

Ein SAS-Makro für die Überprüfung der Balance zwischen ausgewählten Fällen und Kontrollen nach dem Matching von Beobachtungsdaten

Benjamin Mayer
Institut für Epidemiologie und
Medizinische Biometrie
Universität Ulm
Schwabstr. 13
89075 Ulm
benjamin.mayer@uni-ulm.de

Simone Nill
Institut für Epidemiologie und
Medizinische Biometrie
Universität Ulm
Schwabstr. 13
89075 Ulm
simone.nill@uni-ulm.de

Zusammenfassung

Beobachtungsdaten stellen eine wertvolle Datenquelle für die medizinische Forschung dar. Dem Charme solcher Daten, welche unter der Prämisse einer entsprechenden Pflege der betreffenden Datenquelle im Allgemeinen schnell und kostenminimal zur Verfügung stehen, steht die Tatsache gegenüber, dass im Rahmen entsprechender Studien die notwendigen Maßnahmen ergriffen werden müssen, das Problem potentiellen Confoundings angemessen zu berücksichtigen. Dies kann sowohl bei der Festlegung des konkreten Studiendesigns als auch erst im Zuge der Datenanalyse bereits geschehen. In einem früheren Beitrag wurde für diesen Zweck bereits ein SAS-Marko vorgestellt, das die Durchführung eines zweistufigen Matching-Algorithmus ermöglicht, um eine möglichst gute Vergleichbarkeit von Fall- und entsprechenden Kontrollpatienten zu erzeugen. Dieses Makro konnte insbesondere auch mit der Anforderung umgehen, dass die Fall- und Kontrollpatienten aus verschiedenen Datenquellen stammen. Ist das Matching erst einmal durchgeführt, so stellt sich in einem zweiten Schritt die Frage nach der Güte des Matchings, d.h. wie gut die gefundenen Kontrollpatienten den Fallpatienten tatsächlich entsprechen. Dies ist weniger relevant für solche Variablen, für die eine 100%ige Übereinstimmung erreicht werden muss, sondern insbesondere für solche, die beispielweise im Rahmen eines danach geschalteten Propensity Score (PS)-Matchings betrachtet wurden. Um den Erfolg eines durchgeführten Matchings automatisiert überprüfen zu können, wurde ein SAS-Marko implementiert, das einen so genannten Balance-Check vornehmen kann. Hierbei liegt der Fokus auf einer relativ einfachen Gegenüberstellung der verwendeten Matching-Variablen in den späteren Fall- und Kontrollgruppen in Abhängigkeit deren Skalierung (metrisch oder kategorial). Zudem erfolgt die Berechnung der so genannten standardisierten Differenz, welche auf Grundlage literaturbasierter Simulationsergebnisse entsprechende Richtwerte erreichen sollte, damit man von einem erfolgreichen Matching sprechen kann. Die Anwendung dieses Balance-Check-SAS-Makros erfolgt unter Verwendung eines realen Datensatzes aus der Kardiologie.

Schlüsselwörter: Common support, Optimal matching, Propensity Score, Standardisierte Differenz

1 Einleitung

Bei Beobachtungsstudien handelt es sich um nicht experimentelle bzw. nicht interven-tionelle Studien. Ein spezieller Studientyp auf Basis von Beobachtungsdaten ist die Fall-Kontroll-Studie. Dabei handelt es sich um retrospektive Studien, welche aus zwei Grup-pen besteht, nämlich der Fall- und Kontrollgruppe. Hierbei bezeichnet die Gruppe der Fälle eine Population von Probanden oder Patienten, die eine bestimmte Krankheit oder Eigenschaft vorweisen. Im Vergleich dazu ist bei der Kontrollgruppe diese Krankheit bzw. Eigenschaft nicht vorhanden. Oftmals unterscheiden sich die beiden Gruppen initial hinsichtlich ihrer Stichprobengröße [1]. Fall-Kontroll-Studien eignen sich vor allem für die Ermittlung der Ursachen von seltenen Erkrankungen oder von Erkrankungen mit sehr langen Latenzzeiten. Außerdem stehen bei Fall-Kontroll-Studien die Beobach-tungsdaten häufig schneller und kostengünstiger zur Verfügung. Ein Nachteil dagegen ist, dass unter Umständen nicht alle Informationen zu den Risikofaktoren der Probanden vorliegen. Ein weiterer Nachteil ist, dass die Patienten nicht randomisiert in die Fall- oder Kontrollgruppe eingeteilt werden. Dies kann zur Folge haben, dass sich die Patien-ten zwischen den Gruppen nach ihren Merkmalen unterscheiden und somit die Ergeb-nisse durch Confounder verzerrt werden [2].

In Beobachtungsstudien generell können potentielle Confounding-Effekte mit den Me-thoden Stratifizierung, multivariate Analysen und Matching reduziert werden. Mit der Stratifizierung kann der Effekt eines Confounders untersucht werden. Dazu kann die Stichprobe bzgl. des Confounders in Subgruppen unterteilt werden, in denen die Auswertung erfolgt [3]. Bei multivariaten Analysen wird der Zusammenhang zwischen der Zielgröße (abhängige Variable) und den verschiedenen Einflussgrößen bzw. Störgrößen (unabhängige Variablen) untersucht. Mit einem Matching-Ansatz kann in Fall-Kontroll-Studien die Strukturgleichheit zwischen der Fall- und Kontrollgruppe erreicht werden. In diesem Zusammenhang ist es wichtig, dass alle Variablen, welche einen Einfluss auf die Zuordnung zur Fall- und Kontrollgruppe haben, bekannt sind und dokumentiert wurden. In einem früheren Beitrag [4] wurde ein SAS-Makro vorgestellt, das ein so ge-nanntes Optimal Matching durchführen kann, um Fälle und Kontrollen möglichst pas-send zuordnen zu können. Ein generell anzustrebendes Ziel aller Matching-Ansätze ist die Erreichung einer guten Qualität in der Zuordnung von Fällen und Kontrollen. Damit das Matching zu valide auswertbaren Paaren (oder Triplets, etc.) führt, muss der so ge-nannte Common Support überprüft und für angemessen befunden werden. In Propensity Score (PS)-basierten Matching-Verfahren beschreibt der Common Support beispiele-weise den Bereich, in dem PS-Werte sowohl in der Fall-, als auch der Kontrollgruppe vorkommen. In diesem Beitrag soll ein SAS-Makro vorgestellt werden, das die Bewer-tung der Qualität eines durchgeführten Matchings ermöglicht. Die Anwendung dieses Balance-Check-Makros wird anhand eines medizinischen Datenbeispiels aus der Kardi-ologie demonstriert.

2 Material und Methoden

2.1 Matching

Beim Matching werden sogenannte „statistische Zwillinge“ gesucht. Die Kontrollen werden so ausgewählt, dass sie in bestimmten Merkmalen den Fällen ähneln. Die Zuordnung möglichst ähnlicher Probanden führt somit zu einer Abnahme der Verzerrung, die aus Selektionsprozessen entstanden sein könnten [5,6]. Zuordnungsprozesse, bei denen genau eine Person aus der Fallgruppe zu einer aus der Kontrollgruppe zugewiesen wird, ist unter dem Begriff 1:1-Matching bekannt. Wohingegen beim 1:k-Matching jedem Probanden der behandelten Gruppe mehrere Probanden aus der Kontrollgruppe als Matchingpartner zugeordnet werden. Der Vorteil von 1:k-Matching gegenüber 1:1-Matching ist, dass die Power aufgrund der größeren Fallzahl zunimmt. Ein Nachteil hingegen ist, dass unter Umständen schlechtere Zuordnungen der Matchingpartner hinsichtlich deren Merkmale entstehen. Im Wesentlichen können die derzeit verfügbaren generellen Matching-Ansätze in zwei Klassen eingeteilt werden.

In der Praxis sind viele verschiedene Matching-Ansätze verfügbar. So werden beispielsweise beim so genannten exact matching jedem Probanden aus der Fallgruppe ein Proband aus der Kontrollgruppe zugeordnet, welcher genau die identischen beobachtbaren Merkmale aufweist, so dass dieses Verfahren vor allem bei metrisch skalierten Merkmalen problematisch ist und die Angabe von Akzeptanzbereichen erfordert [7,8]. Ein weiteres Problem stellt die Anzahl der Matchingvariablen dar. Je mehr Variablen berücksichtigt werden, desto höher steigt die Wahrscheinlichkeit, dass nicht für alle Probanden ein passender „Zwilling“ gefunden wird. Bei Distanz-basierten Matching-Verfahren erfolgt die Zuordnung der Partner über die Berechnung von Distanzmaßen. Beim Mahalanobis Matching werden beispielsweise diejenigen Probanden einander zugeordnet, welche die kleinste sogenannte Mahalanobis-Distanz aufweisen [9]. Als Alternative zur Mahalanobis-Distanz kann die Propensity Score (PS)-Distanz betrachtet werden. Der PS ist definiert als die bedingte Wahrscheinlichkeit, mit der ein Ereignis unter gegebenen Kovariaten (X) auftritt. Individuen, die einen ähnlichen PS aufweisen, können miteinander verglichen werden, da deren beobachtbare Merkmale bzgl. der Intervention relativ ähnlich sind [7,10]. Der PS kann auf Basis einer logistischen Regression in SAS mit der Prozedur PROC LOGISTIC geschätzt werden. Unabhängig von der Wahl des konkreten Distanzmaßes können die etablierten Matching-Ansätze eingeteilt werden in verschiedene Varianten, wie die Zuordnung von Fall- und Kontrollprobanden erfolgt. Unter dem Begriff Greedy-Matching (z.B. Nearest Neighbor Matching, Caliper-Matching und entsprechende Kombinationen), wird generell ein Verfahren beschrieben bei dem Kontrollpersonen als Matchingpartner gesucht werden, die mit den Fällen eine hohe PS-Übereinstimmung haben. Mit diesem Verfahren werden sukzessiv Kontrollen den Fällen zugeteilt, wobei nicht verhindert werden kann, dass nach und nach schlechtere Matchingpartner gefunden werden. Dies beruht auf der Tatsache, dass keine Rücksicht auf die vorherigen Entscheidungen genommen wird [7]. Ein weiterer Ansatz ist durch das so genannte Optimal Matching gegeben. Beim Greedy Matching werden nach und nach

Fälle denjenigen Kontrollen zugeordnet, welche die kleinste Distanz zueinander haben, wobei keine Rücksicht auf die vorherigen Entscheidungen genommen wird. Dabei kann es passieren, dass ein Matchingpartner aus der Kontrollgruppe besser zu einem Fall gepasst hätte, dieser aber bereits vergeben ist. Das Greedy Matching ist also kein optimales Verfahren. Wird nicht nur die momentan kleinste Distanz betrachtet, sondern auch die minimale totale Distanz, können häufig bessere Matchingpartner gefunden werden. Solch ein Verfahren wird Optimal Matching genannt [7]. In einem früheren Beitrag zur KSFE-Konferenz wurde ein entsprechendes SAS-Makro vorgestellt, das einen Optimal Matching-Ansatz umsetzen kann [4].

Nachdem das Matching durchgeführt wurde ist eine zentrale Frage, wie gut das Matching funktioniert hat, d.h. wie gut die Balance zwischen Fällen und Kontrollen ist. Im Falle eines PS-Matchings kann z.B. der so genannte Common Support betrachtet werden, also das Ausmaß der Überlappung der beiden PS-Verteilungen in der Fall- und der Kontrollgruppe (Abbildung 1).

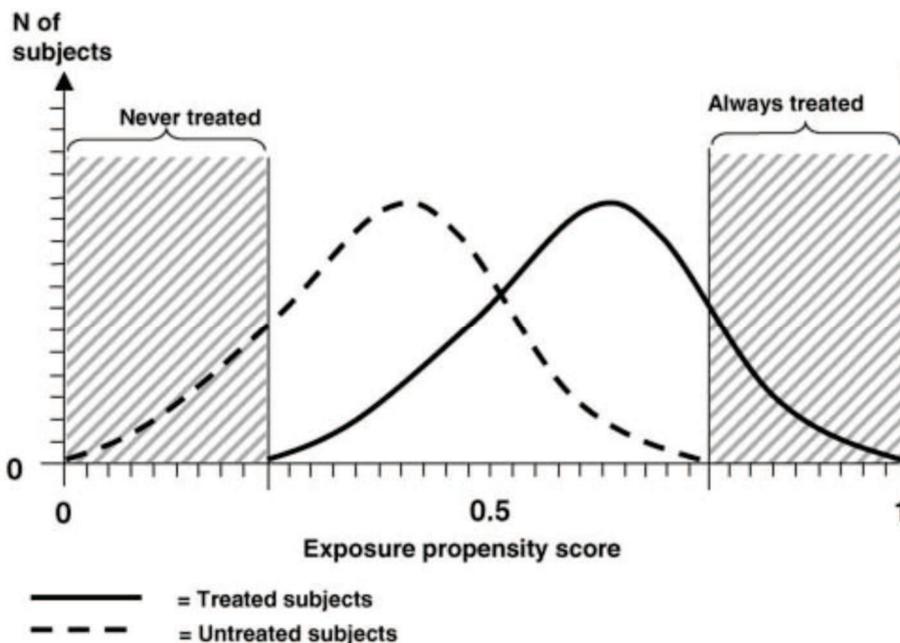


Abbildung 1: Common Support der PS-Verteilung [11]

Im folgenden Abschnitt werden die wichtigsten Konzepte der Überprüfung der Balance zwischen Fällen und Kontrollen vorgestellt

2.2 Bewertung der Matching-Güte

Zur Bewertung der Balance zwischen Fällen und Kontrollen hinsichtlich der für das Matching verwendeten Kovariaten muss diese sowohl vor dem Matching, als auch danach betrachtet werden. Hierzu kann beispielsweise die absolute standardisierte Differenz (ASD) in Abhängigkeit des Merkmalstyps berechnet werden. Für die ASD gibt es jedoch kein einheitliches Kriterium, für welchen Wert die Balancierung der Kovariaten

noch gegeben ist. Rosenbaum und Rubin [12] bezeichnen einen Wert von $d > 0.2$ als „großen“ Unterschied zwischen den Gruppen. Normand et al. [13] haben wiederum vorgeschlagen, dass ein Wert von $d < 0.1$ auf ein vernachlässigbares Ungleichgewicht der Kovariate zwischen Fall- und Kontrollgruppe hinweisen könnte [14]. Diese Werte gehen auf Cohen [9] zurück, welcher bei der Analyse der Effektstärke eine standardisierte Differenz von 0.2 als „kleinen“ Effekt bezeichnet hat. Laut Cohen bedeutet $d = 0.1$, dass die Verteilungen der beiden Gruppen um 7.7% nicht überlappen.

2.2.1 Balance Check bei stetigen Kovariablen

Bei stetigen Variablen wird die ASD vor dem Matching berechnet auf Basis der jeweiligen Mittelwerte in den Vergleichsgruppen, sowie deren gemeinsamer Standardabweichung [15]:

$$d_{vor} = \frac{|\bar{X}_1 - \bar{X}_0|}{S}$$

Für die Standardabweichung sind in der Literatur unterschiedliche Varianten zu finden. Nach Cohen [9] ist sie definiert als

$$S = \sqrt{\frac{s_1^2 + s_0^2}{2}}$$

mit s_1 und s_0 als Standardabweichungen der Fall- und Kontrollgruppe. In dieser Formel werden jedoch unterschiedliche Stichprobengrößen nicht berücksichtigt. Da beim Matching unverbundene Stichproben vorliegen, kann davon ausgegangen werden, dass die Stichprobengröße der Fallgruppe N_1 und der Kontrollgruppe N_0 häufig variieren. In der Formel nach Hedges [16] ist dieser Sachverhalt berücksichtigt:

$$S = \sqrt{\frac{(N_1 - 1)s_1^2 + (N_0 - 1)s_0^2}{N_1 + N_0 - 2}}$$

Die ASD für stetige Variablen wird nach dem Matching dann berechnet über

$$d_{nach} = \frac{|\bar{X}_1 - \bar{X}_0^*|}{S}$$

Dabei lässt sich der Mittelwert der Kontrollgruppe nach dem Matching im allgemeinen berechnen durch [15]

$$\bar{X}_0^* = \frac{1}{N_1} \sum_{i=1}^{N_1} \left(\frac{1}{m_i} \sum_{j=1}^{m_i} X_{0ij} \right)$$

2.2.2 Balance Check bei binären Kovariablen

Für binären Kovariablen wird die ASD berechnet unter Verwendung der Prävalenzen P_1 und P_0 in Fall- und Kontrollgruppe über

$$d_{vor} = \frac{|\hat{P}_1 - \hat{P}_0|}{\sqrt{\frac{\hat{P}_1(1-\hat{P}_1) + \hat{P}_0(1-\hat{P}_0)}{2}}}$$

Analog dazu wird die ASD nach dem Matching berechnet. Alternativ können Odds Ratios bei dichotomen Variablen betrachtet werden [17]. Falls der Wert des Odds Ratios nahe bei 1 liegt, kann davon ausgegangen werden, dass es keinen relevanten Unterschied zwischen Fall- und Kontrollgruppe gibt im Hinblick auf die untersuchte Kovariable.

2.2.3 Balance Check bei kategorialen Kovariablen mit mehr als 2 Ausprägungen

Im Falle kategorialer Kovariablen mit mehr als zwei Ausprägungen (k in $1, \dots, K$; $K > 2$) wird die ASD ebenfalls berechnet unter Verwendung der Prävalenzen P_1 und P_0 in Fall- und Kontrollgruppe, jedoch wird zusätzlich die Mahalanobis-Distanz betrachtet [18]:

$$d_{vor} = \sqrt{(\hat{P}_1 - \hat{P}_0)^T \Sigma^{-1} (\hat{P}_1 - \hat{P}_0)}$$

Hierbei sind die Prävalenzen der Fall- und Kontrollgruppe gegeben durch

$$\begin{aligned} \hat{P}_1 &= (\hat{P}_{12}, \dots, \hat{P}_{1R})^T \\ \hat{P}_0 &= (\hat{P}_{02}, \dots, \hat{P}_{0R})^T \end{aligned}$$

Die Varianz-Kovarianz-Matrix Σ wird dabei berechnet über

$$\Sigma_{kl} = \begin{cases} \frac{\hat{P}_{1k}(1-\hat{P}_{1k}) + \hat{P}_{0k}(1-\hat{P}_{0k})}{2}, & \text{falls } k = l \\ \frac{\hat{P}_{1k}\hat{P}_{1l} + \hat{P}_{0k}\hat{P}_{0l}}{2}, & \text{falls } k \neq l \end{cases}$$

Dieses Vorgehen hat jedoch den Nachteil, dass für die Berechnungen eine Referenzkategorie festgelegt werden muss und diese die ASD beeinflusst. Als Alternative kann

auch ein Chi-Quadrat-Test berechnet werden und ein Kontingenzkoeffizient ausgegeben werden (z.B. Cramér oder Pearson). Als weitere Alternative bietet sich eine zum binären Fall analoge Vorgehensweise an, bei der die ADS für jede Ausprägung der Kovariablen separat berechnet werden

$$d_{vor,nach} = \frac{1}{N_1 + N_0} \sum_{k=1}^K \frac{(N_{1k} + N_{0k}) |\hat{P}_{1k} - \hat{P}_{0k}|}{\sqrt{\frac{(N_{1k}-1)\hat{P}_{1k}(1-\hat{P}_{1k}) + (N_{0k}-1)\hat{P}_{0k}(1-\hat{P}_{0k})}{N_{1k} + N_{0k} - 2}}}$$

2.3 Balance Check-Makro in SAS

Die in Abschnitt 2.2 vorgestellten Konzepte einer rechnerischen Überprüfung der Balance zwischen Fällen und Kontrollen auf Basis der ASD wurden in einem SAS-Makro (SAS-Version 9.4) umgesetzt. Dieses Makro erlaubt - auch als Erweiterung des in der Vergangenheit vorgestellten Matching-Makros [4] - einen automatisierten Balance-Check, in dem die ASD vor und nach dem Matching berechnet und ausgegeben wird für alle am Matching beteiligten Variablen. Ein Basisaufruf des Makros sieht wie folgt aus:

```
%MACRO balancecheck(data_nach, data_vor, strata, match, id, name, register, var_kat=NONE,
var_stetig=NONE)
```

Die Übergabeparameter sind im Einzelnen:

data_vor	Datensatz vor dem Matching
data_nach	Datensatz nach dem Matching
strata	Indikatorvariable zur Unterscheidung der Fall- und Kontrollgruppe (1=Fall, 0=Kontrolle)
match	Eindeutige Identifikationsnummer der Matchingpartner
id	Eindeutige Identifikationsnummer der Fall- und Kontrollpatienten
name	Pfadangabe mit Name der PDF-Datei, die automatisch erzeugt wird
register	Angabe der Registervariable, falls mehrere Datenquellen beteiligt
var_kat	Auflistung der kategorialen Matchingvariablen, Default NONE
var_stetig	Auflistung der stetigen Matchingvariablen, Default NONE

Die bei der Ausführung des Makros automatisch generierte pdf-Datei enthält die wesentlichen Informationen, die für einen Balance-Check wie beschrieben notwendig sind. Zum einen werden pro Merkmal, das im Rahmen des Matchings berücksichtigt wurde, die Verteilung in der Fall- und der Kontrollgruppe ausgegeben. Für kategoriale Merkmale bedeutet dies eine Darstellung in absoluten und relativen Häufigkeiten in den einzelnen Kategorien, bei stetigen Merkmalen werden Mittelwert, Standardabweichung, Median, Minimum, Maximum und Quartile für beide Gruppen ausgegeben. Diese Ausgabe erfolgt jeweils für die Datensätze vor und nach dem Matching. Zudem wird die dabei beobachtete Differenz (vor dem Matching vs. nach dem Matching), sowie die je-

weiligen ASD-Werte und Odds Ratios ausgegeben. Abschließend wird die Anzahl der ungleichen Kategorien über alle kategorialen Matching-Variablen dargestellt.

3 Anwendungsbeispiel

Die Anwendung des oben beschriebenen Makros zur Evaluierung der Matching-Güte in Form eines Balance-Checks soll anhand eines Datensatzes aus dem Bereich Kardiologie demonstriert werden. Im Rahmen einer retrospektiven Datenanalyse sollte untersucht werden, welchen Effekt die Gabe des Medikamentes Apixaban bei der so genannten MitraClip-Therapie zur Behandlung einer Mitralklappeninsuffizienz hat. Als Fälle für diese Analyse lagen Patientendaten vor, die während der MitraClip-Therapie das Medikament zur Antikoagulation erhalten haben. Der Pool entsprechender Kontrollpatienten bekam Apixaban nicht während des Eingriffs. Als Matchingkriterien wurden das Geschlecht, das Alter (>75 Jahre vs. ≤ 75 Jahre), die Ätiologie (degenerativ vs. funktionell) und verschiedene klinische Parameter (z.B. linksventrikuläre Auswurffraktion, NYHA-Klasse, usw.) festgelegt. Das Fall-Kontroll-Matching wurde im Verhältnis 1:1 unter Verwendung eines Optimal Matching Ansatzes mit Betrachtung der PS-Distanz durchgeführt. Der letztliche Makro-Aufruf für das hier vorgestellte Beispiel lautet wie folgt:

```
%balancecheck(data_nach=optimal, data_vor=mitra, var_kat=age gender etiology lvf mr_grade sts  
nayha, strata=group, register=register, match=_match, id=patient, name=F:\mitra_bsp_ksfe.pdf);
```

Der Common Support unter Verwendung der PS-Distanz für die Zuordnung von Fällen und Kontrollen war zufriedenstellend unter Verwendung der angegebenen Variablen (Abbildung 2).

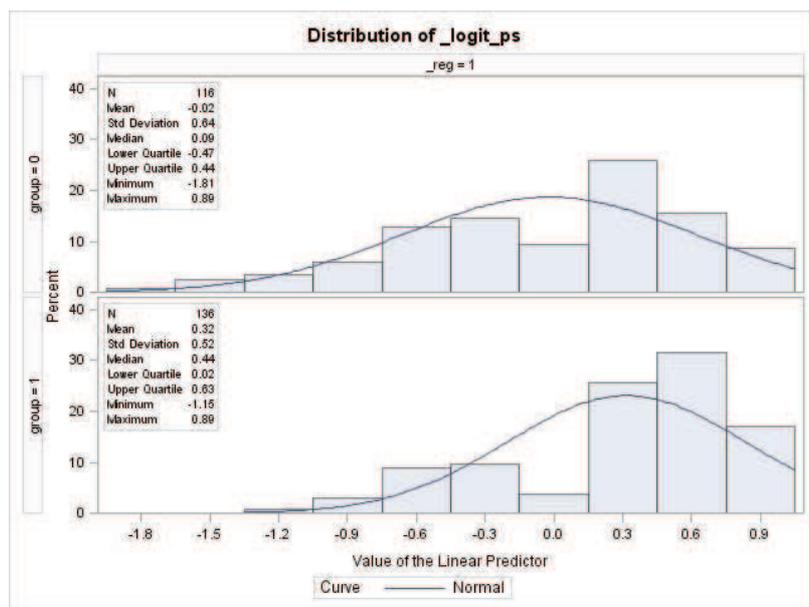


Abbildung 2: Logit PS-Verteilung vor dem optimal-Matching

Der Balance-Check zeigt, mehr als 70% der beteiligten Patienten einen „perfekten“ Matching-Partner gefunden haben und alle kategorialen Matching-Merkmale überein-

stimmen. Bei weniger als 7% aller Patienten betrug der Unterschied zwischen Fällen und Kontrollen höchstens zwei Kategorien (siehe Abbildung 3). Dieses Resultat ist in Anbetracht des relativ großen Bereichs in Abbildung 2, in denen sich die PS-Verteilungen beider Gruppen überlappen, nicht überraschend.

Anzahl der ungleichen Kategorien über alle kategoriellen Matchingvariablen

The FREQ Procedure

<i>summe</i>	<i>Frequency</i>	<i>Percent</i>	<i>Cumulative Frequency</i>	<i>Cumulative Percent</i>
0	83	71.55	83	71.55
1	16	13.79	99	85.34
2	9	7.76	108	93.10
3	4	3.45	112	96.55
4	3	2.59	115	99.14
5	1	0.86	116	100.00

Abbildung 3: Anzahl der ungleichen Kategorien zwischen Fällen und Kontrollen über alle kategorialen Matching-Variablen

Insgesamt kann die Matching-Güte in diesem Datenbeispiel als akzeptabel bezeichnet werden, auch wenn nicht alle Merkmale nach dem Matching eine angestrebte $ASD < 0.1$ erzielen konnten. Dies zeigte sich beispielsweise bei der Variablen Alter (Abbildung 4).

Differenz der Matchingpartner von age nach dem Matching

The FREQ Procedure

age	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	105	90.52	105	90.52
1	11	9.48	116	100.00

register	dvor	dnach	bias	or_vor	or_nach
1	0.26835	0.20379	24.0583	0.56503	0.64673

Verteilung der Variable age vor dem Matching

The FREQ Procedure

group=0

age	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	32	27.12	32	27.12
1	86	72.88	118	100.00

group=1

age	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	54	39.71	54	39.71
1	82	60.29	136	100.00

Verteilung der Variable age nach dem Matching

The FREQ Procedure

group=0

age	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	32	27.59	32	27.59
1	84	72.41	116	100.00

group=1

age	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	43	37.07	43	37.07
1	73	62.93	116	100.00

Abbildung 4: Balance-Check hinsichtlich des Merkmals Alter

Nahezu alle klinischen Matching-Variablen zeigten jedoch eine sehr gute Übereinstimmung, die auf das durchgeführte Fall-Kontroll-Matching zurückgeführt werden kann (Abbildung 5).

Differenz der Matchingpartner von etiology nach dem Matching

The FREQ Procedure

etiology	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	112	96.55	112	96.55
1	4	3.45	116	100.00

register	dvor	dnach	bias	or_vor	or_nach
1	0.19277	0.070284	63.5409	1.48693	1.15390

Differenz der Matchingpartner von Ivf nach dem Matching

The FREQ Procedure

Ivf	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	112	96.55	112	96.55
1	4	3.45	116	100.00

register	dvor	dnach	bias	or_vor	or_nach
1	0.18271	0.070284	61.5322	1.45719	1.15390

Abbildung 5: Balance-Check hinsichtlich des klinischen Merkmale

4 Diskussion

Das in diesem Beitrag vorgestellte SAS-Makro dient zu einer automatisierten Überprüfung der Balance zwischen zwei Kohorten, die im Rahmen eines Fall-Kontroll-Matchings erzeugt wurden [4]. Das Makro berichtet sowohl die Ergebnisse einer Kohorten-spezifischen und Merkmals-abhängigen Deskription aller Matching-Variablen, als auch die zugehörigen ASDs vor und nach dem Matching. Die Güte eines Matchings kann insgesamt als sehr gut bezeichnet werden, wenn alle beteiligten Matching-Variablen nach dem Matching eine $ASD < 0.1$ aufweisen, und als akzeptabel im Falle $ASD < 0.2$. Sollten für einzelne Matching-Kriterien diese Richtwerte nicht erfüllt werden können, sollte auf Basis der zusätzlich ausgegebenen deskriptiven Zusammenfassung vorwiegend inhaltlich bewertet werden, ob die vorliegenden Differenzen zwischen Fall- und Kontrollgruppe (auf das jeweilige Merkmal bezogen) einen nicht akzeptablen Unterschied darstellen. Gegebenenfalls sollte dann erwogen werden die Kriterien für das Matching nochmals zu adaptieren.

Literatur

- [1] J.-B. du Prel, B. Röhrig, G. Hommel und M. Blettner. Choosing Statistical Tests. Part 12 of a Series on Evaluation of Scientific Publications. Deutsches Ärzteblatt International 2010, 107(19), S. 343–348.
- [2] B. Röhrig, J.-B. du Prel, D. Wachtlin und M. Blettner. Types of Study in Medical Research. Part 3 of a Series on Evaluation of Scientific Publications. Deutsches Ärzteblatt International 2009, 106(15), S. 262–268.
- [3] G. P. Hammer, J.-B. du Prel und M. Blettner. Avoiding Bias in Observational Studies. Part 8 of a Series on Evaluation of Scientific Publications. Deutsches Ärzteblatt International 2009, 41, S. 664–668.
- [4] B. Mayer, S. Nill und D. Rothenbacher. Simulation eines zweistufigen Matching-Algorithmus in SAS auf der Basis von Propensity Scores für Beobachtungsdaten aus verschiedenen Registern. KSFE 2017: Proceedings der 21. Konferenz der SAS-Anwender in Forschung und Entwicklung (KSFE), 175-186, Aachen: Shaker Verlag, 2017.
- [5] E. J. Bergstralh und J. L. Kosanke. Computerized matching of cases to controls. Mayo Clinic, Technical Report Number 56, 1995.
- [6] W. G. Cochran. The effectiveness of adjustment by subclassification in removing bias in observational studies. Biometrics 1968, 24(2), S. 295–313.
- [7] M. Coca-Perraillon. Local and global optimal propensity score matching. Proceedings of the SAS® Global Forum 2007 Conference. Cary, NC: SAS Institute Inc., Statistics and Data Analysis: Paper 185-2007, 2007.

- [8] R. B. D'Agostino Jr. Tutorial in biostatistics: Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine* 1998, 17(19), S. 2265–2281.
- [9] J. Cohen. *Statistical power analysis for the behavioral sciences*. 2. Auflage. Hillsdale, NJ [u.a.]: Erlbaum, 1988. ISBN: 0-8058-0283-5.
- [10] W. W. Feng, M. Jun und R. Xu. A method/macro based on propensity score and Mahalanobis distance to reduce bias in treatment comparison in observational study. *Proceedings of the Pharmaceutical SAS® Users Group Conference*. Cary, NC: SAS Institute Inc., Public Health Research: Paper PR05, 2006.
- [11] P. Berta, M. Bossi und S. Verzillo. %CEM: a SAS macro to perform coarsened exact matching. *Journal of Statistical Computation and Simulation* 2017, 87(2), S. 227–238.
- [12] P. R. Rosenbaum und D. B. Rubin. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. In: *The American Statistician* 1985, 39(1), S. 33–38.
- [13] S. T. Normand, M. B. Landrum, E. Guadagnoli, J. Z. Ayanian, T. J. Ryan, P. D. Cleary und B. J. McNeil. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. *Journal of Clinical Epidemiology* 2001, 54(4), S. 387–398.
- [14] H.-P. Altenburg. Über die Umsetzung von Matching-Verfahren mit Hilfe des SAS-Systems. *SAS®: Verbindung von Theorie und Praxis*. *Proceedings der 9. Konferenz der SAS®-Anwender in Forschung und Entwicklung (KSFE)*. Aachen: Shaker, 2005, S. 1–13.
- [15] A. Haviland, D. S. Nagin und P. R. Rosenbaum. Combining propensity score matching and group-based trajectory analysis in an observational study. *Psychological Methods* 2007, 12(3), S. 247–267.
- [16] L. V. Hedges und I. Olkin. *Statistical methods for meta-analysis*. Orlando [u.a.]: Academic Press, 1985.
- [17] R. E. Lanehart, P. R. de Gil, E. S. Kim, A. P. Bellara, J. D. Kromrey und R. S. Lee. Propensity score analysis and assessment of propensity score approaches using SAS® procedures. *Proceedings of the SAS® Global Forum 2012 Conference*. Cary, NC: SAS Institute Inc., *Statistics and Data Analysis*: Paper 314-2012, 2012.
- [18] D. Yang und J. E. Dalton. A unified approach to measuring the effect size between two groups using SAS®. In: *Proceedings of the SAS® Global Forum 2012 Conference*. Cary, NC: SAS Institute Inc., *Statistics and Data Analysis*: Paper 335-2012, 2012.