

Simulation eines zweistufigen Matching-Algorithmus in SAS auf der Basis von Propensity Scores für Beobachtungsdaten aus verschiedenen Registern

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

Dietrich Rothenbacher
 Institut für Epidemiologie und
 Medizinische Biometrie
 Universität Ulm
 Helmholtzstr. 22
 89081 Ulm
 dietrich.rothenbacher@uni-ulm.de

Zusammenfassung

Registerdaten stellen eine wertvolle Datenquelle für die medizinische Forschung dar, die in zunehmendem Maße auch genutzt werden. Dem Nutzen und Charme solcher Daten, welche unter der Prämisse einer entsprechenden Pflege des betreffenden Registers im Allgemeinen schnell und kostenminimal zur Verfügung stehen, steht die Tatsache gegenüber, dass es sich um reine Beobachtungsdaten handelt und daher im Rahmen entsprechender analytischer Studien die notwendigen Maßnahmen ergriffen werden müssen, das Problem potentiellen Confoundings angemessen zu berücksichtigen. In diesem Beitrag wird exemplarisch eine Situation beschrieben, in der ein Fall-Kontroll-Matching im Verhältnis 1:2 durchgeführt werden soll. Der gesamte Algorithmus ist in SAS 9.4 implementiert. Das Matching ist dabei ein zweistufiger Prozess, in dessen Rahmen zunächst eine definierte Anzahl von Hauptmatching-Kriterien zu 100% erfüllt sein muss. Weitere Matching-Variablen können anschließend über einen Propensity Score (PS)-Ansatz zusätzlich berücksichtigt werden. Ein solches Setting könnte in der Praxis von großer Relevanz sein, wenn für bestimmte Sachverhalte eine perfekte Übereinstimmung aus inhaltlichen Gründen angestrebt werden muss (z.B. Übereinstimmung von Krankheitsstadien). Sowohl die Daten der Fälle, als auch der Kontrollen basieren dabei auf erhobenen Registerdaten, die jedoch verschiedenen Quellen entstammen können. Eine Anwendung des vorgestellten Algorithmus erfolgt dann auf der Basis von umfangreichen Beobachtungsdaten (N=30,000) aus dem Bereich der Notfallmedizin. Im Rahmen des Beitrags sollen insbesondere die verschiedenen Matching-Ansätze und die dafür implementierten SAS-Makros zur Durchführung (Mahalanobis matching, optimal matching via PROC OPTNET in SAS 9.4) vorge-

stellt und diskutiert werden, welche für das PS-basierte Matching verwendet und implementiert wurden.

Schlüsselwörter: Caliper; Optimal matching; PROC OPTNET; Propensity Score; Registerforschung

1 Einleitung

Registerdaten stellen eine wertvolle Datenquelle für die medizinische Forschung dar, die in zunehmendem Maße auch genutzt werden. Vor allem im Rahmen der Epidemiologie und der Versorgungsforschung werden häufig Registerdaten verwendet, um entsprechende Fragestellungen zu bearbeiten. Dem Nutzen und Charme solcher teils sehr umfassender Daten, welche unter der Prämisse einer entsprechenden Pflege des betreffenden Registers im Allgemeinen schnell und kostenminimal zur Verfügung stehen, steht die Tatsache gegenüber, dass es sich um reine Beobachtungsdaten handelt und daher im Rahmen entsprechender Studien die notwendigen Maßnahmen ergriffen werden müssen, das Problem potentiellen Confoundings angemessen zu berücksichtigen.

Der Problematik einer eventuellen Ungleichverteilung potentieller Störgrößen auf die eigentlich zu vergleichenden Gruppen kann man mit unterschiedlichen Ansätzen begegnen. So können beispielsweise im Rahmen eines multiplen Regressionsmodells eventuelle Confounder-Variablen als zusätzliche Adjustierungsvariablen mit ins Modell aufgenommen werden, um den primär interessierenden Effektschätzer entsprechend zu korrigieren. Eine weitere Möglichkeit stellen subgruppen-spezifische Analysen dar. Hierbei werden die primär interessierenden Effektschätzer nicht nur auf der Basis des gesamten Datenkollektivs bestimmt, sondern zusätzlich im Sinne einer Sensitivitätsanalyse innerhalb wohldefinierter Subgruppen ermittelt. Die Abgrenzung der Subgruppen erfolgt in Abhängigkeit der aus inhaltlicher Sicht wahrscheinlichsten Störvariablen. Ein abschließender Vergleich aller so ermittelten Effektschätzer ermöglicht dann eine Einschätzung des vorliegenden Verzerrungspotentials.

Eine dritte Möglichkeit dem Confounding-Problem zu begegnen ist die Anwendung von Matching-Verfahren. Dabei werden für Patienten aus einer Gruppe („Fälle“) explizite Matching-Partner aus der anderen Gruppe („Kontrollen“) gemäß vorab festgelegter Kriterien gesucht. Dabei gibt es prinzipiell eine Vielzahl an Möglichkeiten der konkreten Umsetzung des Matching-Prozesses. Auf der einen Seite gibt es sehr strenge Kriterien die beispielsweise vorgeben, dass Patienten aus beiden Gruppen nur dann gematcht werden dürfen, wenn sie bezüglich aller definierten Variablen eine 100%ige Übereinstimmung aufweisen. Dieses Vorgehen kann in manchen Situationen sicherlich inhaltlich gut begründbar und notwendig sein, stellt für die Praxis jedoch meist eine unlösbare Aufgabe dar - vor allem, wenn die Anzahl der Matching-Variablen hoch ist. Weniger stringente Verfahren bedingen nicht auf eine perfekte Übereinstimmung bezüglich der definierten Variablen, sondern basieren vielmehr auf der Berechnung von Distanzmaßen zwischen Fall- und Kontrollpatienten mit dem Ziel diejenigen Patienten zu matchen, die eine möglichst kleine Distanz aufweisen. In Bezug auf die Wahl des Distanzmaßes gibt es wiederum eine Vielzahl von Möglichkeiten. Dieser Beitrag zielt speziell auf die Situ-

ation ab, dass der Propensity Score (PS), welcher aus den vorliegenden Registerdaten berechnet werden kann, als Distanzmaß im Rahmen des Fall-Kontroll-Matchings verwendet wird.

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 [4,7]. 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.

Beim Exact Matching wird jedem Probanden aus der Fallgruppe ein Proband aus der Kontrollgruppe zugeordnet, welcher genau die identischen beobachtbaren Merkmale aufweist. Es muss also bei jedem Matching sichergestellt werden, dass das jeweilige Paar bzgl. aller Variablen die gleichen Ausprägungen besitzt. Die Ausprägungen einer Variablen müssen zu 100% identisch sein, so dass dieses Verfahren vor allem bei metrisch skalierten Merkmalen problematisch ist [6,10]. 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.

Beim Mahalanobis Matching werden diejenigen Probanden einander zugeordnet, welche die kleinste sogenannte Mahalanobis Distanz aufweisen. Dabei wird nicht nur der Abstand zwischen zwei Individuen bzgl. deren Merkmale betrachtet, sondern auch die Varianz jeder Variablen und die Korrelationen mit anderen Variablen. Die Formel für die Mahalanobis Distanz lässt sich aus der multivariaten Normalverteilung herleiten und lautet [13]:

$$D_{ij} = (X_i - X_j)^T \Sigma^{-1} (X_i - X_j)$$

Hierbei ist Σ die wahre oder geschätzte Varianz-Kovarianzmatrix der Fallgruppe, der Kontrollgruppe oder der gepoolten Stichprobe. Häufig wird die Kontrollgruppe zur Berechnung der Varianz-Kovarianzmatrix verwendet [15]. In SAS kann die Varianz-Kovarianzmatrix mit der Prozedur PROC PRINCOMP berechnet werden [14]. Der Vorteil dieser Methode ist, dass die unterschiedlichen Skalierungen der Variablen berücksichtigt werden. Ein großer Nachteil dieser Methode ist, dass sie auf einem mehrdimensionalen

nen Score basiert. Je mehr Variablen zur Berechnung der Mahalanobis Distanz in das Modell aufgenommen werden, desto größer wird die Distanz.

2.2 Propensity Score Matching

2.2.1 Propensity Score

Der PS ist definiert als die bedingte Wahrscheinlichkeit, mit der ein Ereignis unter gegebenen Kovariaten (X) auftritt. Wird beispielhaft als Ereignis die Intervention W betrachtet, so ergibt sich der PS für Individuum i durch

$$e(X_i) = P(W_i=1 | X_i).$$

Individuen, die einen ähnlichen PS aufweisen, können miteinander verglichen werden, da deren beobachtbare Merkmale bzgl. der Intervention relativ ähnlich sind [6, 13]. Der PS hat viele nützliche Eigenschaften, u.a. ist er ein „Balancing Score“. Dieser besagt, dass üblicherweise Fall- und Kontroll-Probanden mit den gleichen PS-Werten die gleiche Verteilung der Kovariaten X haben [11]. Der PS kann auf Basis einer logistischen Regression in SAS mit der Prozedur PROC LOGISTIC geschätzt werden.

2.2.2 Varianten des PS-Matching

Nach der Berechnung des PS können verschiedene Matching-Verfahren angewandt werden, die alle den PS als zugrunde liegendes Distanzmaß verwenden. Insofern können diese Verfahren als eine Alternative zum vorher angesprochenen Mahalanobis-Matching fungieren.

Unter dem Begriff Greedy-Matching, das verschiedene Methoden des PS-Matchings zusammenfasst (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. Für die Genauigkeit der PS-Übereinstimmung kann die Anzahl der Nachkommastellen [9] oder die kleinste PS-Differenz genommen werden (Nearest Neighbor-Ansatz). Zudem ist es wichtig, dass die Reihenfolgen der Kontrollen und Fälle zufällig vorliegen, da zugeordnete Personen nicht erneut gewählt werden. 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 [6]. Beim Nearest Neighbor Matching werden Individuen einander zugeordnet, welche die kleinste PS-Distanz haben. Die Zuordnung ist mit und ohne Zurücklegen möglich. Falls das Verfahren mit Zurücklegen gewählt wird, dann stehen Probanden aus der Kontrollgruppe mehrmals als Matchingpartner für eine Person aus der Fallgruppe zur Verfügung. Der Vorteil dieser Variante ist, dass bessere Matchings gefunden werden können. Ein großer Nachteil hingegen ist, dass unter Umständen nur wenige Probanden von der Kontrollgruppe verwendet werden, obwohl relativ ähnliche Kontrollpersonen noch zur Verfügung stehen. Dies hat zur Folge, dass die Power geringer wird. Der Zuordnungs-

prozess mit Zurücklegen wird empfohlen, wenn eine kleine Stichprobe der Kontrollgruppe vorliegt.

Beim Nearest Neighbor Matching gibt es keine Einschränkung der maximalen Distanz zwischen den PS der Fall- und Kontrollgruppe. Eine große Distanz hat zur Folge, dass die beiden Individuen bzgl. ihrer Kovariaten stark variieren. Aus diesem Grund wird nun ein Toleranzbereich, ein sogenannter Caliper, eingeführt. Dies soll schlechte Matchings verhindern. Hierbei stellt c der vorgegebene Toleranzbereich dar. Für die Wahl des Calipers werden verschiedene Werte empfohlen. Hierzu gibt es keine allgemeingültigen Standards. Je nach Datenlage sind verschiedene Werte ratsam. Cochran und Rubin [2] empfehlen eine Festlegung auf der Basis der berechneten Standardabweichungen der Logit-PS in beiden Gruppen.

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 [6].

2.2.3 PS-Verteilung und Balance Check

Ein wichtiges Ziel des Greedy und Optimal Matchings ist es, eine gute Qualität im PS-Matching zu erreichen. Hier ist sicherzustellen, dass vor dem Matching die Bedingung des so genannten Common Supports erfüllt ist. Der Common Support ist der Bereich, in dem die PS-Werte sowohl in der Fall- als auch in der Kontrollgruppe vorkommen. Folglich entspricht die grau schraffierte Fläche in Abbildung 1 den PS-Werten außerhalb des Common Supports.

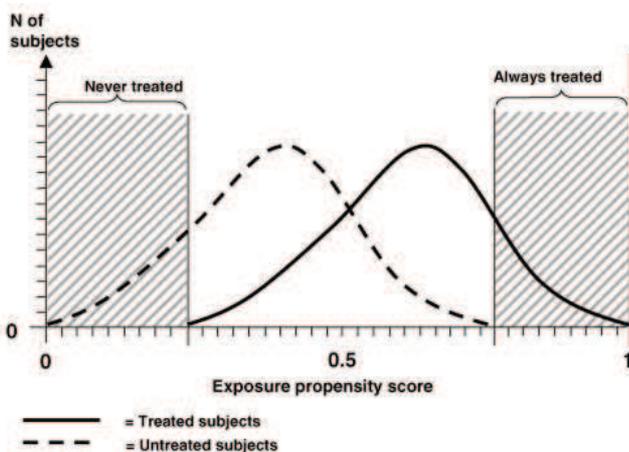


Abbildung 1: Common Support der PS-Verteilung [5]

Je ähnlicher die beiden Kurven verlaufen, desto bessere Matchingpartner können gefunden werden. Dies ist mit der Annahme begründet, dass sich bei Individuen mit hoher

PS-Übereinstimmung, die Verteilung der Kovariaten ähnelt [6]. Hierbei stellt sich die Frage, in welchem Ausmaß die PS-Verteilungen der Fall- und Kontrollgruppe überlappen müssen, damit ein vernünftiges Matching möglich ist. Dazu kann der Anteil der Individuen berechnet werden, welche außerhalb des Common Supports liegen [10]. In der Literatur ist aber kein allgemeingültiger Grenzwert vorhanden, mit dem analysiert werden kann, ab welchem Wert das Matching nicht mehr sinnvoll ist. Alternativ kann die standardisierte Differenz d -in Abhängigkeit des Merkmalstyps- für alle Kovariaten berechnet werden, die für die PS-Schätzung verwendet wurden. Für die (absolute) standardisierte Differenz 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. [8] haben wiederum vorgeschlagen, dass ein Wert von $d < 0.1$ auf ein vernachlässigbares Ungleichgewicht der Kovariate zwischen Fall- und Kontrollgruppe hinweisen könnte [1]. Diese Werte gehen auf Cohen [3] 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.

2.3 PS-Matching in SAS

Zum Thema Matching in SAS sind zahlreiche Makros im Internet frei zugänglich. Hauptsächlich sind Makros zum Greedy Matching vertreten. In SAS selbst kann das Optimal Matching mit der Prozedur PROC OPTNET realisiert werden.

Das Ziel dieses Beitrags war es einen zweistufigen Matching-Algorithmus für Beobachtungsdaten aus verschiedenen Registern zu entwickeln, in dem das Exact Matching mit dem Optimal Matching kombiniert wird. Da die verfügbaren Web-Makros und SAS-Prozeduren diese Voraussetzungen nicht erfüllen, wurde ein eigenes Makro implementiert. Dieses Makro %optimal hat folgende Parameter:

```
%MACRO optimal (data, id, strata, m2, m1=NONE, anz_match=1, distance=EUCLID, caliper=0, class=NONE, weight=NONE, max_dist=NONE, ratio=VAR, register=NONE).
```

Die Übergabeparameter sind im Einzelnen:

data	Eingabedatensatz für das Matching
id	Eindeutige Identifikationsnummer
strata	Indikatorvariable zur Unterscheidung der Fall- und Kontrollgruppe (1=Fall, 0=Kontrolle)
m2	Auflistung der Matchingvariablen für das Optimal Matching
m1	Auflistung der Matchingvariablen für das Exact Matching, Default NONE (kein Exact Matching)
anz_match	Anzahl der Kontrollen, welche jeweils einem Fall zugeordnet werden sollen (1:k-Matching), Default 1 (1:1-Matching)

distance	Bestimmung des Distanzmaßes (ABS: Absolute Distanz, EUCLID: Euklidische Distanz, MAH: Mahalanobis Distanz, PS: Propensity Score Distanz), Default EUCLID
caliper	Maximal zulässige PS-Distanz zwischen zwei Personen der Fall- und Kontrollgruppe, Default 0 (kein Caliper Matching)
class	Angabe der kategorialen Variablen der Matchingvariablen m2 zur Berechnung der Logit-PS mittels logistischer Regression, Default NONE (kein Class Statement)
weight	Gewichtung der Distanz zwischen zwei Personen der Fall- und Kontrollgruppe der jeweiligen Matchingvariablen m2, Default NONE (keine Gewichtung)
max_dist	Maximal zulässige Distanz zwischen zwei Personen der Fall- und Kontrollgruppe der jeweiligen Matchingvariablen m2, Default NONE
ratio	Ratio Matching (CONST: konstantes Ratio, VAR: variables Ratio), Default VAR
register	Angabe der Registervariable, Default NONE (keine verschiedene Registerdaten)

3 Ergebnisse

Die Anwendung des oben beschriebenen Matching-Ansatzes basiert auf einem retrospektiven Datensatz aus der Notfallmedizin und einer fiktiven Fragestellung, welche ein Fall-Kontroll-Matching notwendig macht. Hierzu wurde der ursprüngliche Datensatz von Notarzteinsätzen auf 30,000 Beobachtungen verkleinert. Um die Situation unterschiedlicher Register als mögliche Datenquellen abzubilden, wurden die Patienten per Zufall einem von drei fiktiven Registern zugeordnet. Hierbei wurde zudem die Größe der fiktiven Register variiert, so dass schlussendlich die Register 1 und 3 mit ca. 7,500 Beobachtungen deutlich kleiner waren als Register 2 mit ca. 15,000 Beobachtungen.

Ein übergeordnetes Ziel der Analyse könnte beispielsweise ein Vergleich zweier Patientengruppen mit und ohne Verbrennungen in Bezug auf die Schwere der Bewusstseinsstörung -gemessen über die Glasgow-Coma-Scale (GCS)- sein. Für diesen Zweck kann unter Verwendung des vorgestellten Matching-Makros ein 1:k-Matching vorgenommen werden (1 Fallpatient mit Verbrennung zu k Kontrollpatienten ohne Verbrennung). Die Prävalenz von Verbrennungen in allen 3 Registern war praktisch gleich (zwischen 2.2% und 2.6%).

Als Hauptmatchingkriterien wurden das Geschlecht und die Hauptdiagnose festgelegt. Die Verteilung beider Variablen war wiederum in allen drei Registerquellen vergleichbar. Weitere Matching-Variablen, welche zur Berechnung der PS verwendet wurden, waren Alter, Bewusstseinsstatus, Traumamuster, Polytrauma und Opiatgabe. Der gesamte Matching-Prozess erfolgte in Anlehnung an das in Abbildung 2 dargestellte Schema.

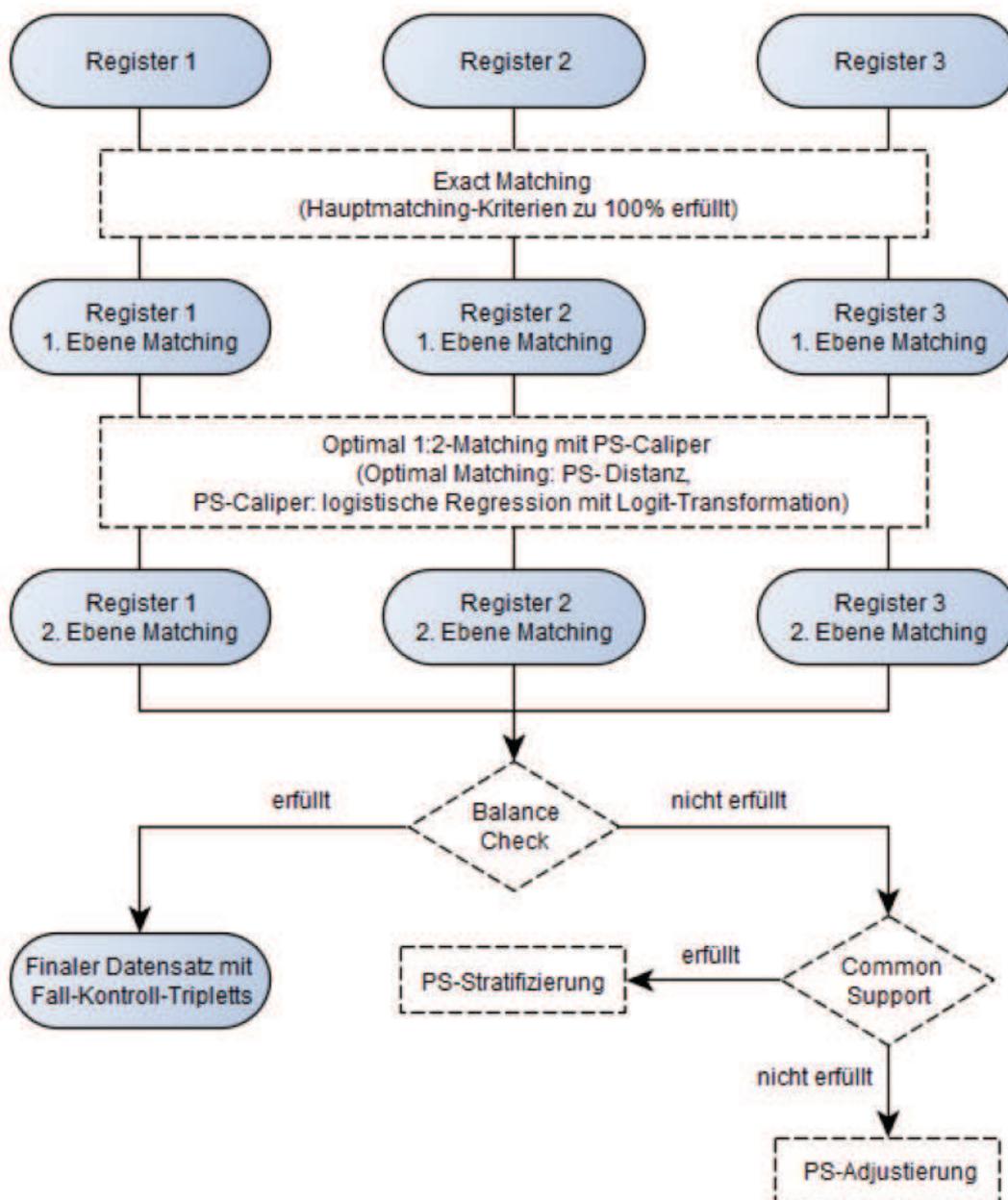


Abbildung 2: Zweistufiger Matching-Algorithmus auf der Basis von PS

Wie dem Algorithmus zu entnehmen ist, spielt der im letzten Kapitel vorgestellte Balance Check eine zentrale Rolle für die Auswahl der letztendlichen Auswertung. Diese basiert nur dann auf den angestrebten Fall-Kontroll-Triplets, wenn der Balance Check erfolgreich war. Alternative Auswertungsstrategien sind durch eine Subgruppen-spezifische Auswertung in PS-Strata bzw. durch ein PS-adjustiertes Regressionsmodell gegeben.

Der letzte Makro-Aufruf für das hier vorgestellte Beispiel lautet wie folgt:

```

%optimal(data=trauma, strata=verbrennung, id=idnr, m2=alter bewusstsein traumamust polytrauma
opiate, m1=regnr geschlecht hauptdiagnosen, distance=PS, caliper=1, anz_match=2, register=regnr,
ratio=CONST, class=bewusstsein traumamust polytrauma opiate);
  
```

Das exakte Matching der ersten Stufe konnte mit den beiden definierten Hauptkriterien gut durchgeführt werden. Alternative Ansätze mit mehr als zwei Hauptkriterien lieferten jedoch sehr schlechte Ergebnisse in dem Sinne, dass für viele Fall-Patienten keine geeigneten Kontrollen mehr gefunden werden konnte. Die Tatsache, dass nach Möglichkeit Fall- und Kontrollpatienten aus demselben Register stammen sollen, wurde durch die Festlegung der Registernummer als m1-Variable realisiert.

Die PS-Berechnung in einem zweiten Schritt erfolgte im Rahmen des %optimal-Makros mittels logistischer Regression auf der Basis der angegebenen m2-Variablen. Als maximal zulässige Distanz zwischen Fällen und Kontrollen wurde ein PS von 1 angegeben. Außerdem wurde die Bedingung eines konstanten Fall-Kontroll-Matchings übergeben (durch die Makro-Variable ratio), d.h. das über anz_match festgelegte 1:2-Matching gilt für alle Fallpatienten.

Der Common Support war in allen drei Registern zufriedenstellend. Stellvertretend für die Analysen des Balance Checks sind in Abbildung 3 die PS-Verteilungen der Fall- und Kontrollgruppen aus Register 3 gegeben.

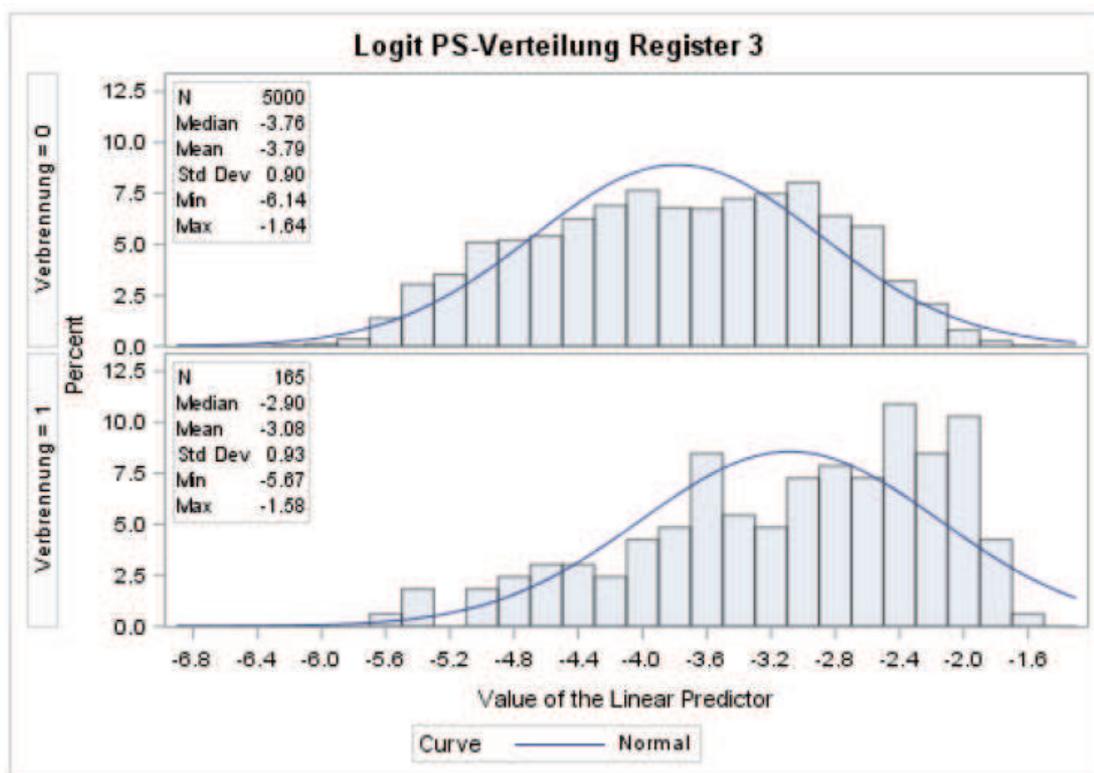


Abbildung 3: Logit PS-Verteilung vor dem optimal-Matching in Register 3

In Erweiterung der allgemeinen Betrachtung zum Common Support auf Basis der PS-Verteilungen kann zudem für jede der PS-relevanten m2-Variablen ein Balance Check erfolgen. Für jede dieser Variablen kann eine standardisierte Differenz zwischen Fall- und Kontrollgruppe vor und nach dem Matching berechnet werden. Wie die Ergebnisse in Tabelle 1 zeigen, konnten die zuvor teils schlechten Werte für d durch das Matching auf einen akzeptablen Wert gebracht werden.

Tabelle 1: Standardisierte Differenzen d vor und nach dem Matching zwischen Fällen und Kontrollen für die eingeschlossenen m2-Variablen

m2-Variable	Register	d_{vor}	d_{nach}
Alter	1	0.97	0.09
	2	1.04	0.09
	3	0.94	0.10
Bewusstsein	1	0.35	0.05
	2	0.36	0.07
	3	0.39	0.02
Traumamuster	1	0.55	0.01
	2	0.53	0.02
	3	0.55	0.01
Polytrauma	1	0.12	0.01
	2	0.09	0.05
	3	0.06	0.03
Opiate	1	0.16	0.05
	2	0.00	0.01
	3	0.05	0.05

4 Diskussion

Das in diesem Beitrag vorgestellte SAS-Makro zur Durchführung eines 1:k-Matchings an Beobachtungs- bzw. Registerdaten ermöglicht eine flexible Steuerung des Zuordnungsprozesses von Fällen und Kontrollen. Es umfasst alle gängigen Parameter und Ansätze, die in der Vergangenheit zum Thema Matching entwickelt und vorgeschlagen wurden. Ein spezieller Fokus dieses Beitrags lag auf der Durchführung eines PS-gestützten Matchings, bei dem der aus den Daten berechnete PS als Distanzmaß zur Festlegung der Fall-Kontroll-Partner fungierte. Alternative Distanzmaße, wie beispielsweise die euklidische Distanz oder die Mahalanobis-Distanz, weisen gegenüber dem PS-Ansatz in manchen Situationen gewisse Vorteile auf, die jedoch nicht pauschal eine besondere Relevanz haben, sondern vielmehr in Abhängigkeit der zugrunde liegenden Daten bzw. Fragestellungen als wichtiger oder unwichtiger eingestuft werden müssen. So ist ein Matching auf Basis der PS-Distanz z.B. sehr gut geeignet,⁷ um insgesamt (d.h. bezogen auf alle betrachteten m2-Variablen) eine gute Balance zwischen Fall- und Kontrollgruppe zu generieren. Trotz eines umfangreichen Probedatensatzes war die Rechenzeit der Prozedur PROC OPTNET, welche den Kern des Makros bildet und ab der Version 9.3 mit Release Version 12.1 zur Verfügung steht, ein zu vernachlässigender Faktor. Für eine noch tiefer gehende Automatisierung des Balance Checks, der jedoch bislang als separates Makro implementiert ist und nicht in dem vorgestellten %optimal-Makro integriert ist, sind jedoch noch weitere Anpassungen der ausführbaren Analysen notwendig.

Literatur

- [1] P. C. Austin. Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research. In: *Communications in Statistics - Simulation and Computation* 2009, 38(6), S. 1228–1234.
- [2] W. G. Cochran und D. B. Rubin. Controlling bias in observational studies: A review. In: *Sankhya: The Indian Journal of Statistics, Series A* 1973, 35(4), S. 417–446.
- [3] J. Cohen. *Statistical power analysis for the behavioral sciences*. 2. Auflage. Hillsdale, NJ: Erlbaum, 1988. ISBN: 0-8058-0283-5.
- [4] A. Deckert. 1:N Matching von Fällen und Kontrollen: Propensity Score vs. PROC SQL. In: *KSFE 2011 - Voneinander lernen. Proceedings der 15. Konferenz der SAS®-Anwender in Forschung und Entwicklung (KSFE)*. Hrsg. von C. Ortseifen, H. Ramroth, M. Weires und R. Minkenber. Aachen: Shaker, 2011, S. 105–119. ISBN: 978-3-8440-0379-6.
- [5] R. J. Glynn, S. Schneeweiss und T. Stürmer. Indications for propensity scores and review of their use in pharmacoepidemiology. In: *Basic & Clinical Pharmacology & Toxicology* 2006, 98(3), S. 253–259.
- [6] S. Guo und M. W. Fraser. *Propensity score analysis: statistical methods and applications*. Los Angeles, London: Sage, 2010. ISBN: 978-1-4129-5356-6.
- [7] C. E. Müller. *Quasiexperimentelle Wirkungsevaluation mit Propensity Score Matching: Ein Leitfaden für die Umsetzung mit Stata*. Saarbrücken, Centrum für Evaluation, CEval-Arbeitspapier 19, 2012.
- [8] 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. In: *Journal of Clinical Epidemiology* 2001, 54(4), S. 387–398.
- [9] L. S. Parsons. Performing a 1:N case-control match on propensity score. In: *Proceedings of the Twenty-Ninth Annual SAS® Users Group International Conference*. Cary, NC: SAS Institute Inc., Posters: Paper 165-29, 2004.
- [10] E. Reinowski. *Matching kleiner Stichproben: Ein Vergleich verschiedener Verfahren*. Martin-Luther-Universität Halle-Wittenberg, Juristische und Wirtschaftswissenschaftliche Fakultät, Dissertation, 2008.
- [11] P. R. Rosenbaum. *Design of observational studies*. New York: Springer, 2010. ISBN: 978-1-4419-1212-1.
- [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] P. R. Rosenbaum und D. B. Rubin. The central role of the propensity score in observational studies for causal effects. In: *Biometrika* 1983, 70(1), S. 41–55.

- [14] SAS Institute Inc. SAS/STAT® 14.1 User's Guide: The PRINCOMP Procedure.
- [15] E. A. Stuart und D. B. Rubin. Best practices in quasi-experimental designs: Matching methods for causal inference. In: Best Practices in Quantitative Methods. Hrsg. von J. W. Osborne. Thousand Oaks, CA: Sage, 2008, S. 155–176. ISBN: 978-1-4129-4065-8.