

Welkom bij de KEMTA Masterclasses

De afdeling klinische epidemiologie en HTA (KEMTA) van het MUMC+ organiseert maandelijkse masterclasses voor een ieder geïnteresseerd in (methoden van) wetenschappelijk onderzoek.

Je vindt onze masterclasses (zowel de presentaties als aankomende onderwerpen) op de KEMTA website: <https://www.mumc.nl/research/infrastructuur-en-ondersteuning/partners/kemta/masterclasses>

<https://www.linkedin.com/company/klinische-epidemiologie-medical-technology-assessment-kemta/>

Klinische predictiemodellen

Sander van Kuijk
Afdeling KEMTA
26 september 2023



Scott Adams

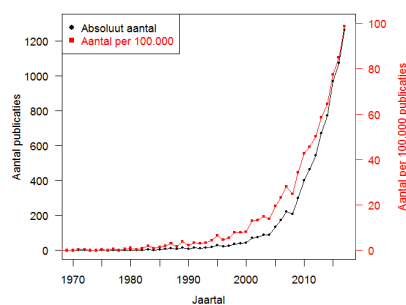
...you can read horoscopes, tea leaves, tarot cards, or crystal balls. Collectively, these methods are known as "nutty methods." Or you can put well-researched facts into sophisticated computer models, more commonly referred to as a complete waste of time.



Inhoud Masterclass

1. Wat is predictie en hoe ontwikkel je een predictiemodel?
2. Predictiemodel performance
3. Interne validatie

Achtergrond



Wat is een (klinisch) predictiemodel?

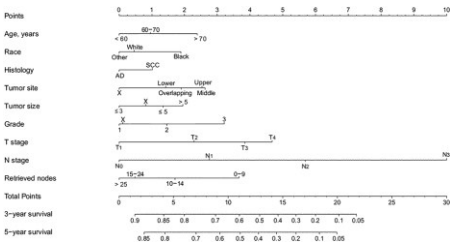
- Een combinatie van predictoren
 - Demografie/ risicofactoren/ prognostische factoren/ klinische observaties/ testuitslagen
- Doel: het berekenen van individuele kans op een (toekomstige) uitkomst
 - Stratificatie van patiënten in hoog- en laagrisico
- Intuitief gepresenteerd
 - Formule?/ nomogram/ online calculator/ app



$$P(\text{herhaling PE}) = 1 / (1 + \exp(-(0.29 - 0.42 * \text{nuchter bloed glucose} + 0.59 * \text{hypertensie} - 0.01 * \text{AD} - 0.41 * \text{groei} + 0.01 * \text{BMI})))$$

$$P(\text{herhaling PE}) = 1 / (1 + \exp(-(0.29 - 0.42 * \text{nuchter bloed glucose} + 0.59 * \text{hypertensie} - 0.01 * \text{AD} - 0.41 * \text{groei} + 0.01 * \text{BMI})))$$

$$P(\text{herhaling PE}) = 1 / (1 + \exp(-(0.29 - 0.42 * 4.3 + 0.59 * 1 - 0.01 * 26 * 7 - 0.41 * 0 + 0.01 * 36))) = 0.165 = 16.5\%$$



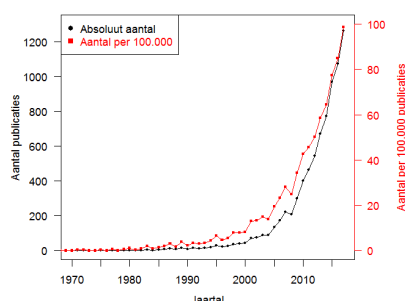
VAGINAL BIRTH AFTER CESAREAN	
Height & weight optional; enter them to automatically calculate BMI	
Maternal age	34 years
Height (range 54-80 in.)	in
Weight (range 80-310 lb.)	lb
Body mass index (BMI, range 15-75)	30 kg/m ²
African-American?	no
Hispanic?	no
Any previous vaginal delivery?	no
Any vaginal delivery since last cesarean?	no
Indication for prior cesarean of arrest of dilation or descent?	yes
Calculate	

Wat maakt predictieonderzoek anders?

- Niet op zoek naar verklaring, maar een goede voorspelling
- Niet geïnteresseerd in individuele predictoren en hun interpretatie, maar een goed werkend geheel
- Resultaten zijn eenvoudig te verifiëren in nieuwe patiënten
- Combinatie van statistische technieken

Stappen in predictie-onderzoek

1. Ontwikkeling
2. (Interne validatie)
3. Externe validatie
4. Impactevaluatie



Stappen in predictie-onderzoek

1. Zoeken beschikbare modellen
2. Externe validatie
3. Impactevaluatie
4. Ontwikkeling
5. (Interne validatie)



~~Stappen in predictie-onderzoek~~ Stappen in model ontwikkeling

1. Ontwikkeling
2. (Interne validatie)
3. Externe validatie
4. Impactevaluatie



RESEARCH

OPEN ACCESS

Check for updates

FAST TRACK

Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal

Laure Wynants,^{1,2} Ben Van Calster,^{2,3} Gary S Collins,^{4,5} Richard D Riley,⁶ Georg Heinze,⁷ Ewoud Schuit,^{8,9} Marc M J Bonten,^{8,10} Darren L Dahly,^{11,12} Johanna A Damen,^{8,9} Thomas P A Debray,^{8,9} Valentijn M T de Jong,^{8,9} Maarten De Vos,^{2,13} Paula Dhiman,^{4,5} Maria C Haller,^{1,14} Michael O Harhay,^{15,16} Liesbet Henckaerts,^{17,18} Pauline Heus,^{8,9} Michael Kammer,^{1,19} Nina Kreuzberger,²⁰ Anna Lohmann,²¹ Kim Luijken,²¹ Jie Ma,² Glen P Martin,^{2,19} David J McLernon,¹ Constanza L Andaur Navarro,¹ Johannes B Reitsma,^{8,9} Jamie C Sergeant,^{24,25} Chunhui Shi,²⁶ Nicole Skoetz,¹⁹ Luc J M Smits,¹ Kym I E Snell,⁶ Matthew Sperin,²⁷ René Spijker,^{8,9,28} Ewout W Steyerberg,⁷ Toshihiko Takada,⁸ Ioanna Tzoulaki,^{29,30} Sander M J van Kuijk,³¹ Bas C T van Bussel,^{1,32} Iwan C C van der Horst,³² Florian S van Royen,⁸ Jan Y Verbakel,^{33,34} Christine Wallisch,^{7,35,36} Jack Wilkinson,²² Robert Wolff,³⁷ Lotty Hooft,^{8,9} Karel G M Moons,^{8,9} Maarten van Smeden⁸

>37,000 hits, 232 modellen!

Stappen in predictie-onderzoek

1. Zoeken beschikbare modellen
2. Externe validatie
3. Impactevaluatie
4. Ontwikkeling
5. (Interne validatie)



Ontwikkeling predictiemodel

- Doel en het studiedesign
- Selectie van potentiële predictoren
- Minimale steekproefgrootte

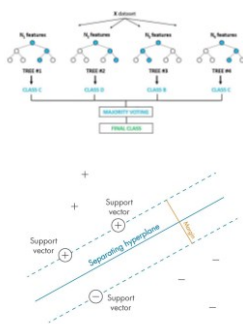
Doel ontwikkeling predictiemodel

- Prognostisch model versus diagnostisch model
- Communiceren individuele kans versus risicofactoren
 - Schatten van kans (in %) op bepaalde uitkomst
 - Kans op pre-eclampsie bij zwangere vrouwen
 - Kans op herhaling van MI na een eerste infarct
 - Indelen in risicogroepen (bijv. laag- en hoog-risico)
 - Risicofactoren voor antibiotische profylaxe
 - Risicofactoren voor preventie myocardi infarct
- Uiteindelijke doel?

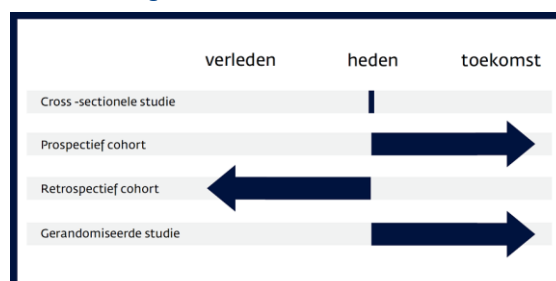
Middel? (logistische) regressie!

...toch liever iets anders?

- Machine learning
 - Random forest
 - Support vector machine
 - Neural network
 - Artificial intelligence
 - Deep learning
 - ...en ga zo maar door...



Studiedesign



Selectie van potentiële predictoren

- Doel van predictorenselectie:
 - Een zo goed mogelijk model met zo min mogelijk predictoren
 - Selectie van variabelen die er echt toe doen
- Categorieën
 - Demografie
 - Type aandoening en ernst
 - (Medische) geschiedenis
 - Comorbiditeiten
 - Functionele status
 - Subjectieve gezondheid (beleving) en QoL
 - Biomarkers
 - Etc.

Causaliteit?

Selectie van potentiële predictoren

- Idealiter op basis van bestaande kennis
 - Artikelen of meta-analyse over risicofactoren
 - Expert-opinie
- Praktische overwegingen predictoren
 - Timing!
 - Kosten, belasting voor patiënt, betrouwbare meetmethode beschikbaar, reguliere meting, etc.
- Nieuwe risicofactoren als onderdeel onderzoeksvraag

Selectie van potentiële predictoren

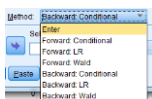
- Stepwise selectiemethoden
 - Stepwise: forward selection of backward elimination

1. Significante predictor

2. Significante verbetering van model fit

- Verandering in Akaike Information Criterion

3. Significante verbetering van model performance



Stepwise selectiemethoden

- Voordelen van stepwise selectie
 - Eenvoudig en objectief, reproduceerbaar
 - Doel om model beknopter te maken wordt snel bereikt
- Nadelen van stepwise selectie
 - 'Testimation bias'
 - Problemen bij collineariteit: arbitraire keuze
 - **Het doet het denkwerk voor ons**
- Toch geautomatiseerde selectie nodig?
 - Backward stepwise elimination met liberale waarde voor alpha (0.2, 0.5)

Steekproefgrootte en power

- Berekening (power = 80%, alpha = 5%, N = ?) past meestal niet bij de onderzoeksvraag
- Té complex model (veel predictoren) doet het waarschijnlijk goed op de eigen data, maar veel slechter in toekomstige patiënten/ deelnemers = **overfitting**
- Uiteindelijk zijn we geïnteresseerd in een model dat het in de **toekomst** goed doet, dus overfitting voorkómen
 - Veel deelnemers voor veel predictoren!

Steekproefgrootte en power

- **'Ouderwetse' stelregel: 10 'events-per-variable'**
 - Let op categorische variabelen, interactietermen, etc.
- Bij een dichotome uitkomst bepaalt de kleinste groep het aantal events!
 - Zeldzame uitkomst: zeer grote steekproef nodig!
- Afwijken van de regel?
 - Graag!

Received: 4 February 2018 | Revised: 13 September 2018 | Accepted: 13 September 2018
DOI: 10.1002/sim.7902

RESEARCH ARTICLE

WILEY *Statistics in Medicine*

Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes

Richard D Riley¹ | Kym HE Snell¹ | Jole Ensor¹ | Danielle L Burke¹ | Frank E Harrell Jr² | Karel GM Moons¹ | Gary S Collins¹

¹Centre for Progress Research, Research Institute for Primary Care and Health Sciences, Keele University, Staffordshire, UK
²Department of Biostatistics, Vanderbilt University School of Medicine, Nashville, Tennessee

When designing a study to develop a new prediction model with binary or time-to-event outcomes, researchers should ensure their sample size is adequate in terms of the number of participants (n) and outcome events (E) relative to the number of predictor parameters (p) considered for inclusion. We propose that the minimum values of n and E (and subsequently the minimum number

Voorbeeld steekproefgrootte

- Predictiemodel voor N2 Lymfekliermetastasen in niet-kleincellig longkanker
- Design: prospectief cohort met 6 potentiële predictoren
 - Leeftijd, geslacht, tumorgrootte, adenocarcinoom (ja/nee), locatie, afwijkende röntgenfoto (ja/nee)
- Minimale steekproefgrootte?
 - 10 *events-per-variable*: $10 \times 6 = 60$ events
 - 60 events / 0.17 (prevalentie) = 353 patiënten



A predictive tool for an effective use of ¹⁸F-FDG PET in assessing activity of sarcoidosis

- Kans op inflammatoire activiteit bij symptomatische patiënten met sarcoïdose
 - Doel: verminderen aantal patiënten dat PET-scan moet ondergaan
 - Middel: uitsluiten patiënten met zeer lage kans
 - Predictoren: verhoogde sIL-2R waarden in het serum en hoog-resolutie CT (volgens classificatiesysteem)
 - Uitkomst: inflammatoire activiteit



Model ontwikkelen in SPSS

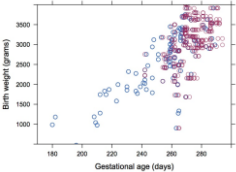
- Ik sla even over:
- Variabelenselectie vóór statistisch modelleren
 - Data-cleaning

Missing data

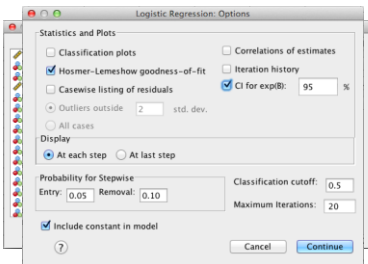
	age	DM	glucose	gestational_age	birthweight	ESG	preterm
1	24.86	4.2	189	680	Yes	No	
2	25.57	5.8	217	1170	No	Yes	
3	25	3.4	195	680	Yes	No	
4	26	5.0	212	680	Yes	No	
5	25	25.14	214	1270	No	No	
6	32	20.44	189	680	No	No	
7	28	27.04	5.1	1180	No	No	
8	28	36.08	5.8	1020	No	Yes	
9	25	32.85	5.9	214	No	No	
10	29	30.72	5.3	221	Yes	No	
11	29	21.11	5.9	237	1020	Yes	
12	30	17.26	5.4	227	1080	Yes	
13	25	22.45	5.5	203	1180	No	No
14	29	24.21	5.3	212	1070	No	No
15	26	23.70	5.7	180	670	Yes	No
16	30	26.39	5.8	203	620	No	No
17	27	24.80	4.4	203	1260	No	No
18	26	34.77	5.2	230	880	Yes	No
19	30	26.22	5.4	203	1030	No	No
20	26	33.46	5.7	201	730	No	Yes
21	28	23.05	5.1	234	1110	No	Yes
22	27	25.03	5.9	202	1080	No	Yes
23	27	24.86	5.5	180	570	Yes	No
24	29	26.71	5.3	217	1480	No	No

Missing data

- Bij > 5% incomplete rijen, effect niet verwaarloosbaar
- Effect op: precisie en eventueel bias
- Oplossing: data-imputatie
 - Multiple imputation?
- Masterclass 2024?



Model schatten in SPSS



Vanregressietabel naar formule

Variables in the Equation							
	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)
Step 1 ^a totalHRCT_score	-.237	.093	6.448	1	.011	1.268	1.056 1.523
lr2afw	1.983	.695	8.135	1	.004	7.267	1.860 28.400
Constant	-.323	.476	.461	1	.497	.724	

a. Variable(s) entered on step 1: totalHRCT_score, lr2afw.

Lineaire deel = $b_0 + b_1X_1 + b_2X_2$

Kans op uitkomst = $P(Y = 1) = \frac{1}{1 + e^{-(\text{Lineaire deel})}}$

Voorbeeld:

Patiënt 1: normale sIL-2R waarden, HRCT classificatie 3

$$\text{Lineaire deel} = -0.323 + 1.983 \cdot 0 + 0.237 \cdot 3 = 0.388$$

$$P(\text{infl. act.}) = 1 / (1 + \exp(-0.388)) = 0.60, \text{ of } 60\%$$

Patiënt 2: afwijkende sIL-2R waarden, HRCT classificatie 4

$$\text{Lineaire deel} = -0.323 + 1.983 \cdot 1 + 0.237 \cdot 4 = 2.608$$

$$P(\text{infl. act.}) = 1 / (1 + \exp(-2.608)) = 0.93, \text{ of } 93\%$$

Hoe goed werkt een predictiemodel?

Scott Adams

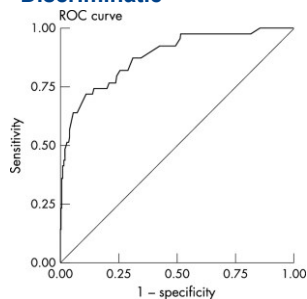
...you can read horoscopes, tea leaves, tarot cards, or crystal balls. Collectively, these methods are known as "nutty methods." Or you can put well-researched facts into sophisticated computer models, more commonly referred to as "a complete waste of time".

**Hoe goed werkt een predictiemodel?**

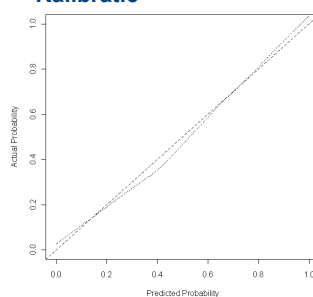
- Statistische 'model fit'
- Vermogen om te discrimineren
- Kalibratie van de voorspellingen

Hoe goed werkt een predictiemodel?

- Statistische 'model fit'
- Vermogen om te discrimineren
- Kalibratie van de voorspellingen

Discriminatie

- Oppervlakte onder de ROC-curve, c-statistic
- Rangschikking van cases/non-cases o.b.v. geschat risico
- Belangrijk voor een groep, bijv. om patiënten in te delen in laag- en hoogrisico groepen

Kalibratie

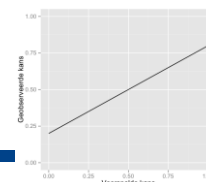
- 'Calibration-in-the-large', calibratie hellingshoek, Hosmer-Lemeshow test, etc...
- Belangrijk voor het individu, bijv. om risico naar patient te communiceren
- Accuraatheid van de predictie

Interne validatie



Interne validatie - doel

- Een schatting maken van de performance van het model voor toekomstige patiënten, correctie van **optimisme**
- Het model aanpassen om toekomstige voorspellingen minder extreem te maken, correctie van **overfitting**
 - Regressiecoëfficiënten dichter bij 0
 - Intercept aanpassen

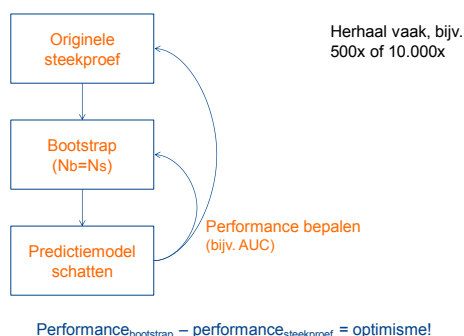


Interne validatie - methoden

- Inefficiënte methoden
 - Split-sample validatie
 - Kruisvalidatie
 - Voordeel: soms heel eenvoudig (in SPSS)
 - Nadeel: veel data nodig!
- Moderne, efficiënte methoden
 - De Bootstrap

De Bootstrap-validatie

- Etymologie: "To pull oneself up by one's own bootstraps" (19e eeuw) (een onmogelijke taak!)
- Het onmogelijke mogelijk maken: het predictiemodel ontwikkelen en valideren met maar één dataset
- Steekproeven trekken uit de eigen data, met teruglegging, en van dezelfde grootte



Reporting guideline, zoals altijd!

"Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis": TRIPOD-statement.

- 22 items, vooral check op compleetheid rapportage
- Uitbreiding voor abstracts en voor geclusterde data
- Kijk op www.equator-network.org voor deze en andere reporting guidelines

Vragen?

