

Neuere Ansätze zur Auswahl von Prädiktionsmodellen

Von Veronika Huber

Gliederung

- Anwendungsbereiche von Prädiktionsmodellen
- Traditionelle Methoden zur Prüfung der Wirksamkeit
- Neuere Ansätze zur Prüfung der Wirksamkeit
- Vor- und Nachteile der neuen Methoden
- Zusammenfassung
- Quellen

Anwendungsbereich von Prädiktionsmodelle

- (Regressions-)Modelle für Vorhersagen oder Diagnosen aus erhobenen Daten
- Im speziellen in der Medizin:
 - Aktueller Krankheitszustand
 - Prognoseabschätzungen
 - Entscheidungshilfen für:
 - Weitere Tests
 - Therapien

Beispiel

- Existierendes (Regressions-) Modell im Vergleich mit
- (Erweiterung des) Modells durch Marker
- An dem konkreten Beispiel:
 - Die Vorhersage von restlichen Tumor Gewebe nach einer Bestrahlung bei Hodenkrebs

Gliederung

- Anwendungsbereiche von Prädiktionsmodellen
- Traditionelle Methoden zur Prüfung der Wirksamkeit
 - Brier score
 - ROC Kurve
 - Einschränkungen der Methoden
- Neuere Ansätze zur Prüfung der Wirksamkeit
- Vor- und Nachteile der neuen Methoden
- Zusammenfassung
- Quellen

Brier score

- Wie nah sind die Vorhersagen am wahren Ergebnis?
- Betrachtet die quadrierten Abstände:

$$Brier = (Y - p)^2$$

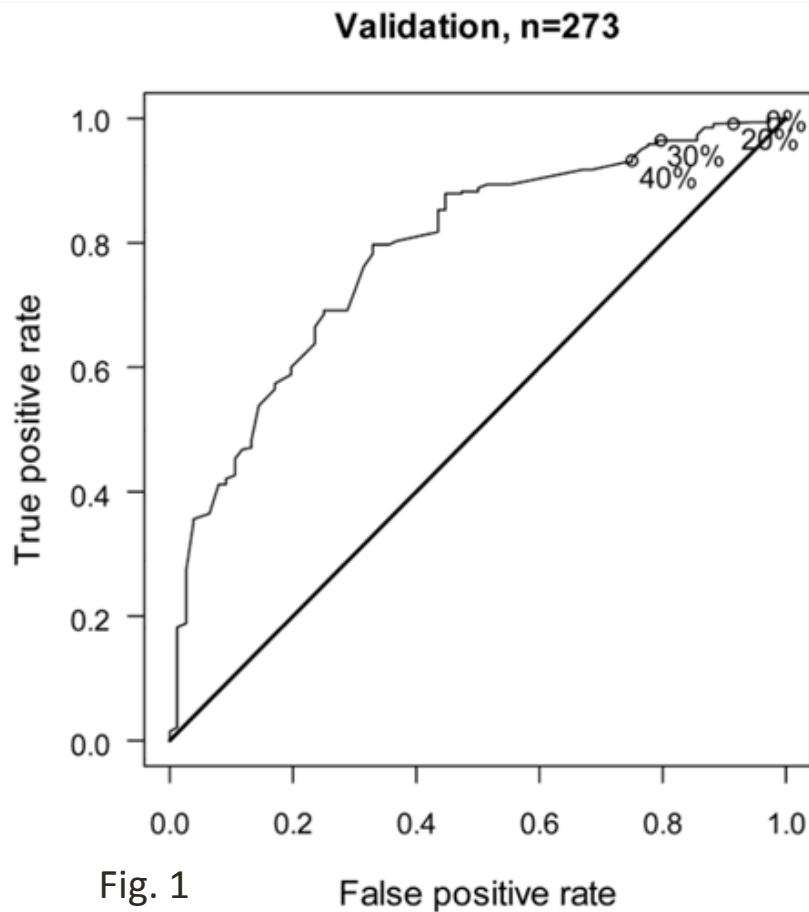
- Y : binäres Ergebnis
- p : vorhergesagte Wahrscheinlichkeit
- 0 perfektes Modell
- 1 informationsloses Modell

Brier score - Beispiel

- $Y = 1$ Sonnenschein
- $Y = 0$ Kein Sonnenschein

- Angenommen $Y = 1$
- $p = 0\%$ Wahrscheinlichkeit für Sonnenschein
 - Brier = $(1 - 0)^2 = 1$
- $P = 70\%$ Wahrscheinlichkeit für Sonnenschein
 - Brier = $(1 - 0.7)^2 = 0.09$
- $p = 100\%$ Wahrscheinlichkeit für Sonnenschein
 - Brier = 0

ROC Kurve



- ROC: Receiver Operating Characteristic
- Wie genau ist die Klassifizierung der Patienten?
- Richtig positiv Rate gegen falsch positivRate
- Je größer die Fläche unter der Kurve (AUC- Area under the Curve), desto besser ist Modell

ROC Kurve - Beispiel

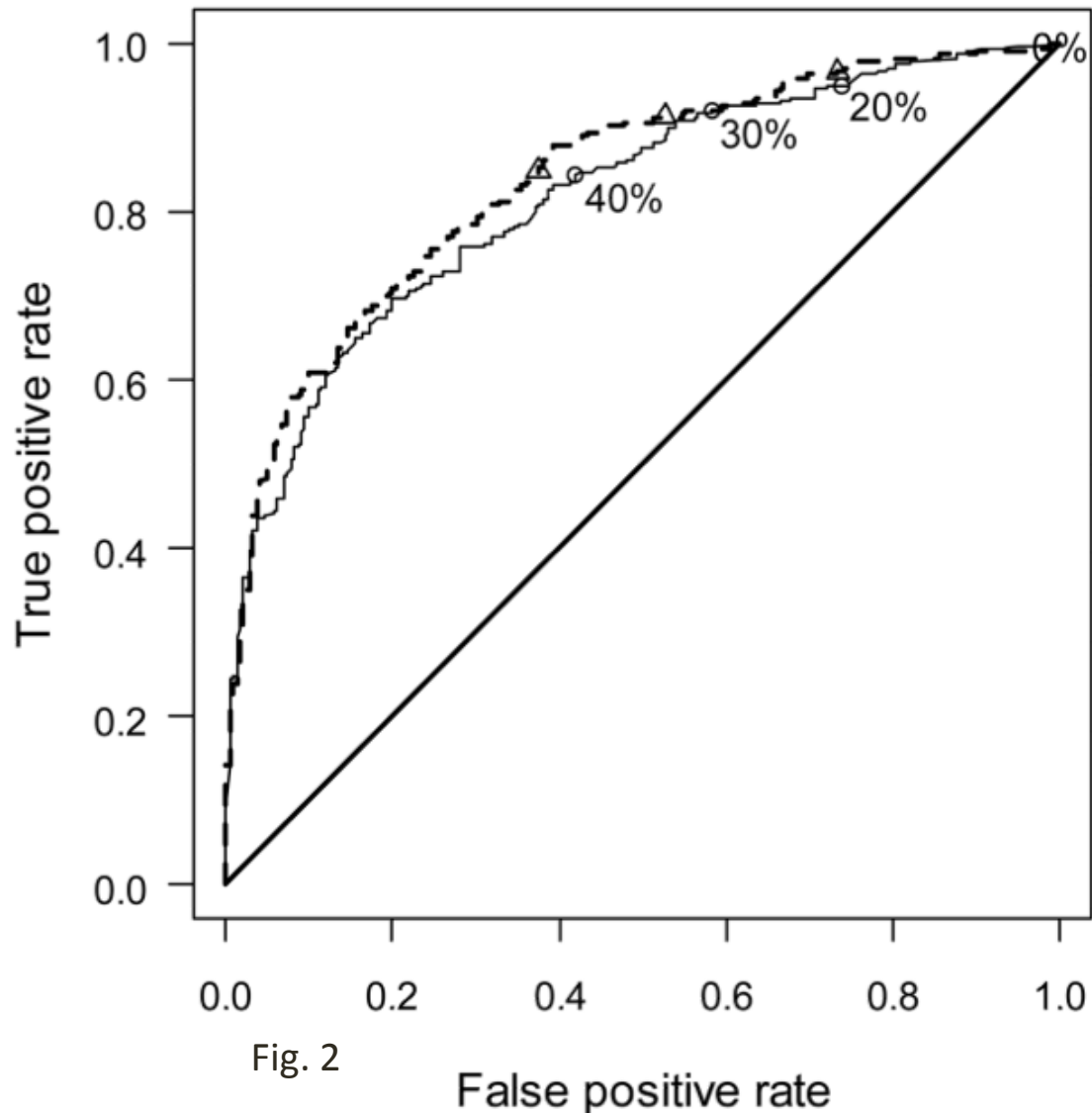


Fig. 2

- Glatte Kurve: Basis Modell
- Gestrichelte Kurve: Modell mit Marker
- Etwas größere Fläche unter der Kurve

Einschränkungen der Methoden

- Nur aus Statistischer Sicht betrachtet
 - Wie genau sind die zahlen
- Angewandt in der Medizin auch andere Faktoren wichtig:
 - Klinischer Nutzen
 - Welches Modell ist wann anzuwenden
- Welches der viele Modelle ist nach Annahme der Vorhersagen mit dem kleinsten möglichen Nachteilen oder größten möglichen Vorteilen verbunden?

Gliederung

- Anwendungsbereiche von Prädiktionsmodellen
- Traditionelle Methoden zur Prüfung der Wirksamkeit
- Neuere Ansätze zur Prüfung der Wirksamkeit
 - Net benefit
 - Entscheidungskurven
 - Reklassifizierung – Cook
 - Reklassifizierung – NRI
- Vor- und Nachteile der neuen Methoden
- Zusammenfassung
- Quellen

Net Benefit - Idee

- unnötigen Behandlungen ins Verhältnis zu den gewonnenen Informationen
- 100 Personen werden auf Krebs getestet
 - nach Biopsien: bei 25 Patienten Krebs bestätigt
 - 75 Personen unnötig behandelt
- Modell mit Markern prognostiziert nur für 72 Personen ein Krebsrisiko
 - nach Biopsien: nur bei 22 Personen Krebs bestätigt
 - aber nur 50 Personen wurden unnötigen behandelt

Net Benefit - Idee

$$\text{Net Benefit} = \text{Nutzen} - (\text{Schaden} \times \text{Ausgleichsfaktor})$$

- Ein Arzt würde 10 Biopsien in Kauf nehmen um 1 Krebspatienten zu finden
 - Ausgleichsfaktor = $1/9$

$$\text{Ausgleichsfaktor: } \frac{\text{Anzahl positiver Funde}}{\text{Anzahl Personen negativer Funde}}$$

- Net Benefit(alle biopsieren) = $25\% - (75\% \times 1/9) = 16,7\%$
- Net Benefit(Modell) = $22\% - (50\% \times 1/9) = 16,4\%$
- Ein höherer Wert ist besser

Net benefit - Angewandt

$$\text{Net benefit} = \frac{\text{richtig positiv}}{N} - \frac{\text{falsch positiv}}{N} \times \frac{p_t}{1 - p_t}$$

- N: Anzahl der Personen mit Risiko
- Grenzwahrscheinlichkeit p_t : ab der es ein Risiko für die Erkrankung gibt.
- Grenzwahrscheinlichkeit in Relation zu Ausgleichsfaktor
 - $p_t = 10\%$ heißt 10 Biopsierte werden toleriert um 1 Krebserkrankung zu finden
 - $p_t = 5\%$ 20 tolerierte Biopsien
 - $100/p_t =$ tolerierte Biopsien

Entscheidungskurve - Idee

- Grafische Darstellung des Net Benefit
- Berechnung der Net Benefit für:
 - Mehrere Modelle
 - Mit jeweils unterschiedlichen Grenzwahrscheinlichkeiten
- X-Achse: Grenzwahrscheinlichkeit
- Y-Achse: Net Benefit

Entscheidungskurve - Beispiel

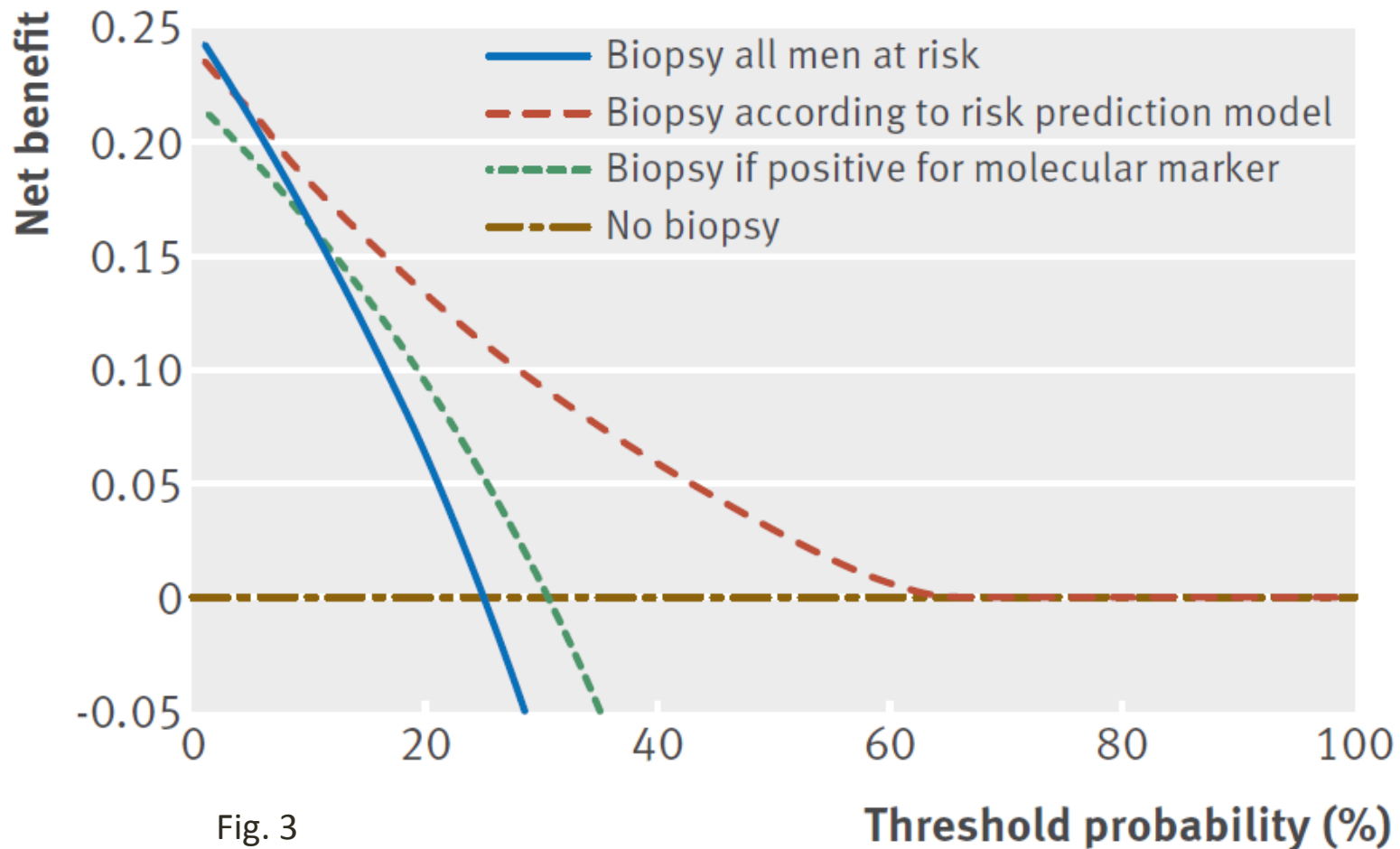


Fig. 3

Entscheidungskurve – Interpretation

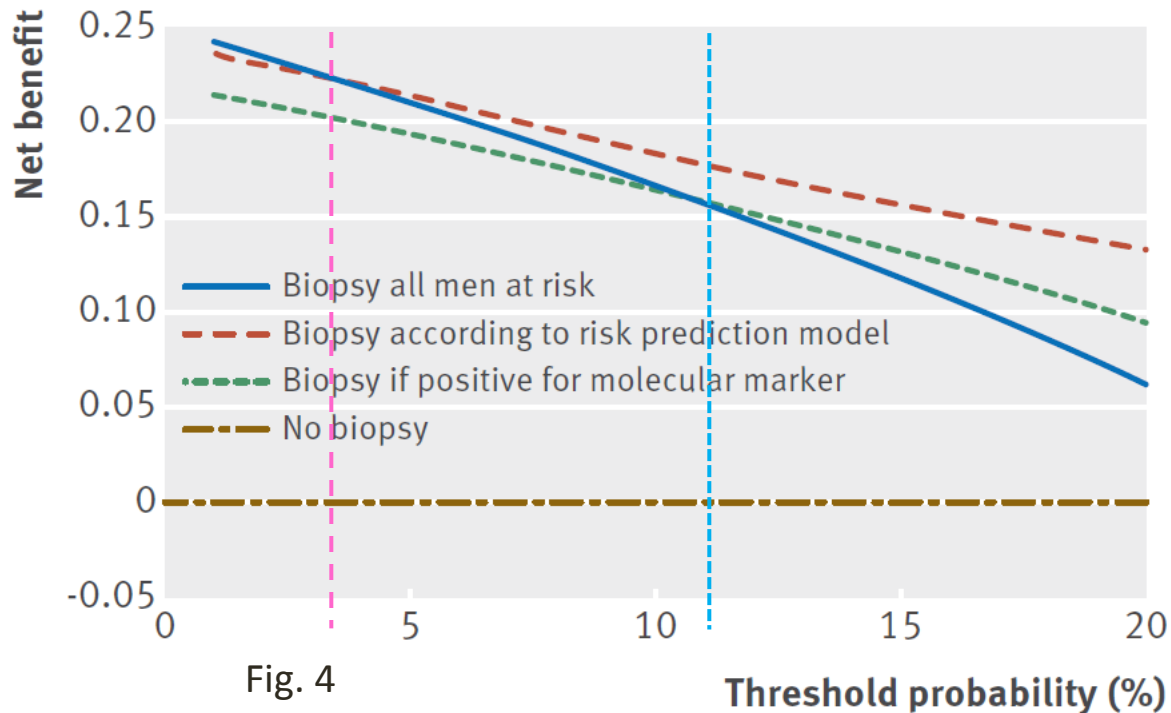


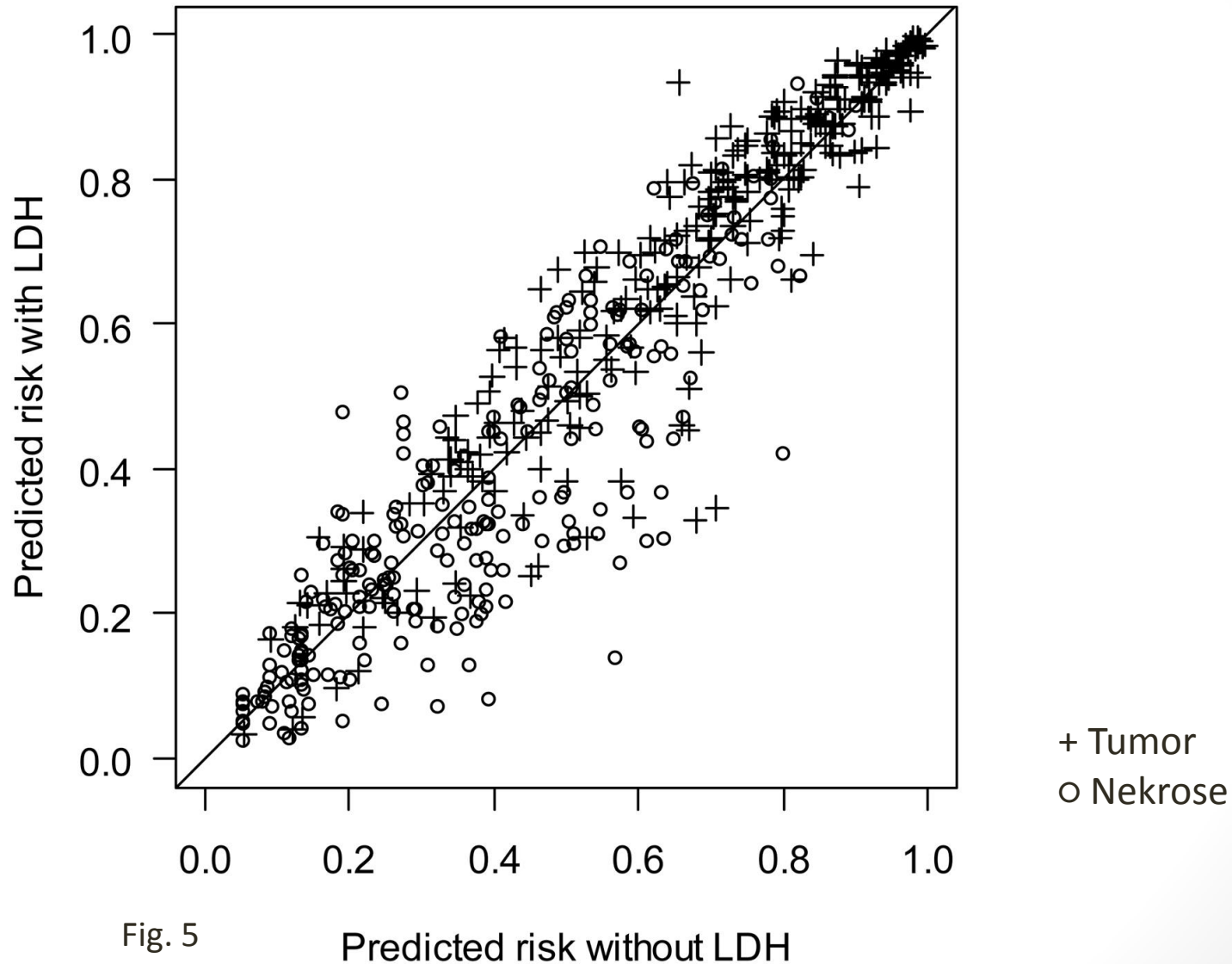
Fig. 4

Threshold probability (%)

- Ab einer Wahrscheinlichkeit von etwa 11% wird das besser klinische Ergebnis durch den Test auf den Marker erreicht

- Unter 4% Wahrscheinlichkeit wäre die beste Strategie alle Männer zu biopsieren
- Ab 4% ist es die beste Strategie dem risk prediction Modell zu folgen

Reklassifizierung - Beispiel



Reklassifizierung – Cook

- Wie viele Subjekte werden in dem neuen Modell mit höherer bzw. niedrigerer Risikowahrscheinlichkeit eingestuft
- Einteilung in Kategorien (klein, mittel, hoch)
- Reklassifizierung in %:

$$R = \frac{n_{höher} + n_{niedriger}}{N}$$

- N : Anzahl der Personen
- $n_{höher}$: höhere Risikokategorie als vorher
- $n_{niedriger}$: niedrigere Risikokategorie als vorher

NRI - Idee

- NRI: Net Reclassification Index
- Erweiterung des Modells von Cook
- Berücksichtigung des Ereignisses:
- Neueinordnung zu höherem Risiko
 - mit Ereignis: Verbesserung
 - ohne Ereignis: Verschlechterung
- Anders herum für Einordnung in geringeres Risiko

NRI – Angewandt

- Kategorialer NRI:
 - Verschiebung zwischen Kategorien
- Kategorie freier NRI:

$$NRI = NRI_{Ereignis} + NRI_{Nicht-Ereignis}$$

$$NRI = \{P(risk_{exp} > risk_{base} | Ereignis) - P(risk_{exp} < risk_{base} | Ereignis)\} - \{P(risk_{exp} < risk_{base} | Nicht - Ereignis) - P(risk_{exp} > risk_{base} | Nicht - Ereignis)\}$$

- $risk_{base}$: Risiko nach Basis Modell
- $risk_{exp}$: Risiko nach erweiterten Modell

Gliederung

- Anwendungsbereiche von Prädiktionsmodellen
- Traditionelle Methoden zur Prüfung der Wirksamkeit
- Neuere Ansätze zur Prüfung der Wirksamkeit

- Vor- und Nachteile der neuen Methoden
 - Net Benefit/ Entscheidungskurve
 - NRI

- Zusammenfassung
- Quellen

Vor- und Nachteile Net Benefit/ Entscheidungskurve

- Anhaltspunkt für Entscheidungen
- Berücksichtigt auch die Nachteile
- Detaillierte Informationen über Präferenzen nicht nötig
 - Übersicht über ein Bereich von Grenzwahrscheinlichkeiten
- zu beachten:
 - Die Wahl der Grenzwahrscheinlichkeit ist oftmals schwierig
 - Keine ausreichenden/ genauen Daten über Nutzen und Nachteile

Vor- und Nachteile NRI

- Möglichkeit zwei Modelle genauer mit einander zu vergleichen als bisher mit der Differenz der AUC Werte
- Gewichtung der Ergebnisse
- Aber der NRI ist noch nicht Fehler frei:
 - Test von Modellen, die mit Markern erweitert wurden die keinen Einfluss auf das Ergebnis haben.
 - NRI war signifikant positiv
 - Während AUC deutlich geringer abwich

Zusammenfassung

- Traditionelle Methoden
 - Mathematische Betrachtung
 - Unterscheidungsvermögen
 - Zuverlässig / Reproduzierbar
- ROC Kurve
- Brier Score

- Für Anwendung in Medizin
 - erweitern durch Klinischen Nutzen
 - Qualität der Entscheidungen
- ROC Kurve erweitern mit NRI
- Entscheidungshilfen durch Net Benefit

Quellen

- Ewout W. Steyerberg, Andrew J. Vickers, Nancy R. Cook, Thomas Gerds, Mithat Gonen, Nancy Obuchowski, Michael J. Pencina, Michael W. Kattan: Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology*. 2010 January; 21(1): 128–138.
- Andrew J Vickers, Ben Van Calster, Ewout W Steyerberg: Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *Research Methods and Reporting*; December 2015
- Margaret S. Pepe, Holly Janes, Christopher I. Li: *Net Risk Reclassification P Values: Valid or Misleading?* Oxford University Press; January 2014
- Kristin Mühlenbruch: Ein Makro zur Berechnung von Diskriminanz- und Reklassifizierungsstatistiken für die Verbesserung eines Prädiktionsmodells bei Anwendung der Cox-Regression: http://saswiki.org/images/f/f6/16._KSFE_2012_-_M%C3%BChlenbruch_-_Ein_Makro_zur_Berechnung_von_Diskriminanz-_und_Reklassifizierungsstatistiken_f%C3%BCr_die_Verbesserung_eines_Pr%C3%A4di ktionsmodells.pdf ; online am 22.05.2017

Diagramm -Quellen

- Fig. 1, Fig. 2 Fig. 3: Ewout W. Steyerberg, Andrew J. Vickers, Nancy R. Cook, Thomas Gerds, Mithat Gonen, Nancy Obuchowski, Michael J. Pencina, Michael W. Kattan: Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology*. 2010 January; 21(1): 128–138.
- Fig. 4, Fig. 5: Andrew J Vickers, Ben Van Calster, Ewout W Steyerberg: Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *Research Methods and Reporting*; December 2015