

Ein Makro zur Berechnung von Diskriminanz- und Reklassifizierungsstatistiken für die Verbesserung eines Prädiktionsmodells bei Anwendung der Cox-Regression

Kristin Mühlenbruch
Deutsches Institut für Ernährungsforschung
Potsdam-Rehbrücke
Arthur-Scheunert-Allee 114-116
14558 Nuthetal
Kristin.Muehlenbruch@dife.de

Zusammenfassung

Die epidemiologische Forschung setzt sich mit der Identifizierung von Determinanten für bestimmte Erkrankungen, vor allem für sogenannte Zivilisationskrankheiten, auseinander. In diesem Zusammenhang hat sich ein Bereich entwickelt, der sich mit der Entwicklung von Risiko-prädiktionsmodellen auseinandersetzt [1]. Die Entwicklung solcher Prädiktionsmodelle ist aber oftmals nur der erste Schritt. Das Ziel ist das beste Vorhersagemodell zu finden und das initiale Prädiktionsmodell durch Hinzunahme von neuen Markern zu verbessern. Diese Verbesserung ist messbar. Zum einen dient die ROC (Receiver Operator Characteristic) Kurve und AUC (Area Under the Curve) zur Berechnung der diskriminativen Fähigkeiten des Prädiktionsmodells. Der Vergleich zweier korrelierter ROC-Kurven ergibt einen p -Wert für die Differenz [2]. Diese von DeLong et al. (1988) entwickelte Methode ist bereits in die SAS Prozedur *proc logistic* implementiert. Eine weitere Möglichkeit, die Verbesserung eines Modells zu berechnen, bieten sogenannte Reklassifizierungsstatistiken. Pencina et al. (2008) schlagen hier vor allem Net Reclassification Improvement (NRI) und Integrated Discrimination Improvement (IDI) vor [3]. Die Weiterentwicklung eines publizierten Makros verfolgte das Ziel, weiterhin *proc logistic* für die Berechnung und den Vergleich von zwei ROC-Kurven zu verwenden und die Berechnung der Reklassifizierungsstatistiken auf der Basis von Survivaldaten durchzuführen. Das Makro soll nun die Möglichkeit bieten, die genannten Berechnungen auf unterschiedliche prospektive Studiendaten anzuwenden.

Schlüsselwörter: Prädiktionsmodell, Diskriminierung, Reklassifizierung, Cox-Regression

1 Hintergrund

In der epidemiologischen Forschung hat sich gegenwärtig ein Bereich manifestiert, der sich mit der Vorhersage von Erkrankungsrisiken beschäftigt. Bereits untersuchte und belegte Zusammenhänge von Determinanten für die Entwicklung chronischer Erkrankungen sind Ausgangspunkt für eine solche Vorhersage. In diesem Zusammenhang werden identifizierte Risikomarker ausgewählt und als Risikofaktoren (unabhängige Variablen) in ein sogenanntes Prädiktionsmodell eingeschlossen. Die Risikofaktoren

können dabei sowohl protektiven, also Risiko senkenden, Einfluss haben als auch Risiko erhöhenden Einfluss. Auf der Basis solcher Modelle wurde im nächsten Schritt das Risiko berechnet, innerhalb der nächsten 5 oder 10 Jahre die Erkrankung von Interesse unter den gegebenen Risikofaktoren zu entwickeln. Vor allem für kardiovaskuläre Erkrankungen wurden Prädiktionsmodelle entwickelt, das wohl bekannteste ist der Framingham Risk Score [4], aber auch im Bereich Typ-2-Diabetes gibt es stetigen Zuwachs [5]. Das Ziel der Entwicklung solcher Prädiktionsmodelle oder Risiko-Scores ist es, vor allem Hochrisikogruppen zu identifizieren und basierend darauf gezielte Präventionsmaßnahmen zu implementieren. Besonders der Transfer in die Praxis mittels Fragebogenversion oder Online-Instrument ist für die Umsetzung des Ziels sehr wertvoll.

1.1 Weiterentwicklung von Prädiktionsmodellen

Der erste Schritt ist nun also die Entwicklung eines Prädiktionsmodells und Bestimmung der Performance, sowie interner oder externer Validität des entwickelten Modells. Hiermit ist eine Grundlage für die Risikovorhersage einer bestimmten Erkrankung unter gegebener Genauigkeit gelegt. Ein weiterführendes Ziel im Bereich Risikoprädiktion ist aber auch, diese Risikovorhersage weiter zu verbessern, welches durch Erweiterungen des Modells erreicht werden soll. In der Regel werden dabei Prädiktionsmodelle, die hauptsächlich aus nicht-invasiven Risikofaktoren bestehen um biochemische oder genetische Marker erweitert, die mit aufwendigeren Messungen und einem höherem finanziellem Aufwand verbunden sind.

Diese Erweiterung soll nun auch auf Relevanz und Signifikanz geprüft werden. Hierfür werden das initiale Prädiktionsmodell (Basismodell) und das erweiterte Modell verglichen. Das Zielkriterium für diesen Vergleich ist die Genauigkeit der berechneten Risiken durch beide Prädiktionsmodelle. Konnte eine Verbesserung der Risikovorhersage durch die Erweiterung des Prädiktionsmodells erreicht werden, ist es im weiteren Vorgehen sinnvoll, den neuen Risikofaktor in das Prädiktionsmodell aufzunehmen. Allerdings ist für die praktische Umsetzung immer auch eine Kosten-Nutzen-Analyse zur Beurteilung heranzuziehen.

1.2 Diskriminierung und Reklassifizierung durch Erweiterungen

Die Methoden, die herangezogen werden, um Erweiterungen hinsichtlich Relevanz und Signifikanz zu beurteilen, sind vor allem die Bestimmung der diskriminativen Fähigkeit mittels Receiver-operator-characteristic (ROC)-Analyse sowie Reklassifizierungsanalysen. Zielkriterium für erstere ist die Fläche unter der ROC-Kurve (ROC-AUC). Die ROC-AUC war über einen langen Zeitraum das Maß der Wahl und Alternativen wurden erst vor kurzer Zeit vorgeschlagen. Die ROC-AUC zu berechnen ist v.a. bei Anwendung der Logistischen Regression für das Prädiktionsmodell intuitiv.

Im Zuge der Verbreitung von Prädiktionsmodellen und der Überprüfung von Modellerweiterungen wurde bei der ROC-AUC ein fundamentaler Nachteil diskutiert: die Fähigkeit kleine Veränderungen zu identifizieren. Der Hintergrund hierzu ist, dass die meis-

ten Prädiktionsmodelle, die angewendet werden, meist bereits über eine gute Diskriminierung verfügen und somit auch kleine Veränderungen möglicherweise relevant sein können. Der Vergleich der ROC-AUCs vom Basismodell und dem erweiterten Modell mittels der von DeLong et al. (1988) entwickelten Methode für den Vergleich zweier korrelierter ROC-Kurven wird mit Hilfe einer Schätzung für die Differenz der ROC-AUCs und eines p-Werts für diese Differenz ausgedrückt [2]. Dieser Vergleich kann aber kleine Unterschiede nicht immer adäquat abbilden.

Vorgeschlagen wurde in diesem Zusammenhang die Reklassifizierungsanalyse erstmals von Nancy Cook [6], bei der eine Einteilung in vordefinierte Risikokategorien sowohl durch das Basismodell als auch durch das erweiterte Modell vorgenommen wird. Diese Kategorisierungen werden gegenübergestellt, um zu schauen, welche Anteile in dieselben Kategorien und v.a. in andere Kategorien eingeordnet wurden. Berechnet wird hieraus der Anteil an Reklassifizierungen (Wechsel der Risikokategorien).

$$\text{Reklassifizierungen (\%)} = \frac{n_{\text{höher kategorisiert}} + n_{\text{niedriger kategorisiert}}}{N}$$

Dieser Ansatz wurde 2008 von Pencina et al. [3] um den besonders wichtigen Gedanken erweitert, dass diese Wechsel in den Kategorien für Fälle und Nicht-Fälle (Kontrollen) unterschiedliche Bedeutungen haben. Während eine Einordnung in höhere Kategorien durch das erweiterte Modell für Fälle eine Modellverbesserung bedeutet, ist dies für Nicht-Fälle (Kontrollen) als Modellverschlechterung zu betrachten. Umgesetzt hat Pencina diesen Gedanken in der Entwicklung des *Net Reclassification Improvement* (NRI). Die folgende Tabelle stellt eine Reklassifizierungstabelle dar, die verdeutlicht, wie Wechsel in den Kategorien in Abhängigkeit vom Fall-Status (Krankheitsstatus) zu betrachten sind.

Tabelle 1: Beispiel für eine Reklassifizierungstabelle

Basismodell	Erweitertes Modell		
	niedriges Risiko	mittleres Risiko	hohes Risiko
niedriges Risiko	keine Veränderung	aufwärts Bewegung Verbesserung (D=1) Verschlechterung (D=0)	aufwärts Bewegung Verbesserung (D=1) Verschlechterung (D=0)
mittleres Risiko	abwärts Bewegung Verbesserung (D=1) Verschlechterung (D=0)	keine Veränderung	aufwärts Bewegung Verbesserung (D=1) Verschlechterung (D=0)
hohes Risiko	abwärts Bewegung Verbesserung (D=0) Verschlechterung (D=1)	abwärts Bewegung Verbesserung (D=0) Verschlechterung (D=1)	keine Veränderung

Basierend auf diesen Auf- und Abwärtsbewegungen für Fälle (D=1) und Nicht-Fälle (D=0) kann mit folgender Formel der NRI berechnet werden.

$$\begin{aligned} NRI &= P(\text{aufwärts} | \text{Fälle}) - P(\text{abwärts} | \text{Fälle}) + P(\text{abwärts} | \text{Nicht-Fälle}) - P(\text{aufwärts} | \text{Nicht-Fälle}) \\ &= NRI_{\text{Fälle}} + NRI_{\text{Nicht-Fälle}} \end{aligned}$$

Der NRI berechnet sich demnach aus den Nettoverbesserungen der Risikovorhersage bei den Fällen (auch $NRI_{\text{Fälle}}$) sowie bei den Nicht-Fällen (Kontrollen; auch $NRI_{\text{Nicht-Fälle}}$).

Ob der Wert des NRI für eine statistisch signifikante Verbesserung steht, lässt sich anhand des p-Werts eines Signifikanztests ablesen. Es wird die Nullhypothese $H_0: NRI=0$ angenommen und der Test basiert auf folgender Teststatistik.

$$Z_{NRI} = \frac{\widehat{NRI}}{\sqrt{\frac{\hat{p}_{\text{aufwärts, Fälle}} + \hat{p}_{\text{abwärts, Fälle}}}{\# \text{ Fälle}} + \frac{\hat{p}_{\text{aufwärts, Nicht-Fälle}} + \hat{p}_{\text{abwärts, Nicht-Fälle}}}{\# \text{ Nicht-Fälle}}}}$$

Ein weiteres Maß, welches Pencina et al. vorgeschlagen haben, ist der *Integrated Discrimination Improvement (IDI)*. Die Berechnung des *IDI* basiert im Gegensatz zum NRI nicht auf den individuellen Risiken und Einordnung in vorgegebene Risikokategorien, sondern betrachtet die mittleren geschätzten Risiken für jedes Modell (alt bezeichnet hier das Basismodell, neu bezeichnet das erweiterte Modell) nach Fall-Status.

$$\widehat{IDI} = (\bar{\hat{p}}_{\text{neu, Fälle}} - \bar{\hat{p}}_{\text{alt, Fälle}}) - (\bar{\hat{p}}_{\text{neu, Nicht-Fälle}} - \bar{\hat{p}}_{\text{alt, Nicht-Fälle}})$$

Auch für den IDI wird ein Signifikanztest durchgeführt, der auf folgender Teststatistik basiert.

$$Z_{IDI} = \frac{\widehat{IDI}}{\sqrt{(\hat{se}_{\text{Fälle}})^2 + (\hat{se}_{\text{Nicht-Fälle}})^2}}$$

2 Entwicklung des Makros %reclassification_phreg

Für die Entwicklung des vorgestellten Makros wurde das Makro %nriidi verwendet, welches von Lars Berglund entwickelt und 2011 von Sundstrom et al. publiziert wurde. Mit diesem Makro konnte basierend auf der logistischen Regression der *NRI* und der *IDI* berechnet werden.

Im folgenden Abschnitt soll zunächst dieses Makro beschrieben werden, um zu zeigen, inwiefern das neue Makro Ergänzungen zum ursprünglichen aufweist.

2.1 Das Makro %nriidi

Mit dem Makro %nriidi ist es möglich, Modellerweiterungen mit Hilfe von Reklassifizierungsmaßen wie dem NRI und IDI zu beurteilen. Die Entwicklung des Makros basiert auf der Publikation von Pencina et al. (2008) und verwendet die dort dargestellten Formeln zur Berechnung dieser Maße sowie *p*-Werte für die Signifikanztests.

Die folgende Tabelle zeigt die Argumente des ursprünglichen Makros *%nriidi*.

Tabelle 1: Argumente des Makros *%nriidi*

Argument	Beschreibung
ds	Input Datensatz für die Analyse (Analysedatensatz)
y	Binäre Variable für den Fall-Status (Erkrankung) (0=Nicht-Fälle/Kontrollen, 1=Fälle)
id	Identifikationsnummer für jeden Probanden
modell1n	Stetige unabhängige Variablen /Risikofaktoren im ursprünglichen Modell (Modell 1); Liste getrennt durch Leerzeichen
modell1c	Kategorielle unabhängige Variablen /Risikofaktoren im ursprünglichen Modell (Modell 1); Liste getrennt durch Leerzeichen
modell2n	Stetige unabhängige Variablen /Risikofaktoren im erweiterten Modell (Modell 2); Liste getrennt durch Leerzeichen
modell2c	Kategorielle unabhängige Variablen /Risikofaktoren im erweiterten Modell (Modell 2); Liste getrennt durch Leerzeichen
nriskcat	Anzahl der zu verwendenden Risikokategorien
riskcutoffs	Cut-off-Werte für die Bildung der Risikokategorien Liste mit Werten in Prozent, die im Makro noch durch 100 dividiert werden; Bsp.: 1 3 5 10
printtable	Option für die Ausgabe der Reklassifizierungstabellen (Y=Yes/N=No)

Mit dem Aufruf des Makros definiert man zunächst den Analysedatensatz, die Outcomevariable mit Fall-Status und die ID-Variable für eine eindeutige Identifikation der Teilnehmer. Die Argumente *modell1n* und *modell1c* bilden zusammen betrachtet das Basismodell zur Berechnung der Risiken, es sind lediglich stetige und kategorielle Variablen getrennt definiert worden. Gleichmaßen wird das erweiterte Modell mit den Argumenten *modell2n* und *modell2c* definiert. Außerdem ermöglicht das Makro eine freie Wahl der Anzahl zu verwendender Risikokategorien, sowie der Cut-off-Werte für diese Kategorien, die in ganzen Prozentwerten der berechneten Risiken anzugeben sind. Abschließend bietet das Makro die Möglichkeit zwischen Ausgabe einer Reklassifizierungstabelle oder Unterdrückung dieser zu wählen.

Zusammengefasst hat das Makro *%nriidi* 10 Argumente, erzeugt 9 Datensätze (Tabelle 2), wobei der Datensatz *nriidi* - der Ergebnisdatensatz - wohl der wichtigste ist. Als SAS-Output wird nur die Reklassifizierungstabelle für Fälle und Nicht-Fälle (Kontrollen) erzeugt, wenn es so angegeben wurde, und der Ergebnisdatensatz gedruckt (*proc print*).

Tabelle 2: Output-Datensätze des Makros %nriidi

Datensatz	Beschreibung
OUT1	Ursprünglicher Datensatz (Analysedatansatz definiert mit Argument ds) plus <i>predicted risks</i> aus <i>proc logistic</i> für Modell1
OUT2	Ursprünglicher Datensatz (Analysedatansatz definiert mit Argument ds) plus <i>predicted risks</i> aus <i>proc logistic</i> für Modell2
OUT	Gemergte Datensätze OUT1 und OUT2 mit <i>predicted risks</i> aus beiden Modellen; Bildung von Risikokategorien mit beiden <i>predicted risks</i> unter Vorgabe von Grenzen (Cut-offs, Argument riskcutoffs=) für die Kategorien (Division des Wertes durch 100) Berechnung der absoluten Differenz zwischen den <i>predicted risks</i> beider Modelle (pred2-pred1)
TEMP2	Datensatz, der Fall-Status und Risikokategorien von Modell 1 und Modell 2 enthält
OUTFR	Mit <i>proc freq</i> ermittelte Häufigkeiten aller Kombinationen vom Risikokategorien für Modell 1 und Modell 2
OUTFR2	Basierend auf OUTFR und TEMP2 werden die Häufigkeiten (absolute und relative) der auf- und abwärts Bewegungen in den Risikokategorien von Modell 1 zu Modell 2 nach Fall-Status berechnet Berechnung der Zielgröße <i>NRI</i> mit SE, z-score, <i>p</i> -Wert

Im Ergebnisdatensatz sind entsprechend Tabelle 3 die Zielparameter mit Standardfehler und *p*-Wert für den Signifikanztest enthalten.

Tabelle 3: Variablen des Ergebnisdatensatzes NRIIDI erzeugt mit dem Makro `%nriidi`

Variablen	Beschreibung
n	Stichprobengröße
nri	Berechneter NRI
se_nri	Standard-Fehler des NRI
p_nri	p -Wert des Hypothesentests für den NRI , H_0 : NRI=0
idi	IDI
se_idi	Standard-Fehler des IDI ¹
p_idi	p -Wert des Hypothesentests für den IDI , H_0 : IDI=0

Der Aufruf bzw. die Zuweisung des Makros wird folgendermaßen durchgeführt (dargestellt ist hier ein Beispielaufruf).

```
%nriidi(ds =gdrs_caseco_42,
        y =fall,
        id =ident,
        modelln = hbalc glucose trigly hdl ggt gpt,
        modellc = sex rauchen,
        model2n = hbalc glucose trigly hdl ggt gpt adiponectin,
        model2c = sex rauchen,
        nriskcat = 5,
        riskcutoffs = 1 2 3 5,
        printtable = Y);
```

2.2 Das Makro `%reclassification_phreg`

Basierend auf dem Makro `%nriidi` wurde das Makro `%reclassification_phreg` entwickelt. Während einige Stellen vom benannten Makro direkt übernommen wurden, ist an anderen Stellen eine neue Programmierung vorgenommen worden. Der entscheidende Unterschied ist die Anwendung der Survivalanalyse mit Cox-Regression. Die logistische Regression wurde demnach nicht mehr für die Berechnung der Risiken verwendet, sondern lediglich für die Berechnung der ROC-AUCs. Außerdem wurde die nun in *proc logistic* eingearbeitete Methode von DeLong et al. (1988) zum Vergleich von ROC-Kurven ergänzt.

¹ Für die Berechnung des Standardfehlers des IDI wird der Standardfehler der mittleren Differenz zwischen den berechneten Risiken von Modell 1 zu Modell2 für Fälle sowie Nicht-Fälle mit *proc means* ermittelt.

Die Argumente des Makros sind in Tabelle 4 zusammengefasst und beschrieben.

Tabelle 4: Argumente des Makros %reclassification_phreg

Argumente	Beschreibung
dataset	Der zu verwendene SAS-Datensatz
id	Identifikationsnummer für jeden Probanden
timevar	Variable, die die Zeitangabe für die Cox-Regression enthält (z.B. Follow-up-Zeit)
outcome	Endpunkt/Zielvariable/Erkrankung/abhängige Variable im logistischen und Cox-Modell
modell1	Unabhängige Variablen / Risikofaktoren im ursprünglichen Modell (Modell 1); Liste getrennt durch Leerzeichen
nvar1	Anzahl an Variablen im ursprünglichen Modell (Modell 1)
roc_title1	Beschriftung der ROC-Kurve für Modell 1 in der Legende der Grafik
modell2	Unabhängige Variablen / Risikofaktoren im erweiterten Modell (Modell 2); Liste getrennt durch Leerzeichen
nvar2	Anzahl an Variablen im erweiterten Modell (Modell 2)
roc_title2	Beschriftung der ROC-Kurve für Modell 2 in der Legende der Grafik
nriskcat	Anzahl der zu verwendenden Risikokategorien
riskcutoffs	Cut-off-Werte für die Bildung der Risikokategorien, z.B. (0.032)(0.054)(0.087)
risk_year	Zu berechnendes Risiko (z.B. 5-Jahresrisiko)
base_muster	Baseline-Kovariatenmuster
printtable	Option für die Ausgabe der Reklassifizierungstabellen (Y=Yes/N=No)

Es wird wie im ursprünglichen Makro zunächst der Analysedatensatz, die ID-Variable und die Outcome-Variable definiert. Zusätzlich ist aber auch zu Beginn die Zeit-Variable definiert, die für die Cox-Regression notwendig ist. Das Basismodell wird gänzlich mit dem Argument Modell1 definiert, ohne zwischen stetigen und kategoriellen Variablen zu unterscheiden. Für das Basismodell wird zusätzlich die Anzahl an unabhängigen Variablen (Risikofaktoren) angegeben und ein Titel für die ROC-Kurve vergeben. Gleichermaßen werden diese Parameter für das erweiterte, also Modell 2 definiert. Für die Risikoberechnung kann wie bei %nriidi die Anzahl an Risikokategorien frei gewählt werden und die Wahl der Cut-off-Werte kann exakter erfolgen, d.h. hier werden absolute Werte mit beliebiger Anzahl an Dezimalstellen angegeben. Entscheidend ist hierbei das Trennen der einzelnen Werte, welches durch Einklammern eindeutig vorgegeben ist. Außerdem kann für die Berechnung der Risiken ein bestimmter Zeitraum definiert werden, also ein 5-Jahresrisiko oder 10-Jahresrisiko, sowie das Baseline-Kovariatenmuster für die Berechnung der Baselinesurvivorfunktion. Abschließend wurde die Option für die Wahl der Ausgabe der Reklassifizierungstabellen übernommen.

Das Makro basiert auf der Zuweisung von 15 Argumenten, zwei verschiedenen Regressionsmodellen und damit einhergehend auch zwei Methoden der Evaluation von Modellerweiterungen und gibt ergänzend auch graphische Darstellungen der ROC-Analysen aus. Es werden insgesamt 26 SAS-Datensätze erzeugt, die in Tabelle 5 ausführlich dargestellt sind.

Tabelle 5: Output-Datensätze des Makros %reclassification_phreg

Datensatz	Beschreibung
ROCDIFF	Kontrastschätzer (Differenz) der ROC-AUC von Modell1 und Modell2 mit SE, unterer und oberer Konfidenzgrenze, Chi-Quadrat-Statistik und p-Wert der Chi-Quadrat-Statistik (DeLong Test)
ROCASS	ROC-AUCs für die beiden Modelle mit SE, untere und oberer Konfidenzgrenzen
OUT1_1	Parameterschätzer der Cox-Regression für Modell1
OUT1_2	Transponierter Datensatz OUT1_1 zur weiteren Verwendung der Parameterschätzungen (in Spalten)
INRISKS1	Enthält das Baseline-Kovariatmuster (Argument base_muster) für Modell 1
RISKSCORE1	Mit Cox-Regression berechnete Survivorfunktion für Modell1 mit unterer und oberer Konfidenzgrenze für steigende Zeitvariable (Argument timevar=) unter Verwendung des Baseline-Kovariatmusters
RISKSCORE1_1	Beschränkung des RISKSCORE1-Datensatzes auf eine vorgegebene Zeit (Argument timevar) Bestimmung des absoluten Baseline-Risikos
OUT_1	Aneinanderfügen des Input-Datensatzes (Analysedatensatzes) mit OUT1_2 Jede Beobachtung (Zeile) enthält die Parameterschätzer der Cox-Regression für Modell1
PRED_1	Berechnung der absoluten Risiken für jede Beobachtung unter Verwendung - der Parameterschätzer für die Variablen in Modell1 multipliziert mit der Ausprägung der Variable (OUT_1) (>>Exponent der Baseline-Survivorfunktion) - des absoluten Baseline-Risikos (Baseline-Survivorfunktion)
OUT2_1	Parameterschätzer der Cox-Regression für Modell2
OUT2_2	Transponierter Datensatz OUT2_1 zur weiteren Verwendung der Parameterschätzungen (in Spalten)
INRISKS2	Enthält das Baseline-Kovariatenmuster (Argument base_muster) für Modell 2
RISKSCORE2	Mit Cox-Regression berechnete Survivorfunktion für Modell2 mit unterer und oberer Konfidenzgrenze für steigende Zeitvariable (Argument timevar) unter Verwendung des Baseline-Kovariatmusters
RISKSCORE2_1	Beschränkung des RISKSCORE2-Datensatzes auf eine vorgegebene Zeit (Argument timevar) Bestimmung des absoluten Baseline-Risikos

OUT_2	Aneinanderfügen des Input-Datensatzes (Analysedatensatzes) mit OUT2_2 Jede Beobachtung (Zeile) enthält die Parameterschätzer der Cox-Regression für Modell2
PRED_2	Berechnung der absoluten Risiken für jede Beobachtung unter Verwendung - der Parameterschätzer für die Variablen in Modell2 multipliziert mit der Ausprägung der Variable (OUT_2) (>>Exponent der Baseline-Survivorfunktion) - des absoluten Baseline-Risikos (Baseline-Survivorfunktion)
OUT	Gemergter Datensatz aus PRED_1 und PRED_1 (Ausschluss der Spalten mit Parameterschätzern)
TEMP2	Datensatz, der Fall-Status und Risikokategorien von Modell 1 und Modell 2 enthält
TEMP3	Erzeugt aus OUT, enthält binäre (0,1)Variablen (up, down), die nach pos./neg. Differenz der Risiken beider Modelle eine Aufwärts- bzw. -Abwärtsbewegung beim absoluten Risiko bedeuten
OUTFR_3	Erzeugt aus <i>proc freq</i> von TEMP3, enthält Häufigkeiten der Aufwärts- bzw. -Abwärtsbewegung beim absoluten Risiko
OUTFR_4	Transponierter Datensatz OUTFR_3 mit n für Aufwärts- bzw. -Abwärtsbewegung beim absoluten Risiko nach Fall-Status
OUTFR	Erzeugt aus <i>proc freq</i> von OUT zur Ermittlung der Häufigkeiten der Risikokategorien aus Modell1 kreuzklassifiziert mit Risikokategorien aus Modell2 (Basis für die ausgegebene Reklassifizierungstabelle)
OUTFR2	Basierend auf OUTFR und OUTFR4 werden die Häufigkeiten (absolute und relative) der auf- und abwärts Bewegungen in den Risikokategorien von Modell 1 zu Modell 2 nach Fall-Status berechnet Berechnung der Zielgröße <i>NRI</i> mit SE, <i>z</i> -score, <i>p</i> -Wert Zusätzliche Berechnung <i>NRI_{Fälle}</i> , <i>NRI_{Nicht-Fälle}</i> , stetiger <i>NRI</i>
IDI	Erzeugt aus <i>proc means</i> für OUT Berechnung der mittleren <i>predicted risks</i> für beide Modelle und die Differenz, jeweils nach Fall-Status
IDI2	Berechnung der Zielgröße <i>IDI</i> mit SE, <i>z</i> -score, <i>p</i> -Wert Zusätzliche Berechnung des relativen <i>IDI</i>
NRIIDI	Gemergter Datensatz aus OUTFR2 und IDI2 Enthält berechnete Zielgrößen (Ergebnisdatsatz)

Der Ergebnisdatsatz NRIIDI beinhaltet *NRI* und *IDI* jeweils mit Standardfehler und *p*-Wert für den Signifikanztest und ergänzend den stetigen *NRI*, der nicht auf Risikokategorien beruht, sondern auf Veränderungen der absoluten Risiken durch die Modellerweiterungen, sowie die Einzelkomponenten des *NRI* für Fälle und Nicht-Fälle.

Tabelle 6: Variablen des Ergebnisdatsatzes NRIIDI erzeugt mit dem Makro *%reclassification_phreg*

Variablen	Beschreibung
n	Stichprobengröße
nri	Berechneter NRI
nri_case	Berechneter NRI für Fälle
nri_noncase	Berechneter NRI für Nicht-Fälle
nri_c	Berechneter stetiger NRI (ohne Kategorien)
se_nri	Standard-Fehler des NRI
p_nri	p -Wert des Hypothesentests für den NRI , H_0 : NRI=0
idi	Berechneter IDI
rel_idi	Berechneter relativer IDI
se_idi	Standard-Fehler des IDI
p_idi	p -Wert des Hypothesentests für den IDI , H_0 : IDI=0

Die SAS-Outputs sind ebenfalls umfangreicher als beim ursprünglichen Makro %nriidi. Es werden die Ergebnisse der logistischen Regression mit ROC-Kurvenvergleich (DeLong-Test) und Plots der ROC-Kurven, Ergebnisse der Cox-Regression mit Parameterschätzern, sowie Reklassifizierungstabellen (optional) ausgegeben. Der gedruckte Ergebnisdatsatz ist nur um wenige Parameter erweitert worden.

Der Aufruf bzw. die Zuweisung des Makros wird folgendermaßen durchgeführt (dargestellt ist hier ein Beispielaufruf).

```
%reclassification_phreg( dataset=gdrs,
                        id=ident, outcome=fall,
                        timevar=fup_zeit,
                        modell=alter_basis groesse taille hypertonie ,
                        nvar1=4,
                        roc_title1='Model 2',
                        model2=alter_basis groesse taille hypertonie
                        alkohol_3 aktiv_2 rauch_2 rauch_4 wgrainbread
                        coffee redmeat,
                        nvar2=11,
                        roc_title2='Model 2 plus Modifiable Risk
                        Factors',
                        nriskcat=5,
                        riskcutoffs=(0.0088)(0.0237)(0.0630)(0.1621),
                        risk_year=5,
                        base_muster=0 0 0 0 0 0 0 0 0 0 0 0 ,
                        printtable=Y);
```

Zusammenfassend berücksichtigt das Makro %reclassification_phreg die implementierte Methode des ROC-Kurvenvergleichs von DeLong bei nur einem Aufruf der Prozedur *proc logistic*, sowie die Anwendung der Risikoberechnung basierend auf unterschiedlichen prospektiven Studiendaten. Es ist möglich die absoluten Risiken für unterschiedlich umfangreiche Modelle zu berechnen, da der lineare Prädiktor aus den Parameterschätzern und den unabhängigen Variablen für jedes Modell neu definiert wird.

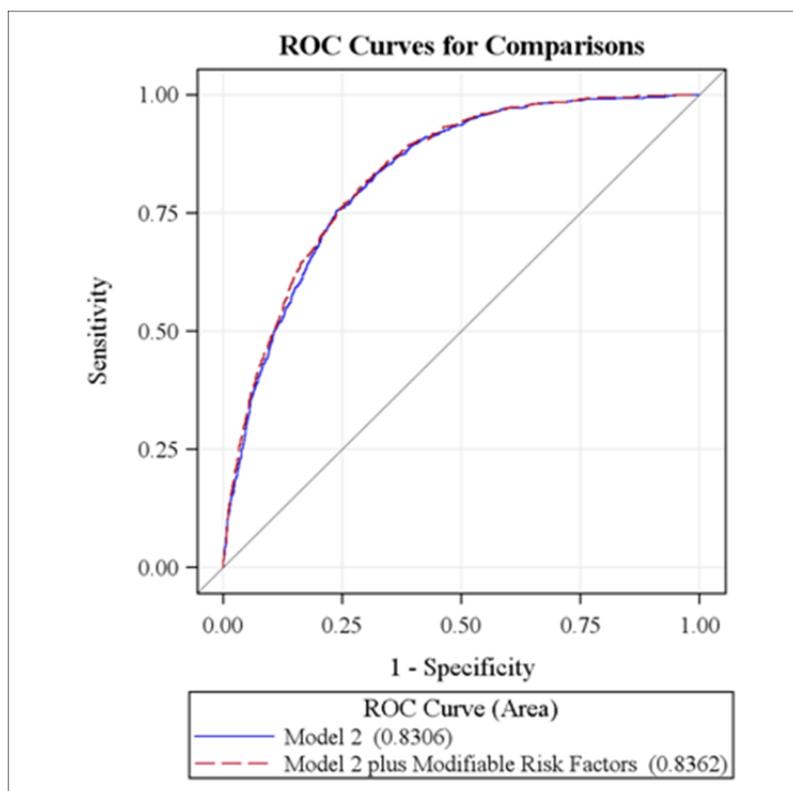
Außerdem ist die Angabe der Cut-off-Werte präziser möglich, da die absoluten Werte, die sich manchmal so aus den Daten ergeben, angegeben werden können.

3 Anwendungsbeispiel für %reclassification_phreg

Zur Verdeutlichung der Funktion dieses Makros soll eine Beispielanalyse gezeigt werden.

Das Ziel der Analyse soll sein, den prädiktiven Wert für die Vorhersage des Typ-2-Diabetes innerhalb der nächsten 5 Jahre von modifizierbaren Risikofaktoren zusätzlich zu einem Basismodell bestehend aus Angaben zum Alter, der Körpergröße, dem Hypertonie-Status sowie Taillenumfang zu bestimmen. Es wird also ein Basismodell mit 4 Variablen mit einem erweiterten Modell mit 11 Variablen verglichen (s. Zuweisung des Makros auf S. 11). Es werden 5 Risikokategorien verwendet und die Cut-offs sind definiert als 0.0088, 0.0237, 0.0630, 0.1621.

Die folgenden Seiten zeigen eine Auswahl der gesamten SAS-Ausgabe.



ROC Association Statistics							
ROC Model	Mann-Whitney				Somers' D (Gini)	Gamma	Tau-a
	Area	Standard Error	95% Wald Confidence Limits				
Model 2	0.8306	0.00615	0.8185	0.8426	0.6611	0.6611	0.0431
Model 2 plus Modifiable Risk Factors	0.8362	0.00605	0.8243	0.8480	0.6723	0.6723	0.0438

ROC Contrast Estimation and Testing Results by Row						
Contrast	Estimate	Standard Error	95% Wald Confidence Limits		Chi-Square	Pr > ChiSq
Model 2 plus Modifiable Risk Factors - Model 2	0.00560	0.001820	0.002030	0.00917	9.4719	0.0021

Cases(1)/Controls(0)= case	Established risk factors + new risk factors					Total N
	1:<0.0088	2:0.0088-0.0237	3:0.0237-0.063	4:0.063-0.1621	5:>=0.1621	
	N	N	N	N	N	
Established risk factors						
1:<0.0088	25	6	.	.	.	31
2:0.0088-0.0237	8	124	25	.	.	157
3:0.0237-0.063	.	26	258	38	.	322
4:0.063-0.1621	.	.	28	203	21	252
5:>=0.1621	.	.	.	16	71	87
Total	33	156	311	257	92	849

Cases(1)/Controls(0)= non-case	Established risk factors + new risk factors					Total N
	1:<0.0088	2:0.0088-0.0237	3:0.0237-0.063	4:0.063-0.1621	5:>=0.1621	
	N	N	N	N	N	
Established risk factors						
1:<0.0088	9723	596	1	.	.	10320
2:0.0088-0.0237	924	5943	568	.	.	7435
3:0.0237-0.063	.	718	3761	275	.	4754
4:0.063-0.1621	.	.	321	1180	61	1562
5:>=0.1621	.	.	.	44	203	247
Total	10647	7257	4651	1499	264	24318

n	nri	nri_casen	nri_noncase	nri_c	se_nri	p_nri
25167	0.034942	0.014134	0.020808	0.19216	0.015460	0.023810

idi	rel_idi	se_idi	p_idi
.003815448	0.073674	.000959408	.000069826

Literatur

- [1] Tzoulaki I, Liberopoulos G, Ioannidis JP. Use of reclassification for assessment of improved prediction: an empirical evaluation. *Int J Epidemiol* 2011;40(4):1094-105.
- [2] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44(3):837-45.
- [3] Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., et al. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27(2):157-72; discussion 207-12.
- [4] Wilson PW, D'Agostino RB, Levy D, et al. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;97(18):1837-47.
- [5] Buijsse B, Simmons RK, Griffin SJ, et al. Risk assessment tools for identifying individuals at risk of developing type 2 diabetes. *Epidemiol Rev* 2011;33(1):46-62.
- [6] Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 2007;115(7):928-35.