeeld verbeterde de AUC met 0,01 door het toevoegen van hdl aan een model voor het voorspellen van HVZ. Deze toename is klinisch moeilijk te interpreteren. Wel zal een toename in AUC vaak samengaan met een positieve NRI en positief verschil in netto-nut.<sup>6,12,25</sup>

Bij gebruik van een enkel afkappunt is de NRI zoals gezegd 2 maal de toename in AUC. Gezien deze simpele mathematische relatie is het opmerkelijk dat de NRI snel populair geworden is in de grote medische tijdschriften,

## LEERPUNTEN

- De kwaliteit van een biomarker voor diagnostiek of prognose wordt bepaald door te analyseren wat de verschillen zijn tussen het wél en niet toepassen van de marker.
- Vaak worden deze verschillen gemeten aan de hand van de oppervlakte onder de ROC-curve ('area under the curve', AUC).
- Nieuwe maten voor de kwaliteit van een biomarker zijn de 'net reclassification improvement' (NRI) en het 'netto-nut'.
- De maten AUC en NRI verschillen niet wezenlijk van elkaar, omdat er een wiskundig verband tussen deze 2 is.
- Het netto-nut ('net benefit') maakt een betere interpretatie mogelijk van het nut van een biomarker voor diagnose of prognose.

en dat de NRI beschouwd wordt als een 'betere maat' dan de toename in AUC.<sup>13</sup> De NRI wordt overigens vaak voor meerdere afkappunten berekend, bijvoorbeeld 10% en 20% risico op HVZ.

Het netto-nut ('net benefit') daarentegen is echt een andere maat dan toename in AUC, doordat een expliciet gewicht wordt toegekend aan de toename in terecht-positieven en de afname in fout-positieven. Uit de berekening van het netto-nut blijkt in ons voorbeeld dat bij ruim 300 personen hdl-cholesterol gemeten moet worden om 1 extra persoon met HVZ te detecteren, gegeven een behandeldrempel tussen 5 en 20%. Deze uitdrukking is makkelijker te interpreteren dan een toename in AUC van 0,012 of een NRI van 0,046. Immers, als we die hdl-cholesterolbepaling bij 300 personen te veel vinden om die ene extra persoon met HVZ te vinden, heeft hdl kenmerkend te weinig toegevoegde waarde in het predictie-model.<sup>25</sup>

Voor een meer objectieve beoordeling van een biomarker moeten we bepaalde normen stellen, bijvoorbeeld voor de kosteneffectiviteit van de bepaling van een marker. Als er geen direct bewijs uit gerandomiseerde studies is, kunnen we onze toevlucht nemen tot besliskundige modellering van een strategie met en zonder gebruik van een marker. Zo werd recent onderzocht wat de kosten en medische effecten zijn van het bepalen van CRP om de wachttijd tot operatie te optimaliseren.<sup>26</sup> Het bepalen

van CRP was niet kosteneffectief als een grens werd gehanteerd van £ 20,000 per voor kwaliteit gecorrigeerd levensjaar (QALY).

## KANTTEKENINGEN

Er komen steeds meer biomarkers beschikbaar, die de belofte inhouden van een steeds verder gepersonaliseerde geneeskunde. Grondige validatie in nieuwe patiëntengroepen is noodzakelijk, aangezien vele in eerste instantie veelbelovende markers hun belofte later niet kunnen waarmaken.<sup>27</sup> Wij hebben ons in dit artikel geconcentreerd op de rol van biomarkers in predictiemodellen voor het identificeren van een hoog-risicogroep. Markers hebben echter ook andere rollen. Zo kunnen bepaalde biomarkers aangeven of een behandeling kans van slagen heeft. Een voorbeeld hiervan is de test op humane epidermale groeifactorreceptor 2 (HER2-test) voor behandeling van vrouwen met borstkanker met trastuzumab.<sup>28</sup>

## CONCLUSIE

De 'net reclassification improvement' (NRI) is niet veel anders dan de 'receiver operating characteristic'(ROC)-analyse aan de hand van de oppervlakte onder de curve (AUC). In de maat van het netto-nut daarentegen wordt beter rekening gehouden met de klinische context. We voorspellen dat maten als netto-nut in de komende jaren ruim gebruikt zullen worden naast de AUC, gegeven de snel voortschrijdende medisch-technologische ontwikkelingen en het toenemende belang van kosteneffectiviteit in de zorg.

Dr. Hanneke Takkenberg, arts-epidemioloog bij het Thoraxcentrum van het Erasmus MC en 2 anonieme referenten gaven waardevol commentaar op het manuscript.

Belangenconflict: geen gemeld. Financiële ondersteuning: geen gemeld.

Aanvaard op 28 augustus 2012

Citeer als: Ned Tijdschr Geneeskd. 2012;156:A5029

[Meer op www.ntvg.nl/klinischepraktijk](http://www.ntvg.nl/klinischepraktijk)

## LITERATUUR

- 1 Tzoulaki I, Liberopoulos G, Ioannidis JP. Assessment of claims of improved prediction beyond the Framingham risk score. *JAMA*. 2009;302:2345-52.
- 2 Kavousi M, Elias-Smale S, Rutten JH, et al. Evaluation of newer risk markers for coronary heart disease risk classification: a cohort study. *Ann Intern Med*. 2012;156:438-44.
- 3 Glasziou PP, Irwig LM. An evidence based approach to individualising treatment. *BMJ*. 1995;311:1356-9.
- 4 Kent DM, Hayward RA. Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification. *JAMA*. 2007;298:1209-12.

- 5 Hlatky MA, Greenland P, Arnett DK, et al. Criteria for evaluation of novel markers of cardiovascular risk: a scientific statement from the American Heart Association. *Circulation*. 2009;119:2408-16.
- 6 Steyerberg EW, Pencina MJ, Lingsma HF, Kattan MW, Vickers AJ, Van Calster B. Assessing the incremental value of diagnostic and prognostic markers: a review and illustration. *Eur J Clin Invest*. 2012;42:216-28.
- 7 Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;27:157-72.
- 8 Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med*. 2011;30:11-21.
- 9 Samenvattingskaartje NHG-standaard 'Cardiovasculair risicomangement' (eerste herziening) M84. Utrecht: Nederlands Huisartsen Genootschap; 2012.
- 10 Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97:1837-47.
- 11 Rutten JH, Mattace-Raso FU, Steyerberg EW, et al. Amino-terminal pro-B-type natriuretic peptide improves cardiovascular and cerebrovascular risk prediction in the population: the Rotterdam study. *Hypertension*. 2010;55:785-91.
- 12 Steyerberg EW, Van Calster B, Pencina MJ. Performance measures for prediction models and markers: evaluation of predictions and classifications. *Rev Esp Cardiol*. 2011;64:788-94.
- 13 Cook NR, Ridker PM. Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. *Ann Intern Med*. 2009;150:795-802.
- 14 Steyerberg EW, Pencina MJ. Reclassification calculations for persons with incomplete follow-up. *Ann Intern Med*. 2010;152:195-7.
- 15 Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21:128-38.
- 16 Steyerberg EW. *Clinical prediction models: a practical approach to development, validation, and updating*. New York: Springer; 2009.
- 17 Rousson V, Zumbo T. Decision curve analysis revisited: overall net benefit, relationships to ROC curve analysis, and application to case-control studies. *BMC Med Inform Decis Mak*. 2011;11:45.
- 18 Hilden J. , geraadpleegd op september 19 2012.
- 19 Peirce CS. The numerical measure of success of predictions. *Science*. 1884;4:453-4.
- 20 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26:565-74.
- 21 Kassirer JP, Pauker SG. The toss-up. *N Engl J Med*. 1981;305:1467-9.
- 22 Steyerberg EW, Vickers AJ. Decision curve analysis: a discussion. *Med Decis Making*. 2008;28:146-9.
- 23 Greenland S. The need for reorientation toward cost-effective prediction. *Stat Med*. 2008;27:199-206.
- 24 Dorresteijn JAN, Wassink AMJ, Steyerberg EW, van der Graaf Y, Visseren FLJ. Behandeffect bij individuele patiënten voorspellen. *Methodologie van onderzoek 6: predictiemodel op basis van klinische trials*. *Ned Tijdschr Geneesk*. 2012;156:A4510.
- 25 Baker SG, Van Calster B, Steyerberg EW. Evaluating a new marker for risk prediction using the test tradeoff: an update. *Int J Biostat*. 2012;8:1-37.
- 26 Henriksson M, Palmer S, Chen R, et al. Assessing the cost effectiveness of using prognostic biomarkers with decision models: case study in prioritising patients waiting for coronary artery surgery. *BMJ*. 2010;340:b5606.
- 27 Kaiser J. Biomarker tests need closer scrutiny. *Science*. 2012;335:1554.
- 28 Hamilton A, Piccart M. The contribution of molecular markers to the prediction of response in the treatment of breast cancer: a review of the literature on HER-2, p53 and BCL-2. *Ann Oncol*. 2000;11:647-63.
- 29 Vickers AJ, Cronin AM, Begg CB. One statistical test is sufficient for assessing new predictive markers. *BMC Med Res Methodol*. 2011;11:13.