



# Methodische Erfordernisse an Studien zur Prognose

7. arriba-Symposium 19.04 2024 Berlin

[www.uni-marburg.de/fb20/allgprmed](http://www.uni-marburg.de/fb20/allgprmed) · © 2023 AMR · Seite 1

1

## ÜBERBLICK

- ▶ Überblick zur Entwicklung und Validierung von (prognostischen) Vorhersagemodellen
  - ▶ Anforderungen an (externe) Validierungsstudien
  - ▶ Kenngrößen der prognostischen Aussagekraft
- ▶ Berücksichtigung zusätzlicher Risikomarker (nicht-traditionelle Risikofaktoren)
- ▶ Kenngrößen zum Vergleich

[www.uni-marburg.de/fb20/allgprmed](http://www.uni-marburg.de/fb20/allgprmed) · © 2023 AMR · Seite 2

2

## DIE SCHLECHTE NACHRICHT: ES HERRSCHT KEIN MANGEL AN VORHERSAGEMODELLEN

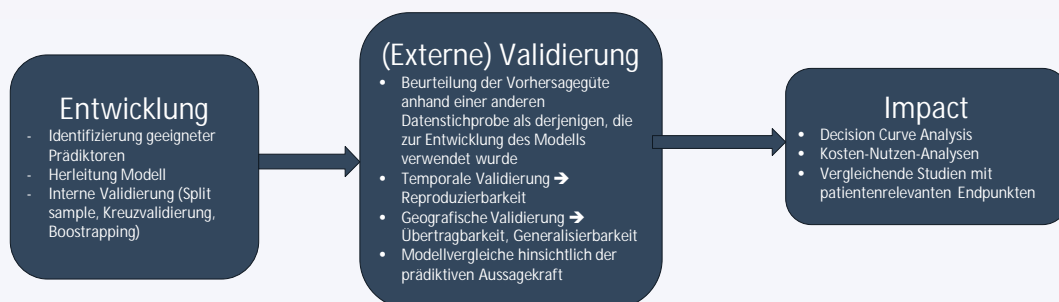
- ▶ 363 Vorhersagemodelle zur individuellen Risikoabschätzung zukünftiger kardiovaskulärer Ereignisse in der Allgemeinbevölkerung (Stand 2013)
- ▶ 118 (33%) machen eine Aussage zum 10 Jahres-Risiko für koronarer Ereignisse
- ▶ Für 132 Vorhersagemodelle (36%) lag eine externe Validierung vor
- ▶ Für 191 (53%) wurde mindestens 1 Kenngröße für die prädiktive Aussagekraft (Diskriminierung, Kalibrierung) berichtet.
- ▶ Für 57 Vorhersagemodelle (16%) ist die Datenbasis ungeeignet (RCT, Fall-Kontroll-Studie, Querschnittstudie)

Damen JAAG, Hoof L, Schuit E, Debray TPA, Collins GS, Tzoulaki I, u. a. Prediction models for cardiovascular disease risk in the general population: systematic review. BMJ. 2016;353:i2416.

www.uni-marburg.de/fb20/allgprmed. © 2023 AMR. Seite 3

3

## ENTWICKLUNG UND EVALUIERUNG KLINISCHER VORHERSAGEMODELLE



Cowley LE, Farewell DM, Maguire S, Kemp AM. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. Diagn Progn Res. Dezember 2019;3(1):16..

www.uni-marburg.de/fb20/allgprmed. © 2023 AMR. Seite 4

4

JAMA | Original Investigation

Effects of the Million Hearts Model on Myocardial Infarctions, Strokes, and Medicare Spending  
A Randomized Clinical Trial

Laura Blue, PhD; Keith Kranker, PhD; Amanda R. Markovitz, ScD; Rhea E. Powell, MD, MPH; Malcolm V. Williams, PhD, MPP; Jia Pu, PhD; David J. Magid, MD, MPH; Nancy McCall, ScD; Allison Steiner, MPH; Kate A. Stewart, SM, PhD; Julia M. Rollison, MPH, PhD; Patricia Markovitch, PhD; G. Greg Peterson, MPA, PhD

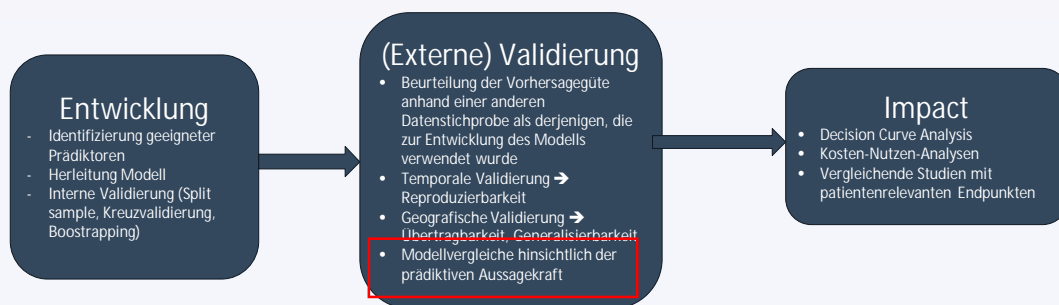
Outcome	Intervention	Control	Difference	
			Unadjusted <sup>b</sup>	Adjusted (90% CI)
<b>Health events (No. per 1000 person-years)</b>				
<b>High- and medium-risk beneficiaries</b>				
First-time myocardial infarctions, strokes, and TIAs (from Medicare claims only) (primary outcome)	15.8	17.0	HR, 0.92	HR, 0.97 (0.93 to 1.0)
First-time myocardial infarctions, strokes, or TIAs, including deaths due to CHD or cerebrovascular disease	18.7	20.3	HR, 0.92	HR, 0.96 (0.93 to 0.99)
All-cause deaths	28.0	29.7	HR, 0.94	HR, 0.96 (0.93 to 0.98)

re fall und 10-  
(for or

mische Anfälle)

5

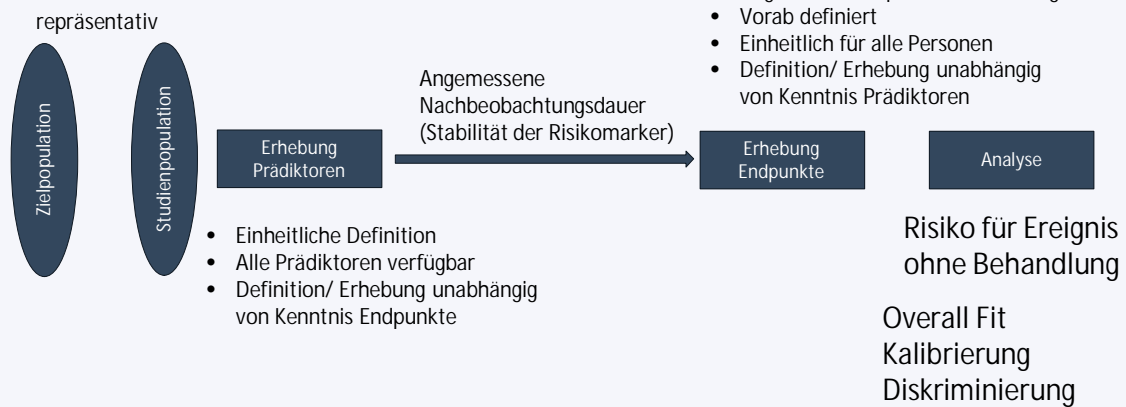
ENTWICKLUNG UND EVALUIERUNG KLINISCHER VORHERSAGEMODELLE



Cowley LE, Farewell DM, Maguire S, Kemp AM. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. *Diagn Progn Res.* Dezember 2019;3(1):16..

6

## EXTERNE VALIDIERUNG PROGNOSTISCHER VORHERSAGEMODELLE – ANFORDERUNG AN DAS STUDIENDESIGN



www.uni-marburg.de/fb20/allgprmed. © 2023 AMR. Seite 7

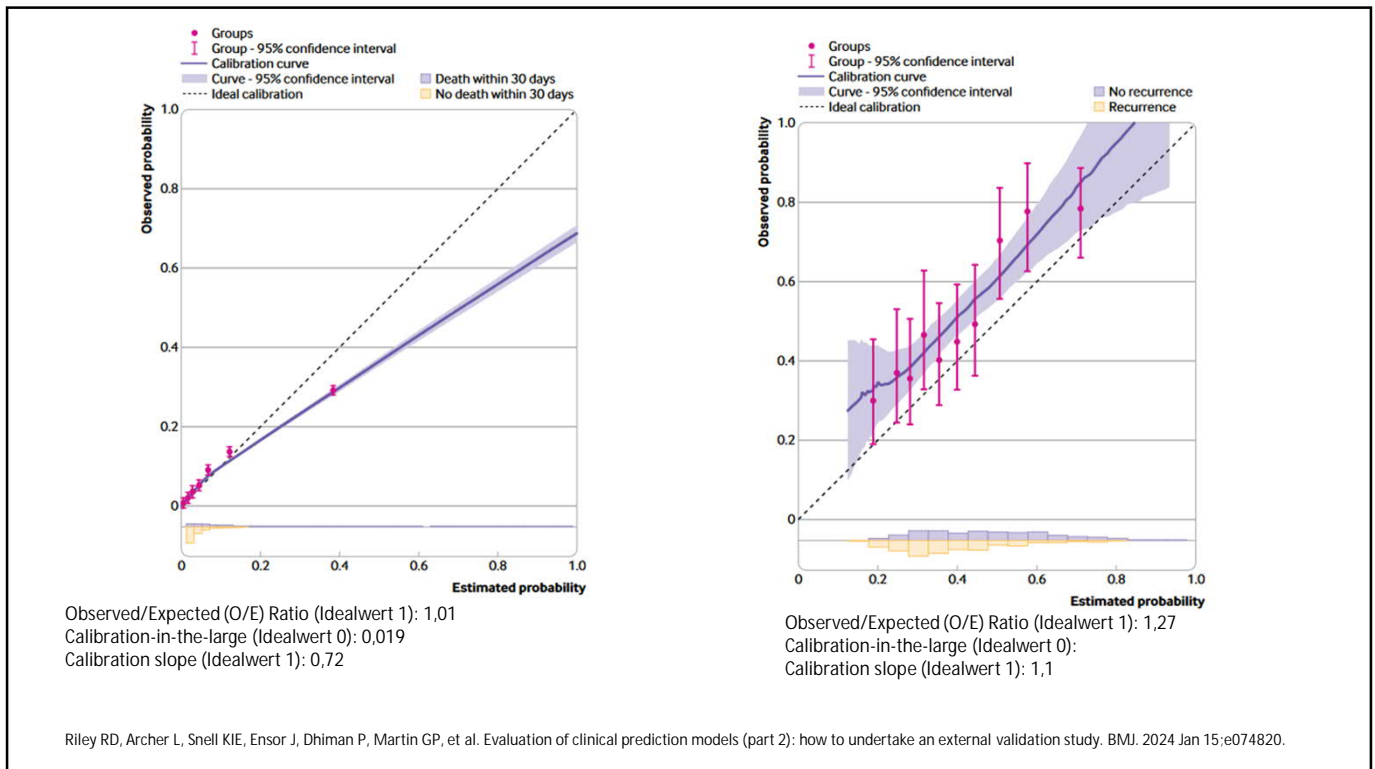
7

## KENNGRÖßEN - KALIBRIERUNG

- ▶ Wie gut stimmt die vom Modell vorhergesagte Wahrscheinlichkeit mit der beobachteten Wahrscheinlichkeit überein?
- ▶ Observed/Expected (O/E) Ratio (Idealwert 1)
- ▶ Calibration-in-the-large (Idealwert 0)
- ▶ Calibration slope (Idealwert 1)

www.uni-marburg.de/fb20/allgprmed. © 2023 AMR. Seite 8

8



9

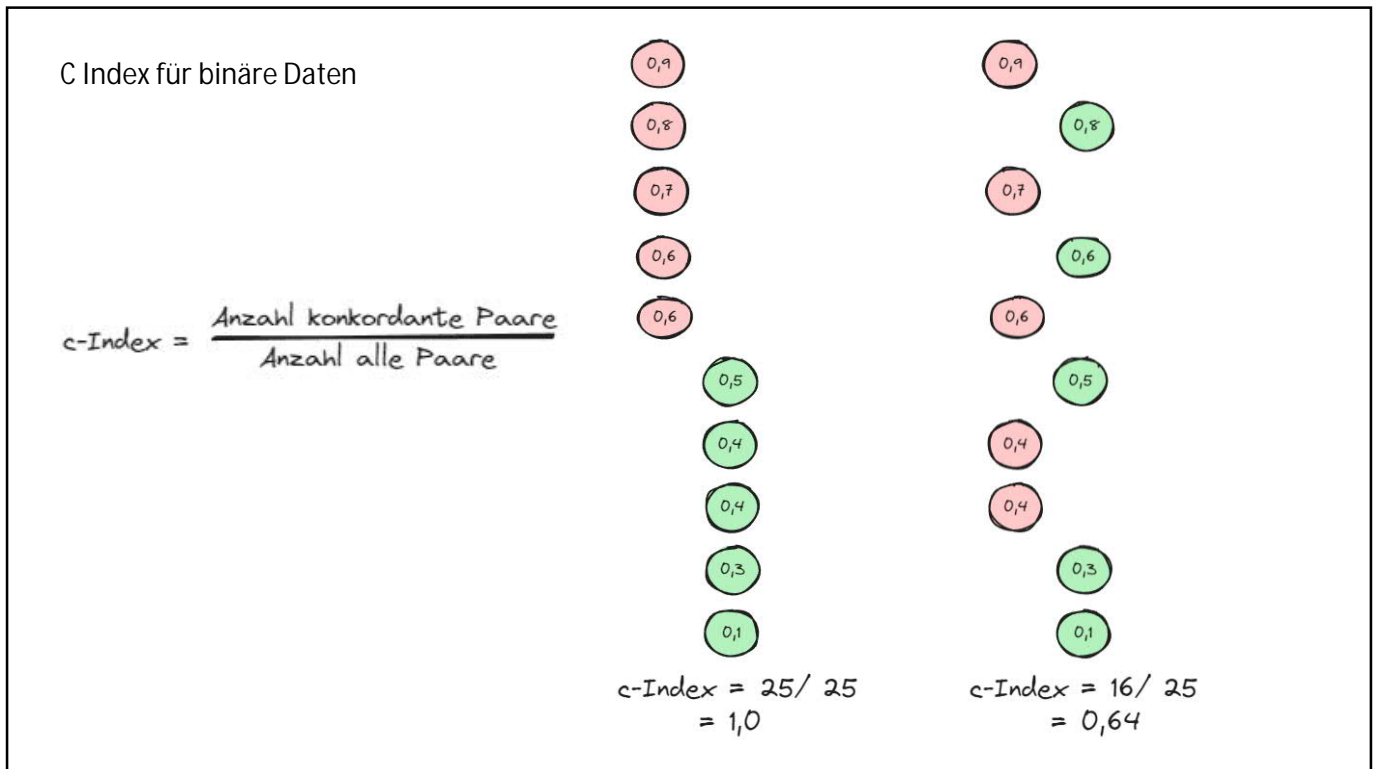
## KENNGRÖßEN - DISKRIMINIERUNG

- ▶ Wie gut unterscheidet das Modell zwischen Gesunden und Kranken bzw. Personen mit und ohne zukünftiges Ereignis?
- ▶ Concordance Index (c Index, c Statistik) – Wahrscheinlichkeit, dass die vom Modell vorhergesagte Wahrscheinlichkeit einer zufällig ausgewählten Person mit Ereignis größer ist als die vom Modell vorgeseigte Wahrscheinlichkeit einer zufällig ausgewählten Person ohne Ereignis.
- ▶ Perfekte Diskriminierung (c-Index = 1): Modell weist allen Fällen einen höheren Wert zu als den Nicht-Fällen.
- ▶ C-Index = 0,5 → Das Modell diskriminiert so gut wie ein Wurf mit einer fairen Münze.

Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation. 2007 Feb 20;115(7):928–35.

www.uni-marburg.de/fb20/allgmed. © 2023 AMR. Seite 10

10



11

AMR  
Institut für Allgemeinmedizin

Philipps  
Universität  
Marburg

## DISKRIMINIERUNG – WAS IST GUT?

- ▶ c-Index < 0,60 → schlechte Diskriminierung
- ▶ c-Index 0,60 bis → möglicherweise hilfreiche Unterscheidung
- ▶ c-Index > 0,75 → eindeutig nützliche Unterscheidung
- ▶ Kontextabhängig!

Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ et al. Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. JAMA 2017; 318(14):1377–84. doi: 10.1001/jama.2017.12126.

www.uni-marburg.de/fb20/allgprmed. © 2023 AMR. Seite 12

12

## VERBESSERUNG DER DISKRIMINIERUNG – WAS IST GUT?

- ▶ United States Preventive Services Task Force (2018)
  - ▶ Veränderungen um  $> 0,1$  → „große Änderung“
  - ▶ Veränderungen von  $0,05$  bis  $0,1$  → „moderate Änderung“
  - ▶ Veränderungen von  $0,025$  bis  $0,05$  → „kleine Änderung“
  - ▶ Veränderungen von  $< 0,025$  → „sehr kleine Änderung“
- ▶ Begründung für Schwellenwerte:
  - ▶ Der Grenzwert von  $0,1$  für „groß“ wurde festgelegt, da er 20 % des möglichen Bereichs darstellt.
  - ▶ Eine Änderung der C-Statistik um  $0,025$  entspricht ungefähr einer um 5 % höheren Sensitivität, wenn die Spezifität 50 % beträgt.

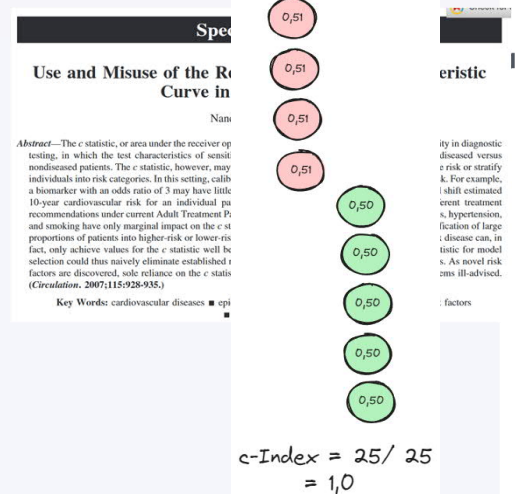
Lin JS, Evans CV, Johnson E, Redmond N, Coppola EL, Smith N. Nontraditional risk factors in cardiovascular disease risk assessment: Updated evidence report and systematic review for the US preventive services task force. JAMA. 2018;320(3):281–97.

www.uni-marburg.de/fb20/allgprmed. © 2023 AMR. Seite 13

13

## KRITIK AN DER C-STATISTIK ZUR BEWERTUNG NEUER RISIKOMARKER

- ▶ Diskriminierung in Vorhersage zukünftiger Ereignisse weniger relevant im Vergleich zur Kalibrierung.
- ▶ Gute Diskriminierung kann mit schlechter Kalibrierung einhergehen.
- ▶ Gute Kalibrierung kann mit schlechter oder mäßiger Diskriminierung einhergehen.



www.uni-marburg.de/fb20/allgprmed. © 2023 AMR. Seite 14

14

## NET RECLASSIFICATION IMPROVEMENT (NRI)

- ▶ Category-based NRI
  - ▶ Overall NRI
  - ▶ Event NRI
  - ▶ Non-Event NRI
- ▶ Continuous NRI
  - ▶ Overall NRI
  - ▶ Event NRI
  - ▶ Non-Event NRI

STATISTICS IN MEDICINE  
*Statist. Med.* 2008; 27:157–172  
 Published online 13 June 2007 in Wiley InterScience  
 (www.interscience.wiley.com) DOI: 10.1002/sim.2929



### Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond

Michael J. Pencina<sup>1,\*</sup>, Ralph B. D'Agostino Sr<sup>1</sup>, Ralph B. D'Agostino Jr<sup>2</sup>  
 and Ramachandran S. Vasan<sup>3</sup>

<sup>1</sup>Department of Mathematics and Statistics, Framingham Heart Study, Boston University, 111 Cummington St., Boston, MA 02215, U.S.A.  
<sup>2</sup>Department of Biostatistical Sciences, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157, U.S.A.  
<sup>3</sup>Framingham Heart Study, Boston University School of Medicine, 73 Mount Wayte Avenue, Suite 2, Framingham, MA 01702-5803, U.S.A.

## NET RECLASSIFICATION IMPROVEMENT (NRI)

Events		Modell 1		
		Negativ	Positiv	
Modell 2	Negativ	40	0	45
	Positiv	15	40	55
		55	45	100
Event NRI	10%			

Sensitivität Modell 1: 45%  
 Sensitivität Modell 2: 55%  
 Differenz: SeM2 – SeM1: 10%

Overall NRI: 10% + -6% = 4 bzw. 0.04

Non-Events		Modell 1		
		Negativ	Positiv	
Modell 2	Negativ	350	120	470
	Positiv	180	350	530
		530	470	1000
Event NRI	-10%			

Spezifität Modell 1: 53%  
 Spezifität Modell 2: 47%  
 SpM2 – SpM1: -6%



## NRI PROBLEME

- ▶ Overall NRI ist die Summe von zwei Anteilswerten mit unterschiedlichen Nennern
- ▶ Overall NRI kann nicht(!) als „Nettoprozensatz der korrekt reklassifizierten Personen“ interpretiert werden.
- ▶ Overall NRI beinhaltet eine implizite Gewichtung der Kosten einer falsch-negativen Reklassifikation (Fälle/ Events herab klassifiziert) bzw. einer falsch-positiven Reklassifikation (Non-Events heraufklassifiziert) in Abhängigkeit von Prävalenz der Fälle:
  - ▶ Im Beispiel Kosten für falsch-negative  $1/0,09 = 11,1$ ; Kosten für falsch-positiv  $1/0,91 = 1,1$
  - ▶ Wenn Prävalenz 50%: Kosten für falsch-negative = Kosten für falsch-positiv
  - ▶ Kann nicht als Anteils- oder Prozentwert interpretiert werden
- ▶ Overall NRI eigentlich nicht interpretierbar
- ▶ “NRI statistics are unhelpful (at best) and misleading (at worst).” (Kerr, 2023)

Leening MJG, Vedder MM, Witteman JCM, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. Ann Intern Med. 2014 Jan 21;160(2):122–31.  
Kerr KF. Net Reclassification Index Statistics Do Not Help Assess New Risk Models. Radiology. 2023 Mar 1;306(3):e222343.

www.uni-marburg.de/fb20/allgprmed. © 2023 AMR. Seite 17

17

## NRI BEISPIEL

- ▶ Entwicklung und Validierung eines Knöchel-Arm-Index-Risikomodells zur Vorhersage koronarer Ereignisse
- ▶ IPD Meta-Analyse (18 Kohorten, davon 24.375 Männer und 20.377 Frauen)
- ▶ Berechnung Framingham Risk Score auf Grundlage berichteter Koeffizienten (FRS Variablen: Alter, Geschlecht, systolischer und diastolischer Blutdruck, Gesamt- und High-Density-Lipoprotein-Cholesterin, Rauchen und Diabetes)
- ▶ Der ABI wurde in vier Gruppen eingeteilt: 0,90; 0,91–1,10; 1,11–1,40; >1,40.
- ▶ Die Studien wurden getrennt nach Geschlecht in zwei Gruppen unterteilt, dann erstens Erstellung eines Entwicklungs- und internen Validierungsdatensatzes und zweitens einen externen Validierungsdatensatzes

Events		FRS Basismodel			
		< 10%	10-20%	> 20%	
Erweitertes Modell FRS mit ABI	<10%	217	5	0	222
	10-19%	46	17	3	66
	>=20%	0	14	15	29
		263	36	18	317

Non-Events		FRS Basismodel			
		< 10%	10-20%	> 20%	
Erweitertes Modell FRS mit ABI	<10%	5074	136	0	5210
	10-19%	432	312	42	776
	>=20%	4	68	84	156
		5500	516	156	6142

Fowkes F, Murray G, Butcher I, Folsom A, Hirsch A, Couper D, u. a. Development and validation of an ankle brachial index risk model for the prediction of cardiovascular events. Eur J Prev Cardiol. März 2014;21(3):310–20.

www.uni-marburg.de/fb20/allgprmed. © 2023 AMR. Seite 18

18

NRI BEISPIEL



► Entwicklung und Validierung eines Knöchel-Arm-Index-Risikomodells zur Vorhersage koronarer Ereignisse

Events		FRS Basismodel			
		< 10%	10-20%	> 20%	
Erweitertes Modell FRS mit ABI	<10%	217	5	0	222
	10-19%	46	17	3	66
	>=20%	0	14	15	29
		263	36	18	317
Event NRI	0,164%				

Non-Events		FRS Basismodel			
		< 10%	10-20%	> 20%	
Erweitertes Modell FRS mit ABI	<10%	5074	136	0	5210
	10-19%	422	312	42	776
	>=20%	4	68	84	156
		5500	516	156	6142
Non-Event NRI	- 0,051				

$$Event\ NRI = \frac{(46+0+14)-(5+0+3)}{317} = \frac{52}{317} = 0,164 \cong 16,4\%$$

$$NonEvent\ NRI = \frac{(136+0+42)-(422+4+68)}{6142} = \frac{-316}{6142} = -0,051 \cong -5,1\%$$

$$Overall\ NRI = 0,164 + -0,051 = 0,113 \cong 11,3 \quad (\text{Publikation: 9,6 durch Schätzung auf Grundlage Kaplan – Meier})$$

www.uni-marburg.de/fb20/allgmed. © 2023 AMR. Seite 19

NRI BEISPIEL



	Männer		Frauen	
	Basismodell 1 FRS als Score	Basismodell 1 + ABI	Basismodell 1 FRS als Score	Basismodell 1 + ABI
C Index	0,672	0,685	0,578	0,690
Overall NRI	4,3%		9,6%	
	Basismodell 2 Risikofaktoren	Basismodell 2 + ABI	Basismodell 2 Risikofaktoren	Basismodell 2 + ABI
	C Index	0,683	0,690	0,788
Overall NRI	2,0%		1,1%	

- „These results suggest that the impact of the ABI is not a fixed phenomenon but is influenced by how well the base risk factor model performs.“
- Im Gegensatz zu rangbasierten Statistiken wie dem C-Index wird der NRI durch eine schlechte Kalibrierung eines Modells beeinflusst.
- Wenn man einen direkten Vergleich zwischen einer Basismodell unter Verwendung der veröffentlichten Koeffizienten und eines erweiterten Modells durchführt, dessen Koeffizienten aus den untersuchten Daten entwickelt wurde, könnte eine positive NRI allein durch eine schlechte Kalibrierung des Basismodells verursacht sein
- Rekalibrierung des Basismodells dann u.U. sinnvoller als die Berücksichtigung eines zusätzlichen Risikomarkers.

Leening MJG, Vedder MM, Wittman JCM, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. Ann Intern Med. 2014 Jan 21;160(2):122-31.

## KATEGORIENFREIE METHODEN

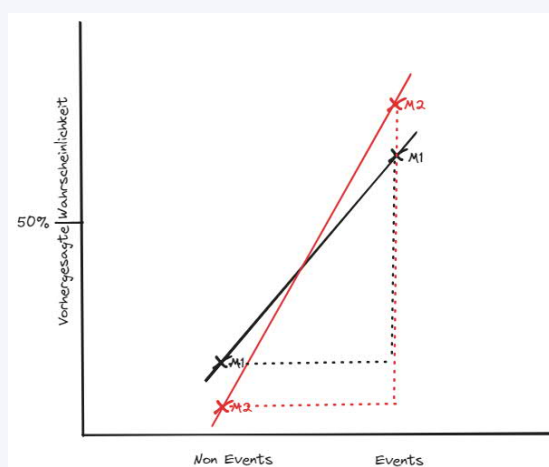
- ▶ Bereits für kategorienbasierte NRI gilt: Die Schätzungen der NRI und ihrer Komponenten steigen mit der Anzahl der Kategorien.
- ▶ Kategorienfreie NRI(0)
  - ▶ Event NRI =  $\frac{\text{Anzahl Events mit höherem Risiko} - \text{Anzahl Events mit niedrigerem Risiko}}{\text{Anzahl Events}}$
  - ▶ Eher Maß für Assoziation statt Modellverbesserung. Das Verhalten kann iratisch sein, wenn der neue Risikomarker nicht normalverteilt ist
- ▶ Integrated Discrimination Improvement (IDI)
  - ▶  $\text{IDI} = (\overline{P_{\text{Erweitert,Events}}} - \overline{P_{\text{Basis,Events}}}) - (\overline{P_{\text{Erweitert,NonEvents}}} - \overline{P_{\text{Basis,NonEvents}}})$ ,
  - ▶ wobei  $\overline{P_{\text{neu,Events}}}$  der Mittelwert vorhergesagten Wahrscheinlichkeiten des erweiterten Modells für die Personen mit Event ist.
  - ▶ Hängt von der Ereignisrate ab. Die Werte sind niedrig und möglicherweise schwer zu interpretieren

Leening MJG, Vedder MM, Wittman JCM, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. *Ann Intern Med.* 2014 Jan 21;160(2):122–31.

www.uni-marburg.de/fb20/allgprmed. © 2023 AMR. Seite 21

21

## INTEGRATED DISCRIMINATION IMPROVEMENT (IDI)



www.uni-marburg.de/fb20/allgprmed. © 2023 AMR. Seite 22

22

## „CLINICAL NRI“

- ▶ Betrachtung nur eines durch das Basismodell definierten Stratum (z.B. Personen mit moderatem Risiko nach Basismodell)
- ▶ → führt zu einer systematischen Verzerrung im Sinne einer Überschätzung der zusätzlichen Aussagekraft des neuen Risikomarkers sowohl bei Reklassifizierung als auch Diskriminierung

Leening MJG, Vedder MM, Witteman JCM, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. *Ann Intern Med.* 2014 Jan 21;160(2):122–31.  
 Bell KJL, White S, Hassan O, Zhu L, Scott AM, Clark J, u. a. Evaluation of the Incremental Value of a Coronary Artery Calcium Score Beyond Traditional Cardiovascular Risk Assessment: A Systematic Review and Meta-analysis. *JAMA Intern Med.* 1. Juni 2022;182(6):634.

www.uni-marburg.de/fb20/allgmed. © 2023 AMR. Seite 23

23

- ▶ Population/ Setting: Rotterdam-Studie, Niederlande, 5933 Personen ohne kardiovaskuläre Vorerkrankung (Durchschnittsalter: 69,1 Jahre [SD: 8,5])
- ▶ Basismodell: Traditionelle KHK-Risikofaktoren (Alter, Geschlecht, systolischer Blutdruck, Behandlung von Bluthochdruck, Gesamt- und High-Density-Lipoprotein-Cholesterinspiegel, Rauchen und Diabetes)
- ▶ Erweiterte Modelle: Berücksichtigung verschiedener neuer Risikomarker (NT-pro-BNP, Koronarer Kalziumscore, Intima-Media-Dicke der Karotisarterien, ...)

### ORIGINAL RESEARCH

Annals of Internal Medicine

#### Evaluation of Newer Risk Markers for Coronary Heart Disease Risk Classification

A Cohort Study

Maryam Kavousi, MD, MSc; Suzette Elias-Smale, MD, PhD; Joost H.W. Rutten, MD, PhD; Maarten J.G. Leening, MD, MSc; Rozemayn Vliegenhart, MD, PhD; Germaine C. Verweert, MSc; Gabriel P. Kreutz, MD, PhD; Matthijs Oudkerk, MD, PhD; Moniek P.M. de Maat, PhD; Frank W.G. Leebeek, PhD; Francesco U.S. Mattace-Raso, MD, PhD; Jan Lindemans, PhD; Albert Hofman, MD, PhD; Ewout W. Steyerberg, PhD; Aad van der Lugt, MD, PhD; Anton H. van den Meiracker, MD, PhD; and Jacqueline C.M. Witteman, PhD

www.uni-marburg.de/fb20/allgmed. © 2023 AMR. Seite 24

24

Appendix Table 2. Global Model-Fit Statistics, for Total Population and by Sex

Extended Model	Likelihood Ratio Test Statistic (Chi-Square)*		
	Total Population	Men	Women
FRS + NT-proBNP	58.7†	35.6†	25.7†
FRS + vWF antigen	2.2	2.3	0.1
FRS + fibrinogen	7.2†	11.4†	0.1
FRS + CKD	5.4†	7.8†	0.3
FRS + leukocyte count	14.1†	19.8†	0.8
FRS + CRP†	4.8†	NA	NA
FRS + homocysteine	5.3†	5.9†	0.7
FRS + uric acid	0.2	0.1	0.0
FRS + CAC score†	60.9†	41.2†	22.6†
FRS + cIMT	5.2†	1.8	3.6
FRS + PAD	3.7†	5.6†	0.1
FRS + PWV	1.2	5.2†	0.9

CAC = coronary artery calcium; cIMT = carotid intima-media thickness; CKD = chronic kidney disease; CRP = C-reactive protein; FRS = Framingham risk score; NA = not applicable; NT-proBNP = N-terminal fragment of pro-brain natriuretic peptide; PAD = peripheral arterial disease; PWV = pulse wave velocity; vWF = von Willebrand factor.  
 \* The increase in model fit after extending the base model (including conventional risk factors) with each new marker.  
 † Statistically significant improvement in model fit (P < 0.05) as evaluated by likelihood ratio chi-square.

Table 3. Discriminative Ability of Risk Markers

Marker	Change in c-Statistic (95% CI)*	NRI (95% CI) for Total Population, %†
NT-proBNP	0.02 (0.01 to 0.04)	7.6 (2.8 to 12.5)
vWF antigen	0.00 (0.00 to 0.00)	0.4 (-1.7 to 2.5)
Fibrinogen	0.00 (0.00 to 0.01)	2.9 (-0.2 to 6.0)
CKD	0.00 (0.00 to 0.00)	2.7 (-0.2 to 5.7)
Leukocyte count	0.01 (0.00 to 0.02)	1.5 (-1.5 to 4.6)
CRP†	0.00 (-0.01 to 0.00)	2.0 (-2.3 to 6.4)
Homocysteine	0.00 (0.00 to 0.00)	-0.3 (-3.0 to 2.3)
Uric acid	0.00 (0.00 to 0.00)	0.8 (-0.5 to 2.1)
CAC score†	0.05 (0.02 to 0.06)	19.3 (12.5 to 26.2)
cIMT	0.00 (0.00 to 0.00)	1.6 (-1.1 to 4.4)
PAD	0.00 (0.00 to 0.00)	0.6 (-1.8 to 2.9)
PWV	0.00 (0.00 to 0.00)	0.00 (-2.1 to 2.1)

Table 4. Coronary Heart Disease Risk Reclassification for the Framingham Intermediate-Risk Group After Extending the Model With Each Marker

Extended Model	Persons With Event, %*		Persons Without Event, %*		NRI (95% CI), %†
	Up	Down	Up	Down	
FRS + NT-proBNP	29.0	16.4	9.8	30.2	33.0 (23.4 to 42.6)
FRS + vWF antigen	4.4	3.9	2.0	5.5	4.0 (-0.2 to 8.1)
FRS + fibrinogen	10.6	4.8	5.9	10.3	10.2 (4.5 to 15.9)
FRS + CKD	10.0	3.2	5.4	8.4	9.8 (4.4 to 15.1)
FRS + leukocyte count	11.5	6.4	6.0	10.2	9.3 (3.2 to 15.4)
FRS + CRP†	12.3	10.4	4.9	12.2	9.2 (0.2 to 18.0)
FRS + homocysteine	7.6	7.8	3.7	8.6	4.7 (-0.9 to 10.3)
FRS + uric acid	0.8	0.0	0.5	2.3	2.6 (1.0 to 4.2)
FRS + CAC score†	37.0	13.0	18.7	34.0	39.3 (26.8 to 51.7)
FRS + cIMT	6.0	4.2	4.3	7.1	4.6 (-0.1 to 9.3)
FRS + PAD	6.8	2.1	4.2	6.8	7.3 (2.9 to 11.7)
FRS + PWV	4.4	2.8	2.1	3.7	3.2 (-0.6 to 7.1)

25

Events		Basismodel			
		< 10%	10-20%	> 20%	
Modell mit CAC Score	<10%	71	19	0	90
	10-20%	50	75	9	134
	> 20%	4	55	62	121
		125	149	71	345
Event NRI	0,2347				

Non-Events		Basismodel			
		< 10%	10-20%	> 20%	
Modell mit CAC Score	<10%	2015	262	17	2294
	10-20%	315	364	60	739
	> 20%	16	144	140	300
		2346	770	217	3333
Non-Event NRI					

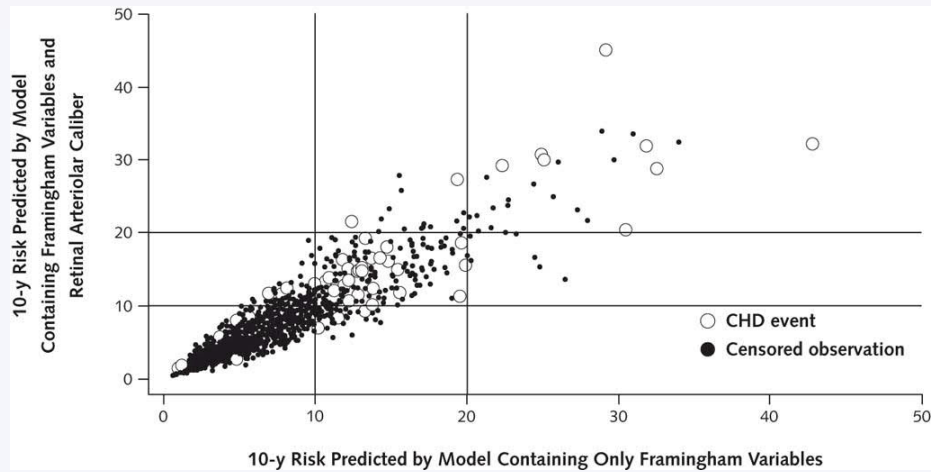
$$Event\ NRI = \frac{(50+4+55)-(19+0+9)}{345} = \frac{81}{345} = 0,234$$

$$NonEvent\ NRI = \frac{(262+17+60)-(315+16+144)}{3333} = \frac{-136}{3333} = -0,04$$

$$Overall\ NRI = 0,234 - 0,04 = 0,194$$

26

## GRAFISCHE ALTERNATIVE ZUR REKLASSIFIKATIONSTABELLE



www.uni-marburg.de/fb20/allgmed. © 2023 AMR. Seite 27

27

## FAZIT

- ▶ Zur Bewertung der zusätzlichen Aussagekraft eines neuen Risikomarkers sollten mehrere Kenngrößen gemeinsam betrachtet werden.
- ▶ Kenngrößen der Diskriminierung und Kalibrierung für beide Modelle sollten betrachtet werden, vorzugsweise aus Daten einer externen Validierung
- ▶ Wenn ein zusätzlicher Risikomarker nicht mit dem Outcome assoziiert ist oder nicht zu einem Anstieg der AUC führt, sollte ein positiver NRI nicht erwartet werden.
- ▶ Bei der Bewertung der NRI müssen eine Reihe von Faktoren berücksichtigt werden.
  - ▶ Die Overall NRI erlaubt keine Bewertung im Hinblick auf die klinische Relevanz.
  - ▶ Größen hängen ab von Anzahl und Lage der Schwellenwerte.
  - ▶ Die gewählten Schwellenwerte sollen klinisch bedeutsam sein.
  - ▶ Konfidenzintervalle (vorzugsweise basierend auf Bootstrapping) statt Hypothesentests
- ▶ Die Einschätzung des tatsächlichen klinischen Nutzens ist/ bleibt schwierig.

 Leening MJG, Vedder MM, Wittman JCM, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. *Ann Intern Med.* 2014 Jan 21;160(2):122–31.

 Mihaescu R, van Zitteren M, van Hoek M, Sijbrands EJG, Uitterlinden AG, Wittman JCM, u. a. Improvement of risk prediction by genomic profiling: reclassification measures versus the area under the receiver operating characteristic curve. *Am J Epidemiol.* 1. August 2010;172(3):353–61.

www.uni-marburg.de/fb20/allgmed. © 2023 AMR. Seite 28

28

Dr. rer.medic. Stefan Bösner, MScN  
Arzt für Allgemeinmedizin

Philipps-Universität Marburg  
Institut für Allgemeinmedizin  
Karl-von-Frisch-Strasse 4  
35043 Marburg

MAIL [haasenri@staff.uni-marburg.de](mailto:haasenri@staff.uni-marburg.de)  
TEL +49 (0)6421 28651-20  
FAX +49 (0)6421 28651-21  
WEB <http://www.uni-marburg.de/fb20/allgprmed>