

Untersuchung der Modellgüte einer Cox-Regression: Neue Möglichkeiten in PROC PHREG

Sandra Müller, Benjamin Mayer, Rainer Muche
Institut für Epidemiologie und Medizinische Biometrie
Universität Ulm
Schwabstr. 13
89075 Ulm
sandra.mueller@uni-ulm.de

Zusammenfassung

Zur Analyse von Überlebenszeitdaten mittels Cox-Regression steht in SAS die Prozedur PROC PHREG zur Verfügung. Diese wurde erst vor kurzem um wichtige Funktionen ergänzt. Seit SAS 9.4 M4 können für rechtszensierte Daten Konkordanz-Statistiken, ROC-Kurven und eine zeitabhängige AUC berechnet werden. In diesem Beitrag werden speziell diese neuen Maße zur Beurteilung der Prognosegüte eines Cox-Modelles kurz vorgestellt und ihre Verwendung veranschaulicht.

Die prädiktive Genauigkeit eines Modells lässt sich durch die Übereinstimmung von beobachtetem und prädiziertem Outcome ausdrücken. Bei einer logistischen Regression ist die Konkordanz-Statistik (C-Statistik) ein gängiges Maß hierfür. Ihr liegt der Ansatz zugrunde, dass Personen mit einem bestimmten Outcome eine größere prädizierte Wahrscheinlichkeit für dieses Outcome haben als Personen, bei denen dieses Outcome nicht beobachtet wurde. In diesem Fall spricht man bei den beiden Personen von einem konkordanten Paar, andernfalls von einem diskordanten Paar. Die C-Statistik entspricht somit dem Anteil an konkordanten Paaren an allen Paaren, von denen bei genau einer Person das Outcome beobachtet wurde. Harrell und Uno definieren jeweils eine Verallgemeinerung dieses Konkordanzmaßes für rechts-zensierte Überlebenszeitdaten. Beide Maße sind jetzt in SAS über die Option CONCORDANCE im PROC PHREG Statement verfügbar. Analog zur logistischen Regression kann auch für rechtszensierte Daten für jeden Zeitpunkt eine ROC-Kurve erzeugt werden. In SAS ist dies nun über die Optionen PLOTS=ROC und ROCOPTIONS im PROC PHREG Statement sowie über das ROC Statement möglich. Über alle möglichen Zeitpunkte betrachtet lässt sich hieraus auch die zeitabhängige AUC, die Fläche unterhalb der ROC-Kurve, berechnen. Diese kann analog über PLOTS=AUC angefordert werden. SAS stellt hier vier verschiedene Möglichkeiten der Berechnung zur Verfügung.

In der Literatur existieren noch weitere Statistiken zur Beurteilung der Modellgüte von Cox-Modellen und andere Definitionen von zeitabhängigen ROC-Kurven, die bisher noch nicht in SAS enthalten sind. Die nun hinzugefügten C-Statistiken nach Harrell und Uno sowie zeitabhängige ROC-Kurven und die AUC sind jedoch eine sinnvolle Erweiterung von PROC PHREG, die hoffentlich in Zukunft noch weiter ergänzt wird.

Schlüsselwörter: Prognosegüte, Survival Analyse, Cox-Regression, PROC PHREG, C-Statistik, AUC, ROC

1 Einleitung

Zur Analyse von Überlebenszeitdaten mittels Cox-Regression steht in SAS die Prozedur PROC PHREG zur Verfügung. Diese wurde erst vor kurzem um wichtige Funktionen ergänzt. Seit SAS 9.4 M4 können für rechtszensierte Daten Konkordanz-Statistiken, ROC-Kurven und eine zeitabhängige AUC berechnet werden. In SAS 9.4 M5 wurden zudem cause-specific Competing Risks ergänzt. In diesem Beitrag sollen speziell die in SAS neuen Maße zur Beurteilung der Prognosegüte eines Cox-Modells kurz vorgestellt und ihre Verwendung anhand eines realen Studienbeispiels veranschaulicht werden.

2 C-Statistiken nach Harrell und Uno

Bei einer logistischen Regression ist die c-Statistik ein wichtiges Maß der Prognosegüte. Ist D_i der Eventindikator und π_i die prädizierte Wahrscheinlichkeit für das Eintreten eines Events für Beobachtung i , so gilt:

$$c = P(\pi_i > \pi_j \mid D_i = 1, D_j = 0)$$

Die c-Statistik ist somit ein Maß der Übereinstimmung der Wahrscheinlichkeitsprognosen und des tatsächlichen Eintretens eines Events. Sie kann als Anteil der konkordanten Paare an allen möglichen Paaren innerhalb des Datensatzes geschätzt werden.

Aufgrund der auftretenden Zensierung ist eine direkte Erweiterung auf Überlebenszeitdaten nicht möglich. Es gibt jedoch verschiedene Ansätze hierfür, unter anderem von Harrell und Uno.

Harrell [2] unterscheidet zwischen brauchbaren und unbrauchbaren Paaren. Unbrauchbar sind für ihn alle Paare ohne Event, mit gleicher Eventzeit ohne Zensierung oder mit einem Event und einer Zensierung mit kürzerem Follow-up. In diesen Fällen ist aufgrund der Zensierung keine Unterscheidung in konkordant oder diskordant möglich. Bezeichnet Z_i den Kovariatenvektor, T_i die Eventzeit und C_i die Zensierungszeit von Person i , so gilt:

$$c_{Harrell} = P(\beta'Z_i > \beta'Z_j \mid T_i < T_j, T_i < \min(C_i, C_j))$$

Harrells c-Statistik ist daher abhängig von der zugrundeliegenden Zensierungsverteilung.

Uno [3] definiert seine c-Statistik für einen fest gewählten Zeitpunkt τ mit $P(C > \tau) > 0$ durch:

$$c_{Uno} = P(\beta'Z_i > \beta'Z_j \mid T_i < T_j, T_i < \tau)$$

Er schätzt diese mittels Gewichtung mit der inversen Zensierungswahrscheinlichkeit und erhält damit ein von der Zensierungsverteilung unabhängiges Maß.

Beide c-Statistiken sind jetzt über die Option CONCORDANCE im PROC PHREG-Statement verfügbar. Die Standardfehler werden über die Angabe SE ausgegeben. Für

Unos c wird dieser durch „Perturbation Resampling“ geschätzt, weshalb hier auch ein Start-wert (SEED), die Anzahl der Stichproben (ITER) und das verwendete Signifikanzniveau (ALPHA) spezifiziert werden können. [1]

```
PROC PHREG DATA=data CONCORDANCE=[HARRELL|UNO] (SE);
CLASS is_male;
MODEL time*event(0)= is_male eq_vas_t1 bvmt_trial1_t1 fkv_f1_t1;
RUN;
```

Abbildung 1: C-Statistiken mit der CONCORDANCE Option in PROC PHREG

Harrell's Concordance Statistic						
Quelle	Schätzung	Standard Fehler	Vergleichbare Paare			
			Konkordanz	Diskordanz	Tied in Prädiktor	Tied in Zeit
Model	0.5755	0.0186	70773	52136	537	83

Abbildung 2: Harrells c im Output von PROC PHREG

Uno's Concordance Statistic		
Quelle	Schätzung	Standard Fehler
Model	0.5820	0.0298

Abbildung 3: C-Statistik nach Uno im Output von PROC PHREG

SAS bietet außerdem die Möglichkeit, die paarweisen Differenzen in Unos c von Submodellen zu berechnen. Die gewählten Modelle werden jeweils in einem ROC-Statement angegeben. Durch DIFF in der Option CONCORDANCE werden die Differenzen angefordert. Zusätzlich können auch Werte für SEED, ITER und ALPHA angegeben werden. Neben dem Schätzer der Differenz von Unos c werden bei der Angabe von SE zudem der Standardfehler und eine approximativ normalverteilte Teststatistik mit zugehörigem p-Wert für den Test auf keine Differenz angezeigt. [1]

```
PROC PHREG DATA=data CONCORDANCE=UNO(SE SEED=123 ITER=100 ALPHA=0.05 DIFF);
CLASS is_male;
MODEL time*event(0)= is_male eq_vas_t1 bvmt_trial1_t1 fkv_f1_t1;
ROC 'ohne fkv_f1_t1' is_male eq_vas_t1 bvmt_trial1_t1;
RUN;
```

Abbildung 4: Differenz in Unos c mit der Option DIFF

Differences in Uno's Concordance Statistic					
Quelle	_Quelle	Schätzung	Standard Fehler	Chi-Quadrat	Pr > ChiSq
ohne fkv_f1_t1	Model	-0.0085	0.0246	0.12	0.7308

Abbildung 5: Differenz in Unos c im Output von PROC PHREG

3 Zeitabhängige ROC-Kurven und AUC

Ähnlich wie bei der logistischen Regression kann für jeden Zeitpunkt t innerhalb der maximalen Beobachtungszeit eine **zeitabhängige ROC-Kurve** generiert werden. Hierzu müssen die Sensitivität und Spezifität des Modells in Abhängigkeit eines Cutpoints c geschätzt werden. In PROC PHREG stehen vier Methoden zur Verfügung:

- Das Nächste-Nachbarn-Verfahren (NNE) nach Akritas [5] ist der Default und erzeugt eine geglättete Kurve mittels einer Nächste-Nachbarn-Kernfunktion. Der Anteil der nächsten Nachbarn, die in die Berechnung einfließen sollen, kann hierbei über einen Parameter bestimmt werden.
- Chambless und Diao [6] schlagen eine rekursive Berechnung (RECURSIVE) ähnlich dem Kaplan-Meier-Schätzer vor. Sie ist besonders für große Datensätze geeignet.
- Die Bedingte-Kaplan-Meier-Methode (KM) nach Heagerty et al. [7] beruht auf den Schätzern der bedingten Überlebensfunktionen $S(\cdot | \beta'Z > c)$ und $S(\cdot | \beta'Z \leq c)$.
- Das „Inverse Probability of Censoring Weighting“ (IPCW) nach Uno et al. [8] verwendet analog zu Unos c eine Gewichtung mittels der inversen Zensierungswahrscheinlichkeit.

Zeitabhängige ROC-Kurven können über PLOTS=ROC im PROC PHREG-Statement angefordert werden. Innerhalb der ROCOPTIONS werden die zu verwendende Methode und die gewählten Zeitpunkte angegeben. Durch die Option AUC wird zusätzlich die Fläche unter der Kurve ausgegeben. [4]

```
PROC PHREG DATA=data PLOTS=ROC
  ROCOPTIONS (METHOD=[NNE|RECURSIVE|KM|IPCW] AUC AT=182 365 547 730);
CLASS is_male;
MODEL time*event(0)= is_male eq_vas_t1 bvmt_trial1_t1 fkv_fl_t1;
RUN;
```

Abbildung 6: ROC-Kurven mit der PLOTS=ROC Option in PROC PHREG

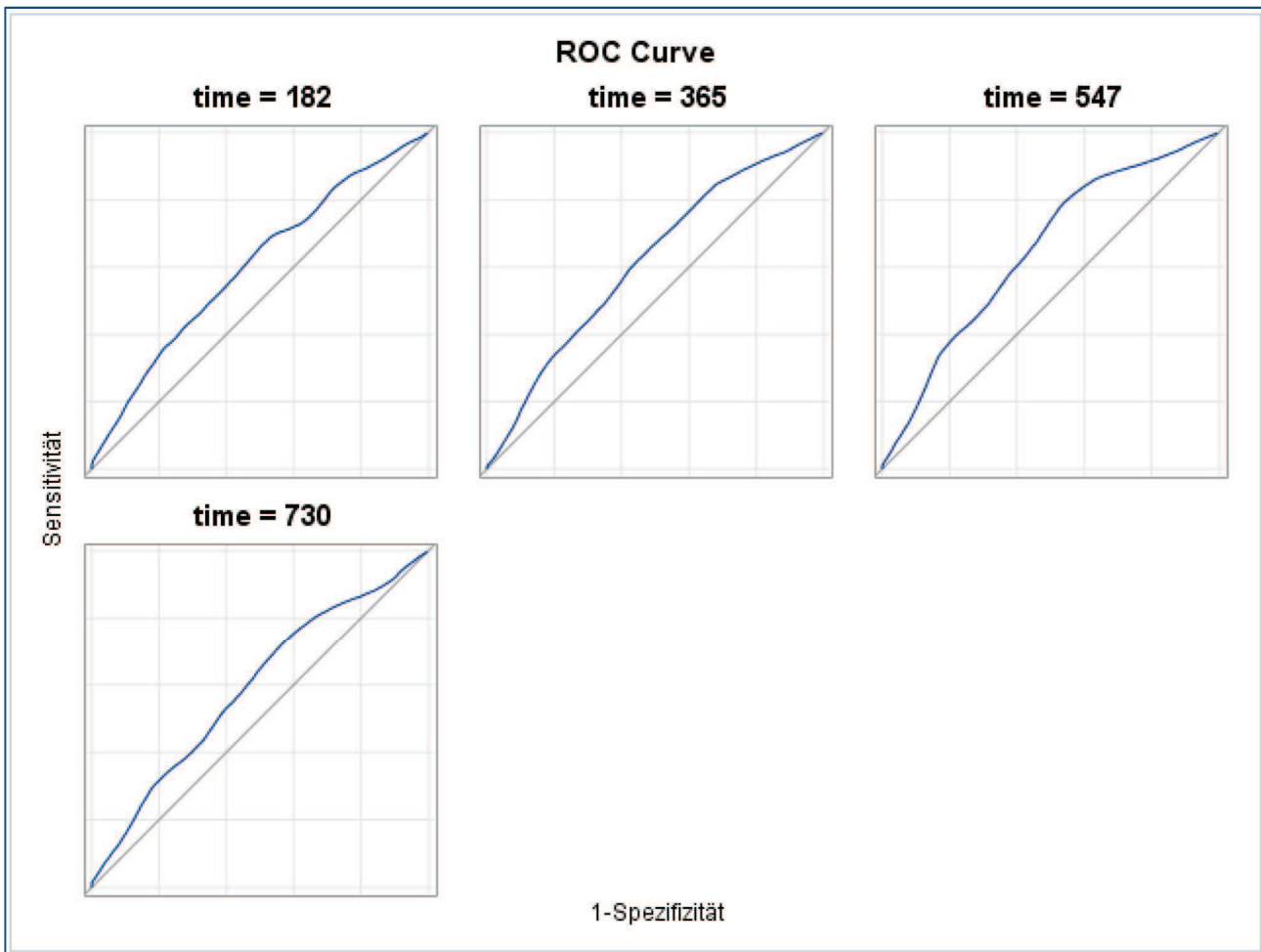


Abbildung 7: ROC-Kurve für die Zeitpunkte 182, 365, 547 und 730 Tage

Time-Dependent Area Under the Curve			
Quelle	ID	time	Schätzung
Model	1	182	0.6041
		365	0.6164
		547	0.6486
		730	0.5932

Abbildung 8: AUC für die Zeitpunkte 182, 365, 547 und 730 Tage

Es ist zudem möglich, die zeitabhängigen ROC-Kurven zweier Modelle zusammen in einer Grafik anzuzeigen. So kann beispielsweise die Auswirkung einer Transformation einer der Einflussgrößen auf die Prognosegüte untersucht werden.

Hierzu werden die linearen Prädiktoren $\beta'Z$ des einen Modells über das OUTPUT-Statement in einem Datensatz gespeichert. Nun werden die zeitabhängigen ROC-Kurven des zweiten Modells wie beschrieben angefordert. Diese können zur besseren Unterscheidung durch die Option ROCLABEL beschriftet werden. Über die Angabe der gespeicherten linearen Prädiktoren im ROC-Statement werden zudem die ROC-Kurven des ersten Modells hinzugefügt und beschriftet. [4]

```

PROC PHREG DATA=data;
CLASS is_male;
MODEL time*event(0)= is_male eq_vas_t1 bvmt_trial1_t1 log_fkv_f1_t1;
log_fkv_f1_t1 = log(fkv_f1_t1);
OUTPUT OUT=data2 XBETA=Y;
RUN;


---


PROC PHREG DATA=data2 PLOTS=ROC
  ROCOPTIONS (METHOD=NNE|RECURSIVE|KM|IPCW] AUC AT=365);
CLASS is_male;
MODEL time*event(0)= is_male eq_vas_t1 bvmt_trial1_t1 fkv_f1_t1
  / ROCLABEL = 'fkv_f1_t1';
ROC 'log(fkv_f1_t1)' pred=Y;
RUN;

```

Abbildung 9: ROC-Kurven zweier Modelle mit dem ROC-Statement in PROC PHREG

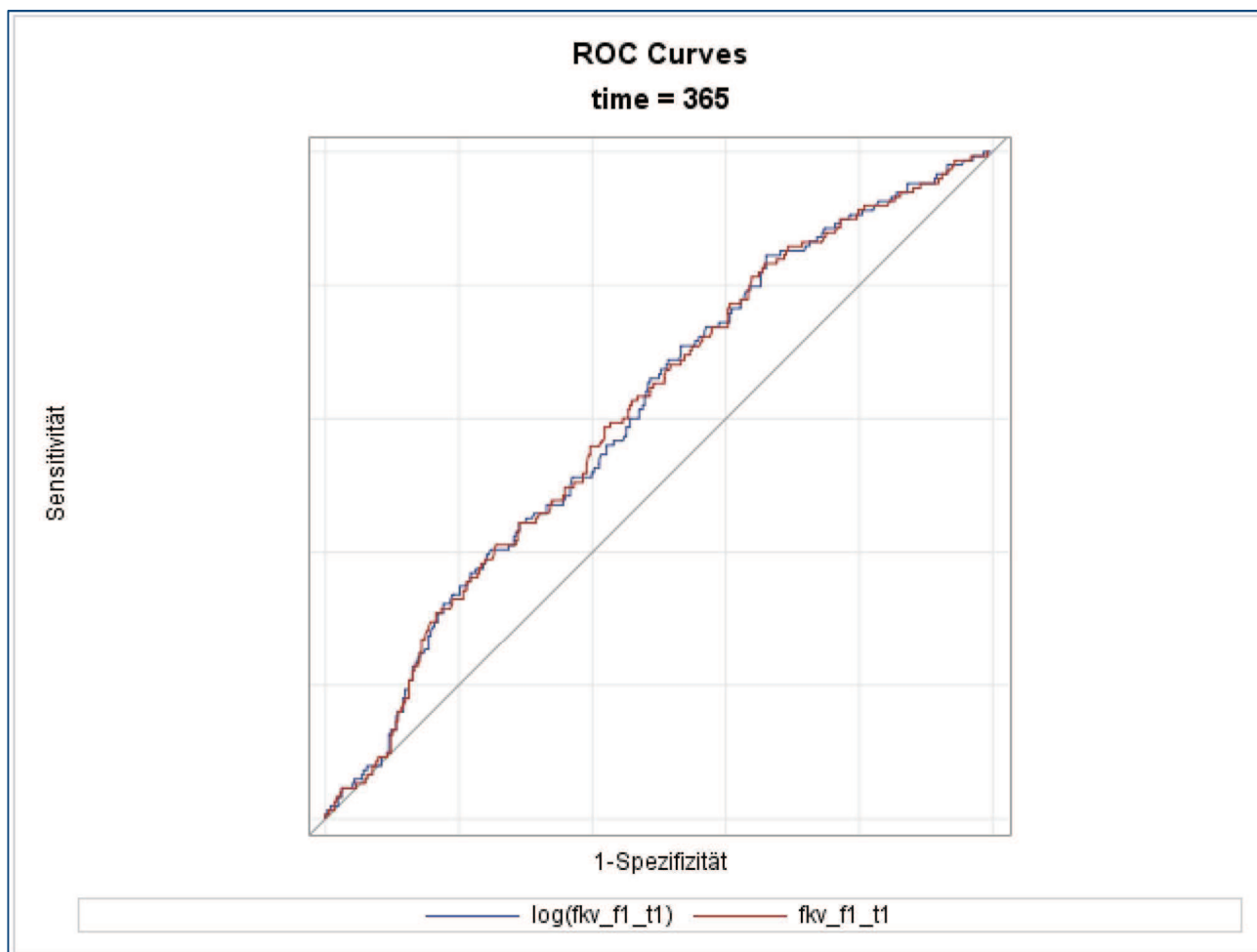


Abbildung 10: ROC-Kurven zweier Modelle für den Zeitpunkt 365 Tage

Time-Dependent Area Under the Curve			
Quelle	ID	time	Schätzung
log(fkv_f1_t1)	1	365	0.6088
fkv_f1_t1	2	365	0.6091

Abbildung 11: AUC zweier Modelle für den Zeitpunkt 365 Tage

Über Integration der Fläche unter der zeitabhängigen ROC-Kurve lässt sich eine **zeitabhängige AUC** berechnen. Diese kann über PLOTS=AUC im PROC PHREG-Statement angefordert werden. Innerhalb der ROCOPTIONS kann die gewählte Methode angegeben werden.

IPCW ist die einzige Methode, für die zusätzlich über die Option CL auch ein punktweises Konfidenzintervall angefordert werden kann. Da dieses über „Perturbation Resampling“ berechnet wird, können auch ein Startwert (SEED), die Anzahl der Stichproben (ITER) und das verwendete Signifikanzniveau (ALPHA) spezifiziert werden. [4]

```
PROC PHREG DATA=data PLOTS=AUC
  ROCOPTIONS (METHOD=[NNE | RECURSIVE | KM | IPCW] );
CLASS is_male;
MODEL time*event(0)= is_male eq_vas_t1 bvmt_trial1_t1 fkv_fl_t1;
RUN;
```

Abbildung 12: AUC über die Zeit mit der PLOTS=AUC Option in PROC PHREG

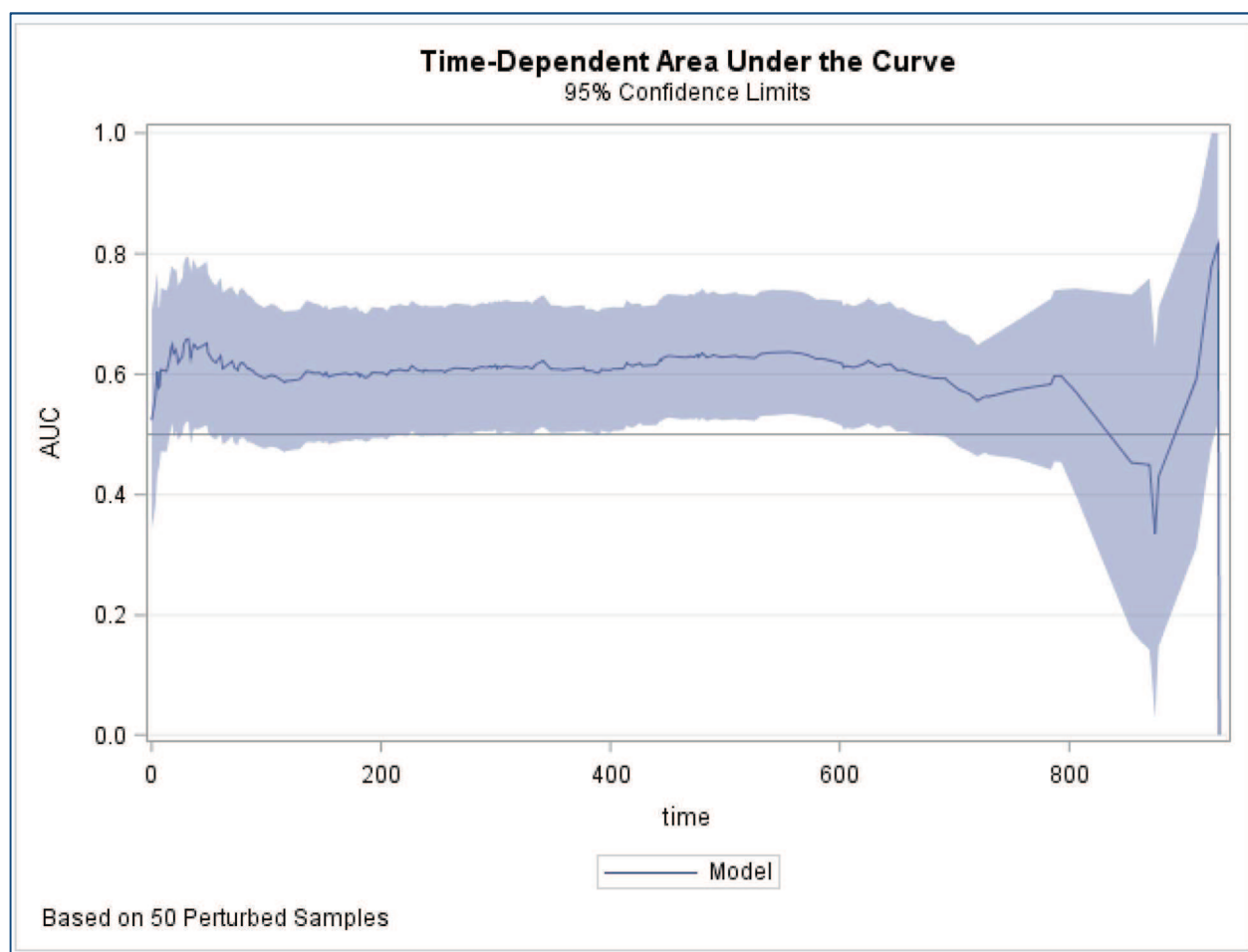


Abbildung 13: AUC über die Zeit mit Konfidenzbereich (IPCW-Methode)

4 Zusammenfassung

In der Literatur existieren noch weitere Statistiken zur Beurteilung der Modellgüte von Cox-Modellen und andere Definitionen von zeitabhängigen ROC-Kurven, die bisher noch nicht in SAS enthalten sind. Die nun hinzugefügten C-Statistiken nach Harrell und Uno sowie zeitabhängige ROC-Kurven und die AUC sind jedoch eine sinnvolle Erweiterung von PROC PHREG, die hoffentlich in Zukunft noch weiter ergänzt wird.

Literatur

- [1] The PHREG Procedure – Concordance Statistics.
http://documentation.sas.com/?docsetId=statug&docsetTarget=statug_phreg_detail_s57.htm&docsetVersion=14.3&locale=de (aufgerufen am 07.08.2018)
- [2] F. E. Harrell Jr., K. L. Lee, D. B. Mark: Tutorial in Biostatistics – Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statist. Med.* 1996, 15 361-387
- [3] H. Uno, T. Cai, M. J. Pencina, R. B. D’Agostino, L. J. Wei: On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Statist. Med.* 2011, 30 1105-1117
- [4] The PHREG Procedure – Time-Dependent ROC Curves.
http://documentation.sas.com/?docsetId=statug&docsetTarget=statug_phreg_detail_s60.htm&docsetVersion=14.3&locale=de (aufgerufen am 07.08.2018)
- [5] M. G. Akritas: Nearest Neighbor Estimation of a Bivariate Distribution Under Random Censoring. *The Annals of Statistics*, 1994, 22(3) 1299–1327
- [6] L. E. Chambless, G. Diao: Estimation of time-dependent area under the ROC curve for long-term risk prediction. *Statistics in Medicine*, 2006, 25 3474–3486
- [7] P. J. Heagerty, T. Lumley, M. S. Pepe: Timedependent ROC curves for censored survival data and a diagnostic marker. *Biometrics*, 2000, 56 337–344
- [8] H. Uno, T. Cai, L. Tian, L. J. Wei: Evaluating prediction rules for t-year survivors with censored regression models. *Journal of the American Statistical Association*, 2007, 102(478) 527–537