

Institut für Medizinische Biometrie und Statistik, Universitätsklinikum Schleswig-Holstein,
Campus Lübeck, Universität zu Lübeck

Randomization In Treatment Arms (RITA): Ein Randomisierungs-Programm für klinische Studien

Randomization In Treatment Arms (RITA): A randomization program for clinical trials

Friedrich Pahlke*, Inke R. König*, Andreas Ziegler

Zusammenfassung

Ein Kernpunkt bei der Durchführung kontrollierter klinischer Studien ist die Randomisierung der Studienteilnehmer auf die verschiedenen Therapiegruppen. Zur korrekten Umsetzung, Dokumentation und Verwendung komplexerer Verfahren ist in der Regel der Einsatz von Software erforderlich. Es wurde ein neues Randomisierungsprogramm entwickelt, welches in diesem Beitrag beschrieben wird. Zunächst werden die Untersuchungen dargestellt, die dem Entwurf und der Implementation vorangingen. Insbesondere wird die Auswahl geeigneter Randomisierungsverfahren diskutiert. Beim Entwurf wurde auf die sehr gute Erweiterungsfähigkeit des Programms sowie seine einfache Handhabung in der Praxis Wert gelegt.

Schlüsselwörter

Randomisierung, kontrollierte klinische Studie, Java, Computer-Programm, Software-Ergonomie

Summary

A central issue in performing controlled clinical trials is the randomization of study participants in different treatment groups. In general, we utilize a software tool for correct implementation, documentation and utilization of more complex procedures. In this paper, we therefore describe the development of a new randomization program. We begin by presenting explorations that have preceded the actual program design and

* Beide Autoren haben zu gleichen Teilen zur Arbeit beigetragen.

implementation. In particular, we discuss the selection of suitable randomization procedures. In the design of the program, we have laid special emphasis on simple handling and the possibility to easily extend the program.

Key words

Randomization, controlled clinical trials, RCT (Randomized Clinical Trial), Java, computer program, software ergonomics

Einleitung

Die randomisierte kontrollierte Studie ist seit vielen Jahren zum Goldstandard der klinischen Forschung geworden (für eine Übersicht siehe z. B. ROSENBERGER und LACHIN, 2002). Die hierbei durchgeführte zufällige Zuteilung der Patienten auf die untersuchten Behandlungsgruppen führt zu verschiedenen Vorteilen (KRACHT, 1992; SCHULZ und GRIMES, 2002a). Zum einen stellt die Randomisierung die geeignetste Möglichkeit dar, Verzerrungen in der Zuordnung zu vermeiden und strukturgleiche Behandlungsgruppen zu erhalten. Durch eine Stratifizierung ist dies nur bedingt möglich, da hier nur bekannte prognostische Faktoren berücksichtigt werden können (SCHÄFER, 1997). Der zweite Vorteil der Randomisierung liegt darin, dass eine Verblindung der Untersucher und Patienten hinsichtlich ihrer Gruppenzugehörigkeit erleichtert wird.

Entsprechend dieser Vorteile wird Randomisierung in der Guideline ICH E6 als die wichtigste Designtechnik zur Reduktion von Verzerrungen hervorgehoben (ICH E6, 1996). Auch im revidierten CONSORT-Statement wird gefordert, dass das Randomisierungsverfahren detailliert in der Publikation beschrieben wird (MOHER et al., 2001).

In der Praxis wird die Randomisierung bei einer klinischen Studie üblicherweise durch die Anwendung von Computerprogrammen unterstützt. Dies verringert zum einen bei einer großen Studie erheblich den Aufwand und macht den Randomisierungsvorgang leichter nachvollziehbar. Zum anderen wurden komplexe Randomisierungstechniken vorgeschlagen, die ohne adäquate Software nicht praxistauglich wären.

Trotz dieses offensichtlichen Bedarfs stehen bislang kaum auf Randomisierung spezialisierte Programme zur Verfügung. Verschiedene bekannte Verfahren sind zwar teilweise in kommerziellen Programm-Paketen implementiert oder werden im Internet zur Verfügung gestellt – eine Übersicht findet sich beispielsweise unter <http://www.sghms.ac.uk/depts/phs/guide/randser.htm>. Der Nutzen dieser Programme ist aber deutlich eingeschränkt. So ist hier beispielsweise nicht die Auswahl zwischen verschiedenen Randomisierungsverfahren oder verschiedenen Zufallszahlengeneratoren möglich.

Aus diesen Gründen wurde entschieden, ein eigenes neues flexibles Randomisierungsprogramm zu entwickeln. In der vorliegenden Arbeit wird die Entwicklung von RITA (**R**andomization **I**n **T**reatment **A**rms) beschrieben, das sowohl gängige als auch seltener eingesetzte Verfahren mit geringer Einarbeitungszeit für kontrollierte klinische Studien zur Verfügung stellt. RITA erlaubt eine sehr flexible Konfiguration der Verfahren. So ist eine patientenweise Randomisierung möglich. Es lassen sich aber auch beliebig viele Patienten en bloc, d. h. auf einmal randomisieren; für jede Gruppe lässt sich der gewünschte Umfang als prozentualer Anteil des Gesamtumfangs oder als tatsächlicher Umfang angeben; außerdem stehen verschiedene Zufallszahlen- und Seed-Generatoren zur Auswahl.

RITA ist aufgrund permanenter Anzeige der gewählten Konfigurationen und eingegebenen Parameter in einem Informationsfenster und integrierter Hilfe leicht zu handhaben. Alle Ergebnisse lassen sich exakt nachvollziehen, da das Programm zum einen das Abspeichern, Öffnen und Drucken der Konfigurationen und Ergebnisse erlaubt, zum ande-

ren die Reproduktion der exakten Ergebnisse durch Eingabe weniger Parameter (Randomisierungsverfahren mit Parametern, Seed, Zufallszahlen-Generator) ermöglicht.

RITA läuft prinzipiell Plattform-unabhängig, also problemlos unter Windows (98, NT, 2000, XP), Linux, Unix und Solaris; erfolgreiche Tests erfolgten bislang unter Windows 98, Windows XP und Windows 2000. Die ansonsten üblichen Mindestanforderungen der Systeme sind auch für RITA ausreichend, so beispielsweise unter Windows 98 ein Prozessor mit 200 MHz bzw. 400 MHz und 128 MB bzw. 192 MB Arbeitsspeicher. Zur zusätzlichen Durchführung von Simulationen ist allerdings ein schnellerer Prozessor mit zwei oder mehr GHz und ein Arbeitsspeicher mit mindestens 512 MB empfehlenswert.

Im ersten Teil der Arbeit werden die verschiedenen Eigenschaften von Randomisierungsverfahren beschrieben, aufgrund derer eine Auswahl getroffen wurde. Anschließend werden die ausgewählten Verfahren vorgestellt und deren prinzipielle Funktionsweise anhand von Beispielen erklärt.

Im zweiten Teil werden die Funktionen beschrieben, die in RITA aufgenommen wurden. Im dritten Teil werden die grundlegenden Anforderungen an ein Randomisierungsprogramm sowie deren Realisierung in RITA dargestellt.

Eigenschaften von Randomisierungsverfahren

Für den Entwurf von RITA wurde zunächst untersucht, welche Verfahren in eine Standard-Software zur Randomisierung aufgenommen werden sollten. Dazu wurden die verschiedenen in Frage kommenden Randomisierungsverfahren hinsichtlich der nachfolgend beschriebenen Eigenschaften untersucht.

Zufällige Verteilung

Allgemein bedeutet Randomisierung, dass jeder Studienteilnehmer mit derselben Wahrscheinlichkeit einer der Therapiegruppen zugeteilt wird. Bei zwei Therapiegruppen könnte das beispielsweise dadurch erreicht werden, dass für jeden neuen Patienten dieselbe Münze geworfen wird. Damit ist das Ergebnis der Randomisierung weder für den Patienten noch für den Studienleiter vorhersagbar.

Das Ziel eines jeden Randomisierungsverfahrens ist es, die Patienten zufällig auf die verschiedenen Therapiegruppen aufzuteilen; nur so wird die methodische Voraussetzung für eine korrekte Bewertung von Diagnostik, Therapie und Prävention erfüllt (KUNDT, 2002).

Balanceverhalten

Balancierung bedeutet, die Umfänge der verschiedenen Therapiegruppen anzugleichen. Das Balanceverhalten eines Randomisierungsverfahrens beschreibt also, wie gut das Verfahren dieses Angleichen beherrscht. Bei einigen Randomisierungsverfahren kann das Balanceverhalten durch die Wahl von Parametern gesteuert werden, damit die Therapiegruppenumfänge am Ende der Randomisierung bzw. zum gewünschten Zeitpunkt den Anforderungen gerecht werden. Dies kann von entscheidender Bedeutung sein, um Festlegungen im Studienplan einzuhalten.

Beispielsweise könnte der Studienplan vorsehen, dass die absolute Differenz der Therapiegruppenumfänge einen bestimmten Wert nicht überschreiten darf, wobei sich die absolute Differenz aus der Subtraktion des Umfangs der kleinsten Therapiegruppe vom Umfang der größten Therapiegruppe ergibt.

Das Balanceverhalten eines Randomisierungsverfahrens ist besonders bedeutsam bei Studien mit geringen Umfängen (vgl. ICH E9, 1998), da hier Unterschiede in den Gruppen-

größen stärker ins Gewicht fallen und größere Einbußen in der Teststärke hingenommen werden müssen als bei großen Stichproben. Allerdings schränken viele Verfahren die zufällige Verteilung ein, um damit ein besseres Balanceverhalten zu erreichen. Diese Einschränkung des Zufalls als Kompromisslösung wird teilweise kritisch gesehen. Insbesondere die weit verbreitete Anwendung von Blockverfahren mit fest gewählter Blocklänge wird kritisiert, da innerhalb der Randomisierungsphase unter bestimmten Umständen Zuweisungen exakt vorhergesagt werden können und somit eine zufällige Zuweisung des Studienteilnehmers auf eine Therapiegruppe nicht mehr vorhanden ist (SCHULZ und GRIMES, 2002b).

Dieser Fall kann beispielsweise auftreten, wenn bei zwei Gruppen (Prüf- und Standardtherapie, P und S) eine Blocklänge von $B = 4$ vorgesehen ist. Angenommen, dreimaliges Randomisieren habe die Gruppen P S P ergeben. Der nächste Patient würde in diesem Fall automatisch der Standardtherapie zugeordnet werden. Noch extremer wäre die Einschränkung, wenn zweimaliges Randomisieren das Ergebnis P P ergeben hätte; nun würden sogar die nächsten beiden Patienten mit Sicherheit der Standardtherapie zugeteilt werden.

Es existieren verschiedene Lösungsvorschläge, die die Möglichkeit ausschließen, dass beispielsweise der Studienleiter innerhalb der Randomisierungsphase einzelne Randomisierungsschritte mit sehr hoher Wahrscheinlichkeit vorhersagen kann. Nur noch vollständig zu randomisieren, ist der radikalste Weg, da er das Balanceverhalten völlig außer acht lässt. Geeigneter erscheint es, Verfahren zu wählen, die keine exakte Vorhersagbarkeit ermöglichen, und bei der Anwendung von Block-Verfahren nur mit langen Blöcken und variablen Blocklängen zu arbeiten, die zufällig ausgewählt werden. Die Vorhersagbarkeit in verschiedenen Randomisierungsverfahren wird unten genauer behandelt.

Ermöglichung unterschiedlicher Therapiegruppenumfänge

In bestimmten Studien ist von Bedeutung, unterschiedliche Therapiegruppenumfänge erhalten zu können. So ist es in manchen Studien leichter, Probanden von einer Studienteilnahme zu überzeugen, wenn ihre Chance, in den Prüfarm zu gelangen, größer als 1 : 1 ist (vgl. z. B. KRACHT, 1992). Auch bei hohen Kosten für eine spezifische Prüftherapie kann es sinnvoll sein, einen Probanden mit einer Wahrscheinlichkeit von mehr als 50% der preiswerteren Therapie zuzuordnen. Entsprechend ist daher für jedes Randomisierungsverfahren zu untersuchen, ob es durch geeignete Wahl der Parameter diese unterschiedlichen Verhältnisse der Umfänge ermöglicht.

Berücksichtigung prognostischer Faktoren

In der ICH E9 (1998) wird empfohlen, in multizentrischen Studien nach Zentren stratifiziert zu randomisieren. Es wird dort ebenfalls darauf hingewiesen, dass in vielen Fällen die Berücksichtigung prognostischer Faktoren von Bedeutung ist, um eine Balancierung innerhalb dieser Faktoren zu erhalten (vgl. auch SCHÄFER et al., 1999). In multizentrischen Studien ist dabei häufig eine frühzeitige Balancierung innerhalb jedes Zentrums wichtig, damit ein vorzeitiges Ausscheiden einer Klinik aus der Studie nicht die Gesamtbalance sowie die Balance in den verschiedenen Strata beeinflusst.

Einige Randomisierungsverfahren ermöglichen innerhalb des Randomisierungsprozesses den Einschluss einer Reihe von bekannten Faktoren. Sie werden als stratifiziert bezeichnet, weil dabei sämtliche Ausprägungsstufen dieser prognostischen Faktoren zu Strata kombiniert werden. Wenn beispielsweise in einer bizenitrischen Studie auch das Geschlecht berücksichtigt werden soll, erhält man die vier Strata männlich in Zentrum₁, weiblich in Zentrum₁, männlich in Zentrum₂ und weiblich in Zentrum₂.

Handhabung

Alle Verfahren erfordern die Auswahl eines Zufallszahlengenerators, eines Seeds oder, falls keiner vorhanden ist, eines Generators für den Seed. Außerdem müssen die Bezeichnungen der Therapiegruppen und damit auch die Therapiegruppenanzahl festgelegt werden.

Bei einigen Randomisierungsverfahren müssen darüber hinaus nur wenige weitere Vorgaben gemacht werden, z. B. die Bestimmung der p -Werte (s. u.) oder ggf. die Definition von Strata. Andere Verfahren hingegen sind aufgrund ihrer Anzahl an Parametern oder ihrer Komplexität schwierig zu handhaben und erfordern deshalb auch im Programm eine ausführliche Hilfe, die den Benutzer unterstützt. Jedes Verfahren ist diesbezüglich zu untersuchen und gegebenenfalls mit Hilfsfunktionen auszustatten, so dass es von allen potentiellen Benutzern problemlos verwendet werden kann.

Randomisierungsverfahren

Für die Implementation wurde eine Reihe stratifizierter und unstratifizierter Randomisierungsverfahren ausgewählt. So sollten einerseits in der Praxis häufig eingesetzte und daher allgemein bekannte Randomisierungsverfahren einbezogen werden. Andererseits wurden auch Verfahren aufgenommen, die möglicherweise wegen der nicht vorhandenen Implementierung weniger verbreitet sind, aber dennoch aufgrund guter Eigenschaften vielversprechend sind.

Vollständige Randomisierung

Das Verfahren der Vollständigen Randomisierung arbeitet für zwei Gruppen nach folgendem Prinzip (BLACKWELL und HODGES, 1957): In jedem Randomisierungsschritt wird, unabhängig von den vorangegangenen Randomisierungsschritten, immer dieselbe, exakt ausbalancierte Münze geworfen. Je nachdem, ob Kopf oder Zahl oben liegt, wird der Patient der einen oder anderen Behandlungsgruppe zugewiesen. Dementsprechend ist die Zuordnung auf eine der Therapiegruppen für jeden Patienten gleich wahrscheinlich und nicht vorhersagbar. Bei beliebig gewählter Anzahl an Therapiegruppen ist die Wahrscheinlichkeit, in eine der Gruppen randomisiert zu werden, in jedem Schritt ebenfalls für jede Gruppe gleich.

Das Verfahren der Vollständigen Randomisierung hat ein unkontrollierbares Balanceverhalten mit möglicherweise großen absoluten Differenzen zwischen den Therapiegruppenumfängen am Ende einer Randomisierung, wobei große Differenzen mit größeren Fallzahlen unwahrscheinlich werden (LACHIN, 1988a). Wenn mindestens 200 Patienten eingeschlossen werden, sind dementsprechend nur geringe Auswirkungen auf die Teststärke der Studie zu erwarten (LACHIN, 1988b). Dadurch dass das Verfahren rein zufällig randomisiert, bietet es natürlich den Vorteil, dass die Unvorhersagbarkeit der einzelnen Randomisierungsschritte bestmöglich gewährleistet ist (KALISH und BEGG, 1985). Daher wird es insbesondere für nicht verblindete (offene) Studien mit großen Fallzahlen empfohlen (SCHULZ und GRIMES, 2002b). Da kein Einfluss auf die Balance genommen wird, können die tatsächlichen Umfänge der einzelnen Therapiegruppen nicht vorab bestimmt werden.

Die Verwendung dieses Verfahrens innerhalb von Strata wirkt sich unter Umständen nachteilig auf das Balanceverhalten aus, da die Randomisierung dann innerhalb von Teilmengen der Gesamtpatientenmenge erfolgt und sich das Balanceverhalten für niedrige Patientenzahlen bei dem Verfahren der Vollständigen Randomisierung verschlechtert. Als Vorteil bleibt festzuhalten, dass das Verfahren sehr einfach zu handhaben ist, da es keine Parameter benötigt. Es ist Spezialfall oder Bestandteil vieler Randomisierungsverfahren.

Beispiel Vollständige Randomisierung

Für eine Studie sollen Patienten rekrutiert und auf die beiden Gruppen Prüftherapie (P) und Standardtherapie (S) aufgeteilt werden. Dazu wird ein Würfel zur Hand genommen, und für jeden neuen Patienten wird einmal gewürfelt. Je nach Augenzahl wird folgendermaßen auf die Therapiegruppen zugewiesen:

Augenzahl 1, 2 oder 3: Zuweisung auf P.

Augenzahl 4, 5 oder 6: Zuweisung auf S.

Für die ersten acht Patienten sei der Würfel wie folgt gefallen: 5, 2, 1, 3, 5, 6, 4, 4. Die Therapiegruppenzuweisungen lauten also: S, P, P, P, S, S, S, S. Nach acht Schritten ergeben sich die Therapiegruppenumfänge dann zu: $n_P = 3$, $n_S = 5$.

Biased Coin Randomisierung BCD(p) von Efron (1971)

Das Biased Coin Design funktioniert für zwei Gruppen wie folgt (FRANE, 1998): Vor jedem Randomisierungsschritt werden die aktuellen Umfänge beider Therapiegruppen umfänge miteinander verglichen. Sind die Umfänge gleich, wird vollständig randomisiert. Ansonsten wird die Wahrscheinlichkeit, in die kleinere Gruppe randomisiert zu werden, auf den Wert p gesetzt. Dabei muss die Wahrscheinlichkeit p größer oder gleich der Wahrscheinlichkeit sein, in die andere Gruppe randomisiert zu werden, also $p \geq 1/2$. Für den Fall $p = 1/2$ entspricht das Biased Coin Design der Vollständigen Randomisierung. Bei beliebig gewählter Anzahl an Therapiegruppen wird p durch einen Vektor $(p_1, p_2, \dots, p_{(\#Gruppen)})$ von Zuweisungswahrscheinlichkeiten ersetzt mit $p_1 \geq p_2 \geq \dots \geq p_{(\#Gruppen)}$ und $\sum p_i = 1$. Bei ungleichen Therapiegruppenumfängen entspricht die Wahrscheinlichkeit, in die kleinste Gruppe randomisiert zu werden, dann p_1 , in die zweitkleinste p_2 usw.

Das Biased Coin Verfahren hat abhängig vom gewählten p -Wert ein sehr gutes Balanceverhalten und ist in der Praxis einfach zu handhaben. Am p -Wert kann der Studienleiter unmittelbar ablesen, mit welcher Wahrscheinlichkeit der nächste Patient in die Therapiegruppe mit dem kleinsten Umfang randomisiert wird. Um die Unvorhersagbarkeit nicht zu gefährden, sollte der p -Wert daher nicht zu nahe bei eins gewählt werden; für zwei Therapiegruppen empfiehlt EFRON (1971) $p = 2/3$. Das Verfahren kann leicht dahingehend angepasst werden, dass unterschiedliche Therapiegruppenumfänge erreicht werden. Prognostische Faktoren werden in diesem Verfahren nicht speziell berücksichtigt, es ist aber die Verwendung des Verfahrens innerhalb von Strata denkbar.

Beispiel Biased Coin Randomisierung

Wie im Beispiel der Vollständigen Randomisierung sollen für eine Studie Patienten rekrutiert und auf die beiden Therapiegruppen P und S verteilt werden. Der p -Wert sei mit $2/3$ fest gewählt. Es wird wieder ein Würfel zur Hand genommen, und für jeden neuen Patienten wird einmal gewürfelt, wobei abhängig von der gewürfelten Augenzahl und den derzeitigen Therapiegruppenumfängen sechs Fälle zu unterscheiden sind:

- (1) $n_P = n_S$ und Augenzahl 1, 2 oder 3: Zuweisung auf P.
- (2) $n_P = n_S$ und Augenzahl 4, 5 oder 6: Zuweisung auf S.
- (3) $n_P < n_S$ und Augenzahl 1, 2, 3 oder 4: Zuweisung auf P.
- (4) $n_P < n_S$ und Augenzahl 5 oder 6: Zuweisung auf S.
- (5) $n_P > n_S$ und Augenzahl 1 oder 2: Zuweisung auf P.
- (6) $n_P > n_S$ und Augenzahl 3, 4, 5 oder 6: Zuweisung auf S.

Tabelle 1: Beispiel zur Biased Coin Randomisierung mit acht Patienten

Patient	1	2	3	4	5	6	7	8
Gewürfelte Augenzahl	5	2	1	3	5	6	4	4
Umfänge (n_P, n_S)	(0, 0)	(0, 1)	(1, 1)	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(3, 4)
Fall	(2)	(3)	(1)	(6)	(2)	(4)	(3)	(3)
Zuweisung	S	P	P	S	S	S	P	P

Für die ersten acht Patienten sei der Würfel wie im Beispiel der Vollständigen Randomisierung gefallen. Es ergibt sich die Tabelle 1. Offensichtlich lauten die Therapiegruppenumfänge nach zehn Schritten dann: $n_P = 4, n_S = 4$.

Urnenmodell $UD(w, \alpha, \beta)$ von Wei (1978a)

Das Urnenmodell arbeitet mit drei zu Beginn festzulegenden Parametern. Für jede Therapiegruppe liegen anfänglich w Kugeln in der Urne, die entsprechend ihrer Gruppenzugehörigkeit markiert sind. Nach der zufälligen Entnahme einer Kugel und Zuweisung des Patienten zur entsprechenden Therapiegruppe wird die Kugel zurückgelegt, und α Kugeln derselben Gruppe werden der Urne hinzugefügt. Weiterhin werden $\beta \geq \alpha$ Kugeln der übrigen Gruppen in die Urne gelegt. Für die Randomisierung des nächsten Patienten wird wieder eine Kugel zufällig entnommen und weiter wie oben beschrieben verfahren.

Die Wahl der Parameter w, α und β erweist sich in der Praxis als recht schwierig, da die optimalen Parameter stark variieren können, z. B. abhängig von der Patientenzahl. Eine Möglichkeit wäre beispielsweise $w = 1, \alpha = 1$ und $\beta = 10$.

Bei kleineren Patientenzahlen erreicht das Urnenmodell eine gute Balance; bei großen Fallzahlen hingegen verhält es sich ähnlich der Vollständigen Randomisierung (WEI und LACHIN, 1988). Dementsprechend ist dann das Balanceverhalten relativ schlecht, trotz guter Reaktion auf unterschiedliche absolute Differenzen. Die p -Werte werden bei diesem Verfahren nach jedem Randomisierungsschritt neu berechnet, und es ist deshalb kaum möglich, die Wahrscheinlichkeit vorherzusagen, auf die verschiedenen Gruppen randomisiert zu werden. Die Unvorhersagbarkeit ist demnach gewährleistet.

Im Gegensatz zur Vollständigen Randomisierung und der Biased Coin Randomisierung, mit denen sich leicht unterschiedliche Therapiegruppenumfänge erzielen lassen, ist dieses im Urnenmodell nicht ohne weiteres möglich. Es existieren verschiedene Ansätze, die dieses Verfahren auch innerhalb von Strata nutzbar machen. Allerdings sind uns keine Untersuchungsergebnisse bekannt, die einen Rückschluss auf das Balanceverhalten im stratifizierten Fall zulassen. Da die Wahl der Parameter genaue Kenntnisse erfordert, erweist sich die Handhabung im Vergleich zu anderen Verfahren als komplizierter.

Beispiel Urnenmodell

Seien zwei Therapiegruppen P und S gegeben und die Parameter $w = 1, \alpha = 1, \beta = 2$. Die Urne sei durch $\{k_P, k_S\}$ gegeben, wobei k_P die Anzahl der Kugeln bezeichnet, die zur Gruppe P gehören und k_S die Anzahl der Kugeln, die zur Gruppe S gehören.

Da $w = 1$, wird die Urne zu Beginn also durch $\{1, 1\}$ repräsentiert (vgl. Abb. 1, 1.). Wird nun zufällig eine Kugel entnommen, die zur Gruppe P gehört (vgl. Abb. 1, 2.), ist die Urne durch $\{0, 1\}$ gegeben (vgl. Abb. 1, 3.) und es erfolgt die Zuweisung des ersten Patienten auf Therapiegruppe P. Danach wird die Kugel zurückgelegt: $\{1, 1\}$ (vgl.

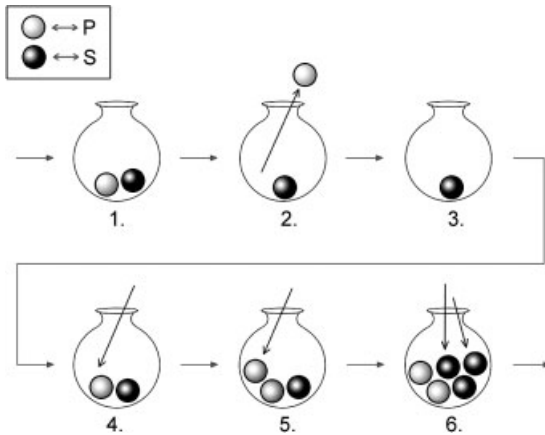


Abbildung 1: Ein Randomisierungsschritt mit dem Urnenmodell

Abb. 1, 4.); $\alpha = 1$ Kugeln derselben Gruppe werden hinzugefügt: $\{2, 1\}$ (vgl. Abb. 1, 5.); $\beta = 2$ Kugeln der übrigen Gruppen werden in die Urne gelegt: $\{2, 3\}$ (vgl. Abb. 1, 6.). Die Wahrscheinlichkeit, dass der nächste Patient in die Gruppe S (mit dem kleineren Umfang) randomisiert wird, ist also 60%.

Permutierte Block-Randomisierung PBR(B) – ohne Strata

Verfahren, die die Zuweisung der Studienteilnehmer zu den verschiedenen Therapiegruppen innerhalb von Blöcken durchführen, werden Blockverfahren genannt (MATTIS und LACHIN, 1988; ZELEN, 1974). Die Blocklänge B kann variabel oder fest gewählt werden und muss ein Vielfaches der Anzahl der Therapiegruppen sein.

Gängige Blocklängen sind z. B. $B = 4, 6$ oder 8 . Zu Beginn wird innerhalb des ersten Blocks randomisiert. Ist der Block voll, wird ein neuer Block verwendet usw. Gefüllte Blöcke sind immer ausbalanciert, d. h. die absolute Differenz ist Null.

Das Verfahren der Permutierten Block-Randomisierung arbeitet grundsätzlich wie zuvor beschrieben, d. h. die Zuweisung der Patienten auf die verschiedenen Therapiegruppen erfolgt innerhalb von Blöcken der Länge B , wobei gefüllte Blöcke immer ausbalanciert sind.

Soll ein neuer Patient auf eine Therapiegruppe randomisiert werden, betrachtet das Verfahren nur den aktuellen Block und berechnet die Wahrscheinlichkeit P_P (analog P_S), in die Gruppe P (analog S) randomisiert zu werden, wie folgt: $P_P = (b - n_P)/(B - n)$. Hierbei ist $b = B/\text{Anzahl Gruppen}$ der maximale Umfang einer Gruppe in einem Block und muss ganzzahlig sein, n_P die Anzahl der Patienten, die bereits im aktuellen Block in die Gruppe P randomisiert wurden, und n die Anzahl der Patienten, die bereits insgesamt im aktuellen Block randomisiert wurden. Die Wahrscheinlichkeit wird ihren Umfängen umgekehrt entsprechend gesetzt, d. h. die Wahrscheinlichkeit, in eine kleine Gruppe randomisiert zu werden, ist größer, in eine größere Gruppe randomisiert zu werden, kleiner. Ist bereits der maximale Umfang einer Gruppe im aktuellen Block erreicht, so erfolgt die Zuweisung auf diese Gruppe mit einer Wahrscheinlichkeit von 0. Wurden im aktuellen Block bisher gleichviele Patienten zu P und S zugewiesen, so wird als Spezialfall eine Vollständige Randomisierung durchgeführt.

Das Verfahren der Permutierten Block-Randomisierung hat ein sehr gutes Balanceverhalten mit einer maximalen absoluten Differenz von $B/2$, wenn die Randomisierung nach

einem halben Block beendet wird. Dem oft kritisierten Nachteil, dass bei bekannter fester Blocklänge die nächsten Zuordnungen zu den Therapiegruppen unter Umständen vorhergesagt werden können, da innerhalb von gefüllten Blöcken Balance erzwungen wird, kann durch den Einsatz variabler Blocklängen und das Verschweigen der tatsächlich gewählten variablen Blocklängen im Studienprotokoll abgeholfen werden. Geplante unterschiedliche Therapiegruppenumfänge sind nach Anpassung des Verfahrens möglich. Die Wahl der Blocklängen erfordert genaue Kenntnisse und muss in der Programmhilfe besonders berücksichtigt werden. Generell gilt, dass eine Blocklänge von zwei Patienten nur bei einer vollständigen Maskierung gewählt werden sollte (LACHIN et al., 1988), bei nicht verblindeten Behandlern werden Blocklängen von mindestens sechs Patienten empfohlen (SCHULZ und GRIMES, 2002a).

Beispiel Permutierte Block-Randomisierung

Seien zwei Therapiegruppen P und S gegeben und die Blocklänge $B = 4$. Der erste Patient im ersten oder einem beliebigen neuen Block sei zufällig auf die Gruppe S randomisiert worden. Die Wahrscheinlichkeit, dass der nächste Patient in die Gruppe P randomisiert wird, berechnet sich nun wie folgt: $P_P = (2 - 0)/(4 - 1) = 2/3$.

Erweiterung: PBR(B) mit variablen Blocklängen

- (1) Es wird eine Reihe von Blocklängen festgelegt, die Verwendung finden dürfen (z. B. 4, 6 und 8).
- (2) Es wird eine der Blocklängen zufällig ausgewählt.
- (3) Es wird innerhalb dieses Blocks randomisiert, bis der Block vollständig ist.
- (4) Falls fertig, stopp; sonst weiter mit (2).

Permutierte Block-Randomisierung innerhalb von Strata

Bei dem Verfahren der Permutierten Block-Randomisierung innerhalb von Strata werden zu Beginn der Randomisierung, wie beim Verfahren der Permutierten Block-Randomisierung ohne Strata beschrieben, Strata gebildet, und dann wird innerhalb dieser nach dem Verfahren der Permutierten Block-Randomisierung randomisiert (ZELEN, 1974).

Bei der Durchführung innerhalb von Strata ist das Balanceverhalten der Permutierten Block-Randomisierung nicht mehr so gut, da es in diesem Fall in jedem Stratum einen letzten Block gibt, der die absolute Differenz mitbestimmt. So beträgt die maximale absolute Differenz bei k Strata $k \cdot B/2$. Dieser Umstand lässt erkennen, dass die Wahrscheinlichkeit für das Auftreten größerer Differenzen zwischen den Stichprobenumfängen insgesamt zunimmt, wenn die Blocklänge und die Strata ungünstig gewählt wurden. Das Verfahren erfüllt die übrigen Eigenschaften wie das Verfahren der Permutierten Block-Randomisierung ohne Strata.

Minimierung von Taves (1974) und Pocock und Simon (1975)

Das Verfahren der Minimierung kombiniert die einzelnen prognostischen Faktoren nicht zu Strata sondern betrachtet sie separat. Nach jedem Randomisierungsschritt errechnet das Verfahren für jede mögliche Verteilung des nächsten Schritts den Mangel an Balance zwischen den Therapiegruppen, wobei nur die für den aktuellen Patienten zutreffenden Ausprägungsstufen der prognostischen Faktoren in die Berechnung mit einbezogen werden.

Im ersten Schritt wird für alle zutreffenden Ausprägungen (z. B. für die prognostischen Faktoren Geschlecht und Zentrum: weiblich, Lübeck) der Mangel an Balance errechnet,

d. h. die absolute Differenz innerhalb der Ausprägungsstufe. Im zweiten Schritt wird durch deren Summation der Gesamt-mangel berechnet, der auftreten würde, falls diese Therapiegruppenzuweisung eintreten würde, es wird also für jede probeweise Therapiegruppenzuweisung ein Gesamt-mangel an Balance kalkuliert.

Ist der Gesamt-mangel an Balance in den Therapiegruppen unterschiedlich, so wird in der ursprünglich von TAVES (1974) vorgeschlagenen Methode dann der Patient so zugewiesen, dass der Gesamt-mangel an Balance nach der Zuweisung so gering wie möglich ist. In der allgemeineren Form (POCOCK und SIMON, 1975) werden hingegen die Gesamt-mängel aller Zuweisungsmöglichkeiten der Größe nach geordnet und ergeben so die Reihenfolge der Wahrscheinlichkeiten, auf eine bestimmte Therapiegruppe randomisiert zu werden. Diese Wahrscheinlichkeiten $p_1, p_2, \dots, p_{(\#Gruppen)}$ werden zu Beginn der Randomisierung festgelegt, wobei $p_1 \geq p_2 \geq \dots \geq p_{(\#Gruppen)}$ und $\sum p_i = 1$ gelten muss. Das heißt, die Wahrscheinlichkeit ist höher, in eine Therapiegruppe mit geringerem Balance-mangel-Wert randomisiert zu werden.

Das Verfahren der Minimierung hat insbesondere für sechs oder mehr prognostische Faktoren ein sehr gutes Balanceverhalten (KUNDT, 2002). Die Unvorhersagbarkeit ist in hohem Maße gewährleistet; auch geplante unterschiedliche Therapiegruppenumfänge kann das Verfahren nach einigen Anpassungen erfüllen. Prognostische Faktoren werden in die Randomisierung miteinbezogen, Wechselwirkungen zwischen ihnen bleiben aber unberücksichtigt. Die Handhabung ist einfach, da außer den p -Werten und den prognostischen Faktoren keine weiteren Parameter benötigt werden. Aufgrund dieser Eigenschaften ist dieses Verfahren besonders häufig praktisch eingesetzt worden.

Beispiel Minimierung

Seien die Faktoren Zentrum mit zwei Ausprägungen (z_1, z_2) und Geschlecht (m, w) sowie zwei Therapiegruppen P und S gegeben. Die Wahrscheinlichkeiten, auf eine bestimmte Therapiegruppe randomisiert zu werden, seien mit $p_1 = 2/3$ und $p_2 = 1/3$ festgelegt.

Das Verfahren der Minimierung betrachtet nun vor jedem Randomisierungsschritt den Faktor Zentrum mit seinen Ausprägungsstufen z_1, z_2 und den Faktor Geschlecht mit den Ausprägungsstufen m, w , und zwar unter Verwendung nachfolgender 2×2 Matrizen. In jedem Matrixfeld befindet sich die Anzahl der Patienten, die mit der in der ersten Spalte bezeichneten Ausprägungsstufe des Faktors Zentrum bzw. Geschlecht in die Therapiegruppe P bzw. S randomisiert wurden. Es seien bisher 20 Patienten wie folgt randomisiert worden (Tabelle 2, oben). Es sei nun der nächste Patient, der randomisiert werden soll, männlich und aus z_2 . Es wird für den nächsten Schritt dann der Mangel an Balance bestimmt, der sich bei der Zuordnung auf P (Tabelle 2, Mitte) und S (Tabelle 2, unten) ergibt. Der Gesamtbalancemangel bei einer Zuweisung auf Therapiegruppe P würde also $(0 + 1) = 1$ betragen, bei einer Zuweisung auf Therapiegruppe S $(2 + 3) = 5$.

Da $1 < 5$, erhält die Therapiegruppe P die Wahrscheinlichkeit $p_1 = 2/3$ und die Therapiegruppe S die Wahrscheinlichkeit $p_2 = 1/3$.

Self-Adjusting-Randomisierung von Nordle und Brantmark (1977)

Das Verfahren der Self-Adjusting-Randomisierung arbeitet nach einer modifizierten Idee des Verfahrens der Minimierung. Dabei wird im ersten Schritt eine Stratifizierungsmatrix aufgebaut, deren Spalten aus den prognostischen Faktoren gebildet werden, die bezüglich ihrer Wertigkeit als wichtiger eingeschätzt werden. Die Zeilen der Matrix werden aus den restlichen prognostischen Faktoren gebildet.

Ein Randomisierungsschritt besteht aus fünf möglichen Einzelschritten. Erfolgt in einem dieser Schritte eine Zuweisung auf eine Therapiegruppe, werden die übrigen Schritte nicht mehr berücksichtigt. In den ersten drei Schritten wird jeweils die absolute Diffe-

Tabelle 2: Beispiel zur Minimierung

nach 20. Randomisierung:				
Gruppe	Faktor Zentrum		Faktor Geschlecht	
	z_1	z_2	m	w
P	4	5	4	5
S	5	6	6	5
21. Randomisierung auf P:				
Gruppe	Faktor Zentrum		Faktor Geschlecht	
	z_1	z_2	m	w
P	4	6	5	5
S	5	6	6	5
Balancemangel		0	1	
21. Randomisierung auf S:				
Gruppe	Faktor Zentrum		Faktor Geschlecht	
	z_1	z_2	m	w
P	4	5	4	5
S	5	7	7	5
Balancemangel		2	3	

renz der Umfänge der Therapiegruppen in der den aktuellen Werten der prognostischen Faktoren entsprechenden Zelle (1. Schritt), Spalte (2. Schritt) oder Zeile (3. Schritt) geprüft. Im vierten Schritt wird die absolute Differenz der Umfänge der Therapiegruppen insgesamt geprüft. Falls die absolute Differenz in einem der Schritte von null verschieden ist, erfolgt eine automatische Zuweisung auf die kleinste Therapiegruppe. Sonst geht es mit dem nächsten Schritt weiter. Im fünften und letzten möglichen Schritt erfolgt die Zuweisung mittels der Vollständigen Randomisierung.

Das Verfahren der Self-Adjusting-Randomisierung hat ein sehr gutes Balanceverhalten, insbesondere für weniger als sechs prognostische Faktoren. Für sechs oder mehr prognostische Faktoren und größere Patientenzahlen wird es etwas schwächer (KUNDT, 2002).

Wenn die Matrixinhalte bekannt sind, ist die Unvorhersagbarkeit unter Umständen nicht mehr gewährleistet, da in den ersten vier Verfahrensschritten eine automatische Zuweisung geschieht. Dem kann durch eine leichte Modifizierung abgeholfen werden: Die mit 100% vorgesehene Zuweisung in einem der ersten vier Schritte ist durch eine geringere Wahrscheinlichkeit zu ersetzen, beispielsweise 90%. Entsprechend könnte dies zu einer Verschlechterung des Balanceverhaltens führen. Die Unvorhersagbarkeit ist dann in einem hohen Maße gewährleistet.

Auch die Eigenschaft der Vorabbestimmung unterschiedlicher Therapiegruppenumfänge kann das Verfahren nach einigen Anpassungen erfüllen. Zwischen den Stratifizierungsfaktoren werden Wechselwirkungen teilweise berücksichtigt (KUNDT, 2002). Die Handhabung ist einfach, da außer den prognostischen Faktoren keine weiteren Parameter benötigt werden.

Tabelle 3: Notation zur Self-Adjusting-Randomisierung

Geschlecht	Zentrum	
	z_1	z_2
m	n_{11P}, n_{11S}	n_{12P}, n_{12S}
w	n_{21P}, n_{21S}	n_{22P}, n_{22S}

n_{ijP} (n_{ijS}) = Anzahl der Patienten, die innerhalb der Zelle c_{ij} in die Gruppe P (Gruppe S) randomisiert wurden.

Beispiel Self-Adjusting-Randomisierung

Gegeben seien die Therapiegruppen P und S und der wichtige prognostische Faktor Zentrum mit den Ausprägungsstufen z_1 und z_2 sowie der weniger wichtige prognostische Faktor Geschlecht mit den Ausprägungsstufen m und w. Es ergibt sich nun die in Tabelle 3 dargestellte Randomisierungsmatrix, in deren Zellen sich die jeweiligen Umfänge n der Therapiegruppen P und S befinden. Solange noch kein Patient randomisiert wurde, sind alle Umfänge gleich null (Tabelle 4, links).

Der erste zu randomisierende Patient sei weiblich und komme aus Zentrum z_1 . Da die absolute Differenz innerhalb der Zelle c_{21} , der ersten Spalte, der zweiten Zeile und der gesamten Matrix gleich Null ist und dementsprechend keiner der ersten vier Schritte der Self-Adjusting-Randomisierung zu einer Zuweisung auf eine Gruppe führt, wird vollständig randomisiert. Das Ergebnis soll in diesem Beispiel S lauten. Dann ergibt sich die Matrix aus Tabelle 4, Mitte.

Wäre der nächste Patient männlich und aus Zentrum z_1 , dann käme es aufgrund der Unbalance in der ersten Spalte zu einer Zuweisung auf Gruppe P und die Matrix sähe wie in Tabelle 4, rechts, aus.

Eine Zusammenfassung der Eigenschaften der verschiedenen Randomisierungsverfahren findet sich in Tabelle 6.

Tabelle 4: Beispiel zur Self-Adjusting-Randomisierung

Geschlecht	vor Randomisierung		nach 1. Randomisierung		nach 2. Randomisierung	
	Zentrum		Zentrum		Zentrum	
	z_1	z_2	z_1	z_2	z_1	z_2
m	(0,0)	(0,0)	(0,0)	(0,0)	(1,0)	(0,0)
w	(0,0)	(0,0)	(0,1)	(0,0)	(0,1)	(0,0)

Tabelle 5: Zuweisungsmatrix

	Zentrum: Lübeck	Zentrum: Kiel
Geschlecht: männlich	Therapiegruppe 2	Therapiegruppe 1
Geschlecht: weiblich	Therapiegruppe 1	Therapiegruppe 2

Tabelle 6: Übersicht über die Randomisierungsverfahren

Verfahren	Unvorhersagbarkeit	Balanceverhalten	Ermöglichung verschiedener Gruppenumfänge	Berücksichtigung prognostischer Faktoren	Handhabung
Vollständige Randomisierung	sehr gut	unkontrollierbar	nein	möglich, nachteilig auf Balance	sehr einfach
Biased Coin Randomisierung	akzeptabel	sehr gut	möglich	möglich, aber Balance unklar	einfach
Urnenmodell	gut	relativ schlecht	unklar	möglich, aber Balance unklar	komplizierter, da optimale Parameter unklar
Permutierte Block-Randomisierung	gut bei variablen Blocklängen	sehr gut, schlechter bei Stratifizierung	möglich	möglich, nachteilig auf Balance	einfach
Minimierung	gut	sehr gut bei vielen Strata, sonst gut	möglich	ja	einfach
Self-Adjusting-Randomisierung	unklar	sehr gut bei wenigen Strata, sonst gut	möglich	ja	einfach
Self-Adjusting-Randomisierung: Modifikation	gut	unklar	möglich	ja	komplizierter, da optimale Parameter unklar

Design: Funktionen von RITA

Folgende Teilfunktionen wurden in RITA aufgenommen:

Randomisierungsverfahren

Wir haben uns dafür entschieden, folgende Verfahren in RITA zu implementieren:

- Vollständige Randomisierung
- Biased Coin Design (EFRON, 1971)
- Minimierung (POCOCK und SIMON, 1975)
- Urnenmodell (WEI, 1978a)
- Permutierte Blockrandomisierung
- Self-Adjusting Randomisierung (NORDLE und BRANTMARK, 1977)

Das Verfahren der Vollständigen Randomisierung wurde wegen seines hohen Bekanntheits- und Verbreitungsgrades mit in RITA aufgenommen. Beim Biased Coin Design war das außerordentlich gute Balanceverhalten in Verbindung mit der Unvorhersagbarkeit entscheidend. Das Verfahren der Minimierung wurde aufgrund seiner weiten Verbreitung aufgenommen. Die ausdrückliche Empfehlung, das Urnenmodell trotz des schlechteren Balanceverhaltens einzusetzen (SCHULZ und GRIMES, 2002b), sowie der hohe Bekanntheitsgrad haben dazu geführt, es ebenfalls in RITA zu implementieren. Beim Verfahren der Permutierten Blockrandomisierung mit variablen Blocklängen (optional) war die hohe Verbreitung ausschlaggebend. Das mächtigste in dem Programm enthaltene Verfahren ist aber das Self-Adjusting Design von Nordle und Brantmark (1977), welches um einige optionale Parameter ergänzt wurde. Hierzu gehört die Eingabe einer geringeren Wahrscheinlichkeit als 100% für Zelle, Spalte, Zeile oder Gesamtheit der Randomisierungsmatrix, um die Unvorhersagbarkeit des Verfahrens zu gewährleisten.

Zufallszahlengeneratoren

Bei den Zufallszahlengeneratoren haben wir uns für MersenneTwister und Drand48 entschieden. MersenneTwister (MT19937) gilt als einer der besten uniformen Pseudo-Zufallszahlengeneratoren. Drand48, ein schneller uniformer Pseudo-Zufallszahlengenerator mittlerer Qualität, wurde in das Programm aufgenommen, weil er sehr verbreitet und bekannt ist. So ist Drand48 beispielsweise auf jedem Unix-Rechner als ANSI C Funktion `drand48()` zu finden oder in der Java-Klassensammlung unter dem Namen `Random()`.

Seedgenerator

Jeder uniforme Pseudo-Zufallszahlengenerator benötigt zu Beginn einen Seed, also einen Startwert (z. B. ein 32-Bit-Integerwert), von dem alle Pseudo-Zufallszahlen rekursiv abhängen. Das bedeutet, dass der Zufallszahlengenerator mittels des Seeds seine erste Pseudo-Zufallszahl bildet und nun als neuen Seed die generierte Pseudo-Zufallszahl nimmt. So ist es möglich, anhand des Seeds alle Zufallszahlen, die für die Randomisierung einer klinischen Studie generiert wurden, zu rekonstruieren, was für die geforderte Reproduzierbarkeit des gesamten Randomisierungsprozesses notwendig ist.

Da von dem Seed alle folgenden Zufallszahlen abhängen, kommt ihm eine besondere Bedeutung zu. Er wird in RITA deshalb als echte Zufallszahl aus dem Internet (www.random.org oder www.fourmilab.ch) geladen. Falls keine Internetverbindung vorhanden ist, wird automatisch ein Pseudo-Seed mit der Java-Methode „`Random()`“ erzeugt, die der wählbaren Funktion `Drand48` entspricht.

Stratafunktionen

Bei der Verwendung von Strata ist es wichtig, dass das Programm auch das Hinzufügen von Ausprägungsstufen bei einem bereits laufenden Randomisierungsprozess unterstützt. Beispielsweise sei eine Reihe von Patienten zweier Zentren randomisiert worden; nun muss es möglich sein, ein drittes Zentrum hinzuzufügen und Patienten aus diesem Zentrum zu randomisieren.

Basisfunktionen

Das Programm unterstützt das Neu Anlegen, Speichern und Öffnen von Studien.

Für jede neu angelegte oder geöffnete Studie öffnet sich ein neues Fenster, so dass mit mehreren Studien parallel gearbeitet werden kann. Die Randomisierungsergebnisse lassen sich jederzeit ausdrucken.

Ein Rechnerproblem während einer laufenden Randomisierung wird von dem Programm beim Neustart erkannt, und es wird die Möglichkeit geboten, die zuletzt bearbeitete Studie exakt wiederherzustellen. Außerdem ist eine Offline-Randomisierung der Patienten möglich. So kann eine Randomisierung auch ohne PC durch- oder fortgeführt werden. Bei stratifizierten Verfahren ist es aufgrund der Abhängigkeiten von den noch unbekanntem Ausprägungsstufen der prognostischen Faktoren des nächsten Patienten nur möglich, eine Zuweisungsmatrix für den nächsten Patienten auszugeben, die beispielsweise wie in Tabelle 5 aussehen könnte. Wäre die nächste Person weiblich und aus Kiel, würde sie Therapiegruppe 2 zugeordnet werden. Das Programm generiert nach jedem Randomisierungsschritt eine solche Matrix, randomisiert also nach jedem Schritt im Hintergrund alle Ausprägungsmöglichkeiten des nächsten Patienten. Diese können ausgedruckt und zur Offline-Randomisierung des nächsten Patienten verwendet werden.

Ist zu Beginn der Studie festgelegt, wie viele Patienten randomisiert werden sollen, können bei unstratifizierten Verfahren auch sämtliche Patienten auf einmal randomisiert werden.

Grundlegende Anforderungen und ihre Realisierung

An das Programm werden verschiedene Anforderungen gestellt, die es zum Teil attraktiver machen, die zum Teil aber auch erforderlich sind, damit die Software eingesetzt werden darf.

Wie die Anforderungen

- Einfache Handhabung
- Programmpflege durch andere Personen
- Gute Erweiterbarkeit (z. B. Hinzufügen von neuen Verfahren)
- Reproduzierbarkeit der Ergebnisse und damit Erfüllung von Richtlinien
- Validierung, Verifikation und Testen des Programms
- Sicherheit

in RITA gelöst wurden, wird im folgenden Teil beschrieben.

Einfache Handhabung

Die grundlegende Handhabung von RITA wird anhand der Screenshots in Abbildung 2 demonstriert. Das Programm ist weitestgehend selbsterklärend. Beim Anlegen einer neuen Studie nimmt der Benutzer schrittweise die notwendigen Einstellungen vor:

- Auswahl eines Randomisierungsverfahrens
- Eingabe der Parameter des Randomisierungsverfahrens (z. B. p -Wert)

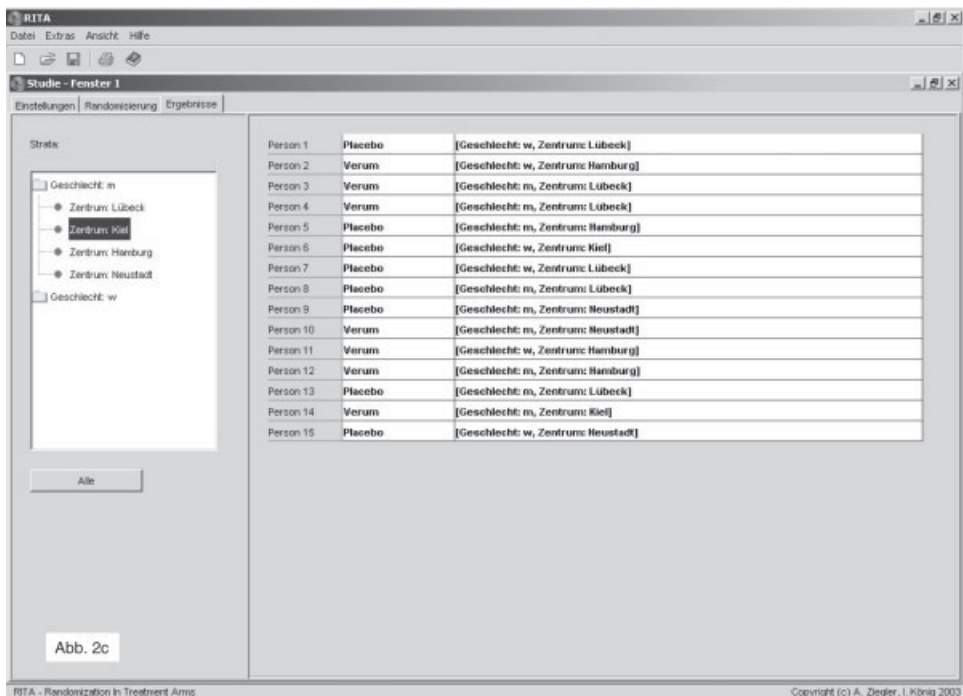
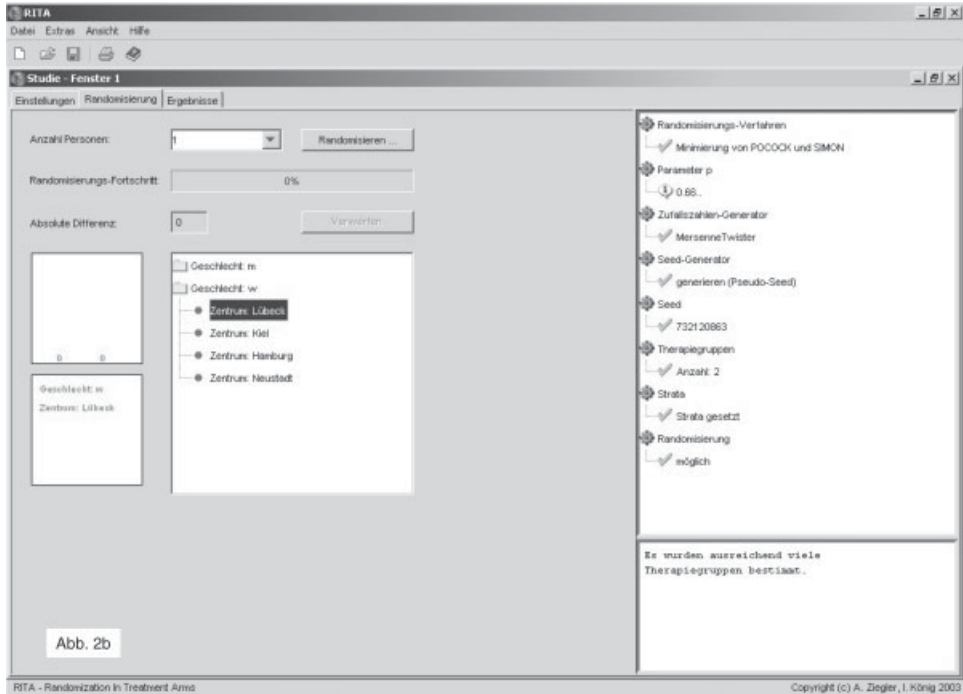


Abbildung 2: Screenshots von RITA

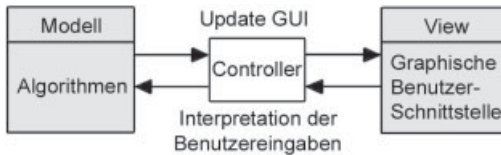


Abbildung 3: Modell-View-Controller Schema

eine reibungslose Einbindung in das Programm zu Eigen. Für die Implementierung der verschiedenen Randomisierungsalgorithmen stehen verschiedene „Toolbox“-Klassen zur Verfügung, die häufig verwendete Operationen, wie z. B. Bestimmung der Umfänge, p -Wert-Berechnung usw. implementiert haben.

Damit ist es möglich, mit geringem Aufwand neue Randomisierungsverfahren hinzuzufügen oder vorhandene Verfahren zu erweitern.

Reproduzierbarkeit der Ergebnisse

Die in einem Design einer kontrollierten klinischen Studie geforderte Transparenz und Nachvollziehbarkeit des Randomisierungsvorgangs und die Reproduzierbarkeit der Ergebnisse muss gewährleistet sein.

RITA leistet dies zum einen durch die Möglichkeiten des Abspeicherns, Öffnens und Druckens der Konfigurationen und Ergebnisse, zum anderen durch die Möglichkeit, die exakten Ergebnisse durch Eingabe weniger Parameter (Randomisierungsverfahren mit Parametern, Seed, Zufallszahlen-Generator) zu reproduzieren. Die Reproduktion kann dabei Schritt für Schritt vorgenommen und so die Randomisierung jedes einzelnen Patienten nachvollzogen werden.

Validierung, Verifikation und Testen des Programms

In verschiedenen Richtlinien wird die Validierung von Software definiert als „confirmation by examination and provision of objective evidence that the particular requirements for a specific intended use are fulfilled“ (CDRH, General Principles of Software Validation; FDA, 820.3; ISO 8402, 1994-04-01). Verifikation hingegen wird beschrieben als „confirmation by examination and provision of objective evidence that specified requirements have been fulfilled“ (FDA, 820.3; ISO 8402, 1994-04-01) oder als „provision of objective evidence that the design outputs of a particular phase of the software development life cycle meet all of the specified requirements for that phase“ (CDRH, General Principles of Software Validation).

Zur Umsetzung der Richtlinien wurde zunächst eine Standard Operating Procedure (SOP) für die Validierung, Verifikation und Testung von RITA erstellt. Es wurde dann für RITA sichergestellt, dass alle geforderten Spezifikationen erfüllt werden. Hierbei wurde im einzelnen geprüft, ob alle Teilfunktionen enthalten und funktionsfähig sind, ob die Handhabung verständlich ist und ob die Ein- und Ausgaberroutinen korrekt funktionieren. Die Funktionsweise wurde auf verschiedenen Plattformen und von verschiedenen wissenschaftlichen und nicht-wissenschaftlichen Mitarbeitern getestet.

Zur weiteren Überprüfung erlaubt RITA jedem Benutzer die Überprüfung des korrekten Laufs des Programms auf dem eigenen System. Dazu werden zu jedem Randomisierungsverfahren Beispiel-Dateien zur Verfügung gestellt, die neben allen notwendigen Parametern auch den Seed enthalten. Unter Eingabe dieser Konstellationen kann der Nutzer dann die zu erhaltenden Ergebnisse verifizieren und somit die Funktionsfähigkeit des Programms bestätigen.

Um zu testen, ob das Programm korrekte Ergebnisse liefert, muss die Arbeitsweise der Randomisierungsalgorithmen vor dem Praxiseinsatz bestätigt werden. So mussten kürzlich beispielsweise mehr als 1000 Frauen erneut rekrutiert und randomisiert werden, weil ein Fehler im Minimierungsalgorithmus zu sehr starker Unbalance zwischen den Therapiegruppen geführt hatte (COMET, 2001).

Um solchen Problemen vorzubeugen, eignen sich in erster Linie Simulationen und der anschließende Vergleich der Simulationsergebnisse mit unabhängigen Ergebnissen. Wir haben für unsere Simulationen die Resultate von KUNDT (2002) zum Vergleich herangezogen. Die Simulationen wurden dafür mit dem in das Programm integrierten Simulator durchgeführt, der so programmiert wurde, dass er jeden Randomisierungsschritt so durchführt, als bediene ein Benutzer das Programm schrittweise. Die Simulationen sind so zwar zeitaufwendig, da der Simulator nicht für sehr schnelle Hintereinanderausführung von Randomisierungsschritten optimiert wurde; dieses Vorgehen gewährleistet aber, dass der Praxiseinsatz des Programms mit simuliert wurde. Ziel der Simulation war, die Wahrscheinlichkeiten für das Auftreten absoluter Differenzen in den Gruppengrößen zu bestimmen. Um diese Wahrscheinlichkeiten mit einer Genauigkeit von 1% bei einer Sicherheit von 95% zu schätzen, wurde jede Simulation einer Randomisierung mit vorgegebenem Randomisierungsverfahren, gegebenen Parametern und vorgegebener Fallzahl n mindestens 10.000-mal durchgeführt. Für jede Simulation einer Randomisierung

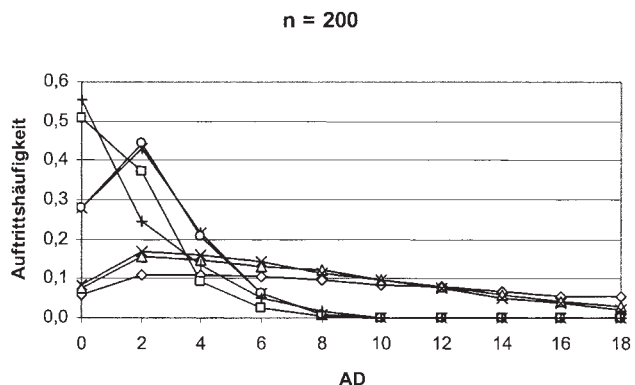
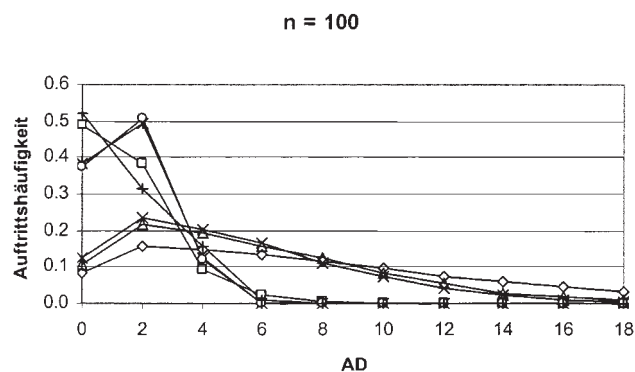


Abbildung 4: Häufigkeit der absoluten Differenzen AD bei verschiedenen Randomisierungsverfahren und Stichprobengrößen

—◇— VR; —□— BCD(2/3); —△— URN(1,1,2); —×— URN(1,1,5);
—*— PBR(B1); —○— PBR(B2); —+— PBR(B)

wurden demnach 10.000-mal n einzelne Randomisierungsschritte durchgeführt. Bei stratifizierten Verfahren wurden die Strata für jeden Patienten zufällig bestimmt. In Abbildung 4 sind die Auftrittshäufigkeiten der absoluten Differenzen 0, 2, 4, . . . , 18 für verschiedene Randomisierungsverfahren dargestellt. Überall erkennt man das schlechte Balanceverhalten der Vollständigen Randomisierung (VR). Das Balanceverhalten des Urnendesigns (URN) wird dem der Vollständigen Randomisierung mit zunehmender Patientenzahl immer ähnlicher und ist im Vergleich zum Biased Coin Design (BCD) und der Permutierten Block-Randomisierung (PBR) relativ schlecht. Der direkte Vergleich mit den Simulationsergebnissen von KUNDT (2002) bestätigt die korrekte Arbeitsweise der Randomisierungsalgorithmen. Es zeigen sich deutlich die beiden von KUNDT (2002) beschriebenen Gruppen besseren (BCD, PBR) und schlechteren Balanceverhaltens (VR, URN).

Sicherheit

In RITA wurden verschiedene Funktionen integriert, die das Verändern gespeicherter Studien durch nicht berechtigte Personen verhindern sollen. So ist es möglich, für jeden Benutzer ein eigenes Konto anzulegen, das mit Benutzername und Kennwort geschützt ist. Nach der gesicherten Anmeldung an diesem Konto kann der Benutzer seine eigenen Dateien einsehen und verändern. Dateien anderer Benutzer können nur eingesehen und nicht verändert werden. Die Verwaltung der Benutzerkonten erfolgt durch einen Administrator mit uneingeschränkten Rechten. Tritt trotzdem ein Schadensfall ein, kann er nachvollzogen werden, da alle Programmaktivitäten in einer Log-Datei festgehalten werden.

Diskussion

Bei der randomisierten kontrollierten klinischen Studie – dem Goldstandard der Therapiestudie – ist eine zufällige Zuteilung der Patienten auf die untersuchten Behandlungsgruppen als die wichtigste Designtechnik zur Reduktion von Verzerrungen anerkannt (ICH E6, 1996). Aufgrund des erheblichen Aufwands in großen Studien, der rechnerischen Komplexität einiger Randomisierungstechniken sowie einer geforderten Nachvollziehbarkeit der Randomisierungsergebnisse wird die Randomisierung in der Praxis üblicherweise durch die Anwendung von Computerprogrammen unterstützt. Allerdings stehen bislang kaum auf Randomisierung spezialisierte Programme zur Verfügung. Daher haben wir in diesem Beitrag die Entwicklung des Programms RITA beschrieben. In diesem sind die Vollständige Randomisierung, das Biased Coin Design (EFRON, 1971), die Minimierung (POCOCK und SIMON, 1975), das Urnenmodell (WEI, 1978a), die Permutierte Blockrandomisierung sowie die Self-Adjusting Randomisierung (NORDLE und BRANTMARK, 1977) implementiert. Das Programm erlaubt eine sehr flexible Konfiguration der Verfahren, beispielsweise ist eine patientenweise oder eine blockweise Randomisierung von Patienten möglich. Verschiedene Zufallszahlen- und Seed-Generatoren können verwendet werden. Eine Reihe von Randomisierungsverfahren, die in der Literatur vorgestellt wurden, sind bislang nicht implementiert. Hierzu zählen beispielsweise die einfache „Random allocation rule“, das „Truncated binomial design“ (BLACKWELL und HODGES, 1957), das „Big Stick Design“ (SOARES und WU, 1983), das „Conditional Markov chain design“ (MATTS und MCHUGH, 1983), das „Dynamic balanced design“ (SIGNORINI et al., 1993), oder das „Marginal urn design“ (WEI, 1978b). Es ist jedoch durch die objektorientierte Programmierweise sehr einfach möglich, solche Verfahren in RITA zu integrieren.

RITA lässt sich problemlos um weitere Funktionen erweitern. Wir haben gezeigt, dass die von RITA erzeugten Ergebnisse nachvollziehbar und reproduzierbar sind, somit gewährleistet RITA eine leichte und sichere Handhabung.

Danksagung

Wir bedanken uns sehr herzlich bei PD Dr. Günther Kundt für hilfreiche Diskussionen zu einer früheren Version dieses Manuskripts.

Literaturverzeichnis

- BLACKWELL, D., HODGES, J. (1957): Design for the control of selection bias, *Ann. Math. Stat.* **28**, 449–460.
- CDRH: General Principles of Software Validation, Final Guidance for Industry and FDA Staff, <http://www.fda.gov/cdrh/comp/guidance/938.pdf>.
- COMET (Comparative Obstetric Mobile Epidural Trial) Study Group UK (2001): Effect on low-dose mobile versus traditional epidural techniques on mode of delivery: a randomised controlled trial, *Lancet* **358**, 19–23.
- EFRON, B. (1971): Forcing a sequential experiment to be balanced, *Biometrika* **58**, 403–417.
- FDA (820.3): Quality system regulation, Part 820, <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?FR=820.3>.
- FRANE, J. W. (1998): A method of biased coin randomization, its implementation, and its validation, *Drug Information Journal* **32**, 423–432.
- ICH E6 (1996): Guideline for good clinical practice, CPMP/ICH/135/95, The European Agency for the Evaluation of Medicinal Products – Human Medicines Evaluation Unit, <http://www.eudra.org/humandocs/humans/ich.htm>.
- ICH E9 (1998): Statistical principles for clinical trials, CPMP/ICH/363/96, The European Agency for the Evaluation of Medicinal Products – Human Medicines Evaluation Unit, <http://www.eudra.org/humandocs/humans/ich.htm>.
- ISO 8402 (1994-04-01): Quality management and quality assurance – Vocabulary, <http://home.netvigator.com/~timwwh01/ISO-8402.htm>.
- KALISH, L. A., BEGG, C. B. (1985): Treatment allocation methods in clinical trials: a review, *Stat. Med.* **4**, 129–44.
- KRACHT, M., 1992: *Randomisierung*. In: Adam J. (ed) *Statistisches Know-how in der medizinischen Forschung*. Ullstein Mosby GmbH, Berlin, 172–210.
- KUNDT, G. (2002): Randomisierungsverfahren für kontrollierte klinische Studien. Eine Analyse ihres Balanceverhaltens und ihr Einfluss auf die Auswertung klinischer Studien mittels Permutationstests. Shaker Verlag, Aachen.
- LACHIN, J. M. (1988a): Properties of simple randomization in clinical trials, *Controlled Clin. Trials* **9**, 312–26.
- LACHIN, J. M. (1988b): Statistical properties of randomization in clinical trials, *Controlled Clin. Trials* **9**, 289–311.
- LACHIN, J. M., MATTS, J. P., WEI, L. J. (1988): Randomization in clinical trials: conclusions and recommendations, *Controlled Clin. Trials* **9**, 365–74.
- MATTS, J. P., LACHIN, J. M. (1988): Properties of permuted-block randomization in clinical trials, *Controlled Clin. Trials* **9**, 327–44.
- MATTS, J. P., MCHUGH, R. B. (1983): Conditional Markov chain design for accrual clinical trials, *Biom. J.* **25**, 563–577.
- MOHER, D., SCHULZ, K. F., ALTMAN, D. G. (2001): The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials, *Ann. Intern. Med.* **134**, 657–662.
- NORDLE, O., BRANTMARK, B. (1977): A self-adjusting randomization plan for allocation of patients into two treatment groups, *Clin. Pharmacol. Ther.* **22**, 825–830.
- POCOCK, S. J., SIMON, R. (1975): Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial, *Biometrics* **31**, 103–115.
- ROSENBERGER, W. F., LACHIN, J. M. (2002): *Randomization in Clinical Trials: Theory and Practice*. Wiley, New York.
- SCHÄFER, H. (1997): Anforderungen an eine patientenorientierte klinisch-therapeutische Forschung, *Dtsch. Med. Wochenschr.* **122**, 1531–1536.
- SCHÄFER, H., BERGER, J., BIEBLER, K.-E., FELDMANN, U., GREISER, E., JÖCKEL, K.-H., MICHAELIS, J., NEISS, A., RASPE, H. H., ROBRA, B.-P., SCHUMACHER, M., TRAMPISCH, H.-J., VICTOR, N., WINDE-

- LER, J. (1999): Empfehlungen für die Erstellung von Studienprotokollen (Studienplänen) für klinische Studien, *Informatik, Biometrie und Epidemiologie in Medizin und Biologie* **30**, 141–54.
- SCHULZ, K. F., GRIMES, D. A. (2002a): Generation of allocation sequences in randomised trials: chance, not choice, *Lancet* **359**, 515–9.
- SCHULZ, K. F., GRIMES, D. A. (2002b): Unequal group sizes in randomised trials: guarding against guessing, *Lancet* **359**, 966–70.
- SIGNORINI, D. F., LEUNG, O., SIMES, R. J., BELLER, E., GEBSKI, V. J. (1993): Dynamic balanced randomization for clinical trials, *Stat. Med.* **12**, 2343–2350.
- SOARES, J. F., WU, C. F. J. (1983): Some restricted randomization rules in sequential designs, *Commun. Statist. Theor. Meth.* **17**, 2017–2034.
- TAVES, D. R. (1974): Minimization: a new method of assigning patients to treatment and control groups, *Clin. Pharmacol. Ther.* **15**, 443–53.
- WEI, L. J. (1978a): The adaptive biased coin design for sequential experiments, *Ann. Statist.* **6**, 92–100.
- WEI, L. J. (1978b): An application of an urn model to the design of sequential controlled clinical trials, *J. Am. Stat. Assoc.* **73**, 559–563.
- WEI, L. J., LACHIN, J. M. (1988): Properties of the urn randomization in clinical trials, *Controlled Clin. Trials* **9**, 345–64.
- ZELEN, M. (1974): The randomization and stratification of patients to clinical trials, *J. Chronic. Dis.* **27**, 365–75.

Prof. Dr. Andreas Ziegler, Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Ratzeburger Allee 160, Haus 4, D-23538 Lübeck, Tel.: 04 51-5 00-27 81, Fax: 04 51-5 00-27 99, ziegler@imbs.uni-luebeck.de