

– Zusammenfassung –

Christopher Alexander Büsch
Dr. sc. hum.

Comparison of different additional benefit assessment methods for oncology treatments

Fach/Einrichtung: Medizinische Biometrie
Doktorvater: Prof. Dr. sc.hum. Meinhard Kieser

Bei der Entwicklung neuer Therapien für die Marktzulassung müssen verschiedene Anforderungen erfüllt werden. Nach der Zulassung wird zudem der Zusatznutzen mit bereits etablierten Behandlungen verglichen. Diese Bewertung kann über die Höhe der Erstattung der neuen Behandlung entscheiden und dazu beitragen, die Unsicherheit der Patienten hinsichtlich der Wirksamkeit und Toxizität zu beseitigen.

Im Falle von nicht heilbaren oder fortgeschrittenen Krankheiten wurden drei verschiedene Zusatznutzenbewertungsmethoden entwickelt, welche sich hauptsächlich in der verwendeten statistischen Metrik unterscheiden. Die Europäische Gesellschaft für Medizinische Onkologie (ESMO) und das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) haben Methoden mit einem ordinalen Ergebnis konstruiert. Für die primäre kategoriale Bewertung vergleicht IQWiG die obere Grenze des 95% Hazard Ratio (HR) Konfidenzintervall (KI) (HR^+) gegen Relative Risiko (RR) skalierte Schwellenwerte. ESMO verwendet die untere 95% HR-KI Grenze (HR^-). Die Methode der Amerikanischen Gesellschaft für Klinische Onkologie (ASCO) hat ein stetiges Ergebnis, welches den HR Punktschätzer (PS) verwendet. Es gibt mehrere Kritikpunkte für jede der Methoden. Zum Beispiel könnte die Verwendung des HR-PS dazu führen, dass Studien mit wesentlichem Zusatznutzen bestraft werden, weil die Varianz des Schätzers nicht betrachtet wird. Dagegen berücksichtigen HR^+ und HR^- die Variabilität vom HR-PS und sollten daher mehr Informationen enthalten.

Das Ziel dieser Thesis ist es ein besseres Verständnis der Methoden zu erlangen und die Frage zu beantworten, welche statistische Metrik bessere Eigenschaften besitzt, um den Zusatznutzen zu beurteilen. Um einen besseren Vergleich zwischen den verschiedenen Methoden zu ermöglichen wurde zudem untersucht, welche ESMO und IQWiG Kategorie welcher ASCO Punktzahl entspricht. Um diese Ziele zu erreichen, wurden die oben genannten Methoden mit Hilfe von umfangreichen Simulationsstudien evaluiert und miteinander verglichen.

Die ursprüngliche IQWiG Methode ($IQWiG_{HR}$) und ESMO zeigen einen großen positiven Zusammenhang bei moderaten Behandlungseffekten. ASCO/ $IQWiG_{HR}$ sowie ASCO/Mod- $IQWiG_{HR}$ (vorgeschlagene Änderung von $IQWiG_{HR}$, welche HR- anstatt RR-skalierte Schwellenwerte verwendet) zeigen hingegen einen großen positiven Zusammenhang über die gesamte Breite von Behandlungseffekten. Zudem vergibt ESMO die maximale Kategorie exzessiv über die gesamte Breite von Behandlungseffekten (in den meisten Szenarien über 80% der Fälle) und kann daher zwischen kleinen und großen Behandlungseffekten nicht unterscheiden. ASCO und $IQWiG_{HR}$ haben ein konservativeres Verhalten.

Verschiedene verletzte Annahmen wie über-/unterpowerete Studien, informative Zensierungen und nicht proportionale Hazards werden von ESMO nicht bestraft. Zudem sind die verwendeten Schwellenwerte von ESMO zu liberal, was zu hohen falsch-positiven Raten und einfach zu erreichender maximaler Kategorie führt. ASCO und $IQWiG_{HR}$ weisen dahingegen ein erwünschtes Verhalten auf. In den meisten Fällen mit verletzten Annahmen ist die maximale Kategorie bei Mod- $IQWiG_{HR}$ nicht exzessiv erhöht und daher eine bessere Lösung.

In anderen Szenarien, bei denen keine Annahmen verletzt sind, ist diese Methode aber zu konservativ, was zu einer niedrigen richtig-positiven Rate und geringer Rate an maximaler Kategorie führt. Bei Vorliegen von verschiedenen Ereigniszeitverteilungen, welche die proportionale Hazard Annahme einhalten und daher keine Veränderung der Resultate

gewünscht ist, zeigt ESMO weiterhin eine erhöhte Rate an maximaler Kategorie. ASCO, IQWiG_{HR}, und Mod-IQWiG_{HR} weisen nicht dieses unerwünschte Verhalten auf. Daher kann man schlussfolgern, dass ESMO die liberalste Methode ist.

Nichtsdestotrotz ist HR^- , welche ESMO unter anderem verwendet, die beste statistische Metrik um den Zusatznutzen zu beurteilen, wenn geeignete Schwellenwerte verwendet werden. Selbst HR-PS hat bessere Eigenschaften als HR^+ .

Um einen besseren Vergleich zwischen den verschiedenen Methoden zu ermöglichen, wurden ASCO Grenzwerte vorgeschlagen, welche den ESMO und IQWiG Kategorien entsprechen: Eine ASCO Punktzahl größer als 17, 20 und 23 entsprechen den ESMO Kategorien 2, 3 und 4. Eine ASCO Punktzahl größer als 20 (23) und 36 (45) entsprechen den IQWiG_{HR} (Mod-IQWiG_{HR}) Kategorien "minor", "considerable" und "major added benefit".

Die Ergebnisse der Dissertation zeigen, dass um die richtig-positive und falsch-positive Rate zu optimieren, HR^- anstatt HR^+ verwendet oder zu mindestens die derzeitigen Schwellenwerte angepasst werden sollten. Daher kann diese Forschung als Hilfestellung bei zukünftigen Verbesserungen der Zusatznutzenbewertungsmethoden verwendet werden und trägt zur Weiterentwicklung der Zusatznutzenbewertung bei.