S. Knippschild[1], C. Baulig[1], J. Hirsch[1], F. Krummenauer[1]

Das CONSORT-Statement zur standardisierten Berichterstattung Randomisierter Klinischer Prüfungen – Evidenz durch Transparenz

The CONSORT Statement for Standardized Reporting of Randomized Clinical Trials – evidence as a consequence of transparency

Die CONSORT-Gruppe (Consolidated Standards of Reporting Trials) hat es sich zur Aufgabe gemacht, Methoden zur Vermeidung unzureichender Berichterstattung zu randomisierten Klinischen Studien bereit zu stellen. Eine zentrale Empfehlung der CONSORT-Gruppe ist zusammengefasst im sogenannten CONSORT-Statement, welches eine evidenzbasierte Minimalempfehlung für die standardisierte Berichterstattung randomisierter Studien darstellt. Entlang einer Checkliste aus 25 Kriterien zu methodischen Charakteristika einer randomisierten Klinischen Studie respektive ihres Designs und der Belastbarkeit von Darstellungen ihrer Ergebnisse nebst einem Flussdiagramm zur Dokumentation der Patientenrekrutierung werden Autoren von Studienberichten Standards zur Verfügung gestellt, welche eine vollständige und transparente Berichterstattung sowie eine kritische Bewertung und Interpretation der zugrunde liegenden Studiendaten sicherstellen sollen. Im folgenden Artikel werden die Kriterien des CONSORT-Statements jeweils kurz erläutert und die resultierende Checkliste entlang einer Beispielstudie aus der Zahnärztlichen Implantologie umgesetzt.

Schlüsselwörter: CONSORT; Klinische Studien; Qualitätssicherung; Berichterstattung

In order to solve well-known consequences of the inadequate reporting of randomized clinical trials the CONSORT Group has provided the CONSORT (Consolidated Standards of Reporting Trials) Statement, which summarizes an evidence-based, minimum set of recommendations for the standardized reporting of randomized clinical trials. These standards are presented by means of a check list on 25 items concerning methodological design and result reporting characteristics of a randomized clinical trial as well as by means of a flow chart on the trial's patient recruitment history. The CONSORT statement offers a transparent process and result documentation standard for authors of clinical trial reports in order to document the trial's findings in a comprehensive manner.

Keywords: CONSORT; clinical trial; quality assurance; reporting standards

In der Zahnärztlichen Implantologie hat sich die Anzahl der Publikationen zu randomisierten kontrollierten Studien (RCTs) im Zeitraum 2003–2013 mehr als verdoppelt. Gleichzeitig ist jedoch deutlich zu erkennen, dass hierbei keine einheitliche Berichterstattung besteht. Der Grund hierfür liegt in der Regel an einer Fülle komplexer Originaldaten, die extrem komprimiert werden müssen, wenn sie in einer Fachzeitschrift zur Publikation in Frage kommen sollen. Auch wenn die Verlage die Beachtung vorhandener Reporting Guidelines fordern, erscheint deren Umsetzung also im Ansatz problematisch, nicht zuletzt aufgrund der primär redaktionellen Vorgaben eben dieser Verlage und Zeitschriften. Im Ergebnis wird es für den Leser schwierig, zuweilen unmöglich, Berichte zu Klinischen Studien und deren Verlaufs- und Ergebnis-Charakteristika nachzuvollziehen respektive deren Validität zu beurteilen, Ergebnisse unterschiedlicher Berichte vergleichend zusammenzufassen oder auf der Grundlage einer Studie die faktischen Therapieumsetzungen zu reproduzieren. Immer wieder wird sogar der Vorwurf laut, dass z.B. pharmazeutische Unternehmen, die in einer Klinischen Studie als Auftraggeber fungierten, nur bedingt Interesse an einer transparenten und vollständigen Berichterstattung hätten [1, 7].

[1] Institut für Medizinische Biometrie und Epidemiologie, (Direktor: Prof. Dr. F. Krummenauer), Fakultät für Gesundheit der Universität Witten/Herdecke, Alfred-Herrhausen-Straße 50, 58448 Witten

Alleine zur Vermeidung solcher Mutmaßungen sollten ein hohes Maß an Transparenz und eine nachvollziehbare Beschreibung der Vorgehensweise angestrebt werden, um die Plausibilität und Glaubwürdigkeit der Ergebnisse beim Leser zu stärken. Daher bestand eine dringende Notwendigkeit zur Vereinheitlichung sowie zur Einführung eines Standards zur Dokumentation der Prozess- und Ergebnis-Charakteristika Klinischer Studien [2]. Dies war Anlass, dass in den 1991er Jahren unabhängig voneinander 2 Arbeitsgruppen an einer Empfehlung für eine einheitliche Berichterstattung von Studienergebnissen randomisierter Studien im Paralleldesign arbeiteten. 1996 wurde durch einen Zusammenschluss dieser beiden Gruppen ein gemeinsames Konzept entwickelt, das sogenannte CONSORT-Statement [15], welches bis heute einem ständigen Optimierungsprozess unterliegt. Im Rahmen dieser Weiterentwicklungen erfolgte 2010 die bisher letzte Anpassung der CONSORT-Empfehlungen an die derzeitige Forschungslandschaft.

Das CONSORT-Statement beschreibt eine detaillierte Vorgehensweise zur Dokumentation und damit auch zur Bewertung Prozess- und Ergebnis-determinierender Charakteristika randomisierter Klinischer Studien. Hiermit werden auf der Seite des Publizierenden eine transparente und vollständige Berichterstattung bewirkt und auf Seiten des Lesers die Nachvollziehbarkeit und das Vertrauen in die präsentierten Ergebnisse gestärkt. Für die praktische Umsetzung stehen dem Anwender eine Checkliste von 25 methodischen Kriterien das Design und die Ergebnisdarstellung einer Studie betreffend sowie ein Flussdiagramm zur Historie der Patientenrekrutierung als „Schablone" zur Verfügung, anhand derer die einzelnen Studien-Charakteristika beleuchtet werden können. Im Folgenden werden diese Punkte vorgestellt und an einer aktuellen Publikation zur Zahnärztlichen Implantologie exemplarisch umgesetzt [16].

Kriterienserie „Titel und Zusammenfassung"

1a Kennzeichnung im Titel als randomisierte Studie

Die Auffindbarkeit einer Studienpublikation als Darstellung einer randomisierten Studie über eine elektronische Datenbank ist von deren Kennzeichnung als solche abhängig. Um eine korrekte Identifizierung zu gewährleisten, sollten Autoren das Wort „randomisiert" bereits explizit im Titel verwenden.

1b Strukturierte Zusammenfassung von Studiendesign, Methoden, Resultaten und Schlussfolgerung

Da oftmals leserseitig z.B. in der Sondierungsphase von Artikeln für Metaanalysen aufgrund des Abstracts die Entscheidung gefällt wird, ob ein Artikel im Volltext lesenswert ist, sollte die Zusammenfassung eines Studienberichtes so konkret wie möglich gehalten sein. Dabei sind generelle Informationen zum Ziel und der Durchführung der Klinischen Studie sowie eine kurze Darstellung des Ergebnisses zum primären Endpunkt erforderlich. Für eine umfassende Kurzdarstellung haben *Hopewell* et al. im Umfeld der CONSORT-Gruppe eine Checkliste für die Erstellung eines Abstracts für die Kurzdarstellung einer randomisierten Studie erarbeitet [8].

Kriterienserie „Einleitung"

2 Hintergrund und Ziele

2a Wissenschaftlicher Hintergrund und Begründung der Studie

Die Deklaration von Helsinki fordert, Studien auf der Grundlage einer genauen Literaturrecherche zu basieren, um Teilnehmer nicht einem unnötigen Risiko auszusetzen. Aus diesem Grund ist an dieser Stelle eine Rationale für die Untersuchung zu bringen, welche die Notwendigkeit der Klinischen Studie rechtfertigt.

2b Genaue Fragestellung oder Hypothese

Am Ende der Beschreibung des Hintergrundes für eine Studie steht eine eindeutige Frage, die mit dieser Untersuchung beantwortet werden soll. Oft kann und soll diese Fragestellung als Hypothese parametrisiert werden. Diese Formulierung stellt auch die Grundlage für eine spätere statistische Auswertung dar.

Kriterienserie „Methoden"

3 Studiendesign

3a Beschreibung des Studiendesigns, einschließlich Zuteilungsverhältnis der Patienten zu den Interventionen

Hier werden Informationen bereitgestellt zur Umsetzung des Ziels der Untersuchung in ein konkretes Design respektive der zugrunde liegende konzeptionelle Rahmen der Zielsetzung (Überlegenheits- oder Nicht-Unterlegenheitsstudie; Wahl der Vergleichs-Intervention[en] respektive ggf. Placebo, bei Prüfungen von Arzneimitteln die Phase der Studie etc.).

3b Wichtige Änderungen der Methoden nach Studienbeginn (z.B. Eignungskriterien) mit Gründen

Mit der Erstellung eines Studienprotokolls vor Beginn einer Studie wird sichergestellt, dass die Durchführung einer Untersuchung unter zuvor festgelegten Bedingungen und nach definierten standardisierten Vorgehensweisen erfolgt. Durch nicht absehbare Umstände können auch im Nachhinein noch Änderungen im Verlauf der Studie notwendig werden. Diese sollten gut dokumentiert werden, da sie dem Leser später helfen, die als notwendig erachteten Änderungen nachzuvollziehen und die Auswirkung auf die Gesamtergebnisse der Studie zu interpretieren.

4 Probanden/Patienten

4a Eignungskriterien der Studienteilnehmer

Für eine genaue Beurteilung der Ergebnisse und deren Übertragbarkeit ist eine detaillierte Charakterisierung des einzuschließenden Kollektivs notwendig. Da die Einschlusskriterien vor der Randomisierung angewendet werden, ist dieser Punkt zum einen für die Argumentation der Einhaltung gesetzlicher und ethischer Normen von zentraler Bedeutung und zum anderen für die externe Validität der Studie und damit der späteren Übertragbarkeit von Studienergebnissen aus der Studie auf die Allgemeinheit. Besondere Ein- oder Ausschlusskri-

terien sollten kurz motiviert werden, da diese für die Generalisierbarkeit der Studienergebnisse einschränkend sein können und somit die ethische Legitimation der Studie per se betreffen.

4b Umgebung und Ort, an dem die Datensammlung erfolgt

Zusammen mit den Einschlusskriterien für Studienteilnehmer und der Beschreibung der Studienrelevanten Interventionen ist auch die Darstellung der Umgebung und des Durchführungsortes der Rekrutierung notwendig. Handelt es sich um Studienteilnehmer, die an einem speziellen Ort rekrutiert wurden? Sind Patienten der Primär-, Sekundär- oder Tertiärversorgung in die Studie eingeschlossen worden oder sind möglicherweise spezielle kulturelle oder klimatische Bedingungen zu berichten? Auch die Beschreibung unterschiedlicher Einrichtungen zur Gesundheitspflege sind im Hinblick auf ihre Organisation, Ressourcen und Erfahrungen – nicht zuletzt auch mit Blick auf das Bestehen lokaler Infrastrukturen für Klinische Studien – zu beschreiben (etwa Einbindung einer abteilungsübergreifenden Core Facility wie universitären Klinischen Studienzentren oder einer klinikeigenen Studienzentrale).

5 Behandlung und Therapie

Alle studienbedingten und ggf. auch diesen zugrunde liegende Standardinterventionen müssen zum Zweck deren Nachvollziehbarkeit und Reproduzierbarkeit detailliert beschrieben werden. Hierzu gehört auch, wie und wann die Interventionen durchgeführt wurden. Auf der Grundlage dieser Informationen wird dem Leser beziehungsweise dem Anwender ermöglicht, eine Entscheidung zur Benutzung zu treffen. Eine Erweiterung für nichtpharmakologische und naturheilmittelbasierte Interventionen wurde im Umfeld der CONSORT-Gruppe von *Boutron* et al. und *Gagnier* et al. zur Verfügung gestellt [3, 6].

6 Endpunkte

6a Definition des primären Endpunktes und sekundärer Endpunkte (früher „Zielkriterien“ genannt), einschließlich der Angaben, wie und wann diese erhoben wurden

In Klinischen Studien müssen Merkmale definiert sein, die als Grundlage zur Beantwortung der primären Fragestellung herangezogen werden. Möglicherweise gibt es mehrere interessante Merkmale, die innerhalb einer Klinischen Studie untersucht werden könnten. Es wird allerdings ausdrücklich davon abgeraten, eine primäre Fragestellung auf mehrere parallele und gleichberechtigte primäre Endpunkte oder Teilfragestellungen zu stützen: Konsistent interpretierbare Ergebnisse werden durch parallele primäre Endpunkte erschwert und damit auch schwerer in klare Therapieempfehlungen umsetzbar. Vielmehr sollte eine primäre Fragestellung oder Hypothese formuliert werden. Alle weiteren Fragen können entlang sekundärer Endpunkte und explorativer Analysen berichtet werden. Insbesondere ist eine genaue Dokumentation der Zeitpunkte erforderlich, an denen Messungen erhoben wurden, ebenso eine Beschreibung der zur Endpunkt-Messung verwendeten Instrumente (z.B. Fragebögen zur Lebensqualität). Eine exakte Definition des primären Endpunktes ist somit für die Interpretation des zentralen Studienergebnisses in eben diesem primären Endpunkt unabdingbar.

6b Änderungen der Endpunkte nach Studienbeginn mit Angabe der Gründe

Es gibt verschiedene Gründe, die eine Änderung der zuvor gewählten Endpunkt-Parametrisierungen erforderlich machen. In diesem Fall ist es zwingend, Protokolländerungen genauestens zu dokumentieren, um – auch ungewollte – Fehldarstellungen bei der späteren Berichterstattung zu vermeiden [5].

7 Fallzahlberechnung

7a Wie wurde die Fallzahl berechnet?

Aus wissenschaftlichen und ethischen Gründen ist die Fallzahl einer Klinischen Studie im Vorfeld so zu kalkulieren, dass das erwartete Ergebnis statistisch gesichert aufgedeckt werden kann und zugleich nicht unnötig viele Patienten in die Studie eingeschlossen und einer schon mit geringerer Fallzahl als unterlegen erkennbaren Therapie zugeordnet werden. Zu einer detaillierten Dokumentation gehören dabei mindestens der aufzudeckende Unterschied zwischen den Gruppen, das gewählte Signifikanzniveau und die gewünschte Power [12]. Neben der errechneten Fallzahl pro Gruppe sollten außerdem Details beschrieben werden zur bei Studienplanung erwarteten Rate nicht kooperierender oder im Studienverlauf unvorhersehbar entfallender Studienteilnehmer, welche nicht mehr der Auswertung zugeführt werden können (erwartete drop-out-Rate).

7b Falls zutreffend, Erklärung aller geplanten Zwischenanalysen und Abbruchkriterien

Wurde im Verlauf der Rekrutierungsphase einer Studie eine Zwischenauswertung geplant vorgenommen oder auch ungeplant als notwendig erachtet, sollen Gründe und der Zeitpunkt hierfür unter diesem Punkt erläutert werden. Zu erwähnen ist die Anzahl der Zwischenauswertungen, ob diese vorab designbedingt geplant in die Untersuchung inseriert oder durch andere Umstände erforderlich wurden (Anweisung eines unabhängigen Monitoring-Komitees), ferner wer die Entscheidung hierfür jeweils getroffen hat. Werden Zwischenanalysen in einer Klinischen Studie geplant, sind auch damit verbundene Konsequenzen in Studienprotokoll und Statistischem Analyseplan der Studie zu berücksichtigen und vorab explizit zu beschreiben.

8 Randomisation – Generierung der Sequenz

8a Methode zur Generierung der Zufallszuteilung

Die „randomisierte“ Zuteilung von Studienteilnehmern auf konkurrierende Therapien ist heute als Standard in Klinischen Studien anerkannt, wird aber häufig falsch verstanden beziehungsweise in ihren Anforderungen unterschätzt [13]. Sie verhindert eine Verzerrung der Ergebnisse durch eventuelle Selektion der Studienteilnehmer und si-

chert so auch die Vergleichbarkeit der Interventionsstichproben bezüglich möglicher prognostischer Kriterien in der Auswertung. Die verwendeten Methoden zur Randomisation müssen in einer Publikation eindeutig genug beschrieben werden zur Bewertung, ob die Anforderungen der Zufälligkeit und Nichtvorhersagbarkeit der Therapie-Zuweisung gewährleistet waren. Für die Beschreibung der Methode zur Randomisation sollten Informationen über die Zuordnung (computerbasierte Nummernzuordnung, Tabellen mit Zufallszahlen etc.), das Randomisationsverhältnis und Stratifikationskritieren der Randomisation (etwa nach Zentren im Falle multizentrischer Studien) gegeben werden.

8b Art der Randomisierung; Details jedweder Restriktion (z.B. Blockbildung, Blockgröße)

Wenn eine Randomisation erfolgte, ist eine exakte Beschreibung der Methode erforderlich. Zusätzlich sollen alle vorgenommenen Einschränkungen, z.B. Blockrandomisation (um gleich große Therapiegruppengrößen zu generieren) oder Stratifikationen (um homogene Gruppen im Hinblick auf Baselinecharakteristika mit prognostischer Revelanz zu erzeugen) beschrieben werden [13].

9 Randomisation – Methode zur Zuteilung

Die logistische Umsetzung einer randomisierten Zuteilung unter Wahrung der Nicht-Aufdeckbarkeit der maskierten Therapievorgabe ist genauso wichtig wie die Bereitstellung der Randomisationsliste als Basis der Zuteilung selber. Nur durch die Geheimhaltung der verwendeten Zuteilungssequenz bis zum Therapiebeginn kann ein Vorwissen bei allen Beteiligten (Studienpersonal und den Studienteilnehmern) vermieden werden. Um darzulegen, dass die Umsetzung erfolgreich war und somit ein wichtiger Schritt für die Verallgemeinerung der Studienergebnisse erfolgt ist, muss auch das Protokoll zur Anwendung der Randomisation detailliert beschrieben werden. Konkret sollte z.B. beschrieben werden, ob für die Prüfärzte Umschläge mit der Information zur Therapiegabe bereitgestellt wurden oder über eine zentrale dritte Stelle – z.B. telefonisch oder per Internet-Abruf – patientenweise Vorgaben zur Wahl der Therapie erfolgten. Dies ist auch bei „optisch" nicht unterscheidbaren Prüfpräparaten klar darzustellen, wenn etwa die Zuteilung praktisch über Kennziffern erfolgt, welche sich jeweils auf der jeweiligen Prüfpräparat-Verpackung und in der Zuteilungsvorgabe eines Patienten entsprechen.

10 Randomisation – Umsetzung

Der Prozess einer Randomisierung ist zum einen rein technisch durchzuführen durch die Generation einer zufälligen Sequenz, zum anderen aber auch durch Entwicklung eines Konzeptes zur Geheimhaltung der Zuordnung bis zum Studienende respektive zum Ende der Durchführung durch Studienpersonal. Hier soll dokumentiert werden, dass die Geheimhaltung und damit verbundene zufällige Zuteilung nicht durch den Prozess und die daran beteiligten Personen manipuliert werden konnten (etwa durch Beschreibung von Methoden zur Sicherstellung, dass in offenen Studien die Messung des primären Endpunktes nicht vom Behandler vorgenommen wurde, sondern von für die Therapiezuordnung maskierten Dritten).

11 Verblindung

11a Verblindung

Der Begriff „verblindet" („blinded") oder „masked" bezieht sich auf das Zurückhalten von Informationen über eine per Randomisation zugeordnete Therapie. Um Ergebnisse durch den subjektiven Einfluss von Beteiligten – Studienteilnehmer und Studienpersonal – zu verhindern, sollte in jeder Klinischen Studie eine Verblindung aller Beteiligter angestrebt werden: Nicht nur die Studienteilnehmer können durch das Wissen über die erhaltene Therapie eine – auch ungewollte – Verfälschung bedingen, auch das die Therapien verabreichende und Endpunkte messende Studienpersonal ist hiervor bei aktiver Kenntnis der Therapie nicht geschützt. Nicht zuletzt an der Planung und Umsetzung der statistischen Analyse beteiligtes Personal muss gegenüber der Gabe der Therapie laut Randomisation verblindet bleiben, um Manipulationsvorwürfen vorzubeugen. In Publikationen sollten daher nicht nur die Zielsetzungen der Studie bezüglich einer Verblindung der Therapie genannt werden, sondern diese durch folgende Punkte erläutert und nachvollziehbar gemacht werden: Welche Personengruppen wurden zu welchem Zweck verblindet? Zu welchem Zeitpunkt wurde die Verblindung der einzelnen Personen vorgenommen? Gibt es Personengruppen, die in dieser Studie noch weitere Aufgaben ausüben?

11b Vorgehen zur Gewährleistung der Ähnlichkeit von Interventionen

Ein weiterer Schritt, um Schritte zur Geheimhaltung der Therapiegabe darzulegen, ist die Beschreibung von Maßnahmen zur Gewährleistung der Ähnlichkeit verwendeter Interventionen. Hier sollten deren Charakteristika (z.B. Aussehen und Geschmack von Prüfpräparaten) sowie deren Darreichungsform beschrieben werden, um zu demonstrieren, wie Vorwissen im Hinblick auf die verabreichte beziehungsweise angewendete Therapie vermieden werden sollte in der konkreten Studie.

12 Statistische Methoden

12a Statistische Methoden zum Vergleich der Gruppen

Die Analyse von Daten einer Klinischen Studie kann auf vielfältige Weise erfolgen. Grundsätzliches Ziel ist die Schätzung des Behandlungseffektes zwischen den Stichproben. Für den Leser einer Publikation ist es wichtig, die Grundlagen für die vorgenommene Auswertung zu erfahren (etwa bezogen auf den Umgang mit nicht vorhanden Messwerten, z.B. durch Ausschluss vereinzelter Patienten von der Auswertung oder durch Werteersatz durch sogenannte Imputationsverfahren). Um das Verständnis des Lesers für die originalen Daten zu fördern und somit auch die Validität der Ergebnisse zu belegen, ist eine detaillierte Beschreibung der zur Analyse verwendeten statistischen Methoden erforderlich. Die Bewertung der Signifikanz der Ergebnisse erfolgt hierbei am besten über die Berechnung eines Konfidenzintervalls, alternativ mittels

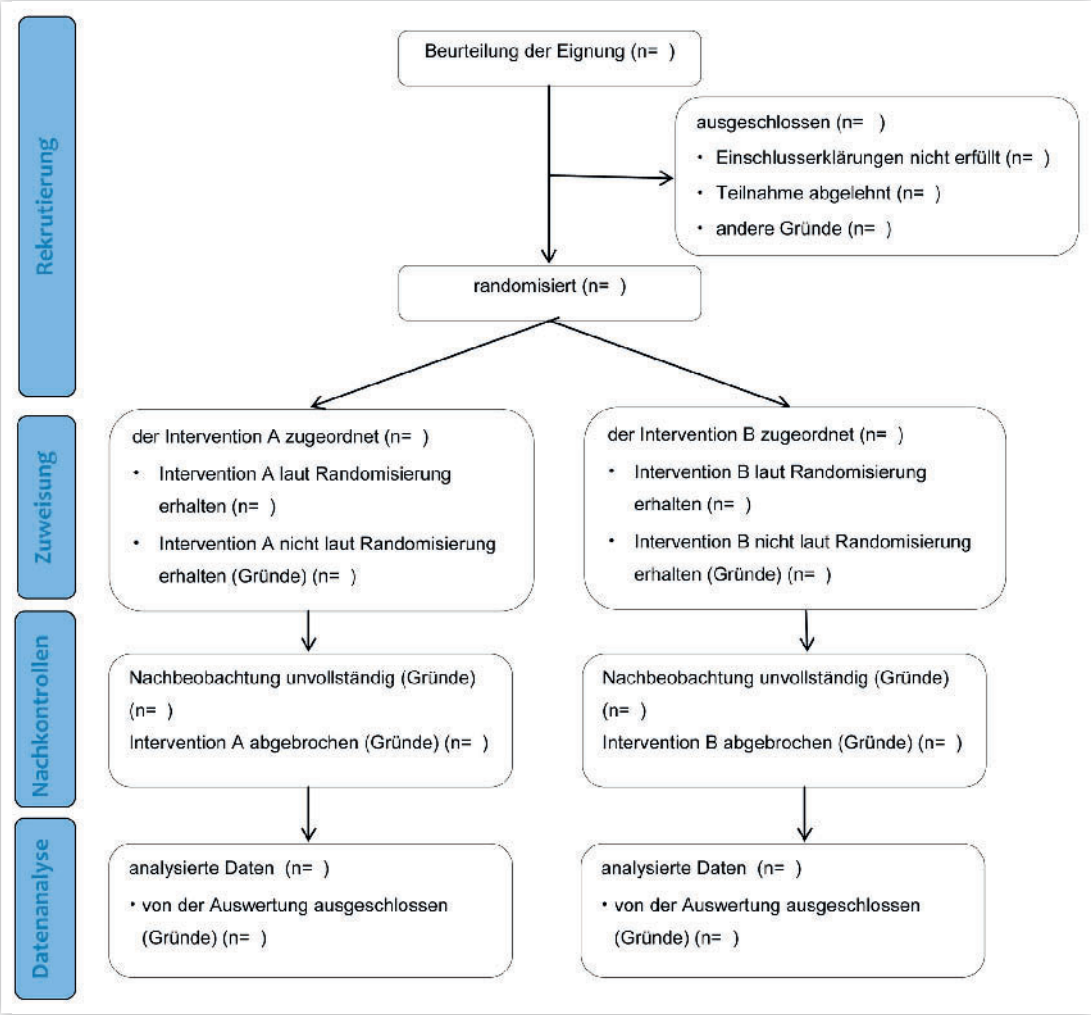

Abbildung 1 Durch das CONSORT-Statement vorgegebenes Flussdiagramm [14] zur Dokumentation der Patientenrekrutierung in einer zweiarmigen randomisierten Klinischen Studie zum Vergleich der Interventionen „A" und „B" durch differenziertes Ausweisen der absoluten Anzahlen von Studienteilnehmern während der gesamten Studienlaufzeit von Screening/Rekrutierung in die Studie über Randomisation (Vorgabe laut Randomisation respektive faktisch vorgenommene Intervention) bis hin und zum Zeitpunkt der Analyse.

Figure 1 CONSORT Flow Diagram as available from the CONSORT statement [14]; documentation of the absolute numbers of study participants in a two-arm randomized clinical trial with regard to the different study phases (screening/recruitment, randomization and analysis) is displayed by means of stratification for the respective trial interventions.

passender deskriptiver Angaben zum Effekt zwischen den Stichproben und dem p-Wert eines zum Skalenniveau des primären Endpunktes passenden Signifikanztests. Wichtig ist hierbei die konkrete Benennung der verwendeten Methoden, der zugrunde liegenden Modellannahmen (ob und warum z.B. von normalverteilten Daten ausgegangen wurde) und den Signifikanzniveauvorgaben.

12b Methoden für zusätzliche Analysen

Wie bereits für die primäre Analyse muss auch die Methodik, z.B. für Subgruppenanalysen und andere, nicht die primäre Fragestellung betreffende Auswertungen im Rahmen der Studienplanung festgelegt werden. Oft wird der Fehler gemacht, dass p-Werte für diverse parallele Analysen generiert werden und somit das Risiko für falsch-positive Ergebnisse erhöht wird. Aus diesem Grund sind Subgruppenanalysen in Ergebnisberichten zu Klinischen Studien als rein explorativ und bestenfalls hypothesengenerierend zu verstehen, sollten also auch als solche gekennzeichnet werden u.a. mit entsprechender Methodenspezifikation.

Kriterienserie „Ergebnisse"

13 Studienteilnehmer: Ein- und Ausschlüsse

13a Ein- und Ausschlüsse

In Klinischen Studien muss unterschieden werden, aus welchen Gründen Patienten nicht mehr an der Untersuchung teilnehmen. Durch die asymmetrische Verringerung der Anzahl randomisierter Studienteilnehmer ist ein Ungleichgewicht zwischen den Stichproben möglich, was eine Verzerrung der Ergebnisse zur Folge haben kann.

Aus diesem Grund ist es wichtig, dem Leser verständlich darzustellen, wie viele Patienten einer zuvor festgelegten Therapie zugeordnet und damit behandelt wurden, wie viele Patienten aber auch – warum – nicht der per Randomisation zugewiesenen Therapie unterzogen wurden oder auch die Therapie wechseln mussten oder durften. Bei der Präsentation der Ergebnisse sollte Wert darauf gelegt werden, die Anzahl der Studienteilnehmer pro Intervention während der gesamten Studienlaufzeit vom Zeitpunkt der Randomisation bis hin zu allen Nachkontrollen klar darzustellen. Beschrieben werden sollten die Anzahl der eingeschlossenen Teilnehmer und die von Studienabbrüchen oder Ausschlüssen. In komplexeren Studien ist es oft schwierig darzulegen, warum einige Patienten die zugewiesene Therapie nicht erhalten haben, aus der Studie ausgeschieden sind oder von der Analyse ausgeschlossen wurden. Meist ist es sinnvoll, diese Informationen in einem Diagramm wiederzugeben, wie es auch von der CONSORT-Gruppe empfohlen wird; eine mögliche Vorlage für ein solches Flussdiagramm zeigt Abbildung 1.

Je nach Studie ist die Beschreibung von Gründen über spezielle Ausschlüsse oder Studienabbrüche von besonderem Interesse. Dabei dient die Benennung der genauen Anzahl von Studienteilnehmern zu den einzelnen Meilensteinen im Verlauf der Studie als zusätzliche Grundlage zur Beurteilung der Ergebnisse:

- Einschluss: Diese Anzahl beschreibt, wie repräsentativ das ausgewählte Kollektiv ist gegenüber dem in Frage kommenden/gescreenten „Pool“.
- Randomisation: Diese Information ist grundlegend für die Beurteilung der effektiven Fallzahl und der darauf anwendbaren Analyse (und deren Power).
- Behandlungszuordnung: Die Anzahl der Probanden ist von Bedeutung für die Interpretation der Ergebnisse u.a. bezüglich Homogenität der Therapiegabe innerhalb der Interventionsstichproben. Hier sollten auch Gründe für nicht verabreichte Behandlungen charakterisiert werden.
- Nachkontrollen: Die Anzahl der Probanden in den Nachkontrollen ist im Hinblick auf die Validität und Interpretation der Ergebnisse wertvoll. Hier sollten Gründe für nicht durchgeführte Nachuntersuchungen berichtet werden.
- Analyse: Die in die Auswertung einbringbare effektive Fallzahl nach Ausfall von Patienten entlang des Studienverlaufs respektive nach Berücksichtigung von Protokollverletzungen (und ggf. dem daraus erwachsenen Ausschluss der betreffenden Patienten aus der Analyse) beschreibt, mit welcher effektiven Power die Studie letztlich in der primären Fragestellung ausgewertet werden konnte und ob die laut methodischer Planung angestrebte Fallzahl auch in allen Interventionsstichproben erreicht wurde.

13b Ausfall von Studienteilnehmern und Ausschlüsse

Zur Charakterisierung von Protokollverletzungen und daraus erfolgten Ausschlüssen ist das unter 13a vorgestellte Flussdiagramm (Abb. 1) nicht immer ausreichend: Wenn Studienteilnehmer nach Randomisation doch wieder ausgeschlossen werden, sollte eine genaue Begründung erfolgen, um einen möglicherweise zwischen den Stichproben asymmetrischen und damit den Stichprobenvergleich verzerrenden Therapiebedingten Effekt in der Studie für den Leser diskutierbar zu machen.

14 Aufnahme und Rekrutierung

14a Rekrutierung der Studienteilnehmer im zeitlichen Kontext

Für Leser, die an einer ähnlichen Fragestellung arbeiten, kann es sinnvoll sein, die zeitliche Einordnung und Abfolge einer Klinischen Studie zu kennen, um ein vollständiges Bild der kompletten Behandlung und der Eingliederung der Studie in die Routinebehandlung eines Patienten zu erfahren. Hierbei sind insbesondere die Zeitabstände der Nachkontrollen zu dokumentieren, sowie das Minimum, Maximum und die mediane Nachuntersuchungszeit der Probanden.

14b Gründe für die Beendung einer Studie

Üblicherweise wird eine Klinische Studie durch das Erreichen einer zuvor definierten Fallzahl oder der erreichten geplanten Studienlaufzeit bzw. Anzahl von Nachuntersuchungen beendet. Finden Zwischenanalysen statt, ist es unbedingt notwendig, die Vorgänge genau zu beschreiben und darzulegen, wer aus welchen Gründen das vorzeitige Ende der Studie veranlasst hat (ob etwa eine Klinische Studie aufgrund eines unerwartet starken Effekts als Ergebnis einer Zwischenanalyse gestoppt wurde oder aufgrund eines als klinisch irrelevant erkannten Effekts zwischen den Stichproben, ob in einer Monitorierung eine die Studie als nicht mehr vertretbar aufzeigende Häufung von unerwünschten Ereignissen erkannt wurde, oder ob durch die Ergebnisse anderer Studien die Durchführung dieser Untersuchung aus ethischen Gründen nicht länger als gerechtfertigt angesehen werden konnte).

15 Patienten-Charakteristika zu Studienbeginn

In randomisierten Studien wird eine Verzerrung durch Selektion der Studienteilnehmer ausgeschlossen. Trotzdem garantiert eine Randomisation nicht die exakte Gleichheit der zu untersuchenden Gruppen bezüglich möglicher Störgrößen. Aus diesem Grund ist die Beschreibung von Teilnehmer-Charakteristika zu Beginn der Studie für den Leser eines Studienberichtes hilfreich zur Einschätzung der faktischen Vergleichbarkeit der Stichproben. Baseline-Charakteristika werden am effektivsten in Form einer Tabelle dargestellt

16 Anzahl der analysierten Studienteilnehmer

Die Anzahl der auswertbaren Daten variiert innerhalb der einzelnen Auswertungen und sollte aus diesem Grund für jede Analyse angegeben werden.

17 Ergebnisse und Schätzmethoden

17a Ergebnispräsentation

Selektive Berichterstattung ist inzwischen ein aktiv diskutiertes Problem [4, 5]. Um diesem entgegen zu wirken, müssen Ergebnisse zum primären Endpunkt und allen sekundären Endpunkten in

den untersuchten Stichproben sachgerecht dargestellt werden. Hierin eingeschlossen ist die Angabe eines Konfidenzintervalls, alternativ eines p-Wertes inklusive eines deskriptiven Maßes für den Effekt zwischen den Stichproben, um die Genauigkeit der Schätzung des Effekts zu charakterisieren.

17b Präsentation von binären Endpunkten

Zu binären Endpunkten sollten sowohl das relative als auch das absolute Effektmaß präsentiert werden [11]: Vor allem bei Stichprobenvergleichen entlang eines primären Endpunktes mit sehr kleinen Auftrittshäufigkeiten kann das Ausmaß des relativen Risikos als alleinige Information irreführend sein. Die Differenz der Häufigkeiten zwischen den Stichproben nebst Konfidenzintervall für diese Differenz hingegen hilft dem Leser bei der Bewertung der klinischen Größenordnung und Relevanz des in der Studie aufgedeckten Effekts.

18 Zusätzliche Analysen

Umfangreiche explorative, d.h. in Subgruppen und entlang mehrerer Endpunkte „suchender" Datenanalysen bedingen falsch-positive Forschungsergebnisse [17]. Aus diesem Grund sollte von ungerichteten explorativen Analysen unter Verwendung von Signifikanzprüfungen abgesehen werden. Vielmehr sollten lediglich zuvor im Studienprotokoll respektive im Statistischen Analyseplan einer Studie festgelegte Auswertungen dieser Art erfolgen. Diese können auch im Nachhinein als glaubwürdig dargestellt werden und es kann im Ansatz schon der Vorwurf einer selektiven Berichterstattung von Ergebnissen verhindert werden. Deskriptive Analysen hingegen sind, sofern klar als solche erkennbar abgehoben von Signifikanzprüfungen der primären Fragestellung, jederzeit als die bestehenden Daten illustrierend anzusehen. Ähnliche Empfehlungen zur Transparenz gelten, wenn Anpassungen der im Studienprotokoll dokumentierten Variablen vorgenommen werden. In diesem Fall sollten sowohl die revidiert konzipierte als auch die initial geplante Analyse durchgeführt und parallel berichtet werden, um den Effekt von Anpassungen auf die Ergebnisse transparent zu machen.

19 Unvorhergesehener Schaden für Studienteilnehmer („harm")

Das Auftreten von unerwünschten Ereignissen hat einen großen Einfluss auf die Entscheidung des Anwenders für oder gegen eine Therapie bzw. einen Eingriff. Nur wenn diesbezügliche Daten explizit präsentiert werden bezüglich Häufigkeiten und Form (ggf. auch bezüglich klinischer Konsequenzen wie Hospitalisierungen etc. sowie der eventuellen kausalen Relation zur Studientherapie), wird dem Leser eine rationale Entscheidung über die Vertretbarkeit der Studientherapien für seine eigenen Patienten ermöglicht. Liegt zudem der Fokus einer randomisierten Klinischen Studie auf unerwünschten Nebenwirkungen oder Komplikationen für die Studienteilnehmer, muss auch die Berichterstattung der Ergebnisse an diese Fragestellung angepasst werden. *Ioannidis* et al. haben im Umfeld der CONSORT-Gruppe eine Erweiterung der Checkliste für die Darstellung „Nachteil-bezogener" Daten in Publikationen entwickelt und zur Verfügung gestellt [9].

Kriterienserie „Diskussion"

20 Studienlimitationen

Nach der Beschreibung der Studienergebnisse müssen diese diskutiert werden, indem alle Limitationen und Ungenauigkeiten der Untersuchung benannt und deren Auswirkungen beschrieben werden. Hierzu gehören z.B. die methodische Charakterisierung verwendeter Instrumente (insbesondere Fragebogenskalen) bezüglich deren Validität und Reliabilität, wenn diese zur Messung von Endpunkten benutzt wurden. Weitere offen zu diskutierende Limitationen können die faktische Umsetzung des geplanten Studiendesigns betreffen, wenn z.B. die Strategien zur Maskierung der Therapie ungewollt durchschaubar wurden oder im Interesse der Patienten teilweise aufgegeben oder in der Umsetzung vereinfacht werden mussten. Limitationen können auch aus der Zusammensetzung der Studienteilnehmer erwachsen. Die Teilnehmer sind zwar per Zufallsauswahl eingeschlossen und somit per se als repräsentative Stichproben zu verstehen, trotzdem kann z.B. durch asymmetrische drop-out-Profile zwischen den Studienarmen deren Vergleichbarkeit der Stichproben eingeschränkt werden.

21 Generalisierbarkeit der Ergebnisse

Die externe Validität bzw. die Generalisierbarkeit von Studienergebnissen ist als Maß zu sehen, wie weit die Ergebnisse einer Klinischen Studie auf andere Menschen jenseits der Studie verallgemeinert werden können. Die interne Validität als Maß unverzerrter Studienumsetzung ist dabei eine grundlegende Voraussetzung und muss auch innerhalb einer Publikation klar dargestellt werden. Vor allem die selbstkritische Diskussion von Limitationen – bewusst bei Planung der Studie im Interesse der Realisierbarkeit oder unbewusst im Nachgang erkannt – kann helfen, Einschränkungen und damit das Ausmaß der Generalisierbarkeit zu bewerten.

22 Interpretation

Den Leser interessiert abschließend, wie die Abwägung von Nutzen und Schaden der aus der Studie abgeleiteten Versorgungsempfehlung eingestuft wurden, ferner wie die Ergebnisse der Untersuchung in Bezug zu anderen publizierten randomisierten Studien zu sehen sind. Hierbei sollte unbedingt darauf geachtet werden, dass die Interpretation als Ergebnis eines vollständigen systematischen Reviews von RCTs erfolgen soll und nicht nur auf Basis einer Auswahl von Studien beschrieben werden darf (die dem Vorwurf einer wissentlichen Selektion von Studien aussetzen könnte).

Kriterienserie „Regulatorische Informationen"

23 Registrierung der Studie

Nicht publizierte Studienergebnisse bedingen selektive Berichterstattung und verfälschen in der Konsequenz die Er-

gebnisse systematischer Reviews. Um dieses Problem zu begrenzen, wurde die Registrierung von Klinischen Arzneimittelstudien inzwischen obligat vorgeschrieben. Eine weitergehende Registrierungsempfehlung für RCTs wurde zudem durch die WHO und Aktivitäten der Cochrane Collaboration realisiert. Mittels Registrierungsnummer können die zentralen Informationen einer jeden RCT eindeutig zugeordnet und der Öffentlichkeit zugänglich gemacht werden. Autoren einer Publikation sollten deshalb die zugewiesene Registrierungsnummer anführen oder eine Begründung für eine unterlassene Registrierung darlegen.

24 Studienprotokoll

Das Studienprotokoll ist die Grundlage für valide Ergebnisse einer Klinischen Studie. Einige Zeitschriften fordern, dass das Studienprotokoll zugänglich gemacht wird. Es ist aber auch für die Registrierung einer Klinischen Studie erforderlich. Die Zugriffsmöglichkeit auf das Studienprotokoll führt zu einem hohen Maß an Transparenz und unterstützt somit das Verständnis für die Untersuchung beim Leser.

25 Finanzierung

Die Information über die Finanzierung einer Klinischen Studie ist ein wichtiger Aspekt, wenn ein Leser den Ergebnisbericht einordnen möchte bezüglich möglicher Interessenskonflikte respektive eventueller Einflussnahmen auf die Darstellungen. Zur Dokumentation sollte die Rolle finanzieller und/oder politischer Träger genau beschrieben werden sowie deren Beteiligung am Studiendesign, der Durchführung und der Auswertung.

Zusammenfassung

Randomisierte und kontrollierte Klinische Studien (RCTs) werden aufgrund ihres Designs in eine hohe Evidenzklasse eingeordnet. Sie weisen ein hohes Maß an interner und externer Validität auf. Die Berichterstattung dieser Ergebnisse ist allerdings nicht einheitlich und somit für den Leser oft schwer umzusetzen. Um das Vertrauen des Lesers in die präsentierten Daten und Ergebnisse zu stärken und einen einheitlichen Standard für die Berichterstattung zu erzielen, wurde das CONSORT-Statement entwickelt und bis heute weiter an die Forschungslandschaft angepasst. Es stellt eine Checkliste und ein Flussdiagramm zur Verfügung und unterstützt damit eine transparente Dokumentation. Diese Vorlagen stellen eine minimal zu bearbeitende Anzahl von Punkten dar, die je nach Studiendesign durch Erweiterungen vervollständigt werden können und müssen.

Umsetzung des CONSORT-Statement in einer konkreten Studie

Im Folgenden sollen entlang einer kürzlich publizierten Studie von *Kim* et al. [10] das CONSORT-Statement in der aktuellsten Form von 2010 angewendet und dessen einzelne Kriterien illustriert werden durch konkrete interpretatorische Umsetzung (Tab. 1). Die Studie vergleicht im 1:1 randomisierten Ansatz 2 Implantate, wobei jedem Studienteilnehmer 2 Implantate des jeweils gleichen Typs eingesetzt wurden (d.h. kein intraindividueller Vergleich vorgenommen wird). Die Studie hat insgesamt 50 Patienten über 12 Monate nach Randomisation beobachtet, davon im Ergebnis 26 Patienten mit dem ersten versus 24 Patienten mit dem zweiten Implantat versorgt (also 52 versus 48 Implantate) im posterioren Quadranten.

Zusammenfassend ist bei dieser Studie auffallend, dass der primäre Endpunkt sowie die sekundären Endpunkte nicht konkret differenziert werden, wenn auch die Untersuchung in der von den Autoren gewählten Interpretation und Gewichtung die Erfolgsrate der Implantate im einjährigen Verlauf in den Vordergrund stellt. Die schon im Vorfeld erwartbar geringen Verlustraten werden in der Publikation (u.a. im Abstract) als „vergleichbar" charakterisiert. Mit Blick auf die im einjährigen Zeitfenster beobachteten Erfolgsraten von 98,1 % versus 97,9 % ist diese Interpretation der klinischen Ergebnisse in jedem Fall legitim. Die statistische Methodenwahl suggeriert jedoch eine auf den Nachweis eines Unterschiedes hin abzielende Studienfragestellung. Im Abstract wird für diesen im Vordergrund berichteten – vermutlich als primär anzusehenden – Endpunkt weder ein p-Wert noch ein Konfidenzintervall berichtet. Für diverse weitere Endpunkte wie die Implantatbeschaffenheit im Verlauf hingegen werden ohne Korrektur für multiples Testen mehrere p-Werte berichtet. Es kann daher nur vermutet werden, dass aufgrund der Planung der Studie auf Nachweis eines Unterschieds und der – absehbar – geringen Ordnung des Unterschiedes im primären Endpunkt nolens volens eine „Äquivalenztendenz"-Aussage auf Basis alleinig deskriptiver Angaben als Kompromiss formuliert wurde. Aufgrund einer nicht belastbaren Fallzahlplanung war zudem die angestrebte Patienten respektive Implantatzahl selbst für eine Studie zum Nachweis eines Unterschieds zwischen den Komplikationsraten nicht ansatzweise suffizient. Wäre zudem mit Blick auf die klinische Rationale der Autoren im Vorfeld ein Studiendesign zum Nachweis von Nicht-Unterlegenheit eines der beiden Implantate parametrisiert worden, hätten sich noch weitaus umfangreichere Fallzahlen ergeben.

Der letztgenannte Aspekt zeigt zugleich jenseits der konkreten Studie einen grundlegenden Nachteil des CONSORT-Statements auf: Die im Statement beleuchteten Kriterien werden primär in der Form „liegt vor" versus „nicht erfolgt" dokumentiert ohne Hinterfragung gemachter Angaben bezüglich deren Korrektheit im Kontext. Es liegt eine ausführlich dokumentierte Fallzahlplanung vor in der betreffenden Studie – diese ist aber sowohl im Ansatz (Planung auf Unterschied statt Planung auf Nicht-Unterlegenheit) wie auch in der konkreten Umsetzung nicht korrekt: Für diese Studie hätte die Planung der Patienten- oder Implantatzahl auf den erwarteten Unterschied zwischen den implantatweisen Auftrittshäufigkeiten des primären Endpunkts ausgelegt werden müssen, nicht aber wie in der Publikation dokumentiert auf die separate Schätzung einzelner implantatweiser Auftrittshäufigkeiten. Auch das von den Autoren in der Publikation explizit abgebildete Flussdiagramm zum Verlauf der Patientenzahlen im Studienablauf ist bei genauerer Betrachtung nicht mit den wirklichen Anforderungen des CONSORT-Statement kompatibel: drop-out-Raten werden nicht oder nur bedingt

	CONSORT-Kriterium		*A randomized controlled clinical trial of two types of tapered implants on immediate loading in the posterior maxilla and mandible [10]*	Einschätzung der Umsetzung des CONSORT-Kriteriums in der Beispiel-RCT
1	Titel und Abstract	**a** Kennzeichnung im Titel als RCT	*"A randomized controlled clinical trial . . ."*	Studie wird als „randomisiert" im Titel gekennzeichnet
		b Strukturierte Zusammenfassung	*"**Purpose**: The aim of this study was to compare clinical outcomes and stability following immediate loading of two types of tapered implants in the partially edentulous posterior maxilla and mandible.* ***Materials and Methods:** A randomized controlled trial with 1 year of follow-up was performed on participants missing two consecutive teeth in a posterior quadrant with tapered implants with a hybrid textured surface. Group 1 received Osstem TSIII HA implants, and group 2 received Zimmer TSV implants. […] Definitive restorations were provided 3 months (mandible) or 6 months (maxilla) later. Outcome measures were survival and success rates, marginal bone level change, implant stability quotient, and peri-implant soft tissue indices.* ***Results:** Fifty participants completed the trial (group 1:52 implants in 26 patients; group 2:48 implants in 24 patients). The success rates were similar — 98.1 % in group 1 and 97.9 % in group 2 — at 12 months after immediate loading, but marginal bone loss was significantly different according to the implant design. Implant stability increased significantly in both arches. There were no significant differences in soft tissue indices between implant systems.* ***Conclusion:** If high primary stability is acquired, tapered implants with hybrid textured surfaces are predictable for immediate loading in the posterior maxilla and mandible. In spite of the influence of implant design on marginal bone loss, all tapered implants showed successful clinical outcomes and stability in immediate loading."*	Strukturierter Abstract ist gegeben, durch folgende Punkte hätte dieser noch transparenter dokumentiert werden können: • Charakterisierung des faktischen Studienziels („Unterschied" versus „Nicht-Unterlegenheit") • Anzahl der randomisierten Patienten • Informationen zum Patienteneinschluss • klare Deklaration des primären Endpunkts • Konfidenzintervall zum Effekt im Endpunkt • Komplikationen/unerwünschte Vorkommnisse • Dokumentation der Registrationsnummer
Einleitung				
2	Hintergrund/Ziele	**a** Wissenschaflicher Hintergrund und Rationale der Studie	*"Although published studies on immediate loading are increasing, most of them are limited to the mandible, especially in the anterior region. […] This result shows that poor bone quality can be considered risk factors […] The statements of the 2003 ITI Consensus Conference noted that immediate restoration or loading was not well documented in the partially dentate maxilla, […] Therefore, immediate loading of implants in the posterior maxilla is yet to be studied. […], irrespective of the implant surface, although implants with modified surfaces showed higher success rates than implants with a machined surface. Many studies stated that primary stability might be a fundamental requirement for immediate loading. […] Implant design and surface modification are considered the two most influential factors for the success of immediate loading in soft bone. An implant with a tapered or threaded form contributes to primary stability by avoiding micro motion, and surface modifications can decrease healing time. […], but tapered implants with different thread designs and a similar surface texture have not yet been compared with respect to clinical outcomes and stability following immediate loading."*	Der Hintergrund der Studie wird unter Berücksichtigung der bestehenden Literatur beschrieben.

	CONSORT-Kriterium		*A randomized controlled clinical trial of two types of tapered implants on immediate loading in the posterior maxilla and mandible [10]*	Einschätzung der Umsetzung des CONSORT-Kriteriums in der Beispiel-RCT
		b genaue Fragestellung oder Hypothese	*"The primary objective of this randomized clinical trial was to evaluate the success of immediately loaded tapered implants in the partially edentulous posterior maxilla and mandible. The secondary objective was to compare clinical outcomes and changes in stability of two types of tapered implant systems with different thread design."*	Es wird zwar ein Endpunkt formuliert, der vermutlich als primär zu verstehen ist; allerdings werden keine konkreten Kriterien genannt, wie dieser „festgemacht" werden soll, bzw. es ist unklar wie „Erfolg dieser Implantate" genau definiert und gemessen werden soll.
Methoden				
3	Studiendesign	**a** Beschreibung des Studiendesigns	*"A randomized clinical trial with a parallel group design … was conducted . . ."*	Eine Dokumentation des Studiendesigns ist erfolgt. Aber es handelt es sich um eine einfach verblindete Studie; dieser Hinweis fehlt hier und findet sich erst später im Text.
		b Änderungen des Studiendesigns		hier nicht zutreffend
4	Probanden/ Patienten	**a** Eignungskriterien	*"A total of 81 participants visited the Section of Dentistry … in response to an advertisement published in a newspaper […]* *Eligible participants had unilateral loss of two consecutive premolar(s) and/or molar(s) in the maxilla or mandible with fully healed sites, a recipient bed with sufficient bone width (≥ 7 mm) and height (≥ 12 mm) to host two 4.5– to 5–, 10-mm implants, and opposing dentition that consisted of natural teeth or a fixed partial denture. Participants were not included in the study if […]"*	Eignungskriterien und Ausschlusskriterien werden ausführlich berichtet.
		b Umgebung/Ort der Studiendurchführung	*"A total of 81 participants visited the Section of Dentistry at Seoul National University Bundang Hospital and Seoul National University Dental Hospital in response to an advertisement published in a newspaper or posters displayed at hospitals. […]* *Surgery was conducted by one oral and maxillofacial surgeon at each hospital who was familiar with the protocol and implant systems. […]* *Follow-up examinations were performed at […]* *During the healing period, clinical evaluations were performed […]"*	Informationen über die Umgebung/den Ort, wo die Studie durchgeführt wurde, sind ausführlich gegeben.
5	Behandlung/ Intervention	Intervention in jeder Studiengruppe	*"**Surgical Procedures*** *Surgery was conducted by […]* *During surgery, primary stability was measured by […]* *The patients received antibiotics […]* ***Prosthetic Procedures*** *Immediately after surgery, impression copings were connected to the implants and silicone impressions were made for provisional crowns. […]"*	Die einzelnen Schritte zur Nachverfolgung der Therapie werden explizit dargestellt.

	CONSORT-Kriterium		*A randomized controlled clinical trial of two types of tapered implants on immediate loading in the posterior maxilla and mandible [10]*	Einschätzung der Umsetzung des CONSORT-Kriteriums in der Beispiel-RCT
6	Endpunkte (Variablen)	**a** Definition der Endpunkte (primär und sekundär)	*"[…] to test the following null hypothesis: Two types of tapered implants with different thread design would show no significant differences in clinical outcomes and stability after immediate loading in the partially edentulous posterior maxilla and mandible.* *During surgery, primary stability was measured by magnetic resonance frequency analysis using a Mentor device (Osstell) and recorded in ISQs. […]* *During the healing period, clinical evaluations were performed […] In addition, periapical radiographs were taken at the time of provisional placement, […] MBL was calculated […] Changes in ISQs during the healing period were compared, and secondary stability of implants was evaluated. Soft tissue parameters […] were assessed according to published criteria. General criteria for implant success […] were followed."*	Die Dokumentation einer Hypothese muss eindeutig auf einen Endpunkt hinweisen. Dessen konkrete Nennung fehlt in der Darstellung bzw. es könnten mehrere Variablen angenommen werden. Im Teil „Material und Methoden" wird keine eindeutige Differenzierung zwischen primären und sekundären Endpunkten mit zugeordneten Messvariablen gegeben.
		b Änderung der Variablen nach Studienbeginn		hier nicht zutreffend
7	Fallzahl	**a** Fallzahlberechnung	*"The required sample size was estimated on the basis of α = .05 (z1: conidence lower limit), power = 0.80 (z2: conidence lower limit), allocation ratio (r) = test (Nt)/control (Nc) = 1, estimated success ratio .97, clinically nondiferential maximum allowance D = 0.1 (Nt or Nc = [(1 + r)p(1 – p)(z1 + z2)2]/r . D2 = [(1 + 1)0.97(0.03)(1.645 + 0.842)2]/(1 . 0.12) ≈ 35.99). A dropout rate of 30 % was assumed, and this increased the number of implants required in each group to 51 (25.5 patients)."*	Hier wird ausführlich eine Fallzahlplanung mit den notwendigen Angaben beschrieben. Diese Planung ist auch in der dargestellten Form per se fehlerfrei, trifft aber auf das eigentlich angestrebte – vergleichende – Studiendesign nicht zu: Es hätte eine auf den Unterschied zwischen den Auftrittshäufigkeiten des primären Endpunktes zwischen den Studienarmen ausgerichtete Fallzahlplanung vorgenommen werden müssen, die selbst bei einer Planung zum Nachweis eines Unterschiedes in den Auftrittshäufigkeiten eine mindestens dreistellige Patientenzahl pro Implantat bedingt hätte. Insofern ist im Ergebnis eine Fallzahlplanung in der Publikation berichtet, diese ist aber faktisch falsch mit klarer Konsequenz für die Belastbarkeit der Studie und der daraus abgeleiteten Aussagen.
		b Zwischenanalysen		hier nicht zutreffend, da nicht im Design eingeplant und auch laut Angaben der Autoren nicht aus Sicherheitsgründen ungeplant notwendig geworden
8	Randomisation/ Sequenzgenerierung	**a** Methode zur Generierung der Zufallszuteilung	*"Participants were randomized into two groups for immediate loading. The implant system was assigned by a randomization list generated by using a random number table …"*	Die Grundidee der Randomisation wird beschrieben.
		b Art der Randomisation	*"… with block sizes of four; each group had an equal number in each block."*	Die Art der Randomisation wird klar beschrieben.
9	Randomisation/Methode der Zuteilung		*<keine Angaben>*	Hier sollte beschrieben werden, wie die Randomisation logistisch umgesetzt wurde, zum Beispiel per Umschlag oder „third party" (telefonisch durch eine Agentur ...).
10	Randomisation/Umsetzung (Teilnehmerweise)		*"Randomization was implemented by a research nurse at Seoul National University Dental Hospital."*	Die Verortung der Umsetzung ist beschrieben.

	CONSORT-Kriterium		*A randomized controlled clinical trial of two types of tapered implants on immediate loading in the posterior maxilla and mandible [10]*	Einschätzung der Umsetzung des CONSORT-Kriteriums in der Beispiel-RCT
11	Verblindung	**a** Verblindung (Methode)	*"The implant system assigned was concealed from the participants until the end of the trial and from the surgeons until the time of surgery."*	Hier bekommt der Leser die Information, dass es sich um eine einfach verblindete Studie handelt. Das ist unter Punkt 3a nicht explizit erfolgt, hier aber zumindest aktiv berichtet.
		b Vorgehen zur Gewährleistung der Ähnlichkeit der Interventionen	*"The two implant systems used in the study had a tapered body with a similar hybrid textured surface: […] and pitch height (Fig 1). The Osstem TSIII HA implant system (Osstem Implant Co; group 1 implants) […] (Zimmer Dental; group 2 implants) had an Ra of 0.98 µm (RBM) and 2.84 µm (HA). The crystallinity of the HA of TSIII HA and TSV surfaces was 97.6 % and 97.0 %, respectively."*	Beschreibung der einzelnen Implantate erfolgt; eine weitergehende Maskierung respektive Sondierungen hierzu werden nicht beschrieben. Da der primäre Endpunkt „Erfolgsrate" in dieser Studie ohnehin kaum einer Behandlerseitigen Manipulation in der Messung unterliegen kann, ist diese Darstellung hier suffizient.
12	Statistische Methoden	**a** Methoden zum Vergleich der Gruppen (primäre Fragestellung)	*"Comparative evaluation of the two implant systems was performed with the chi-square test for categorical variables. The level of significance was set at $p < .05$."*	Die Methode für die Auswertung der primären Fragestellung wird gegeben, wenn auch in einer nicht formal korrekten Form („... zum Signifikanzniveau 0,05, d.h. eine Signifikanz ergäbe sich bei $p <= 0{,}05$."). Die Wahl des Signifikanztests hingegen ist als suboptimal anzusehen: Bei den erwartbar geringen Komplikationsinzidenzen wäre ein exakter Test (nach Fisher) und ein exaktes 95%-Konfidenzintervall zum Vergleich der Auftrittshäufigkeiten im primären Endpunkt sachgerecht.
		b Methoden für zusätzliche Analysen	*„To compare MBL and soft tissue indices between the implant systems, the Mann-Whitney U test was used for continuous variables, after confirmation of normality of the distribution with the Shapiro-Wilk method. Changes in ISQ according to healing time and differences in ISQ according to the direction were compared between the implant systems by a generalized estimating equation unstructured model. Irrespective of normality of the distribution, repeated-measures analysis of variance (ANOVA) in the generalized linear model was also performed for comparison. In cases of within-group significance of differences, paired evaluations were also assessed by means of the Wilcoxon signed-rank test. The level of significance was set at $p < .05$."*	Die Beschreibung der Methodenwahl für explorative Analysen ist gut zusammengefasst und in dieser auch auf die Studie ausgerichteten Form suffizient. Die zu verwendenden Methoden selbst sind suboptimal.
Ergebnisse				
13	Anzahl der Studienteilnehmer	**a** Teilnehmerzahl in jeder Gruppe für Randomisation, Durchführung und Analyse	*"Of the 81 participants who were recruited, 61 satisfied the inclusion criteria and were selected for the study (Fig. 2). The 61 subjects (group 1:30, group 2:31) received 122 implants […] Seven subjects (three in group 1, four in group 2) were excluded from the study because of poor primary stability (ISQ < 60) postoperatively. During the healing period, one subject in group 2 dropped out because of a local inflammatory reaction around the implant. After definitive prosthesis delivery, three more subjects (one in group 1, two in group 2) dropped out because of loss of contact (Fig. 2). Finally, 50 patients (group 1:26 patients; group 2:24 patients) were enrolled in the study and the 1-year follow-up period."*	Es wird die Anzahl von eingeschlossenen und randomisierten Patienten in Textform berichtet. Ein dem CONSORT-Vorschlag angelehntes Flussdiagramm ist vorhanden, welches aber nicht vollständig ist (u.a. nicht die Anzahl der Studienteilnehmer zu allen Terminen ausweist) und nicht die beiden Studienarme separat in der Ausschluss-Chronologie nachvollziehbar darstellt.
		b Verlust von Patienten in jeder Gruppe (lost to follow-up und Ausschlüsse) mit Begründung		

	CONSORT-Kriterium		*A randomized controlled clinical trial of two types of tapered implants on immediate loading in the posterior maxilla and mandible [10]*	Einschätzung der Umsetzung des CONSORT-Kriteriums in der Beispiel-RCT
14	Aufnahme/ Rekrutierung	**a** Zeitpunkte für Aufnahme und Nachkontrollen	*"Of the 81 participants who were recruited, 61 satisfied the inclusion criteria and were selected for the study. The 61 subjects […] received 122 implants between September 2010 and February 2011.* *Follow-up examinations were performed at 2 and 4 weeks; at 2 months and 3 (mandible) or 6 (maxilla) months postoperative; and at 12 months postoperative."*	Die zeitliche Einordnung der Untersuchung wird klar beschrieben.
		b Zeitpunkt und Begründung für Studienende	*"Finally, 50 patients (group 1:26 patients; group 2:24 patients) were enrolled in the study and the 1-year follow-up period."*	Status Quo und Zeitpunkt zum Ende dieser Studie sind klar beschrieben. Weitere Angaben sind hier nicht notwendig, da die Studie nicht unvorhergesehen gestoppt wurde.
15	Demographische und klinische Patienten-Charakteristika zu Studienbeginn		*„The mean age of group 1 patients was 51.6 years, and group 2 patients were, on average, 49.6 years at the time of surgery. The male:female ratios were 12:14 (group 1) and 18:6 (group 2), and there was a borderline difference between the two groups (chi-square test, p = .048). In group 1, 22 implants were inserted into maxillae and 30 into mandibles. In group 2, 24 implants were inserted into both arches."*	Klinische Patienteninformationen werden gegeben. Es wäre jedoch bei der zu klärenden Fragestellung ggf. noch hilfreich, weitergehende soziodemographische Informationen zu berichten in Form prognostisch relevanter Kofaktoren.
16	Anzahl der Teilnehmer, die in die Analyse eingeschlossen wurden		*„Finally, 50 patients (group 1:26 patients; group 2:24 patients) were enrolled in the study and the 1-year follow-up period."*	Die zur Analyse verwendete Teilnehmeranzahl wird berichtet.
17	Ergebnisse und Schätzungen	**a** Ergebnisse für den primären und die sekundären Endpunkte (incl. Konfidenzintervall oder p-Wert)	*"Two implants, both in maxillae, one from each system, showed severe marginal bone loss (2.83 mm in group 1, 3.77 mm in group 2). Therefore, the success rates were 98.1 % in group 1 and 97.9 % in group 2 after 1 year of follow-up."*	Für den primären Endpunkt und die primäre Fragestellung werden nur deskriptive Ergebnisse mittels relativer Häufigkeiten berichtet, eine Signifikanz-Bewertung mittels Konfidenzintervall zum Unterschied zwischen den beiden Auftrittshäufigkeiten des primären Endpunkts wird nicht vorgenommen. Auch der in der Methodenformulierung avisierte Signifikanztest wird nicht berichtet (p-Wert?).
		b Angabe sachgerechter Kenngrößen für binäre Endpunkte (incl. Konfidenzintervall oder p-Wert)	*<keine Angabe>*	Es werden ausschließlich deskriptive Effektmaße in Form von relativen Häufigkeiten angegeben, nicht aber eine hier sinnvolle Number Needed to Treat oder eine Risikodifferenz.
18	Ergebnisse von zusätzliche Analysen		*"Two implants, both in maxillae, one from each system, showed severe marginal bone loss (2.83 mm in group 1, 3.77 mm in group 2). Therefore, the success rates were 98.1 % in group 1 and 97.9 % in group 2 after 1 year of follow-up."*	Ergebnisse explorativer Analysen werden ausführlich und auch methodisch sinnvoll deskriptiv berichtet. Die Abtrennung von konfirmatorischen Ergebnissen zum primären Endpunkt ist jedoch nicht optimal; eine klare Kenntlichmachung als explorative Ergebnisse wird nicht vorgenommen, p-Werte werden nicht als Ergebnisse von explorativen Ergebnissen charakterisiert.
19	unvorhergesehener Schaden für Teilnehmer („harm")		*"Seven subjects (three in group 1, four in group 2) were excluded from the study because of poor primary stability (ISQ < 60) postoperatively. During the healing period, one subject in group 2 dropped out because of a local inflammatory reaction around the implant."*	Die Häufigkeiten von „adverse events" werden nicht explizit dokumentiert respektive Vorkommnisse nicht explizit charakterisiert bezüglich klinischer Relevanz und kausaler Einordnung zur studienbedingten Therapie.

	CONSORT-Kriterium	*A randomized controlled clinical trial of two types of tapered implants on immediate loading in the posterior maxilla and mandible [10]*	Einschätzung der Umsetzung des CONSORT-Kriteriums in der Beispiel-RCT
Diskussion			
20	Limitationen, die durch die Studienplanung oder Durchführung Verzerrungen der Ergebnisse bewirken können	*„This randomized clinical trial revealed a 100 % survival rate […] The null hypothesis tested […] The results of this study partially support rejection of the null hypothesis, because the implants showed significant differences in MBL after 1 year. […]* *Within the limitation of the short-term follow-up period of 1 year, the survival rate […]"*	Die Ergebnisse der Studie werden im Zusammenhang mit der aktuellen Literatur diskutiert. Aber Limitationen bezüglich dieser Untersuchung werden nur ansatzweise genannt und nicht in deren möglicher Konsequenz für die Aussagekraft der präsentierten Daten diskutiert.
21	Generalisierbarkeit der Ergebnisse	*"[…] If high primary stability is acquired, tapered implants with hybrid textured surfaces are predictable for immediate loading in the posterior maxilla and mandible. Although the implant thread design appeared to influence the amount of marginal bone loss, the investigated tapered implants showed successful clinical outcomes and stability following immediate loading."*	In Bezug auf die Verwendung dieser Implantate wird eine Empfehlung ausgesprochen, aber keine Information darüber, wie weit diese Studie validen und Versorgungstypischen Prozessen unterlag und somit Allgemeingültigkeit der Ergebnisse für „übliche Versorgungsabläufe" angenommen werden kann.
22	Interpretation der Ergebnisse im Zusammenhang mit der bestehenden Literatur	*"The success rates for immediate loading in the posterior quadrant were 98.1 % (group 1) and 97.9 % (group 2) in the present study. The survival rate was 100 %. According to previous studies 9, 10, 18, 19 strict patient selection, high primary stability, and splinting are prerequisites for positive outcomes of immediate loading, and long implants are also recommended. There was no difference in primary stability between implant systems with different thread designs. The Zimmer TSV implants (group 2) did not display any advantage of cortical bone anchorage from their design; thus, high primary stability in the group 2 implants could be attributed to their taper. Slightly tapered implant designs have shown better primary stability compared with straight-walled implants, especially in types 3 and 4 bone. Glauser et al. 9 stated …"*	Die Ergebnisse der Studie werden für die unterschiedlichen Variablen und im Zusammenhang mit der aktuellen Literatur explizit diskutiert.
Regulatorische Informationen			
23	Registrierungsnummer der klinischen Studie	*<keine Angabe>*	Registrierungsnummer wird nicht angegeben.
24	Verweis zum Studienprotokoll	*<keine Angabe>*	Ein Verweis über eine Ethik-Voten- oder Studienregistrierungs-bedingte Referenzierung oder Vorab-Publikation des Protokolls sind nicht verfügbar.
25	Finanzierung der Studie/ Interessenskonflikte	*"This study was supported by a grant from the Korea Healthcare Technology R&D Project, Ministry of Health and Welfare, Republic of Korea (A101724). The authors reported no conflicts of interest related to this study."*	Förderer beziehungsweise finanzielle Interessen werden beschrieben. Wie die Finanzierung der Implantate in der Studie erfolgte, wäre jedoch optimalerweise noch zu kommentieren (durch Leistungserstatter im Rahmen der Versorgung, eigene Ausstattung, Anteil der Förderung?) mit Blick auf die Fragestellung der Studie.

Tabelle 1 Anwendung der CONSORT-Kriterien auf eine zweiarmige randomisierte Klinische Prüfung zum Vergleich zweier zahnärztlicher Implantate [10]: Nennung des entsprechenden Kriteriums der Prozess- und Ergebnis-Charakteristik der Studie laut CONSORT-Statement, ggf. gekürzte Zitierung der das Kriterium betreffenden Angaben in der Originalarbeit und Bewertung des Erfülltheitsgrades zu diesem Kriterium auf Basis der verfügbaren Angaben.

Table 1 CONSORT criteria applied to a two-arm randomized clinical trial on the comparison of two dental implants [10]; citations of information provided by the authors in the publication are displayed to demonstrate the extent of availability and correctness of for each of the CONSORT criteria.

implantatweise im Diagramm dargestellt, eine vor der Rekrutierung sicherlich vorgenommene Screeningphase möglicher in Frage kommender Teilnehmer wird nicht dargestellt.

In der Gesamtschau können formal zahlreiche Kriterien des CONSORT-Statements als berücksichtigt bzw. zum Teil sogar ausführlich in der Publikation dokumentiert bezeichnet werden. Gleichzeitig ist hiermit keine Aussage über die Korrektheit und Vollständigkeit der dokumentierten Angaben gegeben. Es muss also bei aller Berechtigung des CONSORT-Statements in der Sicherstellung transparenter Berichterstellung Klinischer Studien auch bei dessen Anwendung in einer Publikation stets neben dem reinen Vorhandensein von Angaben auch deren inhaltliche Belastbarkeit sondiert werden. In diesem Sinne sind auch Gutachter von Publikationen zu Klinischen Studien stets in der Verantwortung, nicht nur das reine Vorliegen von laut CONSORT zu machenden Angaben zu prüfen, sondern zumindest stichprobenhaft auch deren Sinnhaftigkeit im konkreten Kontext.

Korrespondenzadresse

Dr. rer. medic. Stephanie Knippschild
Institut für Medizinische Biometrie und Epidemiologie
Fakultät für Gesundheit der Universität Witten/Herdecke
Alfred-Herrhausen-Straße 50
58448 Witten
Tel.: 02302 926 763
Fax: 02302 926 44 785
stephanie.knippschild@uni-wh.de

Literatur

1. Amiri AR, Kanesalingam K, Cro S, Casey ATH: Does source of funding and conflict of interest influence the outcome and quality of spinal research? Spine J 2014;14:308–314
2. Bartholomäus E: Auf Worte Taten folgen lassen. Dtsch Ärztebl 2013;110:2
3. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, for the CONSORT Group: Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Ann Intern Med 2008; 148:295–309
4. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG: Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004;291:2457–2465
5. Dwan K, Altman DG, Arnaiz JA et al.: Systematic review of the empirical evidence of study publication bias and outcome reporting bias. PLoS One 2008;3:e3081
6. Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C, for the CONSORT Group: Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. Ann Intern Med 2006;144: 364–367
7. Gelberman RH, Samson D, Mirza SK, Callaghan JJ, Pellegrini VD: Orthopaedic surgeons and the medical device industry: the threat to scientific integrity and the public trust. J Bone Joint Surg Am 2010;92:765–777
8. Hopewell S, Clarke M, Moher D et al.: CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. PLoS Med 2008;5:e20
9. Ioannidis JP, Evans SJW, Gotzsche PC et al.: Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med 2004;141:781–788
10. Kim YK, Lee JH, Lee JY, Yi YJ: A randomized controlled clinical trial of two types of tapered implants on immediate loading in the posterior maxilla and mandible. Int J Oral Maxillofac Implants 2013;28:1602–1611
11. Krummenauer F, Baulig C, Al-Nawas B et al.: Risikomaße – kompakt, aber vielfältig interpretierbar. Z Zahnärztl Implantol 2008;24:57–59
12. Krummenauer F, Baulig C, Al-Nawas B et al.: Studiendesigns in der Implantologie (II): Fallzahlplanung – zwischen Hoffen und Wissen. Z Zahnärztl Implantol 2010;26:278, 280, 282, 284
13. Krummenauer F, Baulig C, Hirsch J et al.: Randomisation in Klinischen Studien – durch Zufall zum Erfolg. Z Zahnärztl Implantol 2014;30:60–64, 66
14. Moher D, Hopewell S, Schulz K et al.: CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c869
15. Schulz KF, Altman DG, Moher D: for the CONSORT Group: CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010;8:18
16. The Consort Group: „Consort – Transparent reporting of Trials website“, (Online) Available: http://www.consort-statement.org/checklists/view/32-consort/66-title (Zugriff am 15.12.2014)
17. Tukey JW: Some thoughts on clinical trials, especially problems of multiplicity. Science 1977;198:679–684