

Methodik

Reihe Epidemiologie 8: Verblindung in randomisierten Studien: Wie man verdeckt, wer was erhalten hat

Kenneth F. Schulz*, David A. Grimes

Family Health International, PO Box 13950, Research Triangle Park NC 27709, USA

Lancet 2002, 359: 696–700

Die reiche Geschichte der Verblindung erstreckt sich über zwei Jahrhunderte. Weltweit wissen die meisten Wissenschaftler, was mit Verblindung gemeint ist, doch hinter einem nur allgemeinen Verständnis lauert oftmals Verwirrung. Begriffe wie einfachblind, doppelblind und dreifachblind haben für verschiedene Personen unterschiedliche Bedeutung. Außerdem verwechseln viele medizinische Forscher die Verblindung mit der Geheimhaltung der Randomisierungsliste. Solche Verwechslungen zeigen, dass beide Begriffe missverstanden werden. Verblindung bedeutet, daß Studienteilnehmer, Studienärzte (meist diejenigen, die die Behandlung verabreichen) oder Bewerter (diejenigen, die die Outcome-Daten erheben) die zugeeilte Intervention nicht kennen, sodass sie durch dieses Wissen nicht beeinflusst werden. Im Allgemeinen führt die Verblindung zu einem Rückgang der unterschiedlichen Bewertung von Behandlungsergebnissen (Informationsbias). Sie kann bei den Studienteilnehmern aber auch die Compliance und die Bereitschaft zum Verbleib in der Studie verbessern, während sie die Anwendung zusätzlicher Pflege- oder Therapiemaßnahmen (manchmal als Ko-Intervention bezeichnet) reduziert. Studienärzte und Leser gehen vielfach gutgläubig davon aus, dass eine randomisierte

Studie einfach dann als hochwertig anzusehen ist, wenn sie doppelt verblindet wurde, als ob es sich bei der doppelten Verblindung um die *Conditio sine qua non* einer randomisierten, kontrollierten Studie handelte. Auch wenn die Doppelverblindung (Verblindung von Studienärzten, Studienteilnehmern und Bewertern der Behandlungsergebnisse) auf eine solide Studienplanung schließen lässt, dürfen Studien ohne Doppelverblindung nicht automatisch als minderwertig eingestuft werden. Anstatt sich ausschließlich auf Begriffe wie Doppelverblindung zu verlassen, sollten Wissenschaftler explizit angeben, wer wie verblindet wurde. Wir empfehlen, den Ergebnissen mehr Glauben zu schenken, wenn zumindest die Bewertung der Studienergebnisse verblindet erfolgte, mit Ausnahme von objektiven Endpunkten wie etwa der Mortalität, die kaum Raum für systematische Fehler (Bias) lassen. Wenn Studienärzte ihre Verblindungsmaßnahmen korrekt angeben, kann der Leser sie beurteilen. Leider lässt die Berichterstattung in vielen Artikeln zu wünschen übrig. Wenn in einem Beitrag behauptet wird, es habe eine Verblindung stattgefunden, ohne dies weiter zu erklären, sollte der Leser hinsichtlich ihrer Bias reduzierenden Wirkung skeptisch bleiben.

Die Geschichte der Verblindung in klinischen Studien kann auf einen langen Zeitraum von zwei Jahrhunderten zurückblicken [1]. Weltweit wissen die meisten Wissenschaftler, was damit etwa gemeint ist. Leider lauert hinter einem nur allgemeinen Verständnis oftmals aber Verwirrung. Begriffe wie einfachblind, doppelblind und dreifachblind haben für verschiedene Men-

schen unterschiedliche Bedeutung [2]. Hinzu kommt, dass viele Forscher Verblindung mit der Geheimhaltung der Randomisierungsliste verwechseln. Dass eine solche Verwirrung besteht, deutet darauf hin, dass beide Begriffe missverstanden werden. Doch zwischen ihnen bestehen klare theoretische und praktische Unterschiede. Verblindung verhindert Ascertainment Bias

(Beurteilungsbias, Informationsbias) und schützt die Randomisierungsliste nach der Zuteilung [3,4]. Verfahren zur Geheimhaltung der Randomisierungsliste benutzen Wissenschaftler aber in erster Linie, um Selektionsbias vorzubeugen und um Randomisierungslisten vor der und bis zur Behandlungszuteilung zu schützen. Außerdem lässt sich eine Verblindung in manchen Studien

*Zuschriften an: Dr. Kenneth F. Schulz
E-Mail: KSchulz@fhi.org (K.F. Schulz).

gar nicht erfolgreich durchführen, während die Geheimhaltung der Randomisierungsliste immer erfolgreich umgesetzt werden kann [4,5].

Verblindung stellt einen wichtigen, davon verschiedenen Aspekt randomisierter, kontrollierter Studien dar [3]. Der Begriff Verblindung bedeutet, dass Studienteilnehmer, Studienärzte (meist diejenigen, die die Behandlung verabreichen) oder Bewerter (diejenigen, die die Outcome-Daten erheben) die zugeteilte Intervention nicht kennen, sodass sie durch dieses Wissen nicht beeinflusst werden. Verblindung verhindert Bias in mehreren Phasen einer Studie, sie hat jedoch – je nach den Umständen – unterschiedliche Relevanz. Auch wenn bei den ersten Verblindungsversuchen vielleicht noch Augenbinden benutzt wurden [1], so sind die Vorgehensweisen mittlerweile doch deutlich raffinierter geworden. In diesem Artikel werden wir uns hauptsächlich mit den Merkmalen und Vorteilen der Verblindung befassen.

Potenzielle Effekte von Verblindung

Wenn Studienteilnehmer nicht verblindet werden, kann die Kenntnis der

Gruppenzuteilung das Ansprechen auf die verabreichte Intervention beeinflussen [3]. Teilnehmer, die wissen, dass sie einer Gruppe zugeteilt wurden, die eine neue Therapie erhält, könnten positive Erwartungen hegen oder eine erhöhte Besorgnis entwickeln. Die der Standardtherapie zugeteilten Teilnehmer könnten sich dagegen benachteiligt fühlen oder aber Erleichterung verspüren. Obwohl die Evidenz darauf schließen lässt, dass neue Therapien mit gleich großer Wahrscheinlichkeit sowohl schlechter als auch besser sein können als die Standardbehandlung [6], gehen Studienteilnehmer vermutlich davon aus, dass neue Therapien besser sind als Standardtherapien – neu also im Sinne von verbessert. In jedem Fall können das Wissen um die verabreichte Intervention und die subjektive Wahrnehmung dieser Behandlung die psychischen oder physischen Reaktionen der Teilnehmer beeinflussen. Ferner kann die Kenntnis der Behandlungszuteilung auch einen Einfluss auf die Compliance und die Bereitschaft der Studienteilnehmer zum Verbleib in der Studie ausüben (siehe **Kasten 1**).

Wichtig ist auch die Verblindung der Studienärzte – also derjenigen, die das Studienteam im weitesten Sinne dar-

stellen; dieses Team besteht aus allen Personen (aber nicht ausschließlich), die die Studie geplant haben, die Teilnehmer rekrutieren, die Randomisierung umsetzen, die medizinische Versorgung leisten, die bei der Intervention beratend zur Verfügung stehen, oder die Routinedaten erheben [3]. Zu den Studienärzten, bei denen die Verblindung besonders wichtig ist, zählen alle Behandelnden (wie der behandelnde Arzt oder die Krankenschwester) und Interventionsberater (z.B. jemand, der eine verhaltensbezogene Präventionsbotschaft vermittelt), die während der Studie mit den Teilnehmern in Kontakt kommen. Wenn die Studienärzte unverblindet bleiben, kann sich ihre Einstellung für oder gegen eine Intervention direkt auf die Teilnehmer übertragen [7]. Ihre persönliche Meinung könnte sich auch in der unterschiedlichen Anwendung ergänzender Interventionen im Sinne einer zusätzlichen pflegerischen oder therapeutischen Maßnahme (Ko-Interventionen), in unterschiedlichen Entscheidungen über den Ausschluss von Teilnehmern aus der Studie oder in unterschiedlichen Anpassungen der Arzneimitteldosis (siehe **Kasten 1**) äußern. In Kenntnis der Interventionsgruppenzuteilung

Kasten 1. Potenzielle Vorteile der Verblindung in Abhängigkeit von den erfolgreich verblindeten Personengruppen.

Verblindete Personen

Studienteilnehmer

Mögliche Vorteile

- geringere Wahrscheinlichkeit einer Verzerrung ihres psychischen oder physischen Ansprechens auf die Intervention
- höhere Wahrscheinlichkeit für eine Befolgung der Therapieschemata
- geringere Wahrscheinlichkeit des Bedarfs an zusätzlichen Interventionen
- geringere Wahrscheinlichkeit für ein Ausscheiden aus der Studie ohne verwertbare Daten, was zu Nachbeobachtungsverlusten führen würde

Studienärzte

- geringere Wahrscheinlichkeit der Übertragung von Intentionen oder Einstellungen auf die Studienteilnehmer
- geringere Wahrscheinlichkeit für die unterschiedliche Verabreichung von Ko-Interventionen
- geringere Wahrscheinlichkeit für unterschiedliche Dosisanpassungen
- geringere Wahrscheinlichkeit für die unterschiedliche Handhabung des Ausschlusses von Studienteilnehmern
- geringere Wahrscheinlichkeit, dass sie Teilnehmer in unterschiedlicher Weise zum Verbleib in der Studie ermutigen oder ihnen davon abraten

Bewerter

- geringere Wahrscheinlichkeit einer Bias-bedingten Beeinträchtigung ihrer Outcomebewertungen, vor allem bei subjektiven Zielgrößen

könnten Studienärzte Teilnehmer ferner zum Verbleib in einer Studie ermutigen oder ihnen davon abraten. Vielleicht am wichtigsten ist aber die Tatsache, dass Verblindung dazu beiträgt, die unterschiedliche Bewertung der Behandlungsergebnisse zu reduzieren (oft als Informations-, Beurteilungs- oder Ascertainment Bias bezeichnet) (siehe **Kasten 1**). Wenn zum Beispiel die Bewerter der Behandlungsergebnisse, die die Behandlungszuteilung kennen, glauben, dass eine neue Intervention besser ist als eine bekannte, dann könnten sie die Reaktionen auf diese Intervention großzügiger erfassen. Tatsächlich ergaben in einer placebo-kontrollierten Studie bei Patienten mit Multipler Sklerose [8] die Bewertungen der unverblindeten, nicht jedoch die der verblindeten Neurologen einen auffälligen Interventionsnutzen. Subjektive Zielgrößen wie etwa Schmerzscores sind ein Tummelplatz für Bias [3]. Außerdem können einige objektiv beurteilte Zielgrößen – beispielsweise Salpingitis – voller Subjektivität stecken. Im Allgemeinen jedoch nimmt die Bedeutung der Verblindung als Mittel zur Reduktion von Beurteilungsbias bei weniger subjektiven Zielgrößen ab, da objektive (harte) Outcomes kaum eine Gelegenheit für Bias bieten. Die Erfassung eines harten Outcome wie der Mortalität würde durch die Kenntnis der Intervention kaum beeinträchtigt werden.

Das Lexikon der Verblindung

Als unverblindete (offene oder Open-Label-) Studien bezeichnet man Untersuchungen, in denen alle Beteiligten wissen, wer während der Studie welche Interventionen erhalten hat. Eine Verblindung (Maskierung) bedeutet, dass das Wissen um die Behandlungszuteilung vor den Teilnehmern, Studienärzten oder Bewertern geheim gehalten wird.

Der Begriff einfachblind bedeutet im Allgemeinen, dass eine der drei Personengruppen (normalerweise eher die Studienteilnehmer als die Studienärzte) während der Studie keine Kenntnis von

den Behandlungszuteilungen erhält [9]. Eine einfachblinde Studie könnte theoretisch also auch bedeuten, dass zwar sowohl die Teilnehmer als auch die Studienärzte die Intervention kennen, nicht aber der Bewerter – und das wäre etwas verwirrend.

In einer Doppelblindstudie bleiben meist alle – Studienteilnehmer, Studienärzte und Bewerter – während der gesamten Studie in Unkenntnis der Behandlungszuteilungen [3]. Angesichts der Tatsache, dass es drei Gruppen sind, die nicht über die Zuteilung Bescheid wissen, ist der Begriff doppelblind manchmal etwas irreführend. In der medizinischen Forschung aber wird die Bewertung häufig auch durch einen Studienarzt durchgeführt, sodass der Begriff in diesem Fall völlig korrekt auf zwei Personengruppen Bezug nimmt. Der Begriff dreifachblind bezieht sich in der Regel auf eine Doppelblindstudie, bei der auch die Datenauswertung verblindet erfolgt [10]. Manche Studienleiter bezeichnen Studien aber auch dann als dreifachblind, wenn Studienärzte und Bewerter unterschiedliche Personen sind und beide Gruppen – abgesehen von den Teilnehmern – keine Kenntnis von den Behandlungszuteilungen haben. Der Begriff vierfachblind wird von Studienleitern nur selten verwendet und wenn, dann bezieht er sich auf die Verblindung von Studienteilnehmern, Studienärzten, Bewertern und Statistikern [11]. Fünffachblind müsste demnach bedeuten, dass die Randomisierungsliste verloren gegangen ist und keiner mehr irgendetwas weiß. Mae Wests Behauptung „Zu viel des Guten kann wundervoll sein!“ trifft auf die Verblindung nur bedingt zu.

Im Zusammenhang mit einfacher, doppelter und dreifacher Verblindung zieht sich ein wahrer Begriffswirrwarr durch die Literatur [3], wobei Ärzte, Lehrbücher und Zeitschriftenartikel allesamt mit unterschiedlichen Interpretationen und Definitionen aufwarten [2]. Studienleiter definieren Doppelblindstudien nicht nur nicht einheitlich, sondern sie verschlimmern die Sache vor allem noch dadurch, dass sie in ihren Veröffentlichungen häufig keine eindeutigen Definitionen angeben. Und nun bringen wir – ausgehend von den

ursprünglichen Bemühungen um Verblindung [1] und der einst gebräuchlichen Doppelverblindungsterminologie [12] – durch zwei zusätzliche Definitionen der einfachen und doppelten Verblindung (siehe **Abb. 1**) noch mehr Dunkel ins Spiel. Doch Spaß beiseite: Wenn wir in diesem Artikel den Begriff doppelblind oder verwandte Begriffe benutzen, meinen wir damit, dass Maßnahmen unternommen wurden, um Teilnehmer, Studienärzte und Bewerter gegenüber den Gruppenzuteilungen zu verblinden. Hinsichtlich der Berichterstattung zu randomisierten, kontrollierten Studien fordern wir dringend dazu auf, explizit anzugeben, welche Maßnahmen zur Verblindung welcher Personen ergriffen wurden.

Viel üblicher sind eher knappe Informationen zur Verblindung. Viele Studienleiter versäumen es anzugeben, ob ihre Studie verblindet war oder nicht. So wurde in 51% der 506 Studien über Mukoviszidose [13], 33% der 196 Studien über rheumatoide Arthritis [14] und 38% von 68 dermatologischen Studien [15] nicht angegeben, ob sie verblindet durchgeführt worden waren. Wenn Studien als doppelblind ausgewiesen wurden, fehlte es häufig an weiteren Erläuterungen [14,16–18]. Beispielsweise wurde in nur 14 (45%) von 31 doppelblinden geburtshilflichen und gynäkologischen Studien auf die Ähnlichkeit der Behandlungs- und der Kontrollschemas (etwa Aussehen, Geschmack, Verabreichung) hingewiesen, und nur 5 (16%) enthielten Aussagen, die erkennen ließen, dass die Verblindung erfolgreich war [18].

Maskierung oder Verblindung

Manche sprechen lieber von Maskierung statt Verblindung, beschreiben damit aber dasselbe Vorgehen. Der Begriff Maskierung ist womöglich eher angebracht bei Studien, an denen Patienten mit Sehstörungen teilnehmen, und könnte in Studien, in denen Blindheit als Zielgröße untersucht wird, sogar weniger Verwirrung stiften [3]. Verblindung vermittelt jedoch die starke Botschaft, dass sie Bias verhindern



Abb. 1. Die Autoren: der eine doppelt, der andere einfach verblindet.



Abb. 2. Die Autoren: der eine verblindet, der andere maskiert.

kann. Offensichtlich geht der Begriff Verblindung auf Benjamin Franklin und Kollegen [19] zurück, die den Teilnehmern einer Untersuchung über die angeblichen therapeutischen Wirkungen des Mesmerismus tatsächlich die Augen verbunden hatten, um ihr Urteil durch entsprechendes Wissen nicht zu beeinflussen. Die Vorstellung von Verblindung als der vollständigen Abdeckung der Augen vermittelt eine stärkere Biasprävention als die Maskierung, weil Masken Sehschlitze enthalten können, durch die man sehen kann (siehe Abb. 2). Manchen vermittelt der Begriff

Verblindung auch das Gefühl, es handele es sich dabei um ein sicheres Verfahren. Im Leitfaden der *International Conference on Harmonization (ICH)* [20] zum Beispiel wird vorzugsweise der Begriff Verblindung verwendet. (Die ICH steht für eine intensive trilaterale Zusammenarbeit zwischen den Zulassungsbehörden in Europa, Japan und den USA, mit dem Ziel der Entwicklung gemeinsamer Leitlinien für die Planung, Implementierung und Berichterstattung klinischer Studien). Auch wir ziehen den Begriff Verblindung vor, da er historisch weit zurückreicht, internatio-

nal anerkannt ist, über einen starken Symbolcharakter verfügt und sich auch in den ICH-Leitlinien durchgesetzt hat [3].

Placebos und Verblindung

Manchmal haben Interventionen (Behandlungen) keinen Effekt auf die untersuchten Zielgrößen [3]. Wenn allerdings im Kontext einer gut geplanten randomisierten, kontrollierten Studie den Teilnehmern eine unwirksame Intervention verabreicht wird, kann es gelegentlich zu positiven Effekten auf die Einstellungen der Teilnehmer kommen, was sich wiederum auf die Zielgrößen auswirkt [10]. Wissenschaftler bezeichnen dieses Phänomen als Placebo-Effekt.

Unter einem Placebo versteht man eine pharmakologisch unwirksame Substanz, die Studienärzte den Teilnehmern in der Kontrollgruppe einer Studie verabreichen [3]. Die Verwendung einer Placebo-Kontrollgruppe gleicht den Placebo-Effekt in der Behandlungsgruppe aus und ermöglicht eine unabhängige Beurteilung des Therapieeffekts. Auch wenn Placebos psychologische Wirkungen haben können, werden sie den Studienteilnehmern verabreicht, weil sie ansonsten wirkungslos sind. Als aktives Placebo bezeichnet man ein Placebo, das Symptome oder Nebenwirkungen imitiert (z.B. Mundtrockenheit, Schweißausbrüche), deren Fehlen die Identität der (pharmakologisch) aktiven Testbehandlung andernfalls preisgeben würde. Die meisten Wissenschaftler sind sich darüber einig, dass Placebos den Kontrollpersonen möglichst immer dann verabreicht werden sollten, wenn die Wirkungen einer neuen Therapie für eine Krankheit untersucht werden sollen, für die noch keine wirksame Behandlung existiert [9,10]. Tatsächlich macht Verblindung häufig die Anwendung von Placebos erforderlich.

Normalerweise allerdings wird der Kontrollgruppe eine nachweislich wirksame Standardtherapie verabreicht, wenn es eine solche gibt, um diese mit einer neuen Therapie zu vergleichen [3]. Demnach können Studienärzte zwei aktive Behandlungsgruppen ohne eine

Placebogruppe vergleichen. Aber selbst dann versuchen Studienärzte häufig, eine Verblindung zu erreichen, indem sie die Double-Dummy-Methode anwenden, im Wesentlichen zwei Placebos [11,21]. Um zwei Wirkstoffe miteinander zu vergleichen (der eine in einer blauen Kapsel, der andere in einer roten Kapsel), würden die Studienärzte blaue und rote Placebokapseln vorbereiten. Dann erhielten beide Behandlungsgruppen eine blaue und eine rote Kapsel – die eine wirksam, die andere unwirksam.

Kann eine Verblindung Bias verhindern?

Von manchen Studienleitern, Lesern und Herausgebern wird die Bedeutung der Verblindung hinsichtlich der Prävention von Bias überbetont. Tatsächlich betrachten einige eine randomisierte Studie als hochwertig, wenn sie doppelt verblindet ist, so als ob eine Doppelverblindung die *Conditio sine qua non* einer randomisierten, kontrollierten Studie wäre [3]. So einfach ist es in der Wissenschaft aber leider nicht. Eine randomisierte Studie kann methodisch fundiert und nicht doppelt verblindet sein oder, umgekehrt kann sie doppelt verblindet sein und methodische Schwächen aufweisen. Lasagna [12] hat diesen Gedanken schon vor langer Zeit folgendermaßen ausgedrückt: „Lassen Sie uns das Placebo einmal etwas kritischer unter die Lupe nehmen, denn Placebo und Doppelverblindung haben in unserer Vorstellung wie auch in der Literatur nahezu Fetischstatus erlangt. Man kann problemlos zeigen, dass die Aura der Respektabilität, der Unfehlbarkeit und des wissenschaftlichen Know-how, die sie in den Augen vieler Menschen automatisch zu umgeben scheint, in bestimmten Fällen ganz unverdient ist.“ [12] Auch wenn Doppelverblindung auf einen soliden Studienplan schließen lässt, ist sie nicht der Hauptindikator für die Gesamtqualität einer Studie. Darüber hinaus lassen sich viele Studien gar nicht doppelt verblinden. Solche Studien müssen daher anhand ihrer Vorzüge insgesamt beurteilt werden und nicht

nach einem auf Doppelverblindung beruhenden ungeeigneten Standard. Damit wollen wir jedoch nicht behaupten, dass eine Verblindung irrelevant ist [3]. Intuitiv betrachtet sollte eine Verblindung Bias reduzieren, und die verfügbare Evidenz stützt diesen Eindruck auch. Untersuchungen zur Methodik lassen tendenziell jedoch erkennen, dass eine Doppelverblindung Bias zwar verhindert, zur Vermeidung von Bias im Allgemeinen aber weniger wichtig ist als die adäquate Geheimhaltung der Randomisierungsliste [4,22,23].

Worauf der Leser bei Beschreibungen von Verblindung achten sollte

Wenn Wissenschaftler eine Studie als doppelblind beschreiben, kann der Leser in der Regel davon ausgehen, dass systematische Fehler (Bias) ausgeschaltet wurden. Diese Empfehlung wird durch empirische Evidenz gestützt. Wie in den CONSORT-Leitlinien vorgeschlagen [24,25], sollten Studienleiter aber nicht nur die Begriffe einfachblind, doppelblind oder dreifachblind verwenden, sondern explizit angeben, wer wie verblindet wurde. Wenn zudem behauptet wird, dass Studienärzte, Studienteilnehmer und Bewerter verblindet waren – d.h. Doppelverblindung bestand – dann sollten Informationen zum Verblindungsmechanismus (Kapseln, Tabletten, Film etc.) enthalten sein, zur Ähnlichkeit der Behandlungscharakteristika (Aussehen, Geschmack, Verabreichung) und zur Kontrolle der Randomisierungsliste – also z.B. zum Aufbewahrungsort der Liste während der Studie, zum Zeitpunkt der Offenlegung des Codes für die Auswertung und Angaben zu den besonderen Umständen, unter denen der Code in Einzelfällen gebrochen werden durfte. Solche Zusatzinformationen können die Behauptung, dass doppelt verblindet wurde, stützen oder untergraben (siehe **Kasten 2**) [26–29].

Wenn Wissenschaftler in angemessener Form über ihre Verblindungsmaßnahmen berichten, kann der Leser diese Bemühungen beurteilen. Leider

mangelt es vielen Veröffentlichungen an einer ordentlichen Berichterstattung. Wenn ein Wissenschaftler behauptet, er habe seine Studie verblindet durchgeführt, diese Behauptung aber nicht durch weitere Informationen untermauert, sollte der Leser skeptisch bleiben, was die Wirkung der Verblindung hinsichtlich der Reduktion von Bias betrifft. Beispielsweise wurde für eine Studie [30] über Antibiotikaprophylaxe behauptet, sie sei verblindet gewesen, der Methodenteil des Studienberichts ließ jedoch erkennen, dass kaum eine oder gar keine Verblindung erfolgt war.

Im Idealfall sollten Wissenschaftler auch angeben, ob die Verblindung erfolgreich war. Theoretisch können Studienärzte den Erfolg einer Verblindung beurteilen, indem sie Studienteilnehmer, das beteiligte medizinische Personal oder die Zielgrößenbewerter fragen, welche Intervention ihrer Meinung nach verabreicht wurde (siehe **Kasten 3**). Im Prinzip dürften die Befragten bei erfolgreicher Verblindung mit ihren Antworten nicht viel besser als nur zufällig richtig liegen. Trotz absolut erfolgreicher Verblindung wäre es in der Praxis aber dennoch möglich, dass Studienteilnehmer, medizinisches Personal und Zielgrößenbewerter die Intervention erraten, weil sie über zusätzliche Informationen verfügen. So könnte etwa eine unverhältnismäßig hohe Rate an unerwünschten Wirkungen starke Hinweise auf die Intervention geben. Ungeachtet gewissenhafter Bemühungen um die Doppelverblindung von Studien haben manche Interventionen Nebenwirkungen, die so eindeutig sind, dass ihr Auftreten sowohl den Teilnehmern als auch den medizinischen Betreuern die verabreichte Intervention zwangsläufig verrät [11,24]. Noch wichtiger als Hinweise aufgrund von unerwünschten Wirkungen sind die mit den klinischen Ergebnissen verbundenen Anhaltspunkte. Gewöhnlich freuen sich Wissenschaftler über große klinische Effekte (außer vielleicht in Äquivalenzstudien). Wenn es dazu kommt, könnten das klinische Personal und die Teilnehmer daraus wahrscheinlich – natürlich nicht immer korrekt – ableiten, dass ein Teilnehmer mit einem

Kasten 2. Beschreibungen von Verblindung.

„Weder Patienten, Studienkrankenschwestern, Studienärzte noch sonstiges medizinisches oder Pflegepersonal der Intensivstation hatte während der Studiendauer Kenntnis von den Behandlungszuteilungen. Auch sämtliche statistischen Analysen wurden unter Beibehaltung der Maskierung durchgeführt. Die Randomisierungsstelle war angewiesen, jeden Verdacht auf eine Verletzung der Maskierungsmaßnahmen zu melden. Es wurde kein solcher Fall gemeldet. ... Das Medikament bzw. das Placebo (Präparat ohne aktiven Wirkstoff) für die Infusion mittels Spritzenpumpe oder volumetrischer Pumpe wurde in nicht unterscheidbaren Spritzen oder Beuteln zubereitet“ [26].

„ ... erfolgte doppelblind und placebokontrolliert. ... Weder Patienten noch Ärzte konnten die Placebo- von den Sibutraminkapseln unterscheiden. Der Geschmack der Kapseln war gleich, vorausgesetzt, sie wurden entsprechend der Anweisung im Ganzen geschluckt. ... Die Ergebnisse der biochemischen Analysen lagen vor, bevor der Randomisierungscode am Ende der abgeschlossenen Studie offen gelegt wurde“ [27].

„Die Studie war doppelt verblindet, d.h. weder die Teilnehmerinnen noch das Studienpersonal (darunter auch die Biostatistiker bei Family Health International) wussten, in welcher Gruppe der Nonoxynol-9 Vaginalfilm angewendet wurde. Der Nonoxynol-9 Vaginalfilm enthielt Der Placebo-Film enthielt ... Die Vaginalfilme waren in Aussehen, Verpackung und Beschriftung identisch“ [28].

„Doxycyclin und Placebo lagen in Kapselform vor und sahen gleich aus. ... Der Randomisierungscode verblieb in den USA.“ [Hinweis: Die Studie wurde in Kenia durchgeführt.] „Demnach erfolgten die Verabreichung und alle Bewertungen gegenüber der Behandlungszuteilung verblindet; Studienärzte und Patienten waren auch gegenüber den jeweils aktuellen Studienergebnissen verblindet. Der Code wurde erst nach Abschluss der Datenerhebung offen gelegt“ [29].

Kasten 3. Bewertung des Verblindungserfolgs.

„Wir befragten 126 Beschäftigte danach, welcher Film ihrer Meinung nach das Placebo war. 18% hielten Film A (das Placebo) für das Placebo, 13% dachten, Film B (Nonoxynol-9) sei das Placebo, und 69% konnten sich nicht entscheiden. Von den 68 „peer educators“ (die Mitarbeiter, die am ehesten die Meinung der Studienteilnehmer reflektieren) hielten 16% Film A für das Placebo, 13% Film B und 71% waren unentschieden“ [28].

positiven Ergebnis statt der Kontrollintervention (Standardtherapie) die aktive (neue) Intervention erhalten hat. Wenn sich die aktive (neue) Intervention tatsächlich als vorteilhaft erweist (höchst wünschenswert), dann wären die Folgerungen, die die Befragten daraus ziehen können, öfter als nur zufallsbedingt korrekt [24,31]. Abgesehen von ihren Vermutungen erfragt ein am Ende der Studie durchgeführter Test auf Verblindung in Wirklichkeit womöglich eher das Gefühl der Befragten für unerwünschte Wirkungen oder Wirksamkeit [32,33].

Überdies könnte es sein, dass etwaige Entblindungsbemühungen nur ungerne zugegeben werden, etwa durch präzise Antworten auf entsprechende Nachfragen – mit anderen Worten: Wenn die Gruppenzuteilungen entschlüsselt wurden, könnten die Befragten gegenteilige Antworten zu ihren Entschlüsselungsergebnissen geben, um ihr Vorgehen zu verschleiern. Dieses Problem

sowie die mit unerwünschten Nebenwirkungen und erfolgreichen klinischen Ergebnissen verbundenen Interpretationsschwierigkeiten veranlassen uns, die Nützlichkeit von Verblindungstests unter bestimmten Umständen zu hinterfragen. Studienleiter sollten sich gründlich überlegen, ob es sinnvoll ist, den Erfolg ihrer Verblindungsmaßnahmen zu prüfen; wenn sie es jedoch tun, sollten sie die Ergebnisse etwaiger Beurteilungen angeben. Zumindest sollten sie mitteilen, wenn ihre Verblindungsmaßnahmen versagt haben, wie etwa nicht-identische Placebo- oder Verumzubereitungen. Veröffentlichte Studienberichte enthalten nur selten Beurteilungen von Verblindungsmaßnahmen, wenn doch, dann sollte der Leser die Informationen dazu mit Skepsis beurteilen.

In vielen Studien ist eine Doppelverblindung schwierig oder gar unmöglich. Das trifft beispielsweise auf die meisten chirurgischen Studien zu. Insbesondere

kann etwa eine Studie, in der die mit der Blutentnahme aus dem Ohr oder Daumen verbundenen Schmerzen verglichen werden, nicht doppelblind durchgeführt werden [34]. Aber auch wenn Wissenschaftler ihre Studien nicht als doppelblind o. Ä. beschreiben, kann es sich um wissenschaftlich aussagekräftige Untersuchungen handeln. Abgesehen von der Beurteilung der anderen methodischen Aspekte einer Studie müsste der Leser bewerten, wie viel Bias die fehlende Verblindung zur Folge haben könnte. Der Leser sollte feststellen, ob irgendjemand in der Studie verblindet wurde und welche Vorteile sich daraus ergeben haben könnten (siehe Kasten 1). Tatsächlich ist die Verblindung der Zielgrößenbewerter – selbst in offenen Studien – häufig nicht nur möglich, sondern ratsam [11]. So können etwa Läsionen vor und nach der Behandlung fotografiert und von einem Unbeteiligten bewertet werden [11]. Wir empfehlen, Ergebnis-

sen mehr Glauben zu schenken, wenn die Behandlungsergebnisse von jemandem beurteilt wurden, der keine Kenntnis von den Behandlungszuteilungen hatte.

Allerdings handelt es selbst hierbei nicht um eine absolute Empfehlung. Wie schon bemerkt, lassen einige harte Endpunkte wie die Mortalität nur wenig Raum für Ascertainment Bias. Anders ausgedrückt: Die Verblindung der Bewerter gegenüber harten Endpunkten trägt wohl nicht viel bei.

Fazit

Die Verblindung in klinischen Studien kann auf eine lange Geschichte von zwei Jahrhunderten zurückblicken. Weltweit wissen die meisten Wissenschaftler, was mit Verblindung gemeint ist, doch hinter einem nur allgemeinen Verständnis lauert oftmals Verwirrung. Studienleiter sollten ihre Studie daher nicht nur als einfach-, doppel- oder dreifachblind kennzeichnen, sondern eindeutig angeben, welche Personen in ihrer Studie verblindet und welche nicht verblindet waren. Der Leser sollte, wenn er einen Studienbericht liest und beurteilt, diesbezüglich Klarheit erwarten dürfen.

Wir danken Willard Cates und David L. Sackett für ihre hilfreichen Kommentare zu einer früheren Version dieses Artikels. Ein großer Teil des Materials stammt aus 15 Jahren Lehrtätigkeit im Rahmen des Fakultätsentwicklungsprogramms der Berlex Foundation.

Anmerkung der Redaktion

Die Übersetzung dieses Artikels erfolgte durch Frau Karin Beifuss (Stuttgart), die fachliche Bearbeitung übernahm Frau Gerta Rücker (IMBI-Institut für Medizinische Biometrie und Statistik Universitätsklinikum Freiburg). Beiden sei an dieser Stelle sehr herzlich gedankt.

Literatur

- [1] Kaptchuk TJ. Intentional ignorance: a history of blind assessment and placebo

- controls in medicine. *Bull Hist Med* 1998;72:389–433.
- [2] Devereaux PJ, Manns BJ, Ghali WA, et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA* 2001;285:2000–3.
- [3] Schulz KF, Chalmers I, Altman DG. The landscape and lexicon of blinding in randomized trials. *Ann Intern Med* 2002;136:254–9.
- [4] Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408–12.
- [5] Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA* 1994;272:125–8.
- [6] Chalmers I. What is the prior probability of a proposed new treatment being superior to established treatments? *BMJ* 1997;314:74–5.
- [7] Wolf S. Effects of suggestion and conditioning on action of chemical agents in human subjects: pharmacology of placebos. *J Clin Invest* 1950;29:100–9.
- [8] Noseworthy JH, Ebers GC, Vandervoort MK, Farquhar RE, Yetisir E, Roberts R. The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. *Neurology* 1994;44:16–20.
- [9] Meinert CL. *Clinical trials: design, conduct, and analysis*. New York: Oxford University Press; 1986.
- [10] Pocock SJ. *Clinical trials: a practical approach*. Chichester: Wiley; 1983.
- [11] Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. *BMJ* 2000;321:504.
- [12] Lasagna L. The controlled trial: theory and practice. *J Chronic Dis* 1955;1:353–67.
- [13] Cheng K, Smyth RL, Motley J, O’Hea U, Ashby D. Randomized controlled trials in cystic fibrosis (1966–1997) categorized by time, design, and intervention. *Pediatr Pulmonol* 2000;29:1–7.
- [14] Gøtzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Control Clin Trials* 1989;10:31–56.
- [15] Adetugbo K, Williams H. How well are randomized controlled trials reported in the dermatology literature? *Arch Dermatol* 2000;136:381–5.
- [16] Mosteller F, Gilbert JP, McPeck B. Reporting standards and research strategies for controlled trials: agenda for the editor. *Controlled Clin Trials* 1980;1:37–58.
- [17] DerSimonian R, Charette LJ, McPeck B, Mosteller F. Reporting on methods in clinical trials. *N Engl J Med* 1982;306:1332–7.
- [18] Schulz KF, Grimes DA, Altman DG, Hayes RJ. Blinding and exclusions after allocation in randomised controlled trials: survey of published parallel group trials in obstetrics and gynaecology. *BMJ* 1996;312:742–4.
- [19] Franklin B, Bailly JS, Lavoisier A: (1785) *Rapport des commissaires chargés par le roi, de l’examen du magnetisme animal*. Nice: Chez Gabriel Floteron.
- [20] Department of Health and Human Services, FDA. International conference on harmonisation: guidance on statistical principles for clinical trials. *Fed Regist* 1998;63:49583–98.
- [21] Altman DG. *Practical statistics for medical research*. London: Chapman and Hall; 1991.
- [22] Khan KS, Daya S, Collins JA, Walter SD. Empirical evidence of bias in infertility research: overestimation of treatment effect in crossover trials using pregnancy as the outcome measure. *Fertil Steril* 1996;65:939–45.
- [23] Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998;352:609–13.
- [24] Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663–94.
- [25] Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports or parallel-group trials. *Lancet* 2001;357:1191–4.
- [26] Bellomo R, Chapman M, Finfer S, Hickling K, Myburgh J. Low-dose dopamine in patients with early renal dysfunction: a placebo-controlled randomised trial. Australian and New Zealand Intensive Care Society (ANZICS) Clinical Trials Group. *Lancet* 2000;356:2139–43.
- [27] James WP, Astrup A, Finer N, et al. Effect of sibutramine on weight maintenance after weight loss: a randomised trial. STORM Study Group. *Lancet* 2000;356:2119–25.
- [28] Roddy RE, Zekeng L, Ryan KA, Tamoufe U, Weir SS, Wong EL. A controlled trial of nonoxynol 9 film to reduce male-to-female transmission of sexually transmitted diseases. *N Engl J Med* 1998;339:504–10.
- [29] Sinei SK, Schulz KF, Lamptey PR, et al. Preventing IUCD-related pelvic infection: the efficacy of prophylactic doxycycline at insertion. *Br J Obstet Gynaecol* 1990;97:412–9.
- [30] Baker KR, Drutz HP, Barnes MD. Effectiveness of antibiotic prophylaxis in preventing bacteriuria after multichannel urodynamic investigations: a blind, randomized study in 124 female patients. *Am J Obstet Gynecol* 1991;165:679–81.
- [31] Quitkin FM, Rabkin JG, Gerald J, Davis JM, Klein DF. Validity of clinical trials of antidepressants. *Am J Psychiatry* 2000;157:327–37.

- [32] Sackett DL, Gent M, Taylor DW. Tests for the blindness of randomized trials may not. *Clin Res* 1986;34:711A.
- [33] The Canadian Cooperative Study Group. A randomized trial of aspirin and sulfinpyra-

zone in threatened stroke. *N Engl J Med* 1978;299:53-9.

- [34] Carley SD, Libetta C, Flavin B, Butler J, Tong N, Sammy I. An open prospective randomised trial to reduce the pain of

blood glucose testing: ear versus thumb. *BMJ* 2000;321:20.

Ausschreibung des David Sackett-Preises des Deutschen Netzwerk Evidenzbasierte Medizin

Das Deutsche Netzwerk Evidenzbasierte Medizin (DNEbM e.V.) schreibt erstmalig für das Jahr 2008 den neugeschaffenen David-Sackett-Preis aus. Der Preis zeichnet hervorragende Leistungen auf dem Gebiet der Evidenzbasierten Medizin

und Gesundheitsversorgung (EbM, EbHC) aus. Ausgezeichnet werden können Einzelpersonen, Personengruppen, Institute oder Verbände.

Bewerbungsfrist ist der 1. Dezember 2007.

Evidenz im Blick

Mehr *Informationen* zur Ausschreibung unter http://www.ebm-netzwerk.de/wir_ueber_uns/preise

Langenscheidt Elsevier Wörterbuch Medizin Englisch Englisch-Deutsch/Deutsch-Englisch in Zusammenarbeit mit Elsevier von Dr. Phil. Dr. Med. Stephan Dressler 1. Auflage 2007.

Langenscheidt Verlag München, 623 Seiten Paperback.

Preis: € 19,95 [D], € 20,60 [A], 35,50 CHF
ISBN 978-3-86117-283-3

Mit dem neuen Wörterbuch Medizin Englisch legt der Langenscheidt Fachverlag ein Spezialwörterbuch vor, das in Kooperation mit der Medizin-Redaktion des Elsevier Verlags entstanden ist. Mit rund 35.000 Fachbegriffen und mehr als 35.000 Übersetzungen stellt er eine profunde Grundlage für die internationale Fachkommunikation dar. Enthalten sind

klinisch relevante Fachbegriffe, aktuelle Arzneimittel, Krankheitsbilder und Syndrome mit ca. 1.000 Kurzdefinitionen. Ein Anhang mit Abkürzungen und medizinischen Maßeinheiten bietet zusätzliche Unterstützung. Als Kernbestand enthält das Wörterbuch Stichwörter aus dem Gebiet der klinischen und praktischen Medizin – von *Makuladegeneration* (*macular degeneration*) bis *tonsillitis* (*Mandelentzündung*). Darüber hinaus sind all-

gemeinmedizinisch relevante Stichwörter aus Grundlagenfächern wie Chemie, Physik, Biochemie oder Biologie und aus Nachbargebieten wie Pharmakologie und Biotechnologie sowie die internationalen Freinamen von Medikamenten (INN) vertreten.

Literatur und Rezensionen

