

Kevin Kunzmann

Dr. sc. hum.

Optimal Adaptive Designs for Early Phase II Trials in Clinical Oncology

Fach/Einrichtung: Medizinische Biometrie und Informatik

Doktormutter: Prof. Dr. Annette Kopp-Schneider

Diese Arbeit stellt eine neuartige, schnelle und exakte Art der Berechnung von optimalen zweistufigen Designs für einarmige Studien mit binärem Endpunkt vor. Solche Designs werden häufig in frühen onkologischen Studien verwendet. In diesen Studien wird üblicherweise die Rate untersucht, mit der eine bestimmte Tumorart auf eine neuartige Therapie in der Zielgruppe anspricht (die sogenannte 'response rate' oder Erfolgsrate). Der Therapieerfolg wird dabei häufig gemäß der RECIST Richtlinien bestimmt. Diese neue Methode zur Berechnung solcher optimalen Designs erlaubt die Lösung von Problemen für jede praktisch relevante Studiengröße innerhalb kürzester Zeit. Dies bedeutet eine erhebliche Beschleunigung im Vergleich zu früheren Implementierungen und erleichtert somit die interaktive Untersuchung verschiedener Planungsszenarien. Klinischen Studien müssen üblicherweise unter erheblicher Unsicherheit bezüglich der relevanten Parameter (hier die tatsächliche Erfolgsrate) geplant werden. Dies trifft insbesondere auf Studien in frühen klinischen Phasen eines Entwicklungsprogramms zu, da hier in der Regel zum Planungszeitpunkt kaum datenbasierte Evidenz über die Erfolgsrate vorliegt. Methoden, die diese Planungsunsicherheit berücksichtigen können, sind daher von besonderem Interesse. Die vorliegende Arbeit entwickelt und präsentiert verschiedene Möglichkeiten solche Unsicherheiten in die Studienplanung zu integrieren und erweitert damit die klassische Methode zur Berechnung der benötigten Fallzahl, die üblicherweise von einer festen Punktoption ausgeht. Zu diesem Zweck wurde ein Bayes'scher Ansatz gewählt, der plausible Erfolgsraten gemäß ihrer Vorabwahrscheinlichkeit gewichtet. Im Allgemeinen führt dabei größere a priori Unsicherheit zu optimalen Designs mit weniger variablen Fallzahlen für die zweite Stufe des Designs. Die resultierenden Designs sind also robuster gegenüber Fehlspezifikationen der erwarteten Erfolgsrate. Die in dieser Arbeit entwickelten Konzepte lassen sich von der Situation der binären Endpunkte auf kontinuierliche Endpunkte übertragen. Obgleich dies nicht zentraler Bestandteil der vorliegenden Arbeit ist, wird die Vorgehensweise kurz anhand eines Beispiels erläutert. Kontinuierliche Endpunkte können in frühen onkologischen Studien zum Beispiel dann von Interesse sein, wenn anstelle eines binären Erfolgskriteriums gemäß der RECIST Richtlinien das tatsächliche Tumorzellvolumen (oder dessen Entwicklung im Zeitverlauf) betrachtet werden soll. Neben der optimalen Planung einer Studie ist die effektive Auswertung nach Abschluss der Datenerhebung entscheidend für den Studienerfolg. Die finale statistische Inferenz wird sowohl aus der Bayesianischen als auch aus der frequentistischen Perspektive betrachtet. Im Fall der frequentistischen Betrachtungsweise besteht eine wesentliche Herausforderung in der Definition von p Werten, die mit der Testentscheidung des zugrundeliegenden optimalen Designs in dem Sinne kompatibel sind, dass der p Wert genau dann kleiner als das gewählte Signifikanzniveau ist wenn die Nullhypothese abgelehnt werden kann. In der vorliegenden Arbeit wird argumentiert, dass eine Definition des p Wertes in Bezug auf „Test-kompatible“ Punktschätzer praktische Vorteile hat. Nur wenn p Werte auf Schätzern mit einer solchen Eigenschaft basieren, bleiben die Beziehungen zwischen Punktschätzern, p Werten, und

Konfidenzintervallen erhalten, die aus einstufigen Tests bekannt sind. Für die meisten praktisch sinnvollen Designs ist der klassische Maximum-Likelihood-Schätzer Test-kompatibel, allerdings nicht in allen Situationen. Die vorliegende Arbeit zeigt, wie Test-Kompatibilität eines Schätzers während der Planungsphase überprüft werden kann und beschreibt einen allgemeinen Weg, um kompatible Schätzer zu erhalten. Weiterhin wird argumentiert, dass unverzerrte Schätzer in zweistufigen Designs in der Regel ineffizient sind und nicht zu Test-kompatiblen p Werten führen.

Schließlich betrachtet diese Arbeit auch Methoden, um auf unvorhergesehene Änderungen der Planungsannahmen zu reagieren. Prinzipiell sollten optimale zweistufige Designs zwar nicht während einer laufenden Studie modifiziert werden, da sie die Fallzahl bereits optimal an die beobachteten Zwischenergebnisse anpassen. Eine Designanpassung kann jedoch nötig werden, falls etwa neue Studien-externe Informationen verfügbar werden, die den Planungsprior obsolet machen. Hierbei ist es wichtig nicht-vorgeplante Designanpassungen von der vorgeplanten Anpassung der Stichprobengröße zu unterscheiden. Bei letzterer bleibt das Design selbst unverändert. Die in der vorliegenden Arbeit beschriebenen Methoden decken also alle aus statistischer Sicht relevanten Projektphasen einer frühen einarmigen Studie mit binärem Endpunkt ab.

Software-Implementierungen aller diskutierten Methoden stehen zur Verfügung. Alle in dieser Arbeit vorgestellten Beispiele sind als interaktive „Notebooks“ online verfügbar.

Eine interaktive [Web-App](#) erlaubt das Experimentieren mit optimalen Designs, ohne jegliche Programmierkenntnisse.