

Dorothea Weber
Dr. sc. hum.

Applying Matching Procedures in the Generation and Synthesis of Evidence

Fach/Einrichtung: Medizinische Biometrie u. Informatik

Doktorvater: Prof. Dr. sc. hum. Meinhard Kieser

Der Goldstandard im Rahmen klinischer Studien ist eine doppelt-verblindete, randomisierte Studie. Es gibt jedoch Situationen, die ein solches Design aus ethischen oder praktischen Gründen nicht zulassen. Eine Möglichkeit dennoch einen Vergleich durchzuführen, ist die Verwendung einer externen Kontrollgruppe, diese kann aus einer bereits durchgeführten Studie oder auch aus einem Register stammen. Zieht man eine historische Kontrollgruppe heran, so sind die Studienpopulationen oft nicht vergleichbar. Matchingverfahren können dazu beitragen trotzdem einen Vergleich mit balancierten Patientengruppen durchführen zu können. In dieser Arbeit wird ein adaptives gematchtes Fall-Kontroll-Studiendesign entwickelt, welches eine Zwischenauswertung mit Fallzahlrekalkulation vorsieht. Die Rekalkulation der Fallzahl basiert auf der unteren Grenze des Konfidenzintervalls der Matchingrate bei Zwischenauswertung. Der Schätzer der Matchingrate wird in diesem neuen Ansatz mittels einer dem Bootstrap-ähnlichen Methode auf der Basis von gleich großen Gruppen bestimmt. Dieser Ansatz zeigt gute Ergebnisse in Bezug auf Power und dem Fehler erster Art im Vergleich zu der Herangehensweise unter Verwendung aller Kontrollpatienten bei der Zwischenauswertung. Zudem wurde der Zeitpunkt der Zwischenauswertung untersucht. Simulationen kamen zu dem Ergebnis, dass eine Zwischenauswertung nach Rekrutierung von $1/2$ bis $2/3$ der Zahl an Kontrollpatienten eine gute Wahl ist. Jedoch scheint es, als wäre der Zeitpunkt mehr von der absoluten Zahl an Patienten als vom Anteil abhängig. Ist die Kontrollgruppe jedoch sehr groß ist die zuvor besprochene Methode nicht die beste Wahl. Vielmehr ist es möglich mehr als nur einen Kontrollpatienten pro Interventionspatienten zu matchen. Um diese Anzahl an Matchingpartnern zu bestimmen kann das entwickelte iterative Verfahren verwendet werden. Die Idee basiert auf einer Erhöhung der Matchingpartner und gleichzeitiger Bestimmung der Matchingrate während einer Zwischenauswertung. Die Anzahl wird solange die 1:M Matchingrate oberhalb der 1:1 Matchingrate abzüglich einer definierten Toleranz liegt, erhöht. Die entsprechende 1:M Matchingrate kann dann noch für eine Fallzahlrekalkulation genutzt werden. Das iterative Verfahren ist einfach zu implementieren und lässt eine Kombination mitvielen Studiendesigns zu. Es muss beachtet werden, dass die Anzahl der Matchingpartner maßgeblich von der Überlappung der Patientenkollektive abhängt, d.h. liegt nur wenig Überlappung vor können dementsprechend nur wenige Matchingpartner gefunden werden und umgekehrt. Daraus folgt, dass die studienspezifische Matchingrate für eine Rekalkulation der Fallzahl von Nutzen sein kann, um die Power zu erhöhen. Nicht nur bei der Generierung von Evidenz treten unbalancierte Populationen auf, auch in der Synthese von Evidenz spielt dieses Problem eine Rolle. Eine häufige Situation der Evidenzsynthese ist ein indirekter Vergleich. Dieser ist von Interesse, wenn zwei Medikamente, angenommen A und C, nicht in einem direkten Vergleich gegenübergestellt wurden, jedoch sind Vergleiche der Medikamente A und B sowie C und B verfügbar. Es ist denkbar, dass die Studien AB und CB unterschiedliche Patientenpopulationen betrachten. Liegen in der Situation eines indirekten Vergleichs von einer Studie individuelle Patientendaten vor, kann dieses Ungleichgewicht von Matchingverfahren adressiert werden. Neben der Methode von Bucher ist der Matching Adjusted Indirect Comparison eine etablierte Methode für indirekte Vergleiche, dieser ist jedoch noch nicht hinreichend für praxisrelevante Situationen untersucht. Um diese beiden Methoden gegenüberzustellen und näher in praxisrelevanten Situationen (z.B. verletzte

Annahmen) zu untersuchen wird eine Simulationsstudie durchgeführt. Simulationen führen zu der Vermutung, dass indirekte Vergleiche deutlich zu wenig Power haben, um einen Effekt nachweisen zu können. Wenn sich die Studienpopulationen unterscheiden, Effektmodifikation auftritt und Regressionsmodelle nicht ausreichend adjustiert sind, erreicht man mit dem Matching-Ansatz eine höhere Konfidenzintervallüberdeckung sowie Power. Wird allerdings für Kovariaten gematcht, die den Effekt nicht beeinflussen, so führt das zu einer höheren Verzerrung, die Matchingvariablen sollten deshalb mit Bedacht gewählt werden. Bisher wurde nur eine Studie pro direktem Vergleich verwendet, da Matching Adjusted Indirect Comparisons für diese Situation entwickelt wurden. Es ist jedoch wahrscheinlich, dass mehrere Studien den gleichen Vergleich durchgeführt haben. Kombiniert man Evidenz, sollten immer alle verfügbaren Studien verwendet werden. Um eine Vielzahl von Studien in indirekte Vergleiche einbinden zu können, werden verschiedene Ansätze entwickelt und verglichen. Diese beinhalten alle einen Schritt für die Synthese der Evidenz und einen für den indirekten Vergleich. Der Hauptunterschied zwischen den Ansätzen ist die Reihenfolge dieser beiden Schritte. Berücksichtigt man eine höhere Zahl an Studien kann die gewünschte Power von 80% erreicht werden. Es war allerdings nicht möglich eine Methode zu identifizieren, die in allen betrachteten Szenarien die besten Ergebnisse erzielt. Folglich muss die zugrunde liegende Situation sorgfältig analysiert werden, sowie die Methodik und die Matchingvariablen mit Bedacht gewählt werden.