

Koch EC^{1,2}, Blümle A¹, Antes G¹

Randomisierte und kontrollierte klinische Studien in der Deutschen Zeitschrift für Sportmedizin zwischen 1950 und 2004

Randomised and controlled clinical trials published in the German sports medicine journal "Deutsche Zeitschrift für Sportmedizin" between 1950 and 2004

¹ Deutsches Cochrane Zentrum, Institut für Medizinische Biometrie und Medizinische Informatik, Universität Freiburg

² Klinik für Anästhesie und Intensivmedizin, Klinikum Landkreis Tuttlingen

Zusammenfassung

Randomisierte kontrollierte (RCTs) und kontrollierte klinische Studien (CCTs) bilden die valide Grundlage für die Wirksamkeitsbewertung medizinischer Maßnahmen. Die Erfassung aller Studien zu einer medizinischen Fragestellung ist jedoch kaum möglich, da viele Artikel in nationalen Zeitschriften publiziert werden und als nicht-englischsprachige Publikationen oftmals nicht in elektronischen Datenbanken wie Medline (www.pubmed.gov) erfasst sind. Um die Vollständigkeit der Literaturerfassung zu gewährleisten, sollte eine systematische elektronische Datenbanksuche durch eine Handsuche relevanter medizinischer Zeitschriften ergänzt werden. Ziel der vorliegenden Arbeit war die quantitative und qualitative Charakterisierung von kontrollierten Studien, die in der „Deutschen Zeitschrift für Sportmedizin“ zwischen 1950 und 2004 publiziert worden sind. Material und Methoden: Die „Deutsche Zeitschrift für Sportmedizin“ wurde manuell nach RCTs und CCTs durchsucht. Die identifizierten Studien wurden auf quantitative, qualitative und inhaltliche Aspekte hin untersucht, basierend auf den Richtlinien des CONSORT-Statements. Ergebnisse: 490 kontrollierte Studien wurden identifiziert (182 RCTs; 308 CCTs). 289 (59 %) hatten einen trainingswissenschaftlichen, 143 (29 %) einen klinischen Hintergrund, 58 (12 %) waren fächerübergreifend. Die Evaluation der Studien ergab, dass die CONSORT-Kriterien wie wissenschaftlicher Hintergrund, Art der Intervention, Anzahl der Probanden, statistische Methoden sowie allgemeine Interpretation der Studienergebnisse in den meisten Artikeln (> 90 %) beschrieben waren. Über wichtige Angaben wie Fallzahlbestimmung, Randomisierungsmethode, Geheimhaltung der Behandlungsfolge, Verblindung, Anzahl der ausgewerteten Probanden sowie Signifikanz der Ergebnisse wurde hingegen nur unzureichend (< 10 %) berichtet. Schlussfolgerung: Durch die Anwendung des CONSORT-Statements kann die Qualität der Berichterstattung von RCTs verbessert werden. Dazu ist eine stärkere Wahrnehmung des CONSORT-Statements bei Zeitschriften, Herausgebern und Autoren auch im deutschsprachigen Raum nötig.

Schlüsselwörter: CONSORT, Handsuche, Qualitätsbewertung, Publikationsbias, Sprachbias, Medline

Einleitung

In der Sportmedizin treffen sportwissenschaftliche und medizinische Fragestellungen mit dem gemeinsamen Ziel der Leistungsverbesserung des Sportlers bzw. Patienten sowie der Prävention von Krankheiten aufeinander. Dadurch eröffnet sich ein weites Feld für interessante Fra-

Summary

Randomised controlled (RCTs) and controlled clinical trials (CCTs) are the most valid source for the assessment of the effectiveness of health-care interventions. There is no reasonable chance to identify all available and relevant studies concerning a medical intervention since many articles were published in national journals and, thus they as non-English-speaking publications, they are often not included in electronic databases like Medline (www.pubmed.gov). To be comprehensive, an electronically-driven database search has to be complemented by hand-searching of relevant medical journals. The aim of this study was to assess the quantity and quality of controlled clinical trials, published in "Deutsche Zeitschrift für Sportmedizin" between 1950 and 2004. Materials and Methods: The journal "Deutsche Zeitschrift für Sportmedizin" was searched manually for RCTs and CCTs. Trials identified were examined for aspects of quantity, quality and contents, based on the guidelines of the CONSORT-Statement. Results: 490 controlled trials were identified (182 RCTs; 308 CCTs). 289 (59 %) of them had a training-scientific background, 143 (29 %) a clinical background and 58 (12 %) were multidisciplinary. The evaluation of the studies showed that CONSORT-criteria such as scientific background, kind of intervention, number of participants, statistical methods and interpretation of the results were described in the majority of articles (> 90 %). But important items like sample size calculation, randomisation-sequence generation, allocation concealment, blinding, number of participants analysed and statistical significance of the results were insufficiently (< 10 %) reported. Conclusion: The implementation of the CONSORT-Statement is mandatory to improve the quality of reports of RCTs. Therefore an increasing awareness and publicity of the CONSORT-Statement is needed, especially for journals, publishers and authors, also in the German-speaking areas.

Key words: CONSORT, handsearching, quality assessment, publication bias, language bias, Medline

gestellungen und Untersuchungen bezüglich Leistungsdiagnostik, Trainingswissenschaft, Pharmakawirkung, Rehabilitation und vieles mehr. In Analogie zu anderen Fachbereichen existiert auch in der Sportmedizin eine für den Einzelnen unüberschaubare Vielzahl veröffentlichter Studien – nicht selten mit gegensätzlichen Ergebnissen zur gleichen Problemstellung. Randomisierte kontrollierte Studien (engl. randomised controlled trials, RCTs) gelten

auch in der Sportmedizin als „Goldstandard“, um kausale Zusammenhänge zwischen geprüfter Intervention und beobachtetem Ergebnis zu untersuchen und nachzuweisen (7). Sie bilden die valideste Grundlage für die Wirksamkeitsbewertung medizinischer Maßnahmen und stehen daher auf der höchsten Stufe der Evidenzskala (Tab. 1).

Tabelle 1: Klassifizierung der Validität von Aussagen in Evidenzstufen (Eine feinere Abstufung dieser Klassifikation sind als Oxford Centre for Evidence-based Medicine Levels of Evidence (May 2001) unter http://www.cebm.net/levels_of_evidence.asp zu finden).

Level der Evidenz	Therapiestudien
I	Randomisierte kontrollierte Studien
II	Kohortenstudien
III	Fall-Kontroll-Studien
IV	Fall-Serien
V	Experten

Viele der durch groß angelegte RCTs gewonnenen Erkenntnisse sind inzwischen fester Bestandteil des klinischen Alltags. Oft wird die Verbreitung und Umsetzung von Studienergebnissen jedoch durch eine Reihe von Faktoren verzögert, behindert oder sogar verfälscht. Aufgrund selektiven Publizierens gelangen nicht alle Ergebnisse klinischer Studien an die Öffentlichkeit (6, 14). Darüber hinaus können publizierte Studienergebnisse in der Versorgung nicht immer genutzt und angewendet werden, da vielfältige logistische Probleme einen einfachen, nutzerfreundlichen Zugang zu Bibliotheken und elektronischen Quellen erschweren. In den bestehenden medizinischen Datenbanken, wie z. B. Medline, kann der versierte Kliniker eine Vielzahl von Studien und Übersichtsarbeiten zu einer konkreten Fragestellung identifizieren und sich Zugang zu einem sehr großen Teil der aktuellen Datenlage verschaffen. Ergebnisse aus Studien, die nicht veröffentlicht wurden oder in einer Zeitschrift erschienen sind, die nicht in Medline gelistet ist, bleiben dabei jedoch unberücksichtigt. Bei den in Medline gelisteten Zeitschriften sind die Originalarbeiten i. d. R. über den freien Medline-Zugang Pubmed auffindbar, während Abstracts bzw. Kongressberichte hingegen oft in separaten Supplementbänden publiziert und nicht in die Datenbanken aufgenommen werden. Ebenfalls unterrepräsentiert in Medline sind nicht-englischsprachige Zeitschriften (2,13,18).

Diese Mängel im Publikationsprozess von Studien führen dazu, dass systematische Zusammenfassungen von Studienergebnissen durch vielfältige systematische Verzerrungen (Bias) verfälscht sein können: Selektive Publikation von Studienergebnissen (Publikationsbias) bringt nur einen Teil der Ergebnisse an die Öffentlichkeit (4); interessante und relevante Ergebnisse werden zunehmend in Englisch publiziert und deswegen in verschiedenen Sprachräumen unterschiedlich wahrgenommen (Sprachbias) (8); auch publizierte Ergebnisse sind nicht immer leicht zu finden (Retrievalbias) (6).

Abhilfe schaffen kann hier die systematische manuelle Durchsuchung (im Folgenden als Handsearching (Handsuche) bezeichnet) von medizinischen Zeitschriften nach randomisierten kontrollierten Studien und kontrollierten klinischen Studien (engl. controlled clinical trials, CCTs). Dabei

werden alle Originalarbeiten und Kurzberichte (Abstracts oder Kongressberichte) von veröffentlichten RCTs und CCTs erfasst. Das Handsearching ist gerade bei nicht in Medline gelisteten Zeitschriften von besonderer Bedeutung, da die in ihnen publizierten Studien ohne diese Arbeit praktisch unauffindbar blieben. Ihre Ergebnisse werden infolgedessen nicht in systematische Übersichtsarbeiten mit einbezogen und somit in der Versorgung nicht adäquat berücksichtigt.

Eine weltweite Handsearch-Initiative unter der Leitung der Cochrane Collaboration führte in den letzten zehn Jahren zu einem enormen Zuwachs an elektronisch verfügbaren Studien (12). Durch die Handsearch-Tätigkeit am Deutschen Cochrane Zentrum konnten bisher 17366 deutschsprachige kontrollierte klinische Studien identifiziert werden. Von diesen waren nur 7852 (45,2 %) in Medline registriert (9).

Die identifizierten Literaturzitate werden in das Cochrane Central Register of Controlled Trials (Clinical Trials) (ehemals CENTRAL) aufgenommen, der derzeit größten existierenden Datenbank kontrollierter Studien mit inzwischen über 470000 Literaturzitaten (10). Außerdem leitet die Cochrane Collaboration alle gefundenen RCTs und CCTs zur nachträglichen Kennzeichnung der Medline-Zitate mit einem Publikationstyp (RCT oder CCT) an den Medline-Herausgeber, die National Library of Medicine, weiter und trägt so zu einer effizienteren Indexierung in den elektronischen Datenbanken bei (5). Ziel dieser Studie ist, die in der Deutschen Zeitschrift für Sportmedizin veröffentlichten Publikationen klinischer Studien (RCTs und CCTs) quantitativ und qualitativ zu erfassen. Es wurden verschiedene Elemente der Studienmethodik bewertet, darüber hinaus aber auch die Studienziele und deren klinische Relevanz.

Methodik

Die medizinische Fachzeitschrift Deutsche Zeitschrift für Sportmedizin (seit 1978 unter diesem Namen, zuvor Sportarzt und Sportmedizin (1964-77), Der Sportarzt (1959-63), Sportmedizin (1953-58) und Leibesübungen-Sportarzt-Erziehung (1950-52)), einschließlich der in dieser Zeitschrift veröffentlichten Abstractbände zu Sportmedizinischen Kongressen, wurde am Deutschen Cochrane Zentrum Freiburg (www.cochrane.de) retrospektiv bis hin zum Publikationsbeginn im Jahr 1950 manuell nach Originalartikeln und Abstracts durchsucht, welche mindestens zwei therapeutische oder diagnostische Verfahren im Parallelgruppen- oder Crossover-Design miteinander verglichen. Studien mit randomisierter Interventionszuordnung wurden als RCT klassifiziert. Wurde die Gruppenteilung nicht genauer beschrieben, konnte eine Randomisierung jedoch nicht ausgeschlossen werden oder erfolgte die Zuteilung quasi-randomisiert (z.B. alternierend, durch Münzwurf oder mit Hilfe der Registrierungsnummer des Krankenhauses), wurde die Studie als CCT bezeichnet. Die Klassifizierung in RCTs und CCTs erfolgte nach den Vorgaben des Handsearch-Manuals des US Cochrane Centers (19).

Anschließend wurden die Studien auf ihren Informations-

gehalt geprüft. Beurteilt wurden hierbei Inhalt, Studien- durchführung, Verblindung, Art der Kontrollgruppe, Ver- gleich der Interventionsgruppen, statistische Signifikanz der Ergebnisse sowie demografische und klinische Angaben über das untersuchte Patienten- bzw. Probandenkollektiv (Tabel- le 2). Die Analysekriterien basierten auf den Kriterien des CONSORT-Statements, das 1996 erstmalig publiziert wurde (3). Ziel dieser Richtlinie ist die Verbesserung der Qualität

von Studienberichten von RCTs durch eine standardisierte und transparente Darstellung der Ergebnisse. Das CONSORT- Statement wurde 2001 in einer überarbeiteten Version ver- öffentlicht (1) und 2004 vom Deutschen Cochrane Zentrum ins Deutsche übersetzt (11). Ausgehend von einer minimalen Implementierungsphase von 3 Jahren wurden im Rahmen der vorliegenden Studie Originalartikel, die zwischen 1999 und 2004 publiziert worden waren, hinsichtlich der Bericht-

Tabelle 2: Analysekriterien und –ergebnisse aller identifizierten RCTs und CCTs sowie differen- ziert für die jeweilige Publikationsform (Originalartikel oder Abstract). ©Proband erhielt zeit- gleich sowohl Prüf- als auch Kontrollintervention. * nicht ausgewertet

		alle Studien (n=490)		Original- arbeiten gesamt (n=179)		Abstracts (n=311)	
Studieninhalte	physiologisch	284	58%	86	48%	198	64%
	internistisch	88	18%	49	28%	39	13%
	diagnostisch	42	9%	13	7%	29	9%
	fächerübergreifend	76	15%	31	17%	45	14%
Hintergrund	trainingswissenschaftl.	289	59%	106	59%	183	59%
	klinisch	143	29%	48	27%	95	30%
	fächerübergreifend	58	12%	25	14%	33	11%
Zuteilung	randomisiert	182	37%	66	37%	116	37%
	gleichzeitig	6	1%	4	2%	2	<1%
Teilnehmerzahl	nicht spezifiziert oder quasi-randomisiert (=CCT)	302	62%	109	61%	193	62%
	<20	221	45%	80	45%	141	45%
Studiendesign	≥50	80	16%	35	20%	45	14%
	nicht angegeben	23	5%	2	1%	21	<1%
	parallel	237	48%	95	53%	142	46%
Kontrolle	cross-over	179	37%	75	42%	104	33%
	kombiniert (parallel+cross-over)	5	1%	1	<1%	4	1%
	gleichzeitig	4	1%	1	<1%	3	1%
	nicht angegeben	65	13%	7	4%	58	19%
Verblindung	Intervention	208	43%	74	41%	134	43%
	Placebo	65	13%	28	16%	37	12%
	keine Kontrolle	217	44%	77	43%	140	45%
Zielparameter	einfach	12	2%	6	3%	6	2%
	doppelt	50	10%	29	16%	21	7%
	keine nähere Angaben	1	<1%	0	0%	1	<1%
	gleichzeitig	2	<1%	0	0%	2	<1%
	nicht verblindet	3	<1%	3	2%	0	0%
	nicht möglich	370	76%	124	69%	246	79%
Anzahl der eingeschlossenen Pobjanden = Anzahl der ausgewerteten	nicht angegeben	52	11%	17	10%	35	11%
	objektiv	396	81%	135	76%	261	84%
	subjektiv	27	5%	13	7%	14	4%
Signifikanz der Ergebnisse	gemischt	67	14%	31	17%	36	12%
	zutreffend	55	11%	55	31%	0	0%
	nicht zutreffend	49	10%	44	24%	5	2%
Prospektivität	nicht angegeben	386	79%	80	45%	306	98%
	signifikant	197	40%	107	60%	90	29%
	nicht signifikant	74	15%	27	15%	47	15%
Baseline data	nicht angegeben	219	45%	45	25%	174	56%
	explizit erwähnt	25	5%	6	3%	19	6%
	nicht angegeben	465	95%	173	97%	292	94%
Allocation concealment	angegeben	*	*	106	59%	*	*
	nicht angegeben	*	*	47	26%	*	*
	ungenau	*	*	26	15%	*	*
Allocation concealment	nicht angegeben	490	100%	179	100%	311	100%

erstattung eines jeden CONSORT-Kriteriums und -Unterkriteriums kontrolliert. Durch die Unter- teilung der CONSORT-Punkte 11 und 13 in a), b) und c) waren somit maximal 26 Kriterien zu er- füllen (Tab. 3). Die Originalartikel, die vor 1999 publiziert worden waren sowie alle Abstracts wurden nur auf einige wenige, essentielle, Kriterien hin überprüft (Tab. 2).

Ergebnisse

In den Jahrgängen 1950 bis 2004 der Deut- schen Zeitschrift für Sportmedizin wurden insgesamt 490 kontrollierte Studien (179 Ori- ginalarbeiten und 311 Abstracts) gefunden. Sämtliche durch das Handsearching aufge- fundenen RCTs und CCTs waren nicht in Med- line erfasst. Aus dem gesamten Publikations- zeitraum dieser Zeitschrift (1950 bis heute) waren lediglich zwei Literaturzitate – ein Review und ein Fallbericht – in Medline enthalten. Abbildung 1a zeigt die Anzahl der Originalarbeiten des jeweiligen Publikationsjah- res. Eine Übersicht über die nur unregelmäßig in Abstractbänden veröffentlichten Studien gibt Abbildung 1b. Die Publikationsspitzen – bei den Originalarbeiten im Jahr 1998, bei den Abstracts in den Jahren 1999, 2001 und 2003 – sind durch die Herausgabe zusätzlicher Supplement-Bände bedingt, die Spezialthemen abhandeln (1998) oder Kongressberichte enthalten (1999, 2001, 2003).

Auswertung aller Studien (1950-2004)

Studieninhalte

Über die Hälfte der Studien (58 %) behandel- ten physiologische Fragestellungen (z. B. „Auswirkungen eines Höhentrainings“, „Ver- gleich einer Trainings- mit einer Kontroll- gruppe“), 18 % hatten internistische (z. B. „Metipranolol bei Schützen“), 9 % diagnosti- sche (z. B. „Vergleich verschiedener Laktat- messmethoden“) und 15 % fächerübergreifen- de Inhalte (z. B. „Eiweiß und Krafttraining“). Dabei erfolgten die meisten Studien (59 %) vor einem trainingswissenschaftlichen Hinter- grund; etwa 29 % der Studienergebnisse

Tabelle 3: Dargestellt sind die CONSORT-Kriterien sowie die Anzahl der Studien, die das jeweilige CONSORT-Kriterium erfüllten. Bewertet wurden die zwischen 1999 und 2004 publizierten Originalartikel

CONSORT-Kriterien		Anzahl der Studien (n=21)	
Titel und Abstract	1	Zuordnung zu Therapiegruppen aus Titel und Zusammenfassung erkennlich?	6 28,6%
	Einführung	2	Wissenschaftlicher Hintergrund
Methoden		3	Einschlusskriterien der Probanden/Patienten
	4	Präzise Angaben zu den geplanten Interventionen	21 100,0%
	5	Genauere Ziele, Fragestellung und Hypothesen	20 95,2%
	6	Klar definierte Zielkriterien	18 85,7%
	7	Fallzahlbestimmung	1 4,8%
	8	Erzeugung der Behandlungsfolge	2 9,5%
	9	Geheimhaltung der Behandlungsfolge	0 0,0%
	10	Zuteilungsverfahren in die Gruppen	0 0,0%
	11a	Verblindung der Probanden/Patienten	2 9,5%
	11b	Verblindung derjenigen, die die Intervention durchführten	1 4,8%
	11c	Verblindung derjenigen, die die Zielgröße beurteilten	0 0,0%
	12	Statistische Methoden	20 95,2%
Ergebnisse	13a	Anzahl der Probanden, die die Behandlung erhalten haben	2 9,5%
	13b	Anzahl der Probanden, die die Studie protokollgemäß beendeten	3 14,3%
	13c	Anzahl der Probanden, die in der Analyse des primären Zielkriteriums berücksichtigt wurden	9 42,9%
	14	Nähere Angaben über den Zeitraum der Studienaufnahme/ Nachbeobachtung	3 14,3%
	15	Demographische und klinische Charakteristika aller Gruppen	14 66,7%
	16	Anzahl der bei der Auswertung berücksichtigten Patienten (intention-to-treat-Analyse?)	8 38,1%
	17	Ergebnisse mit Effektgröße und Präzision	19 90,5%
Diskussion	18	Zusätzliche Analysen	5 23,8%
	19	Angabe aller Nebenwirkungen jeder Behandlungsgruppe	1 4,8%
	20	Interpretation der Ergebnisse	19 90,5%
	21	Generalisierbarkeit der Studienergebnisse (externe Validität)	12 57,1%
	22	Vereinbarkeit der Ergebnisse mit der aktuellen Studienlage	20 95,2%

waren von klinischer Relevanz. Allerdings war diese Einteilung schwer objektivierbar, da sich bei einigen Fragestellungen (12 %) Medizin und Trainingswissenschaft überschneiden (z. B. „Sport und Menopause“) (Tab. 2).

Studienmethodik

Von den erfassten Originalarbeiten und Kurzberichten wurden 38 % als RCT und 62 % als CCT klassifiziert. Die durchschnittliche Probandenzahl belief sich auf 35 Studienpatienten. In 45 % der Studien waren weniger als 20, in 16 % der Studien mehr als 50 Teilnehmer eingeschlossen. 48 % der Studien wurden im Parallel-Design, 37 % im Cross-over-Design durchgeführt. In 2 % der Studien waren beide Designs kombiniert, die restlichen 13 % enthielten keine Angaben bezüglich des Studiendesigns. In 43 % der Studien wurde die Interventionsgruppe mit einer unbehandelten Kontrollgruppe verglichen, in 13 % mit einer Placebogruppe. Eine Verblindung der Probanden wurde, zum Teil aufgrund der Fragestellung (z. B. „Training vs. kein Training“), bei rund 76 % der Stu-

dien nicht durchgeführt. Eine eindeutige Verblindung der Studienteilnehmer kam in 13 % der Fälle zur Anwendung. Die Zielparame- ter (Studienendpunkte) waren bei 81 % der Studien objektivierbar – vereinzelt (5 %) aber auch rein subjektiv (z. B. „Schmerz“ oder „Wohlbe- finden“).

Auswertung in den einzelnen Studien

Die Studien wurden auch hinsichtlich ihrer Analyse- methode überprüft. Bei der Intention- to-treat-Analyse werden die Ergebnisse der jeweiligen Patienten in derjenigen Gruppe analysiert, der sie zu Studienbeginn zugeteilt wurden, wohingegen bei einer per-protocol- Analyse nur die Personen in die Analyse ein- geschlossen werden, die protokollgemäß be- handelt wurden (17). In 11 % der kontrollierten Studien war die Anzahl der Studienteilnehmer nach erfolgter Zuteilung zu den Interventionsgruppen identisch mit der Anzahl der ausgewerteten Teilnehmer. Wie die Auswertung der Studien erfolgte - nach in- tention-to-treat oder per-protocol - wurde in keinem Fall explizit erwähnt. In 10 % der Pub- likationen wurde angegeben, dass nicht alle eingeschlossenen Patienten ausgewertet wurden. In den meisten Arbeiten (79 % aller gefundenen Studien; davon 98 % bei den Abstracts und 45 % bei den Originalartikeln) waren die Angaben zu ein- und ausgeschlos- senen Teilnehmern im Verlauf der Studie (zum Zeitpunkt der Aufnahme, der Zuteilung, in der Nachbeobachtung sowie bei Auswertung) un- vollständig. Keine der Studien enthielt ein Flussdiagramm, aus dem dies klar hervorgeht. Angaben zur statistischen Analyse der Ergeb-

nisse fehlten in 45 % der Studien (25 % der Originalartikel, 56 % der Abstracts). Ein statistisch signifikanter Behandlungsunterschied wurde in 40 % der Studien gefun- den, in 15 % der Fälle wurde kein statistisch signifikanter Unterschied beobachtet.

Vollständigkeit der Studienbeschreibung

Um Ergebnisse aus Studien qualitativ bewerten zu könn- en, muss die Berichterstattung über Design, Durch- führung und Analyse vollständig und transparent sein. Zur Bewertung der Vollständigkeit der Studienbeschrei- bung wurden relevante Merkmale der identifizierten Stu- dienberichte evaluiert (Tab. 2):

- Studiengröße: Bei fast allen Studien (95 %) war die Anzahl der Studienteilnehmer angegeben.
- Studiendesign: Bei 87 % der Publikationen war angegeben, ob es sich um eine Parallel- oder Crossover-Studie handelte.
- Studienplanung: Unvollständig berichtet wurde hingegen, wie die Studie geplant wurde und inwieweit die Planung ein- gehalten werden konnte, so dass bei manchen Studien nicht

ausgeschlossen werden kann, dass der Vergleich gegen historische Kontrollen erfolgte. In 95 % der Studien wurde nicht berichtet, ob es sich um eine prospektive Studie handelte.

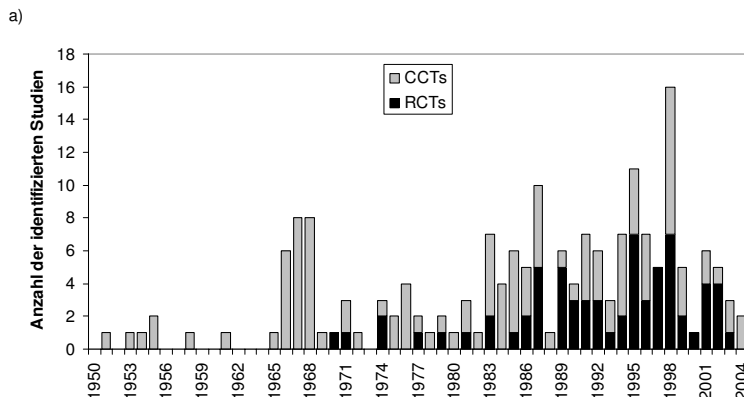
-Demografische Angaben zum Patientenkollektiv: In 59 % der Originalarbeiten waren aussagekräftige demografische und klinische Charakteristika der Probanden zu Studienbeginn aufgeführt. In 15 % der Studien waren die Angaben ungenau (z. B. „Sportstudenten zwischen 19 und 23 Jahren“, „Marathonläufer“), und in 26 % wurde auf eine Beschreibung des Untersuchungskollektivs gänzlich verzichtet.

-Randomisierung: Die Methoden der Erzeugung der Behandlungsfolge sowie deren Geheimhaltung (allocation concealment) waren in keiner der Studien beschrieben (16). Auswertung der Studien auf Übereinstimmung mit dem CONSORT-Statement (1999 bis 2004). Bei 21 der zwischen 1999 und 2004 in der Deutschen Zeitschrift für Sportmedizin publizierten Originalarbeiten handelte es sich um kontrollierte Studien (RCT oder CCT). Die Evaluation der Studienberichte nach den CONSORT-Richtlinien zeigte deutliche Unterschiede hinsichtlich der Vollständigkeit und der Qualität der angegebenen Kriterien (Tab. 3). Wissenschaftlicher Hintergrund, Interventionen, Studienziele und Fragestellung, statistische Methoden sowie allgemeine Interpretation der Studienergeb-

richterstattung: Ein Vergleich zwischen älteren, „vor-CONSORT“ Studien (1950-1998) und neueren, „nach-CONSORT“ Studien (1999-2005) zeigte keine Verbesserung, die sich auf das CONSORT-Statement zurückführen ließe. Einzig die Zuteilung zu den Interventionsgruppen wurde bei neueren Studien öfter (42 %) randomisiert durchgeführt als bei älteren (29 %).

Diskussion

Durch die manuelle Durchsichtung der Deutschen Zeitschrift für Sportmedizin wurde eine Anzahl kontrollierter klinischer Studien gefunden, die überraschend hoch erscheinen mag. Die 490 kontrollierten klinischen Studien (37 % davon randomisiert) sind nicht in Medline enthalten und können deshalb auf dem klassischen Weg durch eine systematische Literatursuche nicht gefunden werden. Durch die Publikation dieser Studienzitate in der Cochrane Library ist ein elektronischer Hinweis auf diese Studien geschaffen worden. Damit sind die Studien aus ihrer faktischen Unerreichbarkeit geholt und global sichtbar gemacht worden. Erst dadurch wird die Nutzung des durch sie erzeugten Wissens möglich, z. B. für die Therapiebewertung in systematischen Übersichtsarbeiten. Mit dieser Arbeit wurden die Studien aus der Sportmedizin den über 17.000 Studien hinzugefügt, die in deutschsprachigen Zeitschriften aus den verschiedensten medizinischen Fachbereichen identifiziert und als Teil der globalen Wissensbasis sichtbar gemacht wurden. Neben dem Auffinden der Studien ist die ausreichende Beschreibung von Studiendesign, -durchführung und -ergebnissen eine notwendige Voraussetzung für die effiziente Nutzung von Studien. Oft fehlen wesentliche Teile der Beschreibung in Publikationen, was die Nutzung der Studienergebnisse erheblich einschränkt. Ob und in welchem Umfang dies auch für die Deutsche Zeitschrift für Sportmedizin gilt, wurde deswegen in der vorliegenden Studie untersucht. Grundlage der Bewertung klinischer Studienberichte ist das CONSORT-Statement, das entwickelt wurde, um die Qualität der Berichterstattung über Studien zu verbessern. Die vorliegende Studie zeigt, dass die Studien bezüglich wesentlicher Eigenschaften oft unvollständig beschrieben waren. Gerade die eindeutige Beschreibung des Studiendesigns ist Grundvoraussetzung für die adäquate, prinzipielle Einstufung einer jeden Studie. Darüber hinaus gibt es einige Punkte der Durchführung und Auswertung, deren Berücksichtigung von besonderer Bedeutung für die Zuverlässigkeit der Aussagen ist. Die Art der Zuweisung der Probanden zur Interventions- oder Kontrollgruppe, die möglichst weitgehende Verblindung aller Beteiligten, eine ausreichende Studiengröße und der angemessene Umgang mit Studienabbrechern sowie die fachgerechte Auswertung sind einige Aspekte, deren Relevanz heute unbestritten ist und die deswegen in eindeutiger Weise beschrieben werden müssen. Die Einhaltung dieser methodologischen Richtmarken bei der Studienplanung, -durchführung und -be-



b)

Jahrgang	Anzahl RCTs	Anzahl CCTs
1966	0	6
1982	4	15
1984	0	15
1997	0	2
1999	30	47
2000	1	2
2001	44	44
2003	40	60
2004	1	0
Summe	120	191
Gesamt		311

Abbildung 1: Anzahl der in den Jahrgängen 1950 bis 2004 identifizierten kontrollierten Studien: a) Originalarbeiten (CCTs grau, RCTs schwarz), b) Abstracts

nisse waren in den meisten Artikeln beschrieben. Angaben zur Fallzahlbestimmung, Randomisierungsmethode, Verblindung der Probanden, zu Studienbetreuern und Studienauswertern sowie zu Interventions-assoziierten unerwünschten Wirkungen waren hingegen die Ausnahme. Von den insgesamt 26 CONSORT-Kriterien und -Unterkriterien waren im Mittel 11,7 pro Studie beschrieben. Keine der Studien beschrieb weniger als neun oder mehr als 18 Kriterien. Auswirkung des CONSORT-Statements auf die Qualität von Studien und deren Be-

wertung: Ein Vergleich zwischen älteren, „vor-CONSORT“ Studien (1950-1998) und neueren, „nach-CONSORT“ Studien (1999-2005) zeigte keine Verbesserung, die sich auf das CONSORT-Statement zurückführen ließe. Einzig die Zuteilung zu den Interventionsgruppen wurde bei neueren Studien öfter (42 %) randomisiert durchgeführt als bei älteren (29 %).

Durch die manuelle Durchsichtung der Deutschen Zeitschrift für Sportmedizin wurde eine Anzahl kontrollierter klinischer Studien gefunden, die überraschend hoch erscheinen mag. Die 490 kontrollierten klinischen Studien (37 % davon randomisiert) sind nicht in Medline enthalten und können deshalb auf dem klassischen Weg durch eine systematische Literatursuche nicht gefunden werden. Durch die Publikation dieser Studienzitate in der Cochrane Library ist ein elektronischer Hinweis auf diese Studien geschaffen worden. Damit sind die Studien aus ihrer faktischen Unerreichbarkeit geholt und global sichtbar gemacht worden. Erst dadurch wird die Nutzung des durch sie erzeugten Wissens möglich, z. B. für die Therapiebewertung in systematischen Übersichtsarbeiten. Mit dieser Arbeit wurden die Studien aus der Sportmedizin den über 17.000 Studien hinzugefügt, die in deutschsprachigen Zeitschriften aus den verschiedensten medizinischen Fachbereichen identifiziert und als Teil der globalen Wissensbasis sichtbar gemacht wurden. Neben dem Auffinden der Studien ist die ausreichende Beschreibung von Studiendesign, -durchführung und -ergebnissen eine notwendige Voraussetzung für die effiziente Nutzung von Studien. Oft fehlen wesentliche Teile der Beschreibung in Publikationen, was die Nutzung der Studienergebnisse erheblich einschränkt. Ob und in welchem Umfang dies auch für die Deutsche Zeitschrift für Sportmedizin gilt, wurde deswegen in der vorliegenden Studie untersucht. Grundlage der Bewertung klinischer Studienberichte ist das CONSORT-Statement, das entwickelt wurde, um die Qualität der Berichterstattung über Studien zu verbessern. Die vorliegende Studie zeigt, dass die Studien bezüglich wesentlicher Eigenschaften oft unvollständig beschrieben waren. Gerade die eindeutige Beschreibung des Studiendesigns ist Grundvoraussetzung für die adäquate, prinzipielle Einstufung einer jeden Studie. Darüber hinaus gibt es einige Punkte der Durchführung und Auswertung, deren Berücksichtigung von besonderer Bedeutung für die Zuverlässigkeit der Aussagen ist. Die Art der Zuweisung der Probanden zur Interventions- oder Kontrollgruppe, die möglichst weitgehende Verblindung aller Beteiligten, eine ausreichende Studiengröße und der angemessene Umgang mit Studienabbrechern sowie die fachgerechte Auswertung sind einige Aspekte, deren Relevanz heute unbestritten ist und die deswegen in eindeutiger Weise beschrieben werden müssen. Die Einhaltung dieser methodologischen Richtmarken bei der Studienplanung, -durchführung und -be-

richterstattung sind Voraussetzung für eine möglichst unverzerrte Einschätzung des untersuchten Behandlungseffektes.

Es überrascht nicht, dass die vom CONSORT-Statement geforderte Vollständigkeit der Studienberichte bei weitem nicht erreicht wurde und in vielen Artikeln selbst grundlegende Aspekte fehlten. Die Beschreibungen ließen in vielen Fällen nicht einmal den sicheren Schluss zu, ob die Kontrollgruppe aus retrospektiv gewonnenen Daten oder aus einer parallel geführten Gruppe stammte. In 95 % der Studien wurde der Umfang zwar angegeben, über eine konkrete Planung des Studienumfangs mit der Überlegung, mit welcher Sicherheit eine vorhandene Überlegenheit der Intervention auch tatsächlich gefunden werden kann (Power-Betrachtung), wurde in keinem der Fälle berichtet. Zu den Schwächen klinischer Studien, die aktuell besondere Beachtung finden, gehört die Wahl des primären Zielkriteriums. Oft entspricht die Darstellung in der Publikation nach Abschluss der Studie nicht dem in der Studienplanung genannten Kriterium (4). Die nachträgliche Auswahl des Zielkriteriums führt zu einem schädlichen Überoptimismus. Dieser Aspekt wird jedoch erwartungsgemäß in keinem Artikel angesprochen.

Bei aller Kritik muss allerdings bedacht werden, dass das CONSORT-Statement relativ jung ist und dass ältere Studien nach neuen Kriterien beurteilt werden, die zur Zeit der Studiendurchführung noch nicht in der Form explizit formuliert waren bzw. noch nicht bis in die Praxis durchgedrungen waren. Für die Verbesserung der Qualität der Studiendurchführung und Berichterstattung ist eine deutlichere Wahrnehmung des CONSORT-Statements bei Zeitschriften, Herausgebern und Autoren auch im deutschsprachigen Raum notwendig. Die Publikation der kürzlich ins Deutsche übersetzten Version in medizinischen Zeitschriften sollte dazu beitragen (11). Warum die Umsetzung so schleppend erfolgt, beruht auf einer Reihe von Gründen, die jedoch in Deutschland bisher nicht systematisch untersucht wurden. Ein Grund ist vermutlich die Abneigung gegen eine weitere gängige Checkliste. Dabei wird übersehen, dass das CONSORT-Statement eine sehr effektive Unterstützung bietet, mit der die Qualität von Artikeln wesentlich verbessert wird und damit für die Autoren auch die Chance einer zügigen Publikation.

Neben den dargestellten Zielen des Handsearchings und den damit verbundenen methodischen Aspekten gibt es ein inhaltliches Nebenprodukt, das bei Herausgebern und Verlagen großes Interesse hervorruft. Die Klassifizierung der kontrollierten Studien nach Fachgebieten liefert ein Bild des publizierten Studiengeschehens der letzten Jahrzehnte, wie es in dieser Form vorher nicht bekannt war. Mit der laufenden Aufarbeitung des Studiengeschehens in den Zeitschriften verschiedener Fachgebiete wird eine empirische Basis aufgebaut, die der Dauerdiskussion um zu wenige und qualitativ schlechte Studien in Deutschland eine rationalere Basis gibt. Die vielen gefundenen Mängel bei der Darstellung bestätigen die gängige Kritik, die überraschend hohen Zahlen aufgefunder Studien zeigen aber auch, dass es ein sehr viel aktiveres Studiengeschehen in Deutschland gibt als vielfach behauptet. Arbeiten wie diese sollten Forschern, Sponsoren, Ethikkommissionen und Zeitschriftenherausgebern die Fakten und die Motivation liefern,

die vorhandene Quantität mit mehr Qualität aufzuwerten. In wie weit sich das bei der schwierigen Lage deutscher Zeitschriften realisieren lässt, werden die nächsten Jahre zeigen (15).

Literatur

1. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gøtzsche PC, Lang T, for the CONSORT group: The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of Internal Medicine* 134 (2001) 663-694.
2. Bassler D, Antes G, Egger M: Non-English reports of medical research. *JAMA* 284 (2000) 1417-24.
3. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz K, Simel D, Stroup DF: Improving the quality of reporting of randomized controlled trials. The Consort Statement. *JAMA* 276 (1996) 637-639.
4. Chan AW: Empirical Evidence for Selective Reporting of Outcomes in Randomized Trials. Comparison of Protocols to Published Articles. *JAMA* 291 (2004) 2457-2465.
5. Dickersin K, Manheimer E, Wieland S, Robinson KA, Lefebvre C, McDonald S and the CENTRAL Development Group: Development of the Cochrane Collaboration's Central Register of Controlled Clinical Trials. *Eval Health Prof* 25 (2002) 38-64.
6. Dickersin K, Scherer R, Lefebvre C: Identifying relevant studies for systematic reviews. *BMJ* 309 (1994) 1286-91.
7. Friedman LM, Furberg CD, DeMets DL: *Fundamentals of Clinical Trials*. Springer Science + Business Media Inc., 1998, Third Edition.
8. Galandi D, Schwarzer G, Antes G: The demise of the randomised controlled trial: bibliometric study of the German-language health care literature, 1948 to 2004. *BMC Medical Research Methodology* 6 (2006) 30.
9. <http://www.cochrane.de/de/handsearch.htm>
10. <http://www.thecochranelibrary.com>
11. <http://www.thieme-connect.de/ejournals/html/dmw/doi/10.1055/s-2004-836117>.
12. McDonald S, Lefebvre C, Antes G, Galandi D, Gøtzsche P, Hammarquist C, Haugh M, Jensen, KL, Kleijnen J, Loep M, Pistotti V, Ruther A: The contribution of handsearching European general health care journals to the Cochrane Controlled Trials Register. *Eval Health Prof* 25 (2002) 65-75.
13. Moher D, Fortin P, Jadad AR, Juni P, Klassen T, Le Lorier J, Liberati A, Linde K, Penna A: Completeness of reporting of trials published in languages other than English: implications for conduct and reporting of systematic reviews. *Lancet* 347 (1996) 363-6.
14. Nieminen P, Isohanni M: Bias against European journals in medical publication databases. *Lancet* 353 (1999) 1592.
15. Ofori-Adjei D, Antes G, Tharyan P, Slade E, Tamber PS: Have online international medical journals made local journals obsolete? *PLoS Me* 3 (2006) 8 - 359.
16. Schultz KF, Grimes DA: Allocation concealment in randomised trials: defending against deciphering. *Lancet* 359 (2002) 614-618.
17. Schulgen G, Schumacher M: Intention-to-treat Analyse, in: Schumacher M, Schulgen G (Hrsg): *Methodik klinischer Studien*. Methodische Grundlagen der Planung, Durchführung und Auswertung. Springer Verlag, Heidelberg, 2002, 147-155.
18. Turp JC, Schulte JM, Antes G: Nearly half of dental randomized controlled trials published in German are not included in Medline. *Eur J Oral Sci* 110 (2002) 405-411.
19. Wieland S, Dickersin K; Manheimer K: Learning to identify and classify reports of controlled trials in healthcare journals. New England Cochrane Center, Providence Office, USA. First presented at the Cochrane Colloquium, Cape Town, October 2000. <http://www.cochrane.us/srch.htm>.

Korrespondenzadresse:
Dr. Anette Blümle
Deutsches Cochrane Zentrum
Institut für Medizinische Biometrie
und Medizinische Informatik
Universitätsklinikum Freiburg
Stefan-Meier-Str. 26
79104 Freiburg
e-Mail: bluemle@cochrane.de