

Tim Friede
Dr. sc.hum.

Methoden zur Bestimmung der Fallzahl in klinischen Studien mit interner Pilotstudie

Geboren am 31.12.1971 in Oldenburg i.O.
Reifeprüfung am 30.05.1991 in Geesthacht / Elbe
Studiengang der Fachrichtung Mathematik vom WS/SS 1992 bis WS/SS 1998
Vordiplom am 08.07.1994 an der Universität Oldenburg
Diplom am 03.12.1998 an der Universität Karlsruhe

Promotionsfach: Medizinische Biometrie und Informatik
Doktorvater: Prof. Dr. N. Victor

Aus ethischen, ressourcen-ökonomischen und zeitlichen Aspekten ist man bestrebt, die Fallzahl in klinischen Studien möglichst klein zu halten. Andererseits ist es im Interesse des Untersuchers, bei Vorliegen eines klinisch relevanten Effektes, die Überlegenheit der experimentellen Therapie mit hoher Wahrscheinlichkeit (z.B. 80% oder 90%) nachzuweisen.

Bei der Planung einer klinischen Studie besteht häufig jedoch eine große Unsicherheit bezüglich der Streuungen in den Behandlungsgruppen. Da das Vorwissen über die Größe der Variabilität aber für die Fallzahlplanung unabdingbar ist, findet die Fallzahlplanung unter großer Unsicherheit statt. Studiendesigns, die eine Fallzahlanpassung auf Grundlage einer im Studienverlauf geschätzten Varianz erlauben, stellen daher eine Lösung dar.

Designs mit Fallzahlanpassung im Studienverlauf lassen sich in Designs mit Zwischenauswertung und Designs mit Data Review (auch Designs mit interner Pilotstudie genannt) unterscheiden. Bei ersteren wird nach Rekrutierung der ersten Patienten der Behandlungseffekt geschätzt und ein statistischer Hypothesentest für den Behandlungseffekt durchgeführt. Diese Designs offerieren die Möglichkeit, die Studie vorzeitig zu beenden. Wenn die Studie mit einem zweiten Abschnitt fortgesetzt wird, können alle Informationen aus dem ersten Abschnitt zur Planung des zweiten verwendet werden. Bei den Designs mit interner Pilotstudie wird nicht der Behandlungseffekt, sondern lediglich Störparameter wie die Varianz innerhalb der Behandlungsgruppen geschätzt, um dann auf Grundlage der Varianzschätzung die Fallzahl anzupassen. Data Reviews können unverblindet oder verblindet durchgeführt werden.

In der vorliegenden Arbeit wird zunächst die Prozedur zur verblindeten Fallzahlanpassung nach Gould und Shih untersucht. Dieses Verfahren nimmt einen nicht unbedeutenden Stellenwert in der Praxis ein, was sich z.B. in der großen Anzahl von Publikationen zu diesem Verfahren ausdrückt. Anhand eines simulierten Datensatzes werden drei wesentliche Schwächen der auf einem EM-Algorithmus basierenden Prozedur offengelegt: (i) Die Schätzungen hängen von der Initialisierung ab, (ii) das Konvergenzkriterium stellt keine Stabilisierung des Algorithmus sicher, (iii) es wird implizit von der Münzwurfrandomisierung ausgegangen, und das Wissen über die durchgeführte Randomisierung wird nicht berücksichtigt. Insgesamt kann vor einer Benutzung des Verfahrens in einem Blind Data Review nur gewarnt werden.

Es werden weitere Verfahren sowohl zur entblindeten wie auch verblindeten Varianzschätzung vorgestellt. Diese Prozeduren werden durch analytische Überlegungen und eine

Simulationsstudie miteinander verglichen. Es zeigt sich, daß die Prozeduren, die eine Entblindung notwendig machen, den verblindeten Prozeduren nicht deutlich überlegen sind. Somit scheint eine Entblindung nicht gerechtfertigt zu sein, zumal bei Berechnungen des tatsächlichen Signifikanzniveaus für die verblindeten Prozeduren keine Überschreitung des nominalen Niveaus beobachtet wird.

Designs mit und ohne der Möglichkeit zum vorzeitigen Studienende werden durch analytische Betrachtungen des Designs nach Wittes und Brittain und des Designs nach Bauer und Köhne verglichen. Bei gleicher erwarteter Power kommt das Design nach Bauer und Köhne in der Regel mit einer geringeren erwarteten Fallzahl aus.

Zusammenfassend können folgende Aussagen gemacht werden. Für eine klinische Studie, bei deren Durchführung eine Zwischenauswertung aus ethischen Gründen nicht zwingend erforderlich ist, sollte man entweder ein Design mit interner Pilotstudie und verblindeter Fallzahlanpassung oder ein Design mit Zwischenauswertung wählen. Das Design mit interner Pilotstudie und entblindetem Data Review hat keine deutlichen Vorteile gegenüber den verblindeten Prozeduren gezeigt. Zudem ist der vormalige Vorteil, die Auswirkungen auf die Fehlerwahrscheinlichkeit 1. Art bei der entblindeten Prozedur zu kennen, nun nicht mehr existent, da nun für die verblindeten Prozeduren nachgewiesen wurde, daß keine praktisch bedeutsame Überschreitung des nominalen Niveaus auftritt. Außerdem unterlag das Design mit interner Pilotstudie und entblindeter Fallzahlanpassung dem adaptiven Design nach Bauer und Köhne bezogen auf die erwartete Fallzahl. Als weiteres Ergebnis der vorliegenden Arbeit halten wir fest, daß die Prozedur zur verblindeten Fallzahlanpassung nach Gould und Shih nicht verwendet werden sollte.