



„Reporting von systematischen Reviews“

Der Fakultät für Medizin und Gesundheitswissenschaften der Carl von Ossietzky Universität
Oldenburg zur Erlangung des Grades und Titels einer

Doktorin der Gesundheitswissenschaften (Dr. rer. medic.)

vorgelegte Dissertation von

Frau Jasmin Helbach

geboren am 28.01.1996 in Dernbach

Erstbetreuer:

Prof. Dr. Falk Hoffmann

Zweitbetreuer:

Prof. Dr. Dawid Pieper

Gutachter*Innen:

Prof. Dr. Rebecca Palm

Prof. Dr. Hajo Zeeb

Tag der Disputation: 21.10.2025

Danksagung

An dieser Stelle möchte ich mich bei den Menschen bedanken, die mich in den letzten Jahren bei der Erstellung meiner Dissertation begleitet und unterstützt haben.

Mein besonderer Dank gilt Falk Hoffmann, der mich auf meinem Weg in die Wissenschaft begleitet hat und mich von Beginn an zum kritischen Denken angeregt hat. Seine hilfreichen Tipps und Anregungen haben mich dazu gebracht, meine eigenen Kompetenzen zu erweitern und wissenschaftlich zu wachsen. Ich danke ihm für seine Unterstützung und die Betreuung dieser Dissertation.

Ebenfalls möchte ich Dawid Pieper danken, der aus der Ferne an all meinen Arbeiten mitgewirkt und stets wertvolle Anregungen gegeben hat.

Ein weiteres großes Dankeschön geht an meine Kolleginnen und Kollegen. Sie haben mich bei der Durchführung meiner Studien tatkräftig unterstützt, mir bei Fragen und Sorgen ein offenes Ohr geboten und sich mit mir über Erfolge gefreut.

Abschließend möchte ich mich bei meiner Familie und meinen Freunden bedanken. Sie haben mich unterstützt, motiviert und aufgefangen, wenn mir alles zu viel wurde. Danke, dass ihr mir immer wieder gesagt habt, dass ihr stolz auf mich seid.

Inhaltsverzeichnis

Abkürzungsverzeichnis	v
Tabellenverzeichnis	vi
Abbildungsverzeichnis	vii
Zusammenfassung.....	viii
Summary.....	ix
Vorwort.....	x
1. Hinführung.....	1
1.1 Evidenzbasierte Medizin.....	1
1.2 Von der Meta-Analyse zu systematischen Reviews	2
1.2.1 Geschichte der Meta-Analyse.....	4
1.2.2 Geschichte des systematischen Reviews.....	5
1.2.3 Traditionelle Reviews	6
1.2.4 Cochrane Reviews.....	7
1.3 Reporting-Leitlinien für systematische Reviews.....	9
1.3.1 QUOROM	9
1.3.2 PRISMA	10
1.4 Aktuelle Entwicklung von systematischen Reviews	10
1.5 Relevanz von Abstracts und die Entwicklung von PRISMA-A.....	12
2. Kurzdarstellung und Ergebnisse der eingeschlossenen Publikationen	14
2.1 Erste Publikation: PRISMA-A Adhärenz in systematischen Reviews	14
2.2 Zweite Publikation: PRISMA-A Adhärenz in Cochrane Reviews bzw. Einfluss der Länge.....	16
2.3 Dritte Publikation: Einfluss der Länge auf die Attraktivität des Abstracts	18
3. Diskussion	20
3.1 Messen der PRISMA-A Adhärenz.....	20
3.2 Anwenden der PRISMA-A Leitlinie	21
3.3 PRISMA-A Adhärenz in Cochrane Reviews	22
3.4 Die Länge von Abstracts von Cochrane Reviews	25
3.5 Die optimale Länge von Abstracts systematischer Reviews.....	26
4. Fazit.....	27
Referenzen.....	29
Anhang.....	39
Anhang 1. Erklärung	40
Anhang 2. Eingeschlossene Einzelarbeiten	41

Abkürzungsverzeichnis

ANOVA	Analysis of Variance (Varianzanalyse)
CDSR	Cochrane Database of Systematic Reviews
CONSORT	Consolidated Standards of Reporting Trials
CONSORT-A	Consolidated Standards of Reporting Trials for Abstracts
GRADE	Grading of Recommendations, Assessment, Development and Evaluation
IQR	Interquartile Range (Interquartilsabstand)
KI	Konfidenzintervall
MECIR	Methodological Expectations of Cochrane Intervention Reviews
NHS	National Health Service (Staatlicher Gesundheitsdienst)
NIH	National Institut of Health (Nationale Gesundheitsinstitute)
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-A	Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Abstracts
PRISMA-P	Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocols
QUOROM	Quality of Reporting of Meta-Analyses
RCT	Randomized controlled trial (Randomisiert kontrollierte Studie)
SD	Standard Deviation (Standardabweichung)
WHO	World Health Organization (Weltgesundheitsorganisation)

Tabellenverzeichnis

Tabelle 1. Evidenzstufen nach David Sackett (1986)	1
Tabelle 2. Die wichtigsten Grundsätze der Cochrane Collaboration	8
Tabelle 3. PRISMA-A Checkliste (2013)	13
Tabelle 4. PRISMA-A Adhärenz in Cochrane vs. Nicht-Cochrane Reviews.....	14

Abbildungsverzeichnis

Abbildung 1. Evidenzpyramide	2
Abbildung 2. Entwicklung von Meta-Analysen und systematischen Reviews im Zeitverlauf.....	3
Abbildung 3. Geschätzte Anzahl der in MEDLINE (über PubMed) indexierten systematischen Reviews pro Jahr (mit 95% KI), 2000-2019 aus Hoffmann et al. 2021.....	11
Abbildung 4 Durchschnittliche Anzahl vollständig berichteter PRISMA-A Items nach Anzahl der Wörter (ohne Cochrane Reviews)	15
Abbildung 5. Flussdiagramm zur zweiten Publikation	16
Abbildung 6. Entwicklung der Länge von Cochrane Review Abstracts von 2000-2022.....	17
Abbildung 7. Flussdiagramm zur dritten Publikation.....	19

Zusammenfassung

Die Zahl der publizierten systematischen Reviews nimmt kontinuierlich zu. Dadurch steigt die Relevanz gut geschriebener Abstracts, denn sie bieten einen schnellen Überblick über das methodische Vorgehen und die wichtigsten Forschungsergebnisse. Zudem hängt die Entscheidung, einen Artikel im Volltext zu lesen, maßgeblich von den im Titel und Abstract präsentierten Informationen ab. Um die Qualität von Abstracts systematischer Reviews zu verbessern, wurde 2013 die PRISMA-Checkliste für Abstracts (PRISMA-A) entwickelt und 2020 aktualisiert. Bisher gibt es jedoch kaum Daten dazu, ob ihre Einführung zu einer Verbesserung der Berichtsqualität von Abstracts geführt hat. Ziel dieser Dissertation ist es, diese Forschungslücke weiter zu schließen und zu untersuchen, inwieweit die PRISMA-A Leitlinie eingehalten wird. Dabei fokussiert die Arbeit vor allem die Unterschiede zwischen Cochrane Reviews und Nicht-Cochrane Reviews sowie dem Einfluss der Abstract-Länge auf die Berichterstattung und Wahrnehmung von Abstracts.

Die vorliegende Dissertation basiert auf drei Einzelstudien. **In der ersten Studie** wurden 534 systematische Reviews zur Effektivität untersucht und anhand der PRISMA-A Checkliste bewertet. Die Studie zeigt, dass die Abstracts im Durchschnitt 5,4 der 12 PRISMA-A Items vollständig berichten, wobei die PRISMA-A Adhärenz zwischen den Items stark variiert (0% bis 98,8%). Abstracts von Cochrane Reviews sind im Durchschnitt etwas besser berichtet als die von Nicht-Cochrane Reviews (6,3 vs. 5,2 Items). Eine stratifizierte Analyse von 435 Nicht-Cochrane Reviews nach Anzahl der Wörter zeigt eine positive Korrelation zwischen der PRISMA-A Adhärenz und der Abstract-Länge ($r = 0,44$, $p < 0,001$). **Die zweite Studie** konzentriert sich ausschließlich auf Cochrane-Reviews. Eine Vollerhebung aller in PubMed indexierten Cochrane-Reviews ($n = 15.188$) zeigte, dass die Anzahl neu publizierter Cochrane-Reviews seit 2013 kontinuierlich zurückgeht. Die Abstracts waren im Median 469 Wörter lang (IQR 389–686 Wörter). Dabei stieg die durchschnittliche Wortanzahl von 353 im Jahr 2000 auf über 800 im Jahr 2022 (30,7% vs. 0,7% ≤ 300 Wörter). Analysen einer zufälligen Stichprobe von 399 Cochrane Reviews Abstracts stratifiziert nach Jahr zeigten, dass die PRISMA-A Adhärenz in den Jahren 2000–2010 und 2011–2015 mit höheren Wortzahlen zunahm, während sich in den Jahren 2016–2022 keine Unterschiede in der PRISMA-A Adhärenz nach Abstract-Länge zeigten. **In der dritten Arbeit**, einer randomisiert-kontrollierten Studie wurden insgesamt 5.397 Forschende im Verhältnis 1:1 zufällig einem kurzen (277 Wörter) oder langen Abstract (771 Wörter) desselben systematischen Reviews, das in zwei verschiedenen Zeitschriften veröffentlicht wurde, zugeordnet. Die Studienteilnahme unterschied sich nicht zwischen den Gruppen (37,8 % vs. 35,0 %, $p = 0,1935$). Die Länge des Abstracts hatte keinen Einfluss auf die Wahrnehmung des Informationsgehalts (85,5% vs. 91,2%; $p=0,0594$) und der Genauigkeit (80,3% vs. 86,3%; $p=0,0868$). Der kürzere Abstract wurde jedoch als attraktiver (60,5% vs. 46,6%; $p=0,0034$) und prägnanter (82,3% vs. 37,9%; $p<0,0001$) bewertet. Drei Viertel der Teilnehmenden befürwortete eine maximale Abstract-Länge von 250 bis 300 Wörtern und etwa die Hälfte befürwortete die Berichterstattung von Informationen zur Finanzierung und Registrierung im Abstract.

Die meisten Abstracts von systematischen Reviews sind nicht vollständig nach PRISMA-A berichtet. Um informative Abstracts zu schreiben, benötigen Autor*innen systematischer Reviews einen gewissen Rahmen an Wörtern. Dabei scheinen 250 bis 300 Wörter ausreichend zu sein. Reporting-Leitlinien wie PRISMA-A sollten diese Länge entsprechend empfehlen, und Journals sollten sie unabhängig vom Studiendesign in ihre Autorenrichtlinien aufnehmen. Dies gilt auch für Cochrane, dessen aktuelle Längenempfehlungen deutlich höher sind.

Summary

The number of published systematic reviews is steadily increasing. This makes abstracts increasingly important, as they provide a quick overview of the methodological approach and the main research findings. Researchers often decide whether to read the full article based on the information presented in the title and abstract. To improve the reporting quality of abstract, PRISMA-A was developed in 2013 and updated in 2020. However, there is a lack of data on the impact of its implementation on the quality of abstract reporting. This dissertation aims to address this research gap by investigating the extent to which PRISMA-A has been adhered to, focusing on differences between Cochrane and non-Cochrane review abstracts, and the impact of abstract length on reporting quality and how abstracts are perceived.

This dissertation is based on three studies. **In the first study**, 534 systematic reviews on effectiveness were examined and evaluated using the PRISMA-A checklist. The mean score of fully reported PRISMA-A items was 5.4 of 12, with adherence varying widely between items (0% to 98.8%). Abstracts of Cochrane reviews received higher mean total scores than non-Cochrane reviews (6.3 vs. 5.2 Items). Stratified analysis of the 453 non-Cochrane reviews by number of words revealed a positive correlation of moderate strength between PRISMA-A adherence and abstract length ($r = 0.44$, $p < 0.001$). **The second study** focuses exclusively on Cochrane reviews. A comprehensive analysis of all Cochrane reviews ($n=15,188$) indexed in PubMed revealed a continuous decrease in the number of Cochrane reviews since 2013. The median word count of the abstracts was 469 (IQR 389–686 words), steadily increasing from 353 words in 2000 to 838 words in 2022 (30.7% vs. $0.7\% \leq 300$ words). Analyses on PRISMA-A adherence in a random sample of 399 Cochrane review abstracts showed a mean score of 6.1 out of 12 fully reported items. Stratified by year, PRISMA-A adherence increased, with higher word counts in 2000–2010 and 2011–2015, while there was no difference in PRISMA-A adherence by abstract length in 2016–2022. **The third paper** was a randomised controlled trial in which 5,397 researchers were randomly assigned at a 1:1 ratio to either a short (277 words) or a long (771 words) abstract of the same systematic review published in two different journals. Trial participation did not differ between groups (37.8% vs. 35.0%; $p=0.1935$). While the short abstract was considered more attractive (60.5% vs. 46.6%; $p=0.0034$) and concise (82.3% vs. 37.9%; $p<0.0001$), the length had no impact on its informativeness (85.5% vs. 91.2%; $p=0.0594$) and accuracy (80.2% vs. 86.3%; $p=0.0868$). Three-quarters of participants supported a maximum abstract length of 250–300 words and approximately half supported reporting information on funding and registration in the abstract.

The majority of abstracts of systematic reviews do not fully adhere to the PRISMA-A guideline. While authors of systematic reviews need a certain number of words to provide informative abstracts, 250–300 words seemed to be sufficient. It is therefore recommended that reporting guidelines such as PRISMA-A include this word count, and that journals incorporate it into their author guidelines, irrespective of the study design. This recommendation also applies to Cochrane, whose current word count recommendations are significantly higher.

Vorwort

Immer wieder stößt man in der täglichen Arbeit auf schlecht berichtete Reviews. Es ist daher besonders wichtig, zu diesem Thema zu forschen, zu publizieren und die eigenen Ergebnisse im aktiven Austausch mit anderen Wissenschaftler*innen zu diskutieren. Mit dem Begriff des systematischen Reviews bin ich zum ersten Mal während meines Bachelorstudiums in Berührung gekommen. Zu diesem Zeitpunkt war der Begriff jedoch noch sehr schwer greifbar und die Tragweite und Bedeutung hat sich mir noch nicht vollständig erschlossen. Erst während meines Masterstudiums und meiner Tätigkeit als studentische Hilfskraft, bei der ich an meiner ersten wissenschaftlichen Publikation über Trends in der Epidemiologie und Berichterstattung von systematischen Reviews¹ mitgearbeitet habe, wurde mir die Relevanz von Reviews vollständig bewusst. Aus den in der Trendstudie gesammelten Daten entstand nicht nur meine Masterarbeit über das Reporting von Restriktionen in systematischen Reviews,² sondern auch meine Leidenschaft für die Methodenforschung.

Systematische Reviews verfolgen das Ziel, die gesamte verfügbare Evidenz zu einer spezifischen Forschungsfrage systematisch zu identifizieren, zu synthetisieren und kritisch zu bewerten.^{3,4} Damit stellen sie einen besonderen Mehrwert für Kliniker*innen, Forscher*innen und politische Entscheidungsträger*innen dar, denn sie bieten einen prägnanten Überblick über den aktuellen Stand der Forschung.⁵⁻⁷ In den letzten Jahrzehnten ist die Anzahl an publizierten systematischen Reviews stark angestiegen.^{1,8-10} Während im Jahr 2000 insgesamt 1.432 systematische Reviews in PubMed indexiert waren, sind es im Jahr 2019 über 25.000, was einer Anzahl von 80 neu publizierten systematischen Reviews pro Tag entspricht.¹ In einigen Forschungsbereichen übersteigt die Anzahl der publizierten systematischen Reviews sogar die Anzahl der neu publizierten Primärstudien.⁹ Dies birgt die Gefahr von "research waste".^{9,11}

Der Wert eines systematischen Reviews bemisst sich an dem methodischen Vorgehen, den gewonnenen Erkenntnissen und der Klarheit der Berichterstattung.^{7,12} Studien zeigen, dass die methodische Qualität von systematischen Reviews oft unzureichend ist und insbesondere hinsichtlich ihrer Berichterstattung deutliches Verbesserungspotenzial aufweist.^{10,13} Aus diesem Grund wurden Berichtsstandards wie die Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Leitlinie¹⁴ entwickelt, die Autor*innen dabei unterstützen sollen, die wichtigsten Aspekte des methodischen Vorgehens und der Ergebnisse ihrer Arbeit angemessen zu berichten. Aufgrund der hohen Anzahl an publizierten systematischen Reviews steigt die Relevanz von gut geschriebenen Abstracts. Denn basierend auf den Informationen, die im Titel und Abstract berichtet werden, entscheiden die Editor*innen, ob die Arbeit für das Journal relevant ist, die Reviewer*innen, ob sie die Begutachtung übernehmen und die Forscher*innen, ob sie die Publikation im Volltext lesen.^{15,16} Aus diesem Grund wurde 2013 eine Erweiterung der PRISMA Leitlinie für Abstracts¹⁷ entwickelt. Allerdings gibt es bislang kaum Studien die untersuchen, ob dies zur Verbesserung der Berichterstattung von Abstracts geführt hat.

In dieser kumulativen Dissertation habe ich mir daher zum Ziel gesetzt, diese Forschungslücke weiter zu schließen und zu untersuchen, inwieweit die PRISMA-A Leitlinie von Autor*innen systematischer Reviews berücksichtigt wird. Darüber hinaus habe ich versucht Faktoren zu identifizieren, die einen Einfluss auf die Berichtsqualität von Abstracts nehmen und Implikationen für die Forschung abzuleiten. Dabei unterscheide ich zwischen Cochrane Reviews und Nicht-Cochrane Reviews und fokussiere mich insbesondere auf den Einfluss der Länge der Abstracts.

Folgende Einzelarbeiten sind Bestandteil dieser Dissertation:

1. Helbach J, Hoffmann F, Pieper D, Allers K. Reporting according to the preferred reporting items for systematic reviews and meta-analyses for abstracts (PRISMA-A) depends on abstract length. *J Clin Epidemiol.* 2022;154:167-177. doi:10.1016/j.jclinepi.2022.12.019
2. Helbach J, Wandscher K, Pieper D, Hoffmann F. Cochrane review abstracts are getting longer, but this has no large impact on the reporting quality. *J Evid Based Med.* 2023;16(3):294-302. doi:10.1111/jebm.12545
3. Helbach J, Wandscher K, Pieper D, Hoffmann F. Readers' attention to shorter versus longer abstracts of systematic reviews: a randomised controlled trial. *BMJ Evid Based Med.* 2025 Oct 17:bmjebm-2024-113613. doi: 10.1136/bmjebm-2024-113613.

Darüber hinaus wird im Rahmen dieser Dissertation auf folgende weitere Arbeiten Bezug genommen, an denen ich als Koautorin beteiligt war:

4. Hoffmann F, Allers K, Rombey T, Helbach J, Hoffmann A, Mathes T, Pieper D. Nearly 80 systematic reviews were published each day: Observational study on trends in epidemiology and reporting over the years 2000-2019. *J Clin Epidemiol.* 2021;138:1-11. doi:10.1016/j.jclinepi.2021.05.022
5. Wandscher K, Helbach J, Pieper D, Hoffmann F. "We used standard Cochrane methods" – observational study on reporting according to PRISMA-A in Cochrane review abstracts between 2016 and 2023. *J Clin Epidemiol.* 2025;181(111713). doi: 10.1016/j.jclinepi.2025.111713

1. Hinführung

1.1 Evidenzbasierte Medizin

Die evidenzbasierte Medizin hat sich in den frühen 1990er Jahren zunehmend etabliert.^{18–20} Ihr Ziel ist die gewissenhafte, explizite und umsichtige Nutzung der derzeit besten wissenschaftlichen Erkenntnisse bei Entscheidungen über die Behandlung einzelner Patient*innen.^{3,4} Maßgeblich für diese Bewegung war David Sackett, ein Pionier und Vater der evidenzbasierten Medizin.²¹ Die Praxis der evidenzbasierten Medizin stützt sich auf drei Säulen und kombiniert die individuelle klinische Expertise von medizinischen Fachkräften, die Werte und Wünsche der Patient*innen mit den aktuell besten verfügbaren externen Erkenntnissen aus der systematischen Forschung.^{3,4,22} Durch die Kombination dieser verschiedenen Perspektiven können Diagnosen effektiver und effizienter gestellt sowie Probleme, Rechte und Präferenzen einzelner Patient*innen im Rahmen klinischer Entscheidungen besser identifiziert und berücksichtigt werden.²³

Der Begriff "beste verfügbare klinische Evidenz" umfasst sowohl die medizinische Grundlagenforschung als auch die patientenorientierte klinische Forschung.³ Letztere fokussiert insbesondere die Wirksamkeit und Sicherheit therapeutischer, rehabilitativer und präventiver Maßnahmen, die Genauigkeit und Präzision diagnostischer Verfahren sowie die Aussagekraft prognostischer Marker.³ Durch die Kombination verschiedener methodischer Verfahren werden neue klinische Daten generiert, die vorhandene wissenschaftliche Evidenz gebündelt und für Kliniker*innen, Patient*innen und politische Entscheidungsträger*innen verständlich aufbereitet.^{4,19} Dabei wird den verschiedenen methodischen Ansätzen eine unterschiedliche Aussagekraft zugeschrieben.^{18,24,25}

Die erste Beschreibung von Evidenzhierarchien wurde 1979 von der kanadischen Task Force on the Periodic Health Examination veröffentlicht.²⁶ Ziel der Task Force war es, evidenzbasierte Empfehlungen für regelmäßige Gesundheitsuntersuchungen zu geben. Zu diesem Zweck entwickelte sie ein System zur Bewertung der Qualität der verwendeten Evidenz, welches bei der Einstufung der Effektivität der betrachteten Maßnahmen und den gezogenen Empfehlungen berücksichtigt werden sollte. Seither wurde das Hierarchiemodell stetig weiterentwickelt^{27,28} und auf die unterschiedlichen Fragestellungen und Herausforderungen diverser Fachgebiete angepasst.²⁵ Die Grundlage für die heutige Forschung und Lehre bildet die von David Sackett in einem Artikel über die Evidenzstufen von Antithrombotika beschriebene fünf-stufige Evidenzhierarchie aus dem Jahr 1986 (Tabelle 1).²⁹

Tabelle 1. Evidenzstufen nach David Sackett (1986)

Level	Typen der Evidenz
I	Große RCTs mit eindeutigen Ergebnissen
II	Kleine RCTs mit unklaren Ergebnissen
III	Kohorten- und Fall-Kontroll-Studien
IV	Historische Kohorten- oder Fall-Kontroll-Studien
V	Fallserien, Studien ohne Kontrollen

RCT= Randomisiert kontrollierte Studie

Um den Ansatz des hierarchischen Denkens zu vereinfachen und grafisch zu veranschaulichen, wurde im Laufe der Zeit das Bild der sogenannten Evidenzpyramide weiterentwickelt (Abbildung 1). Dies illustriert die

Qualität und Quantität der verfügbaren Evidenz in einer aufeinander aufbauender Struktur, bei der die Quantität mit zunehmender Hierarchieebene abnimmt, während die Qualität und klinische Relevanz zunimmt.^{27,30} Eine isolierte Betrachtung der Spitze der Pyramide ohne Bezug zu ihrem Körper und ihrer Basis ist nicht möglich und führt zu missverständlichen und falschen Schlussfolgerungen.^{28,30}

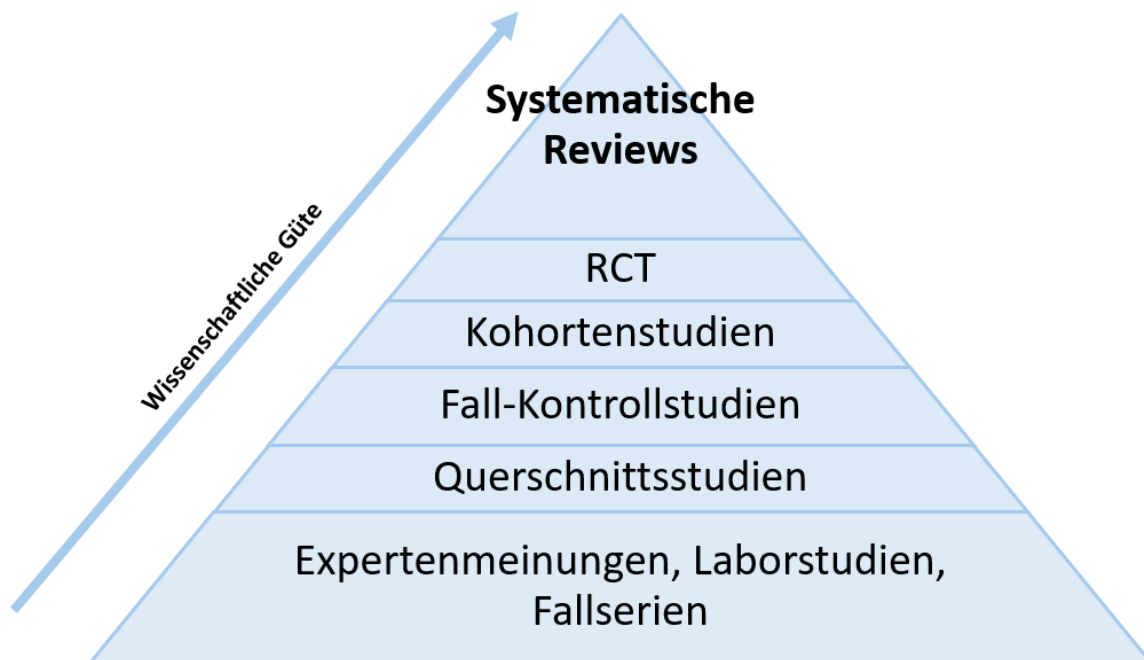


Abbildung 1. Evidenzpyramide

Die evidenzbasierte Medizin stützt sich auf eine breite Basis an Grundlagenforschung, Expertenmeinungen und Fallserien.³¹ Auf dieser untersten Hierarchieebene werden Ideen und Daten generiert, die eine Rechtfertigung für weiterführende, umfassendere und methodisch anspruchsvollere Studien darstellen. Den Körper der Pyramide bilden Beobachtungsstudien, zu denen beispielsweise Fall-Kontroll- und Kohortenstudien zählen. Ihr Ziel besteht in der Erfassung von Trends und Mustern durch Beobachtung natürlicher Verhaltensweisen oder Bedingungen.^{32,33} Randomisierte kontrollierte Studien (RCTs) repräsentieren den höchsten wissenschaftlichen Standard in der Primärforschung. Ihr experimentelles Studiendesign erlaubt die gesicherte Ableitung kausaler Zusammenhänge.^{31–33} Bis in die 2000er Jahre standen RCTs noch an der Spitze der Evidenzhierarchie. Doch mit der zunehmenden Etablierung von Meta-Analysen und systematischen Reviews, die alle verfügbare Evidenz zu einer Forschungsfrage systematisch erfassen, analysieren und für Kliniker*innen, Patient*innen und politische Entscheidungsträger*innen verständlich aufbereiten,^{7,34,35} wurde die Evidenzpyramide überarbeitet und an die heutige Form angepasst.

1.2 Von der Meta-Analyse zu systematischen Reviews

Die Bemühungen, wissenschaftliche Erkenntnisse statistisch zusammenzufassen, Verzerrungen zu reduzieren und Ergebnisse gebündelt zu präsentieren, haben eine lange Geschichte, die sich bis ins 17. Jahrhundert zurückverfolgen lässt.³⁶ Dabei erfolgt die Entwicklung der Meta-Analyse als statistische Methode zur Kombination von unabhängigen Forschungsergebnissen zeitlich vor der Entwicklung des systematischen Reviews (Abbildung 2).³⁷

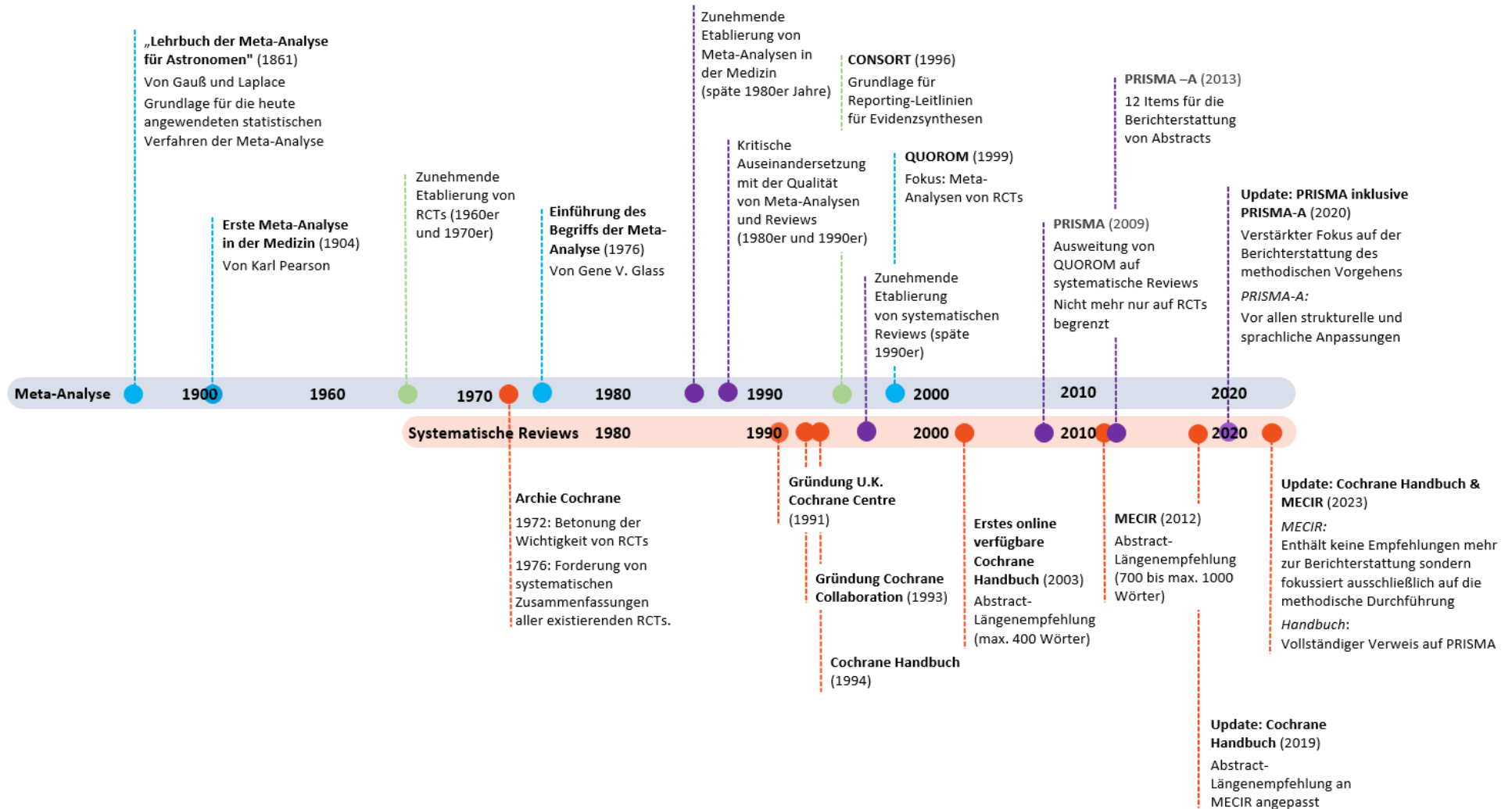


Abbildung 2. Entwicklung von Meta-Analysen und systematischen Reviews im Zeitverlauf

1.2.1 Geschichte der Meta-Analyse

Die Idee, unabhängige Beobachtungen statistisch zu kombinieren, geht auf den französischen Mathematiker Blaise Pascal zurück, der im 17. Jahrhundert mathematische Modelle zu den Gewinnchancen in Glücksspielen entwickelte.^{38,39} Später fanden diese Modelle Anwendung in der Astronomie, wo sie zum Vergleich und zur Kombination von Beobachtungen verschiedener Astronomen genutzt wurden.^{37,38} Die heute in der Meta-Analyse angewandten statistischen Verfahren basieren im Wesentlichen auf den Arbeiten von Gauß und Laplace,³⁶ die erstmals 1861 im „Lehrbuch der Meta-Analyse für Astronomen“ von Airy's⁴⁰ beschrieben wurden. Es dauerte jedoch bis ins 20. Jahrhundert, bis diese Methoden auch ihren Weg in die Medizin fanden und zur Zusammenfassung klinischer Studien eingesetzt wurden.^{36–38}

Die erste so genannte Meta-Analyse in der Medizin wurde 1904 von dem britischen Statistiker Karl Pearson durchgeführt.⁴¹ Pearson analysierte die Wirksamkeit eines Impfstoffs gegen Typhus, indem er die Ergebnisse von elf einschlägigen Studien statistisch zusammenfasste. In den 1920er und 1930er Jahren entwickelte Ronald Fisher die Methoden zur statistischen Synthese von Forschungsergebnissen weiter, und führte das Maximum-Likelihood-Prinzip und die Varianzanalyse (ANOVA) ein.⁴² Das statistische Interesse an Forschungssynthesen hielt bis zum Zweiten Weltkrieg an, und auf der Grundlage von Fishers Arbeiten wurden bis in die 1960er Jahre mehr als ein Dutzend Arbeiten zur Forschungssynthese und Meta-Analyse veröffentlicht,^{37,38} darunter z.B. die Arbeiten von Cochran,⁴³ und Jones und Fiske.⁴⁴

Die Anwendung von Meta-Analysen war zu diesem Zeitpunkt jedoch noch sehr selten und erfuhr erst in den darauffolgenden Jahren einen Bedeutungszuwachs,⁴⁵ was insbesondere auf den starken Anstieg an Forschungsartikeln in den Sozialwissenschaften und der Psychologie zurückzuführen ist.^{36,37} Die Forschenden sahen sich mit der Herausforderung konfrontiert, die wachsende Fülle an Forschungsergebnissen zu organisieren und zu bündeln. Unabhängig voneinander lösten unterschiedliche Forschungsteams diese Herausforderung durch die Anwendung der von Pearson und Fisher beschriebenen statistischen Ansätze zur Kombination von Forschungsergebnissen.³⁷ Ein bedeutender Vertreter dieser Zeit war Gene V. Glass, ein amerikanischer Statistiker und Erziehungspsychologe, der sich intensiv mit der Wirksamkeit von Psychotherapien sowie dem Einfluss der Klassengröße auf die schulischen Leistungen von Kindern und Jugendlichen befasste.^{37,38,46} Im Rahmen der Diskussionen über statistische Herausforderungen⁴⁷ etablierte er den Begriff der Meta-Analyse, den er als "Analyse von Analysen" definierte. Sein Verständnis dessen lässt sich wie folgt zusammenfassen:

"[...] die statistische Analyse einer großen Sammlung von Analyseergebnissen aus Einzelstudien mit dem Ziel, deren Ergebnisse zu integrieren". (übersetzt aus Glass G. 1976; Seite 3)

In der medizinischen Forschung dauerte es noch bis in die späten 1980er Jahre bis sich die Anwendung von Meta-Analysen etablierte.⁴⁵ Besonders beeinflusst wurde diese Entwicklung durch die zunehmende Anzahl an publizierten RCTs⁴⁸ und die Arbeit von Peter Elwood, Archie Cochrane und Kollegen. Das Forschungsteam führte in den 70er Jahren eine randomisierte kontrollierte Studie zum Potential von Acetylsalicylsäure (u.a. Aspirin®) zur sekundären Prävention der Mortalität durch Myokardinfarkt durch.⁴⁹ Ihre Ergebnisse deuteten zwar auf einen positiven Effekt hin, waren aber statistisch nicht signifikant. Als weitere Studien zum Potenzial von Aspirin publiziert wurden, führten Elwood und Cochrane eine Meta-Analyse durch um die Ergebnisse der

einzelnen Studien statistisch zusammenzubringen und zeigten, dass Aspirin tatsächlich einen signifikanten Effekt auf die Mortalität bei Herz-Kreislaufkrankungen hat.⁵⁰

Die Veröffentlichung einiger methodischer Schlüsselwerke, wie beispielsweise „Statistical Methods for Meta-Analysis“ von Hedges und Olkin (1985)⁴⁶ oder die Publikation von Richard Peto von 1987,⁵¹ trug dazu bei, dass die Praxis der quantitativen Forschungssynthese in der medizinischen Forschung zunehmend Anerkennung fand. Heute gilt sie als wichtigste Methode zur Synthese wissenschaftlicher Ergebnisse.

1.2.2 Geschichte des systematischen Reviews

Die signifikante Zunahme der wissenschaftlichen Forschung in der Nachkriegszeit^{52,53} sowie die zunehmende Etablierung von RCTs⁴⁸ führten zu der Erkenntnis, dass die Synthese von Forschungsergebnissen ein weitaus komplexeres Spektrum von Fragen aufwirft als nur die Wahl der statistischen Synthesemethoden.³⁷ Dabei waren Forschende im Bereich der Medizin im Vergleich zu anderen Fachgebieten wie beispielsweise der Psychologie oder den Sozialwissenschaften auf dem Gebiet der Forschungssynthese relative Nachzügler.^{36,37} Obwohl die ersten kritischen Zusammenfassungen von RCTs bereits in den 1960er und 1970er Jahren erstellt wurden, dauerte es bis in die 1990er Jahre, bis Forschungssynthesen im medizinischen Kontext in größerer Zahl erschienen.^{37,54}

Die Etablierung systematischer Reviews wurde durch die Kritik an der klinischen Praxis ausgelöst, wissenschaftliche Erkenntnisse zu vernachlässigen und medizinische Entscheidungen auf unbegründeten Annahmen zu treffen.⁵⁵ Ein besonders großer Kritiker seiner Zeit und zugleich Vater der systematischen Reviews war Archibald (Archie) Cochrane.⁵⁶⁻⁵⁹ Nach dem erfolgreichen Abschluss seines Studiums der Naturwissenschaften begann er im Jahr 1930 mit einem Medizinstudium.^{56,59} Im Jahr 1941 zog Cochrane als Sanitätsoffizier des Royal Army Medical Corps in den Zweiten Weltkrieg und geriet kurze Zeit später in Kriegsgefangenschaft.^{56,59} In dieser Zeit führte er seine erste klinische Studie durch, die er viele Jahre später in seiner Arbeit "Sickness in Salonika: My first, worst and most successful clinical trial"⁶⁰ beschrieb.

Während seiner Zeit im Kriegsgefangenenlager war Cochrane für die medizinische Versorgung vieler Mitgefangener verantwortlich, die er nach eigenen Angaben stets nach bestem Wissen und Gewissen behandelte. In einem späteren Bericht, in dem er diese Zeit reflektierte, beschrieb er jedoch die eigene Angst, durch sein fehlendes Wissen über die beste verfügbare Behandlungsweise Schaden anzurichten und seine Freunde unnötig zu quälen:⁵⁵

„Ich erinnere mich, dass ich damals eine dieser Propagandaschriften über ‚klinische Freiheit und Demokratie‘ las, die für medizinische Offiziere in Kriegsgefangenschaft gedacht waren. Ich konnte es nicht verstehen. Ich hatte eine beträchtliche Freiheit bei der Wahl der Therapie: mein Problem war, dass ich nicht wusste, welche ich wann anwenden sollte. Ich hätte meine Freiheit gerne für ein wenig Wissen geopfert. Ich hatte noch nie etwas von 'randomisierten kontrollierten Studien' gehört, aber ich wusste, dass es keinen wirklichen Beweis dafür gab, dass irgendetwas, was wir anboten, irgendeine Wirkung auf die Tuberkulose hatte, und ich hatte Angst, dass ich das Leben einiger meiner Freunde durch unnötige Eingriffe verkürzte.“ (übersetzt aus Cochrane 1972, Ausschnitt zu finden unter: <https://community.cochrane.org/archie-cochrane-name-behind-cochrane>)

Nach dem Krieg begann Cochrane eine Ausbildung in der präventiven Medizin an der London School of Hygiene and Tropical Medicine.^{23,56,59,61} Dort traf er auf Austin Bradford Hill, einen Pionier der RCTs,⁶² der ihn nachhaltig prägte und sein Bestreben stärkte, die Evidenzbasierung medizinischer Maßnahmen zu optimieren. In seinem 1972 veröffentlichten Buch "Effectiveness and Efficiency: Random Reflections on Health Services"⁵⁵ erörtert Cochrane die Notwendigkeit randomisierter kontrollierter Studien zur Bestimmung einer angemessenen Behandlung. In seiner späteren wissenschaftlichen Laufbahn erkannte er, dass die Ergebnisse von RCTs zu identischen Forschungsfragen oft widersprüchlich sind und forderte die systematische Zusammenfassung aller existierenden RCTs.⁶³

„Es ist sicherlich ein großer Kritikpunkt an unserem Berufsstand, dass wir keine kritische Zusammenfassung aller relevanten randomisierten kontrollierten Studien nach Fachgebieten oder Teilgebieten organisiert haben, die in regelmäßigen Abständen angepasst wird.“ (übersetzt aus Cochrane 1979, Seite 9)

Mit dieser Aussage setzte Archie Cochrane den Grundstein des systematischen Reviews, der bis heute als wichtigster Bestandteil und Goldstandard in der evidenzbasierten Medizin betrachtet wird.^{8,14}

1.2.3 Traditionelle Reviews

Mit der zunehmenden Etablierung von RCTs in den 1960er und 1970er Jahren begannen Forschende die verfügbaren RCTs kritisch zusammenzufassen, um die Wirksamkeit von Behandlungen und Therapien zu bewerten.^{8,37} Traditionell verfolgten Reviews dabei jedoch einen narrativen Ansatz. Das bedeutet, dass die Entscheidungsprozesse, welche Studien integriert werden, welche Daten relevant sind und welche Schlussfolgerungen gezogen werden, kein systematisches Vorgehen aufweisen.^{7,36,64,65} Sie sind vielmehr maßgeblich durch die Erfahrungen und die Expertise der beteiligten Forscher*innen bestimmt.^{36,64,65} Damit geben narrative Reviews zwar einen Einblick in die Thematik, eignen sich jedoch nicht dazu, den aktuellen Stand der Wissenschaft in Bezug auf das Problem vollständig zu erfassen und abzubilden.^{66,65}

In den 1980er Jahren wurden Stimmen laut, die auf die unzureichende Qualität von Reviews hinwiesen und eine kritische Bewertung dieser Arbeiten forderten.^{66–70} Sozialwissenschaftler in den Vereinigten Staaten waren in dieser Hinsicht wegweisend.³⁷ So kritisierte der amerikanische Sozialwissenschaftler David Pillemer (1984) beispielsweise den herkömmlich sehr subjektiven Ansatz von Reviews.⁷¹ Dabei beschreibt er das entstehende Verzerrungspotential durch idiosynkratische Einschätzungen zu Schlüsselfragen, wie beispielsweise der Auswahl der einzubeziehenden Studien und der Interpretation der allgemeinen Schlussfolgerungen. Darüber hinaus kritisiert er, dass die Studien einzeln betrachtet, und Stärken und Schwächen selektiv identifiziert und beiläufig diskutiert werden.⁷¹ Aufbauend auf den Prinzipien und Publikationen aus der Sozialwissenschaft, begannen Gesundheitswissenschaftler wie Oxman und Guyatt (1988) Richtlinien zur Überprüfung der wissenschaftlichen Qualität von Reviews aus dem Gesundheitswesen zu publizieren.⁷²

Ein bedeutendes Werk im Bereich der evidenzbasierten Medizin stellt das Seminarbuch zum "Umgang mit biomedizinischer Literatur" von Kenneth Warren dar.⁷³ Es fungierte als Impulsgeber für die kritische Evaluation und legte den Fokus auf eine sorgfältige Analyse und Bewertung von Publikationen. In diesem Werk betont Edward Kass die Notwendigkeit einer kritischen Evaluierung von Reviews, die in ihrer Methodik und Aussagekraft mit primären Publikationen vergleichbar sein sollten.⁶⁹ Cynthia Mulrow leitete diesen

Evaluationsprozess mit einem wegweisenden Artikel ein,⁷⁰ der 1987 in den „Annals of Internal Medicine“ veröffentlicht wurde und die methodische Qualität von 50 Reviews aus vier großen medizinischen Journals evaluierte. Die Ergebnisse zeigten, dass der Großteil der betrachteten Reviews keine wissenschaftlichen Methoden zur Identifizierung, Bewertung und Zusammenfassung von Informationen verwendet haben. Im selben Jahr wurden weitere einflussreiche Artikel publiziert,^{51,74–76} die einheitlich zu dem Schluss kamen, dass viele systematische Reviews methodische Mängel aufweisen.

Eine Herausforderung bei der Bewertung der methodischen Qualität von Reviews bestand darin, dass verschiedene Begriffe wie beispielsweise "Research Synthesis", "Systematic Review", "Overview" und "Meta-Analyses" verwendet wurden, um den Prozess der wissenschaftlichen Synthese von Forschungsergebnissen zu beschreiben.³⁷ Dabei fehlte lange Zeit eine klare Abgrenzung und die Begriffe wurden teilweise synonym verwendet, obwohl unterschiedliche methodische Ansätze verfolgt wurden.³⁷ So wurde der Begriff "Systematic Review" beispielsweise bereits Ende der 1930er Jahre verwendet, beschrieb damals jedoch noch nicht den stark strukturierten Prozess, den wir heute darunter verstehen.³⁷ Erst Anfang der 2000er Jahre wurden einheitliche Definitionen etabliert, die eine klare Trennung zwischen der Meta-Analyse als statistischem Syntheseverfahren, dem narrativen Review als unsystematische Übersichtsarbeit und dem systematischen Review als strukturiertem Verfahren zur Erfassung und Synthese aller verfügbaren Evidenz ermöglichen^{7,77,78} In den letzten 25 Jahren haben sich verschiedene Definitionen herausgebildet, die dieser Unterscheidung jedoch im Kern treu bleiben.

Die in diese Dissertation einbezogenen Publikationen orientieren sich an der Definition des systematischen Reviews von PRISMA in seiner Erweiterung für die Erstellung von Protokollen (PRISMA-P) aus dem Jahr 2015.⁷⁹ Darin wird ein systematischer Review als ein klar strukturierter Forschungsprozess beschrieben, der darauf abzielt, relevante Studien zu identifizieren, deren Ergebnisse und Qualität kritisch zu bewerten und die wichtigsten Erkenntnisse systematisch zusammenzutragen. Ein systematischer Review im Sinne der PRISMA-P Definition muss drei wesentliche Kriterien erfüllen, um als solcher eingestuft zu werden.⁷⁹ Die Arbeit muss

1. eine klare Fragestellung präsentieren
2. eine systematische Suche durchführen, die darauf abzielt, alle relevante Evidenz zu identifizieren, die den Einschlusskriterien entspricht und das methodische Vorgehen in Bezug auf die Suchstrategie und Studienselektion offenlegen
3. die wichtigsten Charakteristika und Ergebnisse der identifizierten Studien systematisch, z.B. in einer Meta-Analyse oder einer narrativen Synthese, darstellen und zusammenfassen.

1.2.4 Cochrane Reviews

Ein weiterer Meilenstein in der Etablierung von systematischen Reviews war die Arbeit rund um Iain Chalmers in den späten 1980er und frühen 1990er Jahren. Die Erkenntnis, dass systematische Reviews über die Wirksamkeit von Therapieansätzen fehlten, dass Empfehlungen in Lehrbüchern von denen systematischer Reviews abwichen und evidenzbasierte Ansätze lange Zeit unzureichend in die klinische Praxis implementiert wurden,⁸⁰ führte zu einer gesteigerten Aufmerksamkeit der evidenzbasierten Medizin in England.

Im Jahr 1991 reagierte der britische National Health Service (NHS) auf die zunehmende Bedeutung der evidenzbasierten Medizin und die interne Kritik, zu wenig an der Durchführung und Implementierung von

klinischer Forschung beteiligt zu sein, mit der Einrichtung eines Forschungs- und Entwicklungsprogramms.⁸¹ Dieses umfasst die Gründung von zwei Forschungszentren dem NHS Centre for Reviews and Dissemination und dem U.K. Cochrane Centre.^{81,82} Letzteres hat sich zum Ziel gesetzt, wie von Archie Cochrane zwölf Jahre zuvor gefordert, alle existierenden RCTs im Gesundheitsbereich in systematischen Reviews zusammenzufassen.^{63,83,84} Das Cochrane Centre gewann schnell an Popularität und nach einem ersten internationalen Treffen von 77 Forscher*innen unterschiedlicher Nationen auf dem Cochrane Colloquium in Oxford wurde 1993 unter der Leitung von Iain Chalmers die Cochrane Collaboration gegründet.^{85,86}

Tabelle 2. Die wichtigsten Grundsätze der Cochrane Collaboration

No.	Kriterium	Beschreibung
1	Zusammenarbeit	Durch die Förderung von globaler Zusammenarbeit, Teamwork, offener und transparenter Kommunikation und Entscheidungsfindung
2	Auf dem Enthusiasmus der Einzelnen aufbauen	Durch die Beteiligung, Unterstützung und Schulung von Menschen mit unterschiedlichen Fähigkeiten und Hintergründen
3	Vermeidung von Duplikaten	Durch gutes Management, Koordinierung und wirksame interne Kommunikation, um die Wirtschaftlichkeit der Arbeit zu maximieren
4	Minimierung von Verzerrungen	Durch eine Vielzahl von Ansätzen wie wissenschaftliche Strenge, einer breiten Beteiligung und Vermeidung von Interessenkonflikten
5	Auf dem neuesten Stand bleiben	Durch die Verpflichtung, sicherzustellen, dass Cochrane Reviews durch die Identifizierung und Einbeziehung neuer Erkenntnisse aufrechterhalten werden
6	Streben nach Relevanz	Durch die Förderung der Bewertung von Fragen, die für die Menschen, die Entscheidungen über Gesundheit und Gesundheitsversorgung treffen, von Bedeutung sind
7	Förderung des Zugangs	Durch eine weite Verbreitung unserer Ergebnisse, durch die Nutzung strategischer Allianzen und durch die Förderung geeigneter Zugangsmodelle und Bereitstellungslösungen, um den Bedürfnissen der Nutzer weltweit gerecht zu werden
8	Sicherstellung der Qualität	Durch die Anwendung von Fortschritten in der Methodik, die Entwicklung von Systemen zur Qualitätsverbesserung und die Offenheit und das Eingehen auf Kritik
9	Kontinuität	Durch die Sicherstellung, dass die Verantwortung für Überprüfungen, redaktionelle Prozesse und Schlüsselfunktionen beibehalten und erneuert wird
10	Ermöglichung einer breiten Beteiligung	Durch den Abbau von Hindernissen für die Teilnahme und die Förderung der Vielfalt

**Übersetzt aus dem Cochrane Handbuch, verfügbar unter: <https://training.cochrane.org/handbook/current/chapter-i>*

Die heute meist nur als "Cochrane" bezeichnete Organisation ist ein globales Netzwerk, mit dem Ziel die evidenzbasierte Entscheidungsfindung in der Medizin zu fördern. Zu diesem Zweck werden qualitativ hochwertige systematische Reviews zu den Auswirkungen von Interventionen im Gesundheitswesen erstellt,

gepflegt und deren Zugänglichkeit gefördert.⁸⁵ Dabei bildet ihre Arbeit die Grundlage für viele Leitlinien und Strategien, darunter auch solche, die von der Weltgesundheitsorganisation (WHO) erstellt werden.⁸⁷

Cochrane etablierte in kurzer Zeit eine Reihe von Richtlinien und Vorgehensweisen, die das Erstellen und Nutzen von systematischen Reviews standardisieren und die methodische Qualität steigern sowie redundante Forschung bzw. den überflüssigen Verbrauch von Ressourcen reduzieren sollte.⁸⁵ Dabei stützt sich Cochrane auf zehn Grundprinzipien (Tabelle 2) und gründete eine Datenbank (Cochrane Database of Systematic Reviews; CDSR), in der alle von Cochrane publizierten Reviews (nachfolgend als Cochrane Reviews bezeichnet) gesammelt und zugänglich gemacht werden sollten.^{86,88} Mit dem Ziel, hochwertige systematische Reviews zu erstellen, definierte Cochrane von Beginn an klare methodische und strukturelle Anforderungen. Diese entwickelten sich im Zeitverlauf sukzessiv weiter und wurden 1994 erstmals im Cochrane Handbuch dokumentiert.⁸⁹ Im Jahr 2012 wurde zusätzlich zum Handbuch die „Methodological Expectations of Cochrane Intervention Reviews“ (MECIR)⁹⁰ Standards publiziert. Hierbei handelt es sich um methodische Standards, die die Durchführung von Cochrane Reviews leiten sollen. Neben den methodischen Vorgehensweisen definierte MECIR bis 2019 zusätzlich Berichtsstandards, in denen klar festgelegt wurde, welche Informationen in den Reviews berichtet sein müssen.

1.3 Reporting-Leitlinien für systematische Reviews

Im Zuge der kritischen Auseinandersetzung mit der methodischen Güte von Reviews in den 1980er Jahren wurde festgestellt, dass die unzureichende Qualität der systematischen Reviews nicht immer auf eine fehlerhafte methodische Durchführung, sondern häufig auf eine unzureichende Berichterstattung der eingeschlossenen Primärstudien oder der Reviews selbst zurückzuführen war.^{91,92} Oft waren die methodischen Vorgehensweisen nur sehr vage beschrieben, was die Einordnung der Ergebnisse und die Bewertung der methodischen Güte erschwerte.^{13,93,94}

Die im Jahr 1996 publizierte „Consolidated Standards of Reporting Trials“ (CONSORT) Leitlinie⁹⁵ war Ausgangspunkt für die Entwicklung und Etablierung von Reporting-Leitlinien für systematische Reviews. CONSORT ist ein Leitfaden für die Berichterstattung von klinischen Studien und definiert ein Mindestmaß an Informationen, die in RCTs berichtet werden sollten. Die Leitlinie besteht aus einer Checkliste mit 22 Items und einem Flussdiagramm, welches den Weg der Studienteilnehmer*innen von der Rekrutierung bis zur abschließenden Analyse veranschaulicht.⁹⁵ CONSORT ist die erste Reporting-Leitlinie, die über die einzelnen Forschungsgebiete hinaus akzeptiert wurde und legte den Grundstein für die transparente Berichterstattung in RCTs sowie systematischen Reviews.

1.3.1 QUOROM

Im selben Jahr, in dem die CONSORT-Leitlinie veröffentlicht wurde, organisierte das National Institut of Health (NIH) eine Konferenz zur „Quality of Reporting of Meta-Analyses“ (QUOROM).^{96,97} Ziel der Konferenz war es, Expert*innen aus verschiedenen Bereichen zusammenzubringen, um über Möglichkeiten zur Verbesserung der methodischen Qualität von Meta-Analysen zu diskutieren. Der Fokus der Diskussionen lag dabei ausschließlich auf Meta-Analysen von RCTs, andere Studientypen wurden nicht einbezogen. Die Konferenz mündete in der Gründung einer Expertengruppe, darunter David Moher, Kenneth F. Schulz und Douglas G. Altman, die damit beauftragt wurde, eine Reporting-Leitlinie zur Verbesserung der Berichtsqualität von Meta-Analysen zu entwickeln. In Anlehnung an CONSORT für RCTs wurde eine Checkliste

erarbeitet, die sich aus 21 Items zusammensetzt und in fünf Domänen gegliedert ist: Abstract, Einleitung, Methoden, Ergebnisse und Diskussion.⁹⁶ Dabei basierten die entwickelten Items größtenteils auf den Erfahrungen und Empfehlungen der Expertengruppe und waren nur in Teilen wissenschaftlich fundiert. Die Items zielten darauf ab, Fehlerquellen und Verzerrungen von Ergebnissen zu vermeiden und die Reproduzierbarkeit der Meta-Analysen zu gewährleisten.⁹⁶ Darüber hinaus wurde ein Flussdiagramm entwickelt, das den Prozess der Studienidentifizierung veranschaulicht und Informationen über die Anzahl der identifizierten, eingeschlossenen und ausgeschlossenen RCTs sowie Ausschlussgründe beinhaltet.⁹⁶ Nach mehrfacher Testung und Überarbeitung der Checkliste wurde die QUORUM-Leitlinie 1999 veröffentlicht.

1.3.2 PRISMA

Im Jahr 2005 wurde ein erneutes QUORUM-Meeting einberufen, dessen Ziel darin bestand, die Leitlinie zu evaluieren und zu aktualisieren.¹⁴ Eine wesentliche Änderung, die bei diesem Treffen beschlossen wurde, war die Umbenennung von QUORUM in PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). Begründet wurde dieser Schritt mit der sich im Laufe der Zeit verändernden Terminologie von systematischen Reviews und Meta-Analysen sowie dem Bestreben, beide in den Empfehlungen zu berücksichtigen.¹⁴ Es wurden mehrere inhaltliche Überarbeitungen vorgenommen, wobei jedes Item auf Relevanz geprüft wurde. Darüber hinaus wurden neue Items zum Protokoll und zur Einschätzung des Verzerrungsrisikos in den eingeschlossenen Studien erarbeitet, und ergänzende wünschenswerte Items definiert, die von den Autor*innen der systematischen Reviews ausgefüllt werden sollten, sofern sie für die durchgeführte Arbeit relevant erschienen. Alle in PRISMA gegebenen Empfehlungen sind evidenzgestützt und wissenschaftlich begründet. Die finale Version der PRISMA-Leitlinie wurde 2009 veröffentlicht und umfasst eine Checkliste mit 27 Items aus sieben Domänen und eine aktualisierte Version des Flussdiagramms.¹⁴ Ergänzend wurde ein Begleitdokument erstellt und veröffentlicht, in dem alle Items ausführlich beschrieben und anhand von Beispielen erläutert werden.⁹⁸

Im Jahr 2020 wurde eine erneute Überarbeitung und Aktualisierung der PRISMA-Leitlinie vorgenommen.⁹⁹ Mit dem Ziel, vor allem Transparenz und Reproduzierbarkeit zu verbessern, legt das Update vermehrten Fokus auf die Berichterstattung des methodischen Vorgehens. Dabei berücksichtigt sie die methodischen Entwicklungen in den Bereichen der Identifizierung, Auswahl, Bewertung und Zusammenfassung und modifiziert die Struktur bzw. Darstellung der PRISMA Items, um die Umsetzung zu erleichtern.⁹⁹

1.4 Aktuelle Entwicklung von systematischen Reviews

In den vorangegangenen Kapiteln wurde die historische Entwicklung systematischer Reviews umfassend dargestellt. Auch in der Gegenwart ist ein stetiger Wandel in der Methodik, der Epidemiologie und der Berichterstattung systematischer Reviews zu beobachten.^{1,9,100} Bestehende Evidenz hat gezeigt, dass es immer mehr veröffentlichte systematische Reviews gibt, die von schlechter Berichtsqualität sind, irreführende Schlussfolgerungen enthalten oder sich überschneiden und damit zu "research waste" führen.^{9,11,101} Studien zeigen, dass die Anzahl der publizierten systematischen Reviews in den letzten Jahrzehnten stark angestiegen ist,^{9,10,100,102} wobei robuste Daten zum Ausmaß des Anstiegs und zu Veränderungen in den Charakteristika von systematischen Reviews fehlen. Mit dem Ziel, Trends in der Epidemiologie und Berichterstattung veröffentlichter systematischer Reviews aufzuzeigen, haben wir aufbauend auf den Arbeiten von, haben wir aufbauend auf den Arbeiten von Moher et al.¹³ und Page et al.

¹⁰ eine Trendstudie über einen Zeitraum von 20 Jahren (2000 bis 2019) durchgeführt.¹ Die Studie bildet die Grundlage für die Idee dieser Dissertation und ist die Datenbasis der ersten eingeschlossenen Publikation.

Am 1. Januar 2020 führten wir eine Suche in der Datenbank PubMed durch, um alle zwischen 2000 und 2019 indextierten systematischen Reviews zu identifizieren.¹ Anschließend wurde eine zufällige Stichprobe von 4.000 Artikeln gezogen und doppelt unabhängig auf Eignung geprüft. Eingeschlossen wurden alle Artikel, die den PRISMA-P Kriterien⁷⁹ eines systematischen Reviews entsprachen und auf Englisch publiziert waren. Neben allgemeinen Daten wie dem Publikationsjahr, dem Journal, der Anzahl der Autoren und dem Land des korrespondierenden Autors wurden Daten zur Typologie des systematischen Reviews, Anzahl der durchsuchten Datenbanken, Anzahl identifizierter, gescreenter und eingeschlossener Artikel extrahiert. Darüber hinaus wurde untersucht, ob Sprachbegrenzungen vorgenommen, die Qualität der eingeschlossenen Studien bewertet und eine Meta-Analyse durchgeführt wurde. Basierend auf dieser Stichprobe haben wir die Gesamtzahl der veröffentlichten systematischen Reviews pro Jahr mit 95% Konfidenzintervall geschätzt.

Insgesamt identifizierte die Suche 572.871 Artikel die zwischen 2000 und 2019 indextiert wurden.¹ Von unserer Zufallsstichprobe von 4.000 Referenzen haben wir 1.688 im Volltext gelesen und insgesamt 1.132 systematische Reviews eingeschlossen. Unsere Hochrechnungen zeigen, dass die Zahl der indextierten systematischen Reviews über die Jahre hinweg einen kontinuierlichen Anstieg verzeichneten (Abbildung 2). Während im Jahr 2000 noch 1.432 systematische Reviews indextiert wurden, waren es im Jahr 2019 bereits 29.073. Dies entspricht einem Anstieg um den Faktor 20 sowie einer Anzahl von mehr als 2.400 neu indextierten systematischen Reviews pro Monat und mehr als 80 systematischen Reviews pro Tag in 2019.

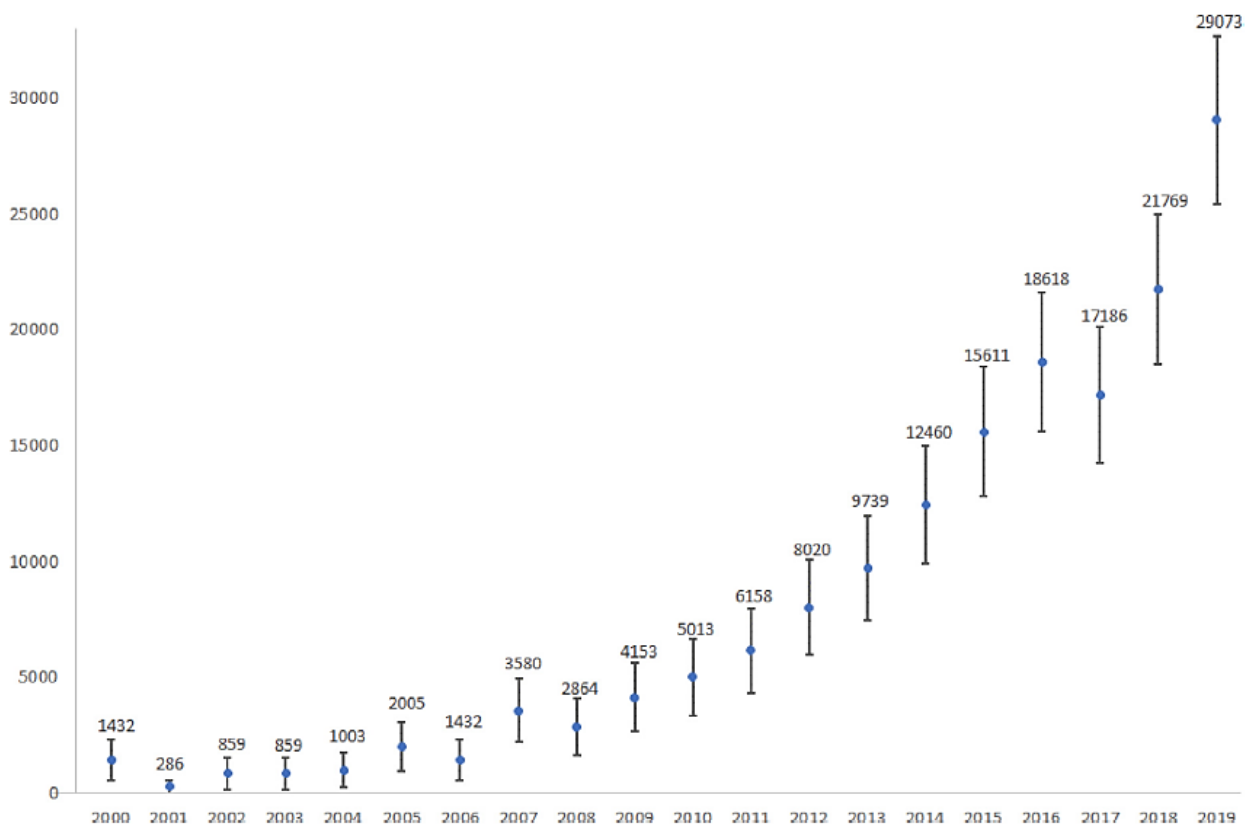


Abbildung 3. Geschätzte Anzahl der in MEDLINE (über PubMed) indextierten systematischen Reviews pro Jahr (mit 95% KI), 2000-2019 aus Hoffmann et al. 2021

Auch die Merkmale der veröffentlichten systematischen Reviews haben sich im Laufe der Zeit verändert.¹ Immer mehr Reviews stammen aus China,^{1,9} die Anzahl der beteiligten Autor*innen ist über die Jahre deutlich gestiegen¹ und die Berichterstattung hat sich verbessert.^{1,103,104} Darüber hinaus ist eine Zunahme der Vielfalt der von Munn et al.¹⁰⁵ eingeführten Review-Typen zu verzeichnen.¹ Während sich Reviews traditionell auf die Bewertung der Wirksamkeit von Interventionen und Therapieansätzen konzentrierten,^{106,107} können heutzutage vielfältigere Forschungsfragen mit systematischen Reviews beantwortet werden.^{105,108,109} Dies spiegelt sich auch in den gefundenen Zahlen unserer Trendstudie wider. Insgesamt stieg die Häufigkeit an anderen Review-Typen seit dem Jahr 2000 stetig an und der Anteil von Reviews zur Wirksamkeit sank von 80% in den Jahren 2000 bis 2004 auf 45% in den Jahren 2015 bis 2019. Auch die Journals, in denen systematische Reviews veröffentlicht werden, werden immer vielfältiger.¹ Dabei ist die Cochrane Database of Systematic Reviews insgesamt das Journal mit dem höchsten Anteil an entsprechenden Publikationen wobei dieser Anteil über die Jahre stark abgenommen hat. Während in den Jahren 2000 bis 2004 noch knapp die Hälfte aller systematischen Reviews von Cochrane publiziert wurden, waren es in den Jahren 2005 bis 2009 noch 15,3% und in den Jahren 2015 bis 2019 lediglich 3,5%.¹

1.5 Relevanz von Abstracts und die Entwicklung von PRISMA-A

Der kontinuierliche Anstieg neu veröffentlichter systematischer Reviews stellt eine Herausforderung für die evidenzbasierte Medizin dar, denn er kollidiert mit den oft begrenzten Ressourcen und zeitlichen Kapazitäten von Forscher*innen und Kliniker*innen.^{110–112} Aus diesem Grund sind Abstracts von besonderer Relevanz. Sie fungieren als primärer Kontaktpunkt zwischen Forschung und Leserschaft und bieten einen schnellen Überblick über das methodische Vorgehen und wichtigste Ergebnisse.^{110–112} Zusammen mit dem Titel ist der Abstract der wichtigste Faktor bei der Entscheidung, ob eine Arbeit im Volltext gelesen wird oder nicht.¹⁷ Darüber hinaus sind Titel und Abstract häufig die einzigen frei zugänglichen Teile eines Manuskriptes. Abstracts sollten entsprechend in der Lage sein, für sich alleine zu stehen^{17,100} und alle wesentlichen Inhalte des Manuskripts in einem informativen, genauen, attraktiven, prägnanten sowie strukturierten Format abzubilden.^{110,113,114}

Mit dem Ziel auch die Qualität von Abstracts zu verbessern, wurde im Jahr 2013 eine Erweiterung von PRISMA spezifisch für Abstracts (PRISMA-A)¹⁷ veröffentlicht. Die Leitlinie soll Autor*innen dabei unterstützen, die wesentlichen Inhalte ihrer Arbeit zu erkennen und transparent zu berichten. Sie definiert insgesamt 12 Items (Tabelle 3), die sich auf den Titel, den Hintergrund, die Methodik, die Ergebnisse, die Schlussfolgerungen, die Finanzierung und Registrierung beziehen. In dem dazugehörigen Statement Paper¹⁷ werden alle Items anhand von Beispielen beschrieben. Im Jahr 2020 wurde die PRISMA-A Leitlinie aktualisiert und in die allgemeine PRISMA Leitlinie integriert.⁹⁹ Dabei wurden die Items inhaltlich kaum modifiziert, jedoch sprachlich so angepasst, dass die Checkliste besser anwendbar ist. In Übereinstimmung mit der allgemeinen PRISMA-Checkliste fokussiert sich auch PRISMA-A seit 2020 verstärkt auf die Berichterstattung des methodischen Vorgehens und ergänzt ein zusätzliches Item zu den verwendeten Synthesemethoden. Da davon auszugehen ist, dass die Implementierung der Anpassungen einige Zeit in Anspruch nimmt und die in dieser Dissertation eingeschlossenen Publikationen überwiegend systematische Reviews enthalten, die vor dem PRISMA-A-Update 2020 publiziert wurden, wurde zur Bewertung der PRISMA-A Adhärenz die Version aus dem Jahr 2013 verwendet.

Tabelle 3. PRISMA-A Checkliste (2013)

Titel	
1. Titel	Identifiziert den Artikel als systematisches Review, Meta-Analyse oder beides
Hintergrund	
2. Ziel	Präzise Angaben zur Fragestellung des Reviews mit Bezug auf Teilnehmer, Intervention, Vergleich oder Zielkriterien
Methoden	
3. Auswahlkriterien	Angabe der Studien- und Berichtsmerkmale, die als Auswahlkriterien verwendet wurden
4. Informationsquelle	Benennung der verwendeten Datenbanken mit Suchzeitpunkt
5. Risk of bias	Angabe der Methoden zur Beurteilung des Risk of bias
Ergebnisse	
6. Eingeschlossene Studien	Anzahl und Art der einbezogenen Studien sowie Teilnehmer und relevante Charakteristika der Studien
7. Ergebnissynthese	Präsentation der Hauptergebnisse, bevorzugt mit Angabe der Anzahl der Studien und Teilnehmer für jedes Ergebnis. Wenn eine Meta-Analyse durchgeführt wurde, sollten Zusammenfassungsmaße und Konfidenzintervalle angegeben werden
8. Effektgrößen	Angaben der Richtung und Größe des Effekts (in Wort und Zahl)
Diskussion	
9. Einschränkungen	Kurze Zusammenfassung der Stärken und Limitationen der Evidenz
10. Schlussfolgerungen	Allgemeine Interpretation der Ergebnisse und/oder Angabe wichtiger Implikationen
Andere	
11. Finanzierung	Angaben von Quellen der finanziellen Unterstützung
12. Registrierung	Angaben zur Registrierung einschließlich Registernummer

**Übersetzt aus Beller et al. 2013¹⁷*

Obwohl die PRISMA-A Leitlinie bereits vor über zehn Jahren veröffentlicht wurde, gibt es bislang nur wenige Daten darüber, inwiefern ihre Einführung zu einer Verbesserung der Berichterstattung von Abstracts geführt hat. Aktuelle Studien deuten jedoch darauf hin, dass Abstracts nach wie vor deutliches Verbesserungspotential aufweisen, die Wortzahl positiv mit der Berichterstattung assoziiert ist und dass strukturierte Abstracts besser berichtet sind als unstrukturierte Abstracts. Die Daten beziehen sich jedoch auf sehr spezifische Forschungsfelder^{115–118} oder Journals^{119–121} und kaum eine Studie unterscheidet zwischen Cochrane und Nicht-Cochrane Reviews.

Ich habe mich daher dazu entschlossen, das Thema meiner Dissertation „Berichterstattung von systematischen Reviews“ weiter zu spezifizieren und mich ausschließlich mit der Berichterstattung von Abstracts zu befassen. Dabei fokussiere ich mich vor allem auf die Unterschiede zwischen Cochrane Reviews und Nicht-Cochrane Reviews und den Einfluss der Abstract-Länge auf die Berichterstattung und Wahrnehmung von Abstracts.

2. Kurzdarstellung und Ergebnisse der eingeschlossenen Publikationen

2.1 Erste Publikation: PRISMA-A Adhärenz in systematischen Reviews

Die erste Publikation¹²² zielte darauf ab, die Berichterstattung von Abstracts von systematischen Reviews anhand der PRISMA-A Leitlinie zu evaluieren und mögliche Unterschiede zwischen Cochrane und Nicht-Cochrane Reviews zu identifizieren.

Methodisch stützte sich die Arbeit auf die oben beschriebene Trendstudie zur Epidemiologie und Berichterstattung von systematischen Reviews.¹ Im Rahmen dieser Studie wurden insgesamt 1.132 systematische Reviews identifiziert, von denen sich 535 mit der Wirksamkeit von Interventionen befassten. Da systematische Reviews zur Wirksamkeit bis heute am häufigsten vorkommen, der Großteil der Cochrane Reviews diesem Typ zugehörig ist¹ und sich PRISMA-A primär auf diesen Typ von Reviews fokussiert, haben wir uns in dieser Arbeit dazu entschieden, uns ausschließlich auf diesen Review-Typ zu fokussieren. Unser primärer Endpunkt war die PRISMA-A Adhärenz. Diese wurde anhand eines Scores von 0 bis 12 Punkten bewertet, wobei wir für jedes vollständig berichtete Item einen Punkt vergeben haben. Um die Anwendung der PRISMA-A Leitlinie zu vereinfachen, haben wir basierend auf dem Statement Paper von Beller et al.¹⁷ eine eigene Operationalisierung entwickelt, die in einigen Items leicht vom Original abwich. Die Daten wurden deskriptiv analysiert und nach Cochrane Reviews vs. Nicht-Cochrane Reviews stratifiziert. Für die Nicht-Cochrane Reviews haben wir zusätzliche Analysen durchgeführt, die nach Publikationsjahr, Länge und Struktur stratifiziert wurden. Da der Anteil von Cochrane Reviews zu gering war um robuste Aussagen treffen zu können, wurden diese aus den zusätzlichen Analysen ausgeschlossen.

Tabelle 4. PRISMA-A Adhärenz in Cochrane vs. Nicht-Cochrane Reviews

Nummer	PRISMA-A Items	Cochrane (n=81)	Nicht-Cochrane (n=453)	Total (n=534)
1	Titel	1 (1,2%)	389 (85,9%)	390 (73,0%)
2	Ziel	81 (100%)	419 (92,5%)	500 (93,6%)
3	Auswahlkriterien	40 (49,4%)	114 (25,2%)	154 (28,8%)
4	Informationsquelle	74 (91,4%)	129 (28,5%)	203 (38,0%)
5	Risk of bias	6 (7,4%)	53 (11,7%)	59 (11,0%)
6	Eingeschlossene Studien	35 (43,2%)	49 (10,8%)	84 (15,7%)
7	Ergebnissynthese*	74 (97,4%)	363 (80,3%)	437 (82,8%)
8	Effektgrößen*	62 (81,6%)	212 (46,9%)	274 (51,9%)
9	Einschränkungen*	57 (75,0%)	186 (41,2%)	243 (46,0%)
10	Schlussfolgerungen	81 (100%)	447 (98,7%)	528 (98,8%)
11	Finanzierung	0 (0%)	0 (0%)	0 (0%)
12	Registrierung	0 (0%)	11 (2,4%)	11 (2,1%)
Gesamtscore				
	Mittelwert (SD)	6,3 (1,3)	5,2 (1,4)	5,4 (1,4)
	Median (IQR)	6,0 [5,0-7,0]	5,0 [4,0-6,0]	5,0 [5,0-6,0]

SD= Standard Deviation (Standardabweichung); IQR = Interquartile Range (Interquartilsabstand); * Der Nenner ist unterschiedlich (n = 528), da das Item bei systematischen Übersichten, in denen keine Studien eingeschlossen sind, nicht anwendbar war.

Insgesamt wurden 81 Cochrane und 454 Nicht-Cochrane Reviews in die Analyse eingeschlossen. Im Durchschnitt wurden 5,4 der maximal möglichen 12 Items vollständig berichtet (Tabelle 4). Die Einhaltung der individuellen Items variierte zwischen 0% und 98,8%. Drei der zwölf Items (Item 2 (Ziel), Item 7 (Zusammenfassung der Ergebnisse) und Item 10 (Interpretation der Ergebnisse)) wurden in mehr als 80% der Abstracts vollständig berichtet. Alle anderen Items wurden in weniger als 55% der Abstracts vollständig berichtet. Dabei waren die Items 6 (Risk of Bias) und 12 (Registrierung) mit 11% bzw. 3% besonders selten vollständig berichtet und das Item 11 (Finanzierung) wurde nie berichtet. Drei der zwölf Items (Items 3 (Einschlusskriterien), 4 (Informationsquelle) und 6 (Eingeschlossene Studien)) setzen sich aus mehreren Teilkomponenten zusammen. Hier zeigte sich, dass diese Items lediglich in 15–40% der Abstracts vollständig erfüllt sind, in mehr als 70% jedoch zumindest teilweise. Zudem konnten wir feststellen, dass Abstracts von Cochrane Reviews in fast allen PRISMA-A Items eine höhere Adhärenz aufweisen. Eine hervorstechende Ausnahme ist Item 1 (Titel). Während 85,9% der Nicht-Cochrane Reviews sich selbst als systematisches Review betiteln, tun dies nur 1,2% der Cochrane Reviews.

Die stratifizierten Analysen nach Länge, Struktur und Jahr, in die ausschließlich Nicht-Cochrane Reviews eingeschlossen wurden, zeigt eine positive Korrelation zwischen der Anzahl an Wörtern und der Anzahl erfüllter PRISMA-A Items ($r = 0,44$; $p < 0,001$). Die Adhärenz steigt dabei linear von 4,5 Punkten im ersten Quartil (≤ 208 Wörter) auf 5,9 Punkte im vierten Quartil (≥ 278 Wörter; Abbildung 4). Diese Tendenz lässt sich auf die meisten der 12 PRISMA-A Items anwenden, insbesondere auf jene, die eine hohe Komplexität aufweisen und mehrere Komponenten beinhalten (z.B. Item 3 (Auswahlkriterien), 4 (Informationsquelle) und 6 (eingeschlossene Studien)). Des Weiteren konnten wir feststellen, dass strukturierte Abstracts mit 5,4 Punkten eine höhere PRISMA-A Adhärenz aufweisen als unstrukturierte Abstracts (4,8 Punkte). Das Publikationsjahr hatte hingegen keinen Einfluss auf die Berichterstattung der Abstracts.

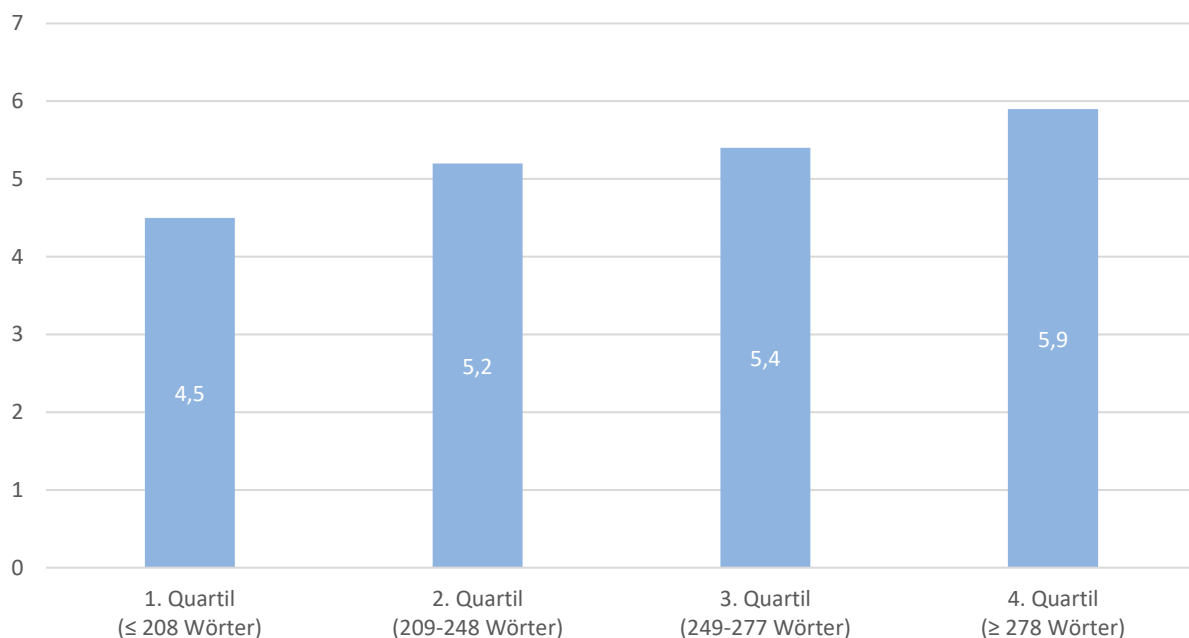


Abbildung 4 Durchschnittliche Anzahl vollständig berichteter PRISMA-A Items nach Anzahl der Wörter (ohne Cochrane Reviews)

Die in dieser Arbeit inkludierten Abstracts von Cochrane Reviews weisen im Median eine Länge von 507 Wörtern auf und sind damit doppelt so lang wie die Abstracts von Nicht-Cochrane Reviews (249 Wörter). Aufgrund der geringen Anzahl an Cochrane Reviews in unserer Stichprobe waren in dieser Arbeit jedoch keine weiteren Aussagen über den Einfluss der Länge von Cochrane Review Abstracts auf deren Berichtsqualität möglich.

2.2 Zweite Publikation: PRISMA-A Adhärenz in Cochrane Reviews bzw. Einfluss der Länge

Die Ergebnisse unserer Trendstudie¹ sowie unserer Publikation zur Berichterstattung von Abstracts¹²² zeigen, dass sich Cochrane Reviews von Nicht-Cochrane Reviews unterscheiden. In Anbetracht der Historie von systematischen Reviews sowie der langjährigen führenden Rolle der Cochrane Collaboration in der Erstellung von systematischen Reviews ist dieses Ergebnis nicht überraschend. Gleichwohl konnte festgestellt werden, dass die Häufigkeit von Cochrane Reviews im Zeitverlauf abgenommen hat¹ und auch Cochrane Reviews deutliche Verbesserungspotenziale in Bezug auf die Berichterstattung von Abstracts aufweisen.¹²² Da die Anzahl eingeschlossener Cochrane Reviews in den beiden zuvor beschriebenen Studien zu gering war, um robuste Aussagen treffen zu können, haben wir uns in unserer zweiten Publikation¹²³ zum Ziel gesetzt, Trends in der Epidemiologie von Cochrane Reviews aufzuzeigen und die Berichterstattung der Abstracts anhand der PRISMA-A Leitlinie in einer größeren Stichprobe zu bewerten.

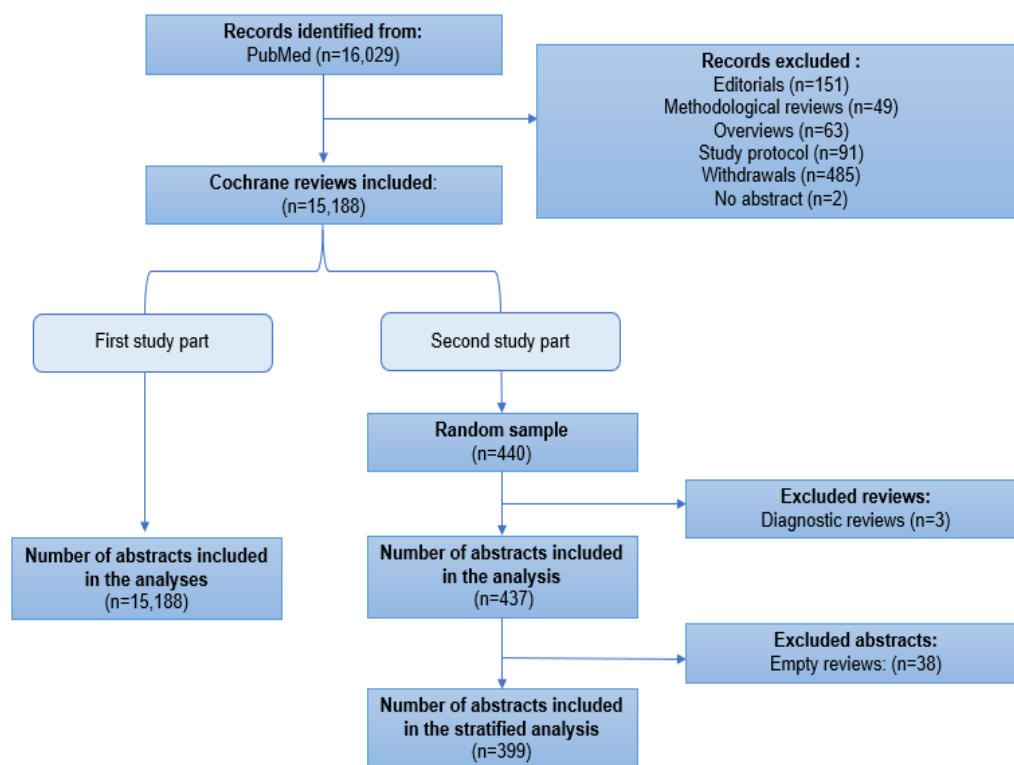


Abbildung 5. Flussdiagramm zur zweiten Publikation

Die Studie basiert auf zwei Teilen (Abbildung 5). Der erste Teil besteht aus einer Vollerhebung aller in PubMed indextierten Cochrane Reviews und zielt darauf ab, Trends in der Anzahl der Publikationen pro Jahr sowie deren Abstract-Länge aufzuzeigen. Dazu wurde PubMed am 18. November 2022 mit dem Suchbegriff „Cochrane Database Syst Rev“[jour]“ ohne jegliche Einschränkungen durchsucht. Eingeschlossen wurden alle in der Cochrane Database publizierten systematischen Reviews. Während Cochrane traditionell

ausschließlich Reviews publiziert hat, hat sich die Datenbank im Laufe der Jahre auch für andere Studientypen geöffnet, darunter Editorials, methodische Reviews und Overviews. Diese Studientypen wurden aus dieser Arbeit ausgeschlossen (Abbildung 5).

Im Rahmen der durchgeführten Vollerhebung wurden für alle Cochrane Reviews das Jahr, der Titel, die Autor*innen, die Cochrane-Datenbank-Nummer sowie der Abstract direkt aus PubMed extrahiert. Des Weiteren wurde die Anzahl an Wörtern im Abstract sowie die Anzahl beteiligter Autor*innen mit dem Statistikprogramm SAS für Windows Version 9.4 (SAS Institute Inc, Cary, North Carolina) bestimmt. Der zweite Teil umfasst eine nach Jahren stratifizierte Zufallsstichprobe von Cochrane Reviews und fokussiert sich auf die PRISMA-A Adhärenz der Abstracts. Die Bewertung der Abstracts erfolgte anhand der in der ersten Publikation entwickelten PRISMA-A Operationalisierung. Analog zur ersten Publikation wird auch in dieser Arbeit für jedes vollständig berichtete Item ein Punkt vergeben (Bewertungsskala: 0 bis 12 Punkte). Eine höhere Punktzahl kann dabei als Indiz für eine bessere Berichterstattung gewertet werden. Um eine Aussage über die Komplexität der Forschungsfrage treffen zu können, wurde für die Zufallsstichprobe zusätzlich noch die Anzahl eingeschlossener Studien extrahiert.

Im Zeitraum zwischen 2000 und 2022 wurden insgesamt 15.188 Cochrane Reviews in PubMed indexiert. Dabei stieg die Zahl von 509 publizierten Reviews im Jahr 2001 auf 999 im Jahr 2013, um anschließend wieder auf 518 im Jahr 2021 zu sinken. Mit Blick auf die Länge der Abstracts konnten wir feststellen, dass die mediane Länge der Abstracts bis 2010 relativ konstant zwischen 387 und 401 Wörtern pro Jahr lag, bevor sie dann kontinuierlich bis auf einen Median von 838 im Jahr 2022 anstieg (Abbildung 6). Der Anteil der Abstracts mit einer Länge von ≤ 300 Wörtern ging von 30,7% im Jahr 2000 auf 6,6% im Jahr 2010 zurück und lag im Jahr 2022 bei unter 1%.

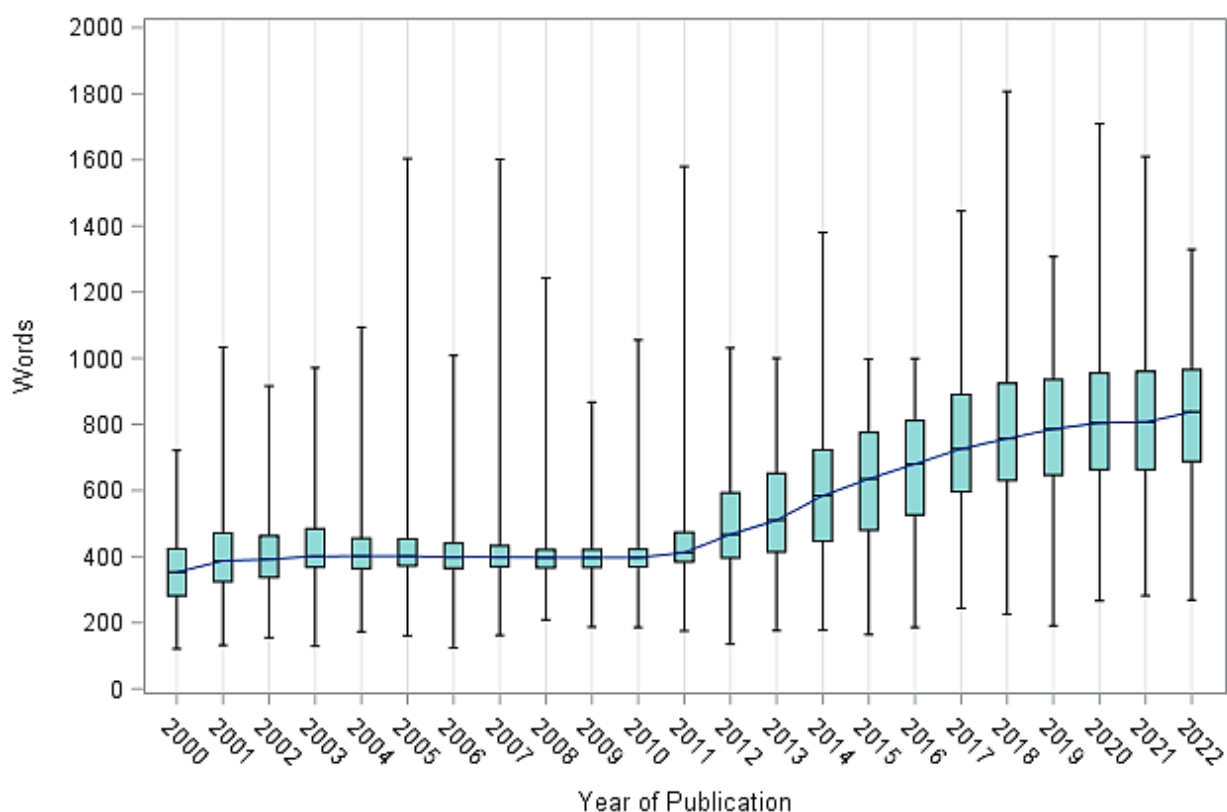


Abbildung 6. Entwicklung der Länge von Cochrane Review Abstracts von 2000-2022

In den zweiten Teil unserer Studie haben wir eine zufällige Stichprobe von insgesamt 427 systematischen Reviews eingeschlossen und anhand von PRISMA-A bewerte. Insgesamt wurden durchschnittlich 6,1 der 12 Items vollständig erfüllt. Dabei stieg die PRISMA-A Adhärenz mit zunehmender Wortzahl von 5,4 Items im ersten Quartil (≤ 398 Wörter) auf 7,1 Items im letzten Quartil (≥ 671 Wörter) an. Stratifiziert nach Jahren zeigt sich allerdings, dass dieser Zusammenhang nur bis 2015 zu finden ist. In den Jahren 2016-2022, in denen über 65% der Studien eine Länge von über 670 Wörtern aufweisen, konnte keine Korrelation zwischen Länge und PRISMA-A Adhärenz festgestellt werden.

2.3 Dritte Publikation: Einfluss der Länge auf die Attraktivität des Abstracts

In den ersten beiden Publikationen^{122,123} konnten wir zeigen, dass die Länge des Abstracts einen Einfluss auf die Berichterstattung nach PRISMA-A hat. Gleichzeitig konnten wir aufzeigen, dass die Abstracts von Cochrane Reviews über die Jahre immer länger werden und Abstract-Längen von bis zu 1.000 Wörtern erreichen.¹²³ Es ist fraglich, ob Abstracts dieser Länge den Kriterien eines gut geschriebenen Abstracts entsprechen und als informativ, genau, attraktiv und prägnant wahrgenommen werden. Bisher gibt es jedoch weder allgemein gültige Empfehlungen zur optimalen Länge von Abstracts systematischer Reviews, noch gibt es Studien, die den Einfluss der Länge auf das Leseverhalten und die Wahrnehmung von Abstracts untersuchen. In meiner dritten Publikation¹²⁴ haben wir uns daher zum Ziel gesetzt, die subjektive Wahrnehmung eines Abstracts aus Sicht der Forschenden abzubilden und zu untersuchen, ob längere Abstracts die Aufmerksamkeit, die dem Artikel von Forschenden geschenkt wird, reduzieren. Weiterhin haben wir untersucht, wie Forschende den Informationsgehalt, die Genauigkeit, die Attraktivität und die Präzision von kurzen bzw. langen Abstracts bewerten.

Zu diesem Zweck haben wir eine zweiarmige, einfach verblindete, randomisierte, kontrollierte Studie durchgeführt, die einem vorab bei ClinicalTrials.gov registrierten Protokoll (NCT06525805) folgt. Eingeschlossen wurden Forschende aus der ganzen Welt, die zwischen dem 1. Januar und dem 26. März in PubMed einen systematischen Review indexiert hatten. Insgesamt wurden 6.230 Referenzen auf ihre Eignung geprüft, von denen 5.397 als systematisches Review klassifiziert und in die Studie eingeschlossen wurden. Für diese wurden die Kontaktdaten der korrespondierenden Autoren manuell extrahiert und im Verhältnis 1:1 randomisiert. Beide Gruppen erhielten ein identisches Anschreiben mit Informationen zur Studie, zum Datenschutz, zur Anonymität und einem Link zur Teilnahme per E-Mail. Durch Anklicken des Links bestätigten die angeschriebenen Forscher*innen automatisch ihr Mitwirken an der Studie und wurden als „Respondents“ gezählt. Je nach Gruppenzugehörigkeit enthielt die erste Seite des Surveys entweder einen kurzen oder einen langen Abstract desselben systematischen Reviews, der in zwei Journals mit unterschiedlicher Länge veröffentlicht worden war (Cochrane Database: 771 Wörter;¹²⁵ BJU International: 277 Wörter¹²⁶).

Unser Ziel war es, das Suchverhalten in PubMed nachzustellen. Wir definierten die Studienteilnahme nach dem Lesen des Abstracts als unseren primären Endpunkt. Diese wurde als Indikator für die Aufmerksamkeit interpretiert, die dem Artikel von Forschenden geschenkt wird und der Entscheidung, diesen im Volltext zu lesen. Die Studienteilnahme wurde berechnet indem die Anzahl der „Teilnehmenden“ (angeschriebene Autor*innen, die nach dem Lesen des Abstracts mit dem Survey fortgefahren sind) durch die Gesamtzahl der „Respondents“ (angeschriebene Autor*innen, die auf den Link in der E-Mail geklickt haben) geteilt wurde. Als sekundäre Endpunkte wurden die von Bahadoran et al.¹¹⁰ beschriebenen vier Indikatoren eines gut

geschriebenen Abstracts (Informationsgehalt, Genauigkeit, Attraktivität und Präzision) sowie allgemeine Abstract Charakteristika betrachtet.

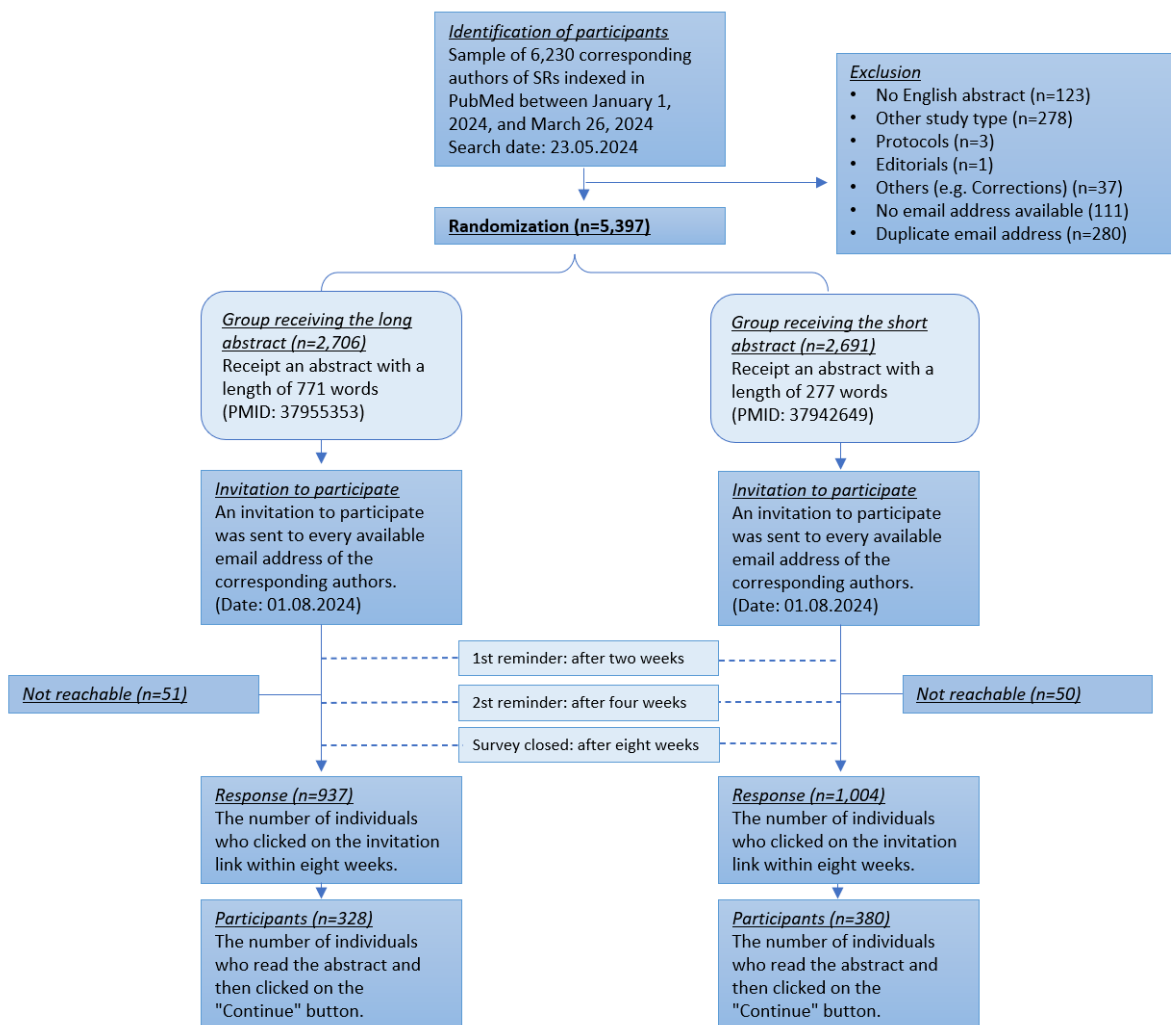


Abbildung 7. Flussdiagramm zur dritten Publikation

Insgesamt 5.397 Forscher*innen, darunter 2.691 in der Gruppe des kurzen Abstracts und 2.706 in der Gruppe des langen Abstracts, wurden per Mail kontaktiert und zur Teilnahme eingeladen (Abbildung 7). Davon folgten 1.941 dem Teilnahmelink. Nach dem Lesen des Abstracts setzten 35,0% der Autor*innen, die den langen, und 37,8%, die den kurzen Abstract erhielten, die Teilnahme an der Studie fort. Dieser Unterschied war statistisch nicht signifikant ($p=0,1935$). Dabei brauchten die Teilnehmer*innen, die den langen Abstract erhielten im Median 77,5 Sekunden um den Abstract zu lesen während die Teilnehmer*innen die den kurzen Abstract erhielten, im Median 66,0 Sekunden brauchten ($p=0,0702$).

Die Länge des Abstracts hatte keinen Einfluss auf die Bewertung des Informationsgehalts (kurzer Abstract: 85,5% vs. langer Abstract: 91,2%; $p=0,0594$) und der Genauigkeit (kurzer Abstract: 80,3% vs. langer Abstract: 86,3%; $p=0,0868$). Kürzere Abstracts wurden jedoch als attraktiver (kurzer Abstract: 60,5% vs. langer Abstract: 46,6%; $p=0,0034$) und prägnanter (kurzer Abstract: 82,3% vs. langer Abstract: 37,9%; $p<0,0001$) bewertet. Drei Viertel der Teilnehmenden befürwortete eine maximale Abstract-Länge von 250 bis 300 Wörtern, und fast alle präferierten ein strukturiertes Format. Interessant war zudem, dass 41,1% bzw. 55,5% befürworteten, dass Informationen zur Finanzierung bzw. zur Registrierung im Abstract berichtet sein sollten.

3. Diskussion

In meiner ersten Publikation¹²² habe ich gezeigt, dass die Abstracts systematischer Reviews deutliches Verbesserungspotential aufweisen und nur 5 von 12 PRISMA-A Items vollständig erfüllen. Dabei berichteten Cochrane Reviews im Median ein Item mehr als Nicht-Cochrane Reviews. Für Nicht-Cochrane Reviews konnten wir einen linearen Anstieg der PRISMA-A Adhärenz mit zunehmender Länge zeigen. Wir schlossen daraus, dass Autor*innen von systematischen Reviews eine bestimmte Anzahl von Wörtern benötigen, um informative Abstracts zu schreiben und empfahlen eine Abstract Länge von 250-300 Wörtern, die sich an den bereits durch CONSORT für Abstracts (CONSORT-A)¹²⁷ etablierten Richtlinien für RCTs orientiert. In meiner zweiten Publikation,¹²³ die sich ausschließlich auf Cochrane Reviews konzentriert, konnte ich einen kontinuierlichen Rückgang an neu publizierten Reviews seit dem Jahr 2013 aufzeigen. Gleichzeitig stieg die mediane Länge der Abstracts von 353 Wörtern im Jahr 2000 auf 828 Wörter im Jahr 2022. Weniger als ein Prozent der 2022 publizierten Cochrane Reviews erfüllten die von uns abgeleitete Empfehlung und waren kürzer als 300 Wörter. Der positive Längeneffekt, der sich bei den Abstracts der Nicht-Cochrane Reviews zeigte, löste sich bei den Abstracts von Cochrane Reviews mit zunehmender Länge auf. In meiner dritten Publikation¹²⁴ habe ich die Erkenntnisse zum Effekt der Länge auf die Wahrnehmung von Abstracts aus der Perspektive von Forschenden erweitert. Dabei hatte die Länge des Abstracts keinen Einfluss auf die Studienteilnahme, die Wahrnehmung des Informationsgehalts und der Genauigkeit. Kürzere Abstracts wurden jedoch als attraktiver und prägnanter wahrgenommen. Darüber hinaus befürworteten drei Viertel der Befragten eine Abstract-Länge von 250 bis 300 Wörtern. Etwa die Hälfte der Befragten befürwortet zudem die Berichterstattung von Informationen zur Finanzierung und Registrierung im Abstract.

3.1 Messen der PRISMA-A Adhärenz

In unserer ersten Publikation haben wir gezeigt, dass im Durchschnitt 5 von 12 PRISMA-A Items in den Abstracts von systematischen Reviews vollständig berichtet wurden.¹²² Dabei wies die Übereinstimmung zwischen den einzelnen Items eine starke Variation auf (0% bis 98,8%). Unsere Ergebnisse unterscheiden sich in Teilen deutlich von denen anderer Studien. Während Maticic et al.¹¹⁷ eine ähnliche Gesamtadhärenz beschreiben, berichten Adobes et al.¹¹⁵ und Tsou et al.¹²⁸ von höheren PRISMA-A Übereinstimmungen. Insbesondere mit Blick auf die Items 3 (Einschlusskriterien), 4 (Informationsquelle) und 6 (Eingeschlossene Studien), welche sich aus mehreren Teilkomponenten zusammensetzen, unterscheiden sich die Studien deutlich. Dabei berichten Maticic et al.¹¹⁷ mit Übereinstimmungen von unter 20% etwas geringere PRISMA-A Adhärenzen. Adobes et al.¹¹⁵ und Tsou et al.¹²⁸ hingegen beschreiben mit Übereinstimmungen von über 60% deutlich höhere Adhärenzen. Diese Diskrepanzen lassen sich einerseits darauf zurückführen, dass die Studien unterschiedliche Forschungsfelder berücksichtigen, und andererseits darauf, dass unterschiedliche Bewertungssysteme genutzt wurden. Während Maticic et al.,¹¹⁷ genau wie wir, ein binäres Punktesystem verwenden, bei dem für jedes vollständig erfüllte Item ein Punkt vergeben wurde, nutzten Adobes et al.¹¹⁵ und Tsou et al.¹²⁸ ein weniger strenges Bewertungssystem und vergaben auch für teilweise erfüllte Items Punkte. Darüber hinaus bietet PRISMA-A bei einzelnen Items einen hohen Grad an Interpretationsspielraum. Während in meiner ersten Publikation¹²² alle Daten von mir alleine extrahiert wurden und lediglich 20% von einem zweiten Reviewer geprüft wurden, haben wir in der zweiten Publikation¹²³ alle Daten doppelt und unabhängig voneinander extrahiert. Trotz umfassender Kalibrierung haben wir für einzelne Items Abweichungen von über 15% dokumentiert.¹²³ Diese Abweichungen stehen im

Einklang mit den Ergebnissen zur Interrater-Reliabilität von Maticic et al.¹¹⁷ und zeigen, dass eine einfache Datenextraktion in Bezug auf PRISMA-A fehleranfällig ist. An dieser Stelle ist darauf hinzuweisen, dass auch andere Reporting Leitlinien wie PRISMA^{14,99}, CONSORT¹²⁷, RECORD¹²⁹ oder STROBE¹³⁰ in Teilen sehr komplex sind und Items definieren, aus denen nicht unmittelbar ersichtlich ist, wie diese genau zu operationalisieren sind. Entsprechend ist davon auszugehen, dass bei der Bewertung dieser Leitlinien ähnliche Herausforderungen bestehen. Um trotz dieser Herausforderung vergleichbare Ergebnisse zur Einhaltung von Reporting Leitlinien zu generieren, sollten in zukünftigen Studien alle Daten doppelt und unabhängig voneinander anhand einer klar formulierten und offengelegten Operationalisierung extrahiert werden. Zudem sollte die Interrater-Reliabilität ermittelt und berichtet werden und eine vorherige Kalibrierung erfolgen.

3.2 Anwenden der PRISMA-A Leitlinie

Die hohe Komplexität der PRISMA-A Leitlinie¹⁷ erschwert nicht nur die Messung der Adhärenz sondern auch deren Anwendung. In unserer ersten Publikation konnten wir zeigen, dass der Großteil der PRISMA-A Items in weniger als 50% der Abstracts vollständig erfüllt wurde.¹²² Dies scheint zumindest in Teilen auf die hohen Anforderungen einiger Items zurückzuführen zu sein. PRISMA-A¹⁷ fordert beispielsweise für Item 6 (Eingeschlossene Studien), dass die Anzahl und Art der eingeschlossenen Studien, deren Teilnehmer sowie die studienrelevanten Merkmale berichtet werden. Während lediglich 15% der betrachteten Abstracts dieses Item vollständig erfüllten, berichteten über 90% der Abstracts unserer Studie das Item zumindest teilweise.¹²² Dabei werden die Anzahl der eingeschlossenen Studien und Studienteilnehmer*innen (52,8%) häufiger berichtet als die Studiencharakteristika (24,6%). Ähnliches zeigt sich bei den Items 3 (Einschlusskriterien) und 4 (Informationsquelle). Dies lässt darauf schließen, dass die Zusammenfassung vieler Informationen zu übergeordneten Items möglicherweise dazu führt, dass einzelne Teilkomponenten systematisch weniger stark berücksichtigt werden oder dass Items von den Autor*innen als erfüllt erachtet werden, obwohl nur Teile davon berichtet wurden. Autor*innen, Editor*innen und Reviewer*innen von systematischen Reviews müssen entsprechend vermehrt für die Wichtigkeit von Abstracts und die Einhaltung der PRISMA-A Leitlinie sensibilisiert werden.

Die letzten beiden Items der PRISMA-A Leitlinie (Item 11 Finanzierung und Item 12 Registrierung) werden in Abstracts von systematischen Reviews so gut wie nie berichtet.¹²² Dies ist vermutlich weniger auf die Autor*innen selbst zurückzuführen, sondern vielmehr auf die Autorenrichtlinien und Publikationsprozesse der Journals. Im Rahmen des Einreichprozesses werden Informationen zur Finanzierung und Registrierung in der Regel standardmäßig erhoben. Sie werden jedoch meist unabhängig vom Abstract gesammelt und im eingereichten Manuskript auf dem Titelblatt oder der letzten Seite der Publikation platziert. Das Problem dabei ist, dass Informationen, die nicht im Titel oder Abstract stehen, nicht in Datenbanken wie PubMed übertragen werden und somit für die Leser*innen nicht direkt einsehbar sind. Finanzierung und Registrierung sind zwei wesentliche Indikatoren für Transparenz, die in jedem Manuskript beschrieben sein sollten. Mit Bezug auf den Abstract haben wir in den ersten beiden Publikationen^{122,123} vermehrt auf die methodische Transparenz Bezug genommen und haben den Items Finanzierung und Registrierung eine weniger große Relevanz zugesprochen. Überraschenderweise haben wir in der dritten Publikation jedoch festgestellt, dass knapp die Hälfte der befragten Autor*innen von systematischen Reviews befürworten, dass Abstracts Informationen zur Finanzierung und Registrierung beinhalten sollen.¹²⁴ Dies steht im Kontrast zur

gegenwärtigen Praxis, in der Informationen zur Registrierung in weniger als 3% der Abstracts und Informationen zur Finanzierung nahezu nie berichtet werden.^{122,123,131} Ein ähnliches Bild zeigt sich auch in anderen Studiendesigns. So empfiehlt CONSORT für Abstracts¹³² von RCTs, analog zu PRISMA-A, dass Informationen zur Finanzierung und Registrierung berichtet werden sollten. Studien zeigen jedoch, dass diese Informationen auch hier oft fehlen.^{133–137} Es wird daher empfohlen, dass Journals ihre Autorenrichtlinien überarbeiten und Informationen zur Finanzierung und Registrierung unabhängig vom Studiendesign in den Abstract aufnehmen.

Insgesamt ist die PRISMA-A Leitlinie sehr komplex und für die Autor*innen systematischer Reviews nur schwer vollständig umsetzbar. Mit dem Ziel, die praktische Anwendung zu erleichtern, wurde PRISMA-A im Jahr 2020⁹⁹ aktualisiert, wobei vor allem strukturelle und sprachliche Anpassungen vorgenommen wurden. Im Gegensatz zur Version von 2013¹⁷ fehlt für das Update jedoch eine detaillierte Beschreibung der Anforderungen für jedes Item, insbesondere derjenigen, die geändert oder neu hinzugefügt wurden. Dies ist vermutlich darauf zurückzuführen, dass PRISMA-A mit dem Update in die allgemeine PRISMA Leitlinie integriert wurde.⁹⁹ Dies unterstreicht einerseits die Relevanz von gut geschriebenen Abstracts, erschwert andererseits aber eine einheitliche Anwendung und Operationalisierung der Items. Anstelle einer ausführlichen Beschreibung und Diskussion zur Relevanz der einzelnen PRISMA-A Items ist zu empfehlen, dass beim nächsten Update wieder mehr Fokus auf die praktische Umsetzung gelegt wird. Während es für die allgemeine PRISMA Leitlinie neben dem Statement Paper⁹⁹ noch ein Explanation and Elaboration Paper inklusive Supplement-Dokument mit zusätzlichen Berichtsbeispielen gibt,¹² gibt es für das PRISMA-A Update 2020 lediglich eine kurze Checkliste, in dem die einzelnen Items ohne Erklärung gelistet sind. Insbesondere für unerfahrene Autor*innen ist es schwierig, gute Abstracts zu schreiben. Hier führen komplexe Leitlinien wie PRISMA-A möglicherweise eher zu Verwirrung als dass sie Unterstützung bieten.

3.3 PRISMA-A Adhärenz in Cochrane Reviews

Unsere Analysen zur PRISMA-A Adhärenz von Cochrane Reviews zeigen, dass diese im Durchschnitt ein PRISMA-A Item mehr berichten als Nicht-Cochrane Reviews (6,1 vs. 5,2 Items).¹²³ Allerdings variiert auch bei Cochrane Reviews die Erfüllung der einzelnen Items stark und nur 4 von 12 Items werden in mehr als 85% der Abstracts berichtet. Item 1 (Titel) und Item 5 (Risk of Bias) werden lediglich in 0,7% bzw. 7,8% berichtet und Item 11 (Finanzierung) und Item 12 (Registrierung) werden in keinem der betrachteten Cochrane Reviews berichtet.¹²³ Damit sind die Ergebnisse mit einer Ausnahme (Item 1; Titel) mit denen von Nicht-Cochrane Reviews vergleichbar.

Dass Cochrane Reviews ein PRISMA-A Item mehr berichten als Nicht-Cochrane Reviews lässt sich vermutlich auf die hohen methodischen Standards sowie die standardisierten Redaktions- und Publikationsprozesse von Cochrane zurückführen. Mit dem Cochrane Handbuch^{89,138} und der MECIR Leitlinie⁹⁰ wurden zwei hochkomplexe Reporting-Standards etabliert, in denen genau festgelegt ist, welche Inhalte wie und wo zu berichten sind. Beide Dokumente verweisen explizit auf PRISMA und somit auch indirekt auf PRISMA-A. Wir haben jedoch festgestellt, dass insbesondere für die Items mit der niedrigsten PRISMA-A Adhärenz (Item 1 (Titel), Item 5 (Risk of Bias), Item 11 (Finanzierung) und Item 12 (Registrierung)) keine, eigene oder abweichende Empfehlungen von Cochrane gegeben werden.

Weniger als ein Prozent der Cochrane Reviews kennzeichnen sich im Titel eindeutig als systematische Reviews.¹²³ Damit weicht der Großteil der Cochran Reviews sowohl von PRISMA^{14,99} als auch von PRISMA-A¹⁷ ab und erhalten in beiden Leitlinien automatisch einen Punkt weniger. Beide Leitlinien empfehlen im ersten Item (Titel) einheitlich, dass sich systematische Reviews bereits im Titel eindeutig als solche kennzeichnen sollten. Dies tun über 85% der Nicht-Cochrane Reviews. Während im Cochrane Handbuch¹⁵⁹ die Einhaltung der PRISMA Leitlinie empfohlen wird, werden im Style Manual¹³⁹ eigene, von PRISMA abweichende Empfehlungen für den Titel definiert. Dabei rät Cochrane explizit von einer Betitelung als systematisches Review ab und bezeichnet diese sogar als überflüssig.¹³⁹ Historisch konnte diese Empfehlung damit begründet werden, dass Cochrane ausschließlich systematische Reviews publiziert und der Name der Datenbank "Cochrane Database of Systematic Reviews" bereits impliziert, dass es sich um ein systematisches Review handelt. Heute ist diese Begründung jedoch nicht mehr haltbar. Cochrane publiziert zunehmend Protokolle, Editorials und Overviews und beschränkt sich damit bewusst nicht mehr nur auf systematische Reviews. Die fehlende Bezeichnung als systematisches Review kann zudem einige nicht unerhebliche praktische Konsequenzen für die Auffindbarkeit von Cochrane Reviews haben.¹⁴⁰ Bei einer Ad-hoc-Suche in PubMed mit dem Begriff "systematic review" im Titel und Abstract werden Cochrane Reviews beispielsweise nur sehr selten gefunden, auch wenn es welche zu der Thematik gibt. Dies wirkt sich möglicherweise negativ auf die Lese- und Zitierhäufigkeit von Cochrane Reviews aus, die laut Goldkuhle et al.¹⁴¹ sowie Rosengaard et al.¹⁴² niedriger ist als die von in anderen hochrangigen Journals veröffentlichten systematischen Reviews. Um die Auffindbarkeit, die Qualität und die Reichweite von Cochrane Reviews zu erweitern, empfehlen wir, dass sich diese zukünftig entsprechend der PRISMA Leitlinien bereits im Titel als systematische Reviews ausweisen.

Wie bei Nicht-Cochrane Reviews war in unserer Studie auch bei Cochrane Reviews das Item 5 (Risk of Bias) mit unter 10% sehr selten und die Items 11 (Finanzierung) und 12 (Registrierung) bei den eingeschlossenen Artikeln nie berichtet.¹²³ Dabei gibt Cochrane selbst keine spezifischen Empfehlungen zum Reporting von Finanzierung und Registrierung. PRISMA-A definiert Cochrane Reviews allerdings explizit als Ausnahme, für die die Berichterstattung der Registrierung als nicht zwingend notwendig erachtet wird.¹⁷ Dies wird darin begründet, dass alle publizierten Cochrane Reviews gemäß den Vorgaben von Cochrane vorab mit einem geprüften Protokoll registriert werden müssen. Darüber hinaus wurden Cochrane Reviews von 2013 bis 2023 zusätzlich automatisch in PROSPERO, einem internationalen Register für systematische Reviews, registriert.¹⁴³ Seit 2023 ist die Registrierung in PROSPERO kein automatisierter Prozess mehr, sondern lediglich eine Empfehlung von Cochrane. In Bezug auf Item 5 (Risk of Bias) bestehen Unstimmigkeiten zwischen den Empfehlungen von PRISMA-A^{17,99}, dem Cochrane Handbuch¹⁴⁴ und den MECIR Leitlinie.⁹⁰ Probleme bei der Konzeption und Durchführung einzelner Studien können die Gültigkeit der Ergebnisse in Frage stellen. Daher ist es ein wichtiger Teil von systematischen Reviews, die Validität einzelner Studien und das Risiko, dass sie den tatsächlichen Interventionseffekt überschätzen, zu bewerten.¹⁷ PRISMA-A fordert entsprechend, dass die zur Bewertung des Verzerrungsrisikos verwendeten Methoden berichtet werden. Im Rahmen dessen wird auf die Vielzahl an Instrumenten verwiesen, die zur Bewertung des Verzerrungsrisikos herangezogen werden können. Es wird daher empfohlen, neben den zu bewertenden Komponenten auch das genutzte Instrument zu berichten.¹⁷ An dieser Stelle sind die Empfehlungen von Cochrane weniger streng. MECIR adressiert das Verzerrungsrisiko seit 2012⁹⁰. Anstelle einer detaillierten Beschreibung der methodischen Vorgehensweisen betont MECIR jedoch, dass es für eine Vielzahl von systematischen Reviews

ausreicht, zu vermerken, dass die Standard-Cochrane-Methoden eingehalten wurden, ohne diese näher zu konkretisieren. Das Cochrane Handbuch¹⁴⁵ wurde erst 2019 um Empfehlungen zur Bewertung des Verzerrungsrisikos erweitert. Dabei wird auf das Bewertungsinstrument "Grading of Recommendations, Assessment, Development and Evaluation" (GRADE)¹⁴⁶ verwiesen, welches jedoch nicht das Verzerrungsrisiko beschreibt, sondern die Gesamtqualität der Evidenz bewertet. Im Jahr 2023 hat Cochrane eine Überarbeitung und Vereinheitlichung der Berichtsempfehlungen vorgenommen. Seitdem enthält MECIR keine Empfehlungen mehr für die Berichterstattung, sondern fokussiert sich ausschließlich auf die methodische Durchführung von Cochrane Reviews. Damit ist das Cochrane-Handbuch seit 2023 die einzige Berichtsleitlinie für Cochrane-Reviews. Seit dem Update verweist dieses vollständig auf PRISMA 2020⁹⁹. Die Überarbeitung wird von uns positiv bewertet und es lassen sich bereits erste Anzeichen für die Umsetzung erkennen. Während Finanzierung und Registrierung bis 2023 nie berichtet wurden, sind Anfang 2025 erste Cochrane Reviews in PubMed indexiert, die diese Informationen im Abstract berichten.^{147–149} Nichtsdestotrotz handelt es sich hierbei aktuell noch um Einzelfälle, weshalb die Implementierung und Einhaltung der Überarbeitung in den kommenden Jahren kritisch zu evaluieren ist.

Es ist evident, dass sich Cochrane Reviews in Form, Umfang und grundlegenden Charakteristika von Nicht-Cochrane Reviews unterscheiden und eine höhere methodische Qualität aufweisen.^{10,150,151} Gleichzeitig zeigen Studien, dass auch Cochrane Reviews Mängel haben.^{152,153} Es fehlen jedoch methodische Studien, die sich explizit mit den Unterschieden zwischen Cochrane- und Nicht-Cochrane Reviews auseinandersetzen und die methodische Güte von Cochrane Reviews kritisch hinterfragen. Dabei beschränken sich aktuelle Studien zur methodischen Qualität häufig ausschließlich auf Nicht-Cochrane Reviews.^{154,155} Darüber hinaus haben wir die Erfahrung gemacht, dass methodische Arbeiten, die sich unter anderem mit Cochrane Reviews befassen, häufig im Rahmen des Peer-Review-Prozesses von Gutachtern bewertet werden, die offenbar eine hohe Affinität zu Cochrane aufweisen. Die Gutachter unserer Arbeiten neigten zum Teil dazu, apodiktische Argumentationen zu verwenden und pauschale Argumente anzuführen, um jegliche Kritik zu entkräften. So haben wir beispielsweise Kommentare erhalten wie:

"Rightly or wrongly, Cochrane might view 'Identify the report as a systematic review' in the title as redundant (item 1)."

"For review groups that always rely on published search strategies (sometimes freely available on their websites), is it reasonable to argue that listing every database and register in each of their review abstracts would be redundant (Item 4)?"

Es mag sein, dass Cochrane selbst einzelne PRISMA-A Items als weniger wichtig erachtet und Inhalte, die in Cochrane internen Standards definiert sind als redundant ansieht. Für Forschende außerhalb von Cochrane erscheint dieses Vorgehen jedoch wenig transparent und nachvollziehbar. Erschwerend hinzu kommt, dass es weltweit über 50 Cochrane-Gruppen gibt, die systematische Reviews zu spezifischen Fachgebieten publizieren. Dabei variiert die Berichterstattung zwischen den einzelnen Cochrane-Gruppen stark und einige Gruppen verweisen auf eigene Standards, die nicht im Cochrane Handbuch beschrieben sind. In vielen Fällen bleibt unklar, was genau unter dem Begriff „Standardmethoden“ zu verstehen ist und welche methodischen Vorgehensweisen darunter zusammengefasst sind.¹³¹ Hinzu kommt, dass es im Bereich der Evidenzsynthesen bekannte Autor*innen gibt, die sowohl für PRISMA als auch für Cochrane Arbeiten publizieren, wobei sie

offenbar unterschiedliche Regeln anwenden. So ist Matthew J. Page, der Erstautor der PRISMA Leitlinie,⁹⁹ beispielsweise Koautor eines Reviews,¹⁵⁶ der nicht vollständig nach PRISMA-A berichtet ist und bereits beim ersten Item (Titel) einen Punkt verliert. Diese Ambivalenz wirft die Frage auf, ob sich die Autor*innen an dieser Stelle selbst dazu entschieden haben, von den eigens entwickelten Leitlinien abzuweichen oder ob die rigorosen Vorgaben von Cochrane eine PRISMA-A konforme Berichterstattung im Titel und Abstract unmöglich machen.

3.4 Die Länge von Abstracts von Cochrane Reviews

Ein weiterer Unterschied von Cochrane Reviews im Vergleich zu Nicht-Cochrane Reviews ist die Länge der Abstracts. Diese stieg kontinuierlich von 353 Wörtern im Jahr 2000 auf 838 Wörter im Jahr 2022.¹²³ Dabei lag die Länge der Abstracts bis 2010 noch konstant bei unter 400 Wörtern. Diese Entwicklung steht im Widerspruch zu den bestehenden Cochrane-Richtlinien, kann jedoch durch die Veröffentlichung der MECIR-Leitlinie im Jahr 2012⁹⁰ zum Teil erklärt werden. Während das Cochrane Handbuch bis 2019 eine maximale Abstract-Länge von 400 Wörtern vorsieht, empfiehlt MECIR seit 2012 eine Abstract-Länge von weniger als 700 Wörtern, erlaubt jedoch Abstracts von bis zu 1000 Wörtern.⁹⁰ Im Jahr 2019 wurden diese Empfehlung auch in das Cochrane Handbuch übernommen.¹⁴⁵ Seit 2017 wird die empfohlene Länge von 700 Wörtern jedoch kontinuierlich überschritten und der längste gefundene Abstract¹⁵⁷ übertrifft mit 1.807 Wörtern alle jemals von Cochrane ausgesprochenen Empfehlungen deutlich. Dabei zeigt sich, dass sich der in Nicht-Cochrane Reviews identifizierte lineare Zusammenhang zwischen Abstract-Länge und PRISMA-A Adhärenz in Cochrane Reviews mit zunehmender Länge auflöst.¹²³ Auffällig ist, dass der Ergebnisteil von Cochrane Reviews deutlich länger ist als der Methodenteil¹³¹ und eine höhere PRISMA-A Adhärenz aufweist.^{123,131} Dies deckt sich mit dem Trend, im Methodenteil der Abstracts von Cochrane Reviews vermehrt auf Standard Methoden zu verweisen, anstatt das verwendete Vorgehen transparent zu berichten.¹³¹ In einer weiteren Studie zu den Trends in der Berichterstattung von Standardmethoden in Cochrane Reviews konnten wir zeigen, dass im Jahr 2023 mehr als die Hälfte der Cochrane Reviews im Abstract auf die Verwendung von Standardmethoden verwiesen.¹³¹ Dabei bezogen sich die beschriebenen Standardmethoden zumeist auf die Datenerhebung und -analyse, zunehmend aber auch auf die durchgeführte Suche. Während es in Bezug auf die Datenerhebung und -analyse sowohl in MECIR als auch im Cochrane Handbuch Empfehlungen zum Berichten von Standardmethoden gibt, fordern beide Leitlinien in Bezug auf die Suche eine transparente Berichterstattung der verwendeten Datenbanken sowie des Suchdatums der letzten Suche. Nichtsdestotrotz steigen die Angaben von Standardmethoden auch hier von 1,6% im Jahr 2016 auf 9,7% im Jahr 2023. Gleichzeitig sank die Länge des Abschnitts zur Suche über die Jahre von 60 Wörtern auf 37,5, während die der Ergebnisse von 275 auf 448 anstieg. Dies unterstreicht die hohe Priorisierung der Ergebnisse in Abstracts von Cochrane Reviews. Die Fülle der präsentierten Informationen stellt die Leserschaft jedoch vor die Herausforderung, die für sie relevanten Ergebnisse zu identifizieren. Dabei werden längere Abstracts als weniger präzise wahrgenommen. Neben der eingeschränkten Auffindbarkeit könnte dies ein weiterer Grund für die niedrigeren Zitationshäufigkeiten von Cochrane Reviews sein.

Insgesamt sind die Abstracts von Cochrane Reviews mehr als dreieinhalb Mal so lang wie die Abstracts von Nicht-Cochrane Reviews.^{123,131} Dabei lassen sich zentrale Argumente für längere Abstracts wie beispielsweise eine hohe Komplexität der Reviews nicht als Begründung anführen. Cochrane Reviews umfassen im

Durchschnitt lediglich die Hälfte der Studien, die Nicht-Cochrane Reviews einschließen (7 vs. 13 Studien),¹²² und sind damit weniger komplex. Darüber hinaus sind selbst die Abstracts von leeren Cochrane Reviews (0 eingeschlossene Studien) im Median mit 318 Wörtern länger als die Abstracts von Nicht-Cochrane Reviews (249 Wörter).¹²³ Ein weiteres häufig angeführtes Argument für das Verfassen von umfassenden Abstracts ist, dass der Abstract nicht selten der einzige frei zugängliche Teil eines Artikels ist. Das mag heute zwar noch für viele Publikationen stimmen, die wissenschaftliche Gemeinschaft hat das Problem jedoch erkannt und diskutiert zunehmend über die Bedeutung und Umsetzung von „Open Science“. Viele Journals, darunter auch Cochrane, haben in den vergangenen Jahren bereits ihre Richtlinien aktualisiert und Open-Access-Strukturen implementiert. Bereits seit 2013 werden die Volltexte aller publizierter Cochrane Reviews spätestens ein Jahr nach Veröffentlichung in der Cochrane Library frei zugänglich gemacht.¹⁵⁸ Darüber hinaus gibt es mit über 100 Ländern mit niedrigem und mittlerem Einkommen nationale Zugangsvereinbarungen, die Artikel für über 3 Milliarden Menschen weltweit unmittelbar zugänglich machen.¹⁵⁸ Es erscheint daher nicht notwendig, in den Abstracts von Cochrane Reviews alle Ergebnisse des Hauptmanuskripts zu berichten. Um eine erhöhte Präzision und Attraktivität zu gewährleisten, sollte Cochrane die eigenen Empfehlungen zur Länge von weniger als 700 und maximal 1000 Wörter überdenken.

3.5 Die optimale Länge von Abstracts systematischer Reviews

Die Funktion eines Abstracts besteht darin, die wesentlichen Inhalte eines Manuskripts in informativer, genauer, attraktiver und prägnanter Form wiederzugeben und Interesse für den Volltext zu wecken.^{110,113,114} Die Länge scheint dabei ein wesentlicher Einflussfaktor zu sein. In unserer dritten Publikation haben wir versucht, den Entscheidungsprozess, einen Artikel im Volltext zu lesen oder nicht, zu reproduzieren.¹²⁴ Wir haben die Hypothese aufgestellt, dass längere Abstracts eine „abschreckende“ und „erschlagende“ Wirkung haben und sich negativ auf diese Entscheidung auswirken. Um diese Hypothese zu überprüfen, haben wir die Studienteilnahme nach dem Lesen des Abstracts als primären Endpunkt definiert. Diese wurde als Indikator der Aufmerksamkeit und als potientes Maß für die Entscheidung einen Artikel im Volltext zu lesen interpretiert. Entgegen unserer Hypothese konnten wir keinen signifikanten Unterschieden in der Studienteilnahme zwischen den Gruppen finden (kurzer Abstract: 35,0 % vs. langer Abstract: 37,8 %). Wir konnten jedoch zeigen, dass kürzere Abstracts häufiger als attraktiver und prägnanter wahrgenommen wurden.¹²⁴ Diese Ergebnisse sind insofern widersprüchlich, als dass die Annahme nahe liegt, dass Abstracts, die als weniger attraktiv empfunden werden, zu einem schnelleren Aufmerksamkeitsverlust und häufigerem Studienabbruch führen. Eine mögliche Erklärung für diesen Widerspruch ist, dass die Abstracts nicht mit der Intention gelesen wurden, eine Entscheidung zu treffen, sondern um weiter an der Studie teilzunehmen. Unser Studienaufbau war demnach nicht geeignet, den Entscheidungsprozess abzubilden, und unser primäres Outcome stellt keine geeignete Messgröße für die Aufmerksamkeit von Forschenden dar.

Die Länge hat in unserer dritten Publikation keinen Einfluss auf den wahrgenommenen Informationsgehalt eines Abstracts. Um informative Abstracts schreiben zu können, brauchen Autor*innen von systematischen Reviews jedoch eine gewisse Anzahl an Wörtern. So steigt die PRISMA-A Adhärenz bei Abstracts von Nicht-Cochrane Reviews mit zunehmender Länge linear an.¹²² Verschiedene Studien,^{121,154} darunter auch unsere erste Publikation,¹²² deuten darauf hin, dass ein Rahmen von 250-300 Wörter für die meisten systematischen Reviews ausreichend ist, um alle wichtigen Informationen zu berichten. Zudem befürworteten mehr als 75% der in der dritten Publikation befragten Autor*innen von systematischen Reviews, dass Abstracts diese Länge

nicht überschreiten sollten.¹²⁴ Während es für systematische Reviews keine Empfehlungen hinsichtlich der optimalen Länge von Abstracts gibt, definiert CONSORT¹⁵⁹ für RCTs seit 2008 eine Länge von 250 bis 300 Wörtern als ausreichend, um adäquat berichten zu können. Da CONSORT historisch gewissermaßen als Ausgangspunkt für die Entwicklung von PRISMA angesehen werden kann und sich beide Leitlinien in Umfang und Schwerpunktsetzung sehr ähneln, erscheint es plausibel, dass diese Länge auch für Abstracts systematischer Reviews ausreichend ist. In einigen Journals ist die zulässige Länge von Abstracts auf unter 250 Wörter begrenzt,^{160,161} was die Autor*innen unabhängig vom Studiendesign vor eine Herausforderung stellt. Um sicherzustellen, dass die Autor*innen ausreichend Platz haben, um alle relevanten Informationen zu berichten, gleichzeitig jedoch gefordert sind sich auf die wesentliche Inhalte zu beschränken, sollten anerkannte Reporting-Leitlinien wie PRISMA-A einheitlich eine Länge von 250 bis 300 Wörtern empfehlen. Diese Längenempfehlung sollte unabhängig vom Studiendesign gelten und in den Autorenrichtlinien der Journals, darunter auch Cochrane, berücksichtigt werden.

4. Fazit

Mit meiner Dissertation, die zwei Beobachtungsstudien zur PRISMA-A Adhärenz und eine randomisiert kontrollierte Studie zur Wahrnehmung von Abstracts umfasst, konnte ich die Forschungslücke im Bereich der Berichterstattung von Abstracts systematischer Reviews weiter schließen. Die wichtigsten Erkenntnisse meiner Arbeit waren, dass Abstracts von systematischen Reviews im Allgemeinen deutliches Verbesserungspotential aufweisen, sich Cochrane Reviews insbesondere in Bezug auf die Länge von Nicht-Cochrane Reviews unterscheiden und dass die Länge sowohl die Berichterstattung als auch die Wahrnehmung von Abstracts beeinflusst. Autor*innen von systematischen Reviews benötigen eine gewisse Mindestlänge, um angemessen berichten zu können. Der positive Effekt der Länge löst sich jedoch mit zunehmender Wortzahl auf. Zudem werden längere Abstracts von Forschenden als weniger präzise und attraktiv wahrgenommen.

Es spricht vieles dafür, dass die von CONSORT 2008 eingeführte Längenempfehlung für Abstracts von RCTs von 250 bis 300 Wörtern auch auf systematische Reviews übertragbar ist. Abstracts sind von besonderer Relevanz, und verfolgen das Ziel die wichtigsten Inhalte des Manuskripts in kurzer und prägnanter Form wiederzugeben. Dieses Ziel gilt losgelöst vom Studiendesign, weshalb wir es als wichtig ansehen, einheitliche Richtlinien zu etablieren, die auf alle Abstracts übertragbar sind. PRISMA-A sollte entsprechend, genau wie CONSORT, eine Abstract-Länge von 250 bis 300 Wörtern empfehlen, und Journals sollten diese Empfehlung in ihre Autorenrichtlinien integrieren. So könnten Autor*innen unabhängig vom Studiendesign dazu befähigt werden, informative Abstracts zu schreiben, und gleichzeitig dazu aufgefordert werden, sich auf die wesentlichen Inhalte zu fokussieren. Diese Empfehlung gilt auch für Cochrane Reviews, deren aktuelle Längenempfehlungen mit 700 Wörtern und maximal 1000 Wörtern deutlich höher sind.

Cochrane Reviews gelten gemeinhin als Goldstandard unter den Reviews. Seit 2013 nimmt die Anzahl neu publizierter Cochrane Reviews jedoch kontinuierlich ab. Gleichzeitig steigt die Anzahl systematischer Reviews, die in anderen hochrangigen Journals publiziert werden. Damit sinkt die Relevanz von Cochrane Reviews und bei einer Suche nach methodisch hochwertigen Reviews wird zunehmend auch nach Nicht-Cochrane Reviews gesucht. Mit Blick auf die Abstracts konnten wir zeigen, dass sich Cochrane Reviews mit ihren zum Teil sehr komplexen Standards möglicherweise selbst im Weg stehen und ihre Auffindbarkeit und

Transparenz reduzieren. Um ihrem guten Ruf auch zukünftig gerecht zu werden und wieder mehr Präsenz in der medizinischen Forschung einzunehmen, sollte Cochrane die eigenen Standards kritisch evaluieren und die Hürden bei der Erstellung und Auffindbarkeit von Cochrane Reviews abbauen.

Nicht nur die Anzahl systematischer Reviews nimmt stetig zu, auch in der Primärforschung ist ein kontinuierlicher Anstieg der Publikationen zu verzeichnen. Entsprechend steigt in beiden Bereichen die Relevanz gut geschriebener Abstracts, nicht nur in der Zielgruppe der Wissenschaftler*innen, sondern auch für anderen Nutzer*innen von Evidenz. Wir müssen entsprechend unseren Blickwinkel erweitern und Studien durchführen, die sich nicht nur auf die verschiedenen Forschungsperspektiven, sondern auch auf die Nutzung und Wahrnehmung von Abstracts außerhalb des Forschungskontextes beziehen. Ziel sollte es sein, übergeordnete Strategien und Implikationen für Abstracts abzuleiten, die unabhängig vom Studiendesign anwendbar sind und die Bedürfnisse aller Nutzer*innen berücksichtigen.

Referenzen

1. Hoffmann F, Allers K, Rombey T, et al. Nearly 80 systematic reviews were published each day: Observational study on trends in epidemiology and reporting over the years 2000-2019. *J Clin Epidemiol*. 2021;138:1-11. doi:<https://doi.org/10.1016/j.jclinepi.2021.05.022>
2. Helbach J, Pieper D, Mathes T, et al. Restrictions and their reporting in systematic reviews of effectiveness: an observational study. *BMC Med Res Methodol*. 2022;22(1):230. doi:10.1186/s12874-022-01710-w
3. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71-72. doi:10.1136/bmj.312.7023.71
4. Guyatt G, Cairns J, Churchill D, et al. Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine. *JAMA*. 1992;268(17):2420-2425. doi:10.1001/jama.1992.03490170092032
5. Oxman AD, Cook DJ, Guyatt GH. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. *JAMA*. 1994;272(17):1367-1371. doi:10.1001/jama.272.17.1367
6. Swingler GH, Volmink J, Ioannidis JPA. Number of published systematic reviews and global burden of disease: database analysis. *BMJ*. 2003;327(7423):1083-1084. doi:10.1136/bmj.327.7423.1083
7. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med*. 1997;126(5):376-380. doi:10.7326/0003-4819-126-5-199703010-00006
8. Aromataris E, Pearson A. The systematic review: an overview. *Am J Nurs*. 2014;114(3):53-58. doi:10.1097/01.NAJ.0000444496.24228.2c
9. Ioannidis JPA. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. *Milbank Q*. 2016;94(3):485-514. doi:10.1111/1468-0009.12210
10. Page MJ, Shamseer L, Altman DG, et al. Epidemiology and Reporting Characteristics of Systematic Reviews of Biomedical Research: A Cross-Sectional Study. *PLoS Med*. 2016;13(5):1-30. doi:10.1371/journal.pmed.1002028
11. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet (London, England)*. 2009;374(9683):86-89. doi:10.1016/S0140-6736(09)60329-9
12. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*. 2021;372:n160. doi:10.1136/bmj.n160
13. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and Reporting Characteristics of Systematic Reviews. *PLoS Med*. 2007;4(3):447-455. doi:10.1371/journal.pmed.0040078
14. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*. 2009;339(7716):332-336. doi:10.1136/bmj.b2535
15. Islamaj Dogan R, Murray GC, Névél A, Lu Z. Understanding PubMed user search behavior through log analysis. *Database (Oxford)*. 2009;2009:bap018. doi:10.1093/database/bap018
16. Annesley TM. The abstract and the elevator talk: a tale of two summaries. *Clin Chem*. 2010;56(4):521-524. doi:10.1373/clinchem.2009.142026
17. Beller EM, Glasziou PP, Altman DG, et al. PRISMA for Abstracts: Reporting Systematic Reviews in Journal and Conference Abstracts. *PLoS Med*. 2013;10(4):e1001419. doi:10.1371/journal.pmed.1001419
18. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet*. 2017;390(10092):415-423. doi:[https://doi.org/10.1016/S0140-6736\(16\)31592-6](https://doi.org/10.1016/S0140-6736(16)31592-6)

19. Timmermans S, Kolker ES. Evidence-based medicine and the reconfiguration of medical knowledge. *J Health Soc Behav.* 2004;45 Suppl:177-193.
20. Knezevic NN, Manchikanti L, Hirsch JA. Principles of Evidence-Based Medicine BT - Essentials of Regenerative Medicine in Interventional Pain Management. In: Navani A, Atluri S, Sanapati M, eds. Springer International Publishing; 2024:35-48. doi:10.1007/978-3-031-50357-3_2
21. Collier R. Dr. David Sackett, a giant among giants (1934-2015). *C Can Med Assoc J = J l'Association medicale Can.* 2015;187(9):640-641. doi:10.1503/cmaj.109-5072
22. Sackett DL. Evidence-based medicine. *Spine (Phila Pa 1976).* 1998;23(10):1085-1086. doi:10.1097/00007632-199805150-00001
23. Shah HM, Chung KC. Archie Cochrane and his vision for evidence-based medicine. *Plast Reconstr Surg.* 2009;124(3):982-988. doi:10.1097/PRS.0b013e3181b03928
24. Evans D. Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *J Clin Nurs.* 2003;12(1):77-84. doi:10.1046/j.1365-2702.2003.00662.x
25. Burns PB, Rohrich RJ, Chung KC. The levels of evidence and their role in evidence-based medicine. *Plast Reconstr Surg.* 2011;128(1):305-310. doi:10.1097/PRS.0b013e318219c171
26. Canadian Task Force on the Periodic Health Examination. The periodic health examination. Canadian Task Force on the Periodic Health Examination. *Can Med Assoc J.* 1979;121(9):1193-1254.
27. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. *Evid Based Med.* 2016;21(4):125 LP - 127. doi:10.1136/ebmed-2016-110401
28. Shaneyfelt T. Pyramids are guides not rules: the evolution of the evidence pyramid. *BMJ Evidence-Based Med.* 2016;21(4):121-122. doi:10.1136/ebmed-2016-110498
29. Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest.* 1986;89(2 Suppl):2S-3S.
30. Milano G. The hierarchy of the evidence-based medicine pyramid: classification beyond ranking. *Joints.* 2015;3(3):101. doi:10.11138/jts/2015.3.3.101
31. Röhrig B, du Prel JB, Wachtlin D, Blettner M. Types of study in medical research: part 3 of a series on evaluation of scientific publications. *Dtsch Arztebl Int.* 2009;106(15):262-268. doi:10.3238/arztebl.2009.0262
32. Munnangi S, Boktor SW. Epidemiology Of Study Design. In: *StatPearls [Internet]*. StatPearls Publishing; 2025. <https://www.ncbi.nlm.nih.gov/books/NBK470342/>
33. Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. *Lancet (London, England).* 2002;359(9300):57-61. doi:10.1016/S0140-6736(02)07283-5
34. Kapoor MC. Types of studies and research design. *Indian J Anaesth.* 2016;60(9):626-630. doi:10.4103/0019-5049.190616
35. Lelgemann M, Donner-Banzhoff N. Wo erhalte ich Antworten auf meine Fragen? In: Kunz R, Ollenschläger G, Raspe H, Jonitz G, Donner-Banzhoff N, eds. *Lehrbuch Evidenz-Basierte Medizin in Klinik Und Praxis.* Deutscher Ärzte-Verlag; 2007.
36. Egger M, Smith GD, O'Rourke K. Introduction: Rationale, Potentials, and Promise of Systematic Reviews. In: *Systematic Reviews in Health Care.* John Wiley & Sons, Ltd; 2001:1-19. doi:<https://doi.org/10.1002/9780470693926.ch1>
37. Chalmers I, Hedges L V, Cooper H. A brief history of research synthesis. *Eval Health Prof.* 2002;25(1):12-37. doi:10.1177/0163278702025001003

38. O'Rourke K. An historical perspective on meta-analysis: dealing quantitatively with varying study results. *J R Soc Med.* 2007;100(12):579-582. doi:10.1177/0141076807100012020
39. Maxwell MR. A Mathematical Perspective on Gambling. *MIT Undergrad J Math.* Published online 1999. <https://api.semanticscholar.org/CorpusID:17485172>
40. Airy GB. *On the Algebraical and Numerical Theory of Errors of Observations and the Combination of Observations.* London, Macmillan; 1861. <http://eudml.org/doc/204681>
41. Pearson K. Report on Certain Enteric Fever Inoculation Statistics. *Br Med J.* 1904;2(2288):1243 LP - 1246. doi:10.1136/bmj.2.2288.1243
42. Fisher R. *The Design Of Experiments.* Edinburgh: Oliver and Boyd; 1935.
43. Cochran WG. The combination of estimates from different experiments. *Biometrics.* 1954;10:101-129. doi:10.2307/3001666
44. Jones L V, Fiske DW. Models for testing the significance of combined results. *Psychol Bull.* 1953;50(5):375-382. doi:10.1037/h0057844
45. Lee WL, Bausell RB, Berman BM. The growth of health-related meta-analyses published from 1980 to 2000. *Eval Health Prof.* 2001;24(3):327-335. doi:10.1177/01632780122034948
46. Hedges L, Olkin I. *Statistical Methods for Meta-Analysis.* Academic Press; 1985.
47. Glas G V. Primary, Secondary, and Meta-Analysis of Research. *Educ Res.* 1976;5(10):3-8. doi:10.3102/0013189X005010003
48. Meldrum ML. A brief history of the randomized controlled trial: From oranges and lemons to the gold standard. *Hematol Oncol Clin North Am.* 2000;14(4):745-760. doi:[https://doi.org/10.1016/S0889-8588\(05\)70309-9](https://doi.org/10.1016/S0889-8588(05)70309-9)
49. Elwood PC, Cochrane AL, Burr ML, et al. A randomized controlled trial of acetyl salicylic acid in the secondary prevention of mortality from myocardial infarction. *Br Med J.* 1974;1(5905):436-440. doi:10.1136/bmj.1.5905.436
50. Elwood P. The first randomized trial of aspirin for heart attack and the advent of systematic overviews of trials. *J R Soc Med.* 2006;99(11):586-588. doi:10.1177/014107680609901121
51. Peto R. Why do we need systematic overviews of randomized trials? *Stat Med.* 1987;6(3):233-244. doi:10.1002/sim.4780060306
52. Quirke V, Gaudillière JP. The era of biomedicine: science, medicine, and public health in Britain and France after the Second World War. *Med Hist.* 2008;52(4):441-452. doi:10.1017/s002572730000017x
53. Valier H, Timmermann C. Clinical trials and the reorganization of medical research in post-Second World War Britain. *Med Hist.* 2008;52(4):493-510. doi:10.1017/s0025727300002994
54. Breuer C, Meerpohl JJ, Siemens W. From standard systematic reviews to living systematic reviews. *Z Evid Fortbild Qual Gesundheitsw.* 2023;176:76-81. doi:<https://doi.org/10.1016/j.zefq.2022.11.007>
55. Cochrane AL. *Effectiveness and Efficiency: Random Reflections on Health Services.* Nuffield Trust; 1972. <https://api.semanticscholar.org/CorpusID:59534091>
56. Stavrou A, Challoumas D, Dimitrakakis G. Archibald Cochrane (1909-1988): the father of evidence-based medicine. *Interact Cardiovasc Thorac Surg.* 2014;18(1):121-124. doi:10.1093/icvts/ivt451
57. Macleod S. Cochrane's problem: psychoanalysis and anejaculation. *Australas psychiatry Bull R Aust New Zeal Coll Psychiatr.* 2007;15(2):144-147. doi:10.1080/10398560701200117
58. Purssell E, McCrae N. A Brief History of the Systematic Review. In: *How to Perform a Systematic*

Literature Review: A Guide for Healthcare Researchers, Practitioners and Students. Springer International Publishing; 2020:5-17. doi:10.1007/978-3-030-49672-2_2

59. Robbins R. Profiles in medical courage: evidence-based medicine and Archie Cochrane. *Southwest J Pulm Crit Care*. 2012;5:65-73. <https://api.semanticscholar.org/CorpusID:81555693>
60. Cochrane AL. Sickness in Salonica: my first, worst, and most successful clinical trial. *Br Med J (Clin Res Ed)*. 1984;289(6460):1726-1727. doi:10.1136/bmj.289.6460.1726
61. Hill MD, Silver FL, Austin PC, Tu J V. Rate of stroke recurrence in patients with primary intracerebral hemorrhage. *Stroke*. 2000;31(1):123-127. doi:10.1161/01.str.31.1.123
62. Randal J. Austin Bradford Hill: A Pioneering Force Behind Clinical Trials. *JNCI J Natl Cancer Inst*. 1999;91(1):11. doi:10.1093/jnci/91.1.11
63. Cochrane AL, Teeling-Smith G, Wells NE. 1931-1971: A critical review, with particular reference to the medical profession. In: *Medicines for the Year 2000*. Office for Health Economics; 1979:1-11.
64. Sukhera J. Narrative Reviews: Flexible, Rigorous, and Practical. *J Grad Med Educ*. 2022;14(4):414-417. doi:10.4300/JGME-D-22-00480.1
65. Haddaway NR, Bethel A, Dicks L V, et al. Eight problems with literature reviews and how to fix them. *Nat Ecol Evol*. 2020;4(12):1582-1589. doi:10.1038/s41559-020-01295-x
66. Light RJ, Pillemer DB. *Summing Up: The Science of Reviewing Research*. Harvard University Press; 1984. <http://www.jstor.org/stable/j.ctvk12px9>
67. Cooper HM. Scientific Guidelines for Conducting Integrative Research Reviews. *Rev Educ Res*. 1982;52(2):291-302. doi:10.3102/00346543052002291
68. Jackson GB. Methods for Integrative Reviews. *Rev Educ Res*. 1980;50(3):438-460. doi:10.3102/00346543050003438
69. Kass E. Reviewing reviews. In: Warren K, ed. *Coping with the Biomedical Literature*. Praeger; 1981:79-91.
70. Mulrow CD. The medical review article: state of the science. *Ann Intern Med*. 1987;106(3):485-488. doi:10.7326/0003-4819-106-3-485
71. Pillemer DB. Conceptual Issues in Research Synthesis. *J Spec Educ*. 1984;18(1):27-40. doi:10.1177/002246698401800105
72. Oxman AD, Guyatt GH. Guidelines for reading literature reviews. *C Can Med Assoc J = J l'Association medicale Can*. 1988;138(8):697-703.
73. Warren K, ed. *Coping with the Biomedical Literature*. Praeger; 1981.
74. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med*. 1987;107(2):224-233. doi:10.7326/0003-4819-107-2-224
75. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. *N Engl J Med*. 1987;316(8):450-455. doi:10.1056/NEJM198702193160806
76. Hampton JR. The use of beta blockers for the reduction of mortality after myocardial infarction. *Eur Heart J*. 1981;2(4):259-268. doi:10.1093/oxfordjournals.eurheartj.a061206
77. Egger M, Smith GD, Altman DG, eds. *Systematic Reviews in Health Care: Meta-Analysis in Context (2nd Ed.)*. BMJ Books; 2001. <https://onlinelibrary.wiley.com/doi/book/10.1002/9780470693926>
78. Feinleib M. A Dictionary of Epidemiology, Fourth Edition - Edited by John M. Last, Robert A. Spasoff, and Susan S. Harris. *Am J Epidemiol*. 2001;154(1):93-94. doi:10.1093/aje/154.1.93-a

79. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4(1):148-160. doi:10.1186/2046-4053-4-1
80. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *JAMA*. 1992;268(2):240-248.
81. Sheldon T, Chalmers I. The UK Cochrane Centre and the NHS Centre for reviews and dissemination: respective roles within the information systems strategy of the NHS R&D programme, coordination and principles underlying collaboration. *Health Econ*. 1994;3(3):201-203. doi:10.1002/hec.4730030308
82. Harrison J. Research and development in the NHS. *Psychiatr Bull*. 1993;17(5):291-292. doi:10.1192/pb.17.5.291
83. Chalmers I, Dickersin K, Chalmers TC. Getting to grips with Archie Cochrane's agenda. *BMJ*. 1992;305(6857):786-788. doi:10.1136/bmj.305.6857.786
84. Green S, McDonald S. Cochrane Collaboration: more than systematic reviews? *Intern Med J*. 2005;35(1):3-5. doi:https://doi.org/10.1111/j.1445-5994.2004.00747.x
85. Chalmers I. The Cochrane collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Ann N Y Acad Sci*. 1993;703:155-156. doi:10.1111/j.1749-6632.1993.tb26345.x
86. Cumpston M, Flemyng E, Thomas J, Higgins J, Deeks J, Clarke M. Chapter I: Introduction. In: Higgins J, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.5*. Cochrane; 2024. <https://training.cochrane.org/handbook/current/chapter-i>
87. Tovey D. The role of The Cochrane Collaboration in support of the WHO Nutrition Guidelines. *Adv Nutr*. 2014;5(1):35-39. doi:10.3945/an.113.004895
88. Clarke M. The Cochrane Collaboration. In: Machin D, Day S, Green S, eds. *Textbook of Clinical Trials*. John Wiley & Sons, Ltd; 2006:39-46. doi:https://doi.org/10.1002/9780470010167.ch3
89. Chandler J, Hopewell S. Cochrane methods--twenty years experience in developing systematic review methods. *Syst Rev*. 2013;2:76. doi:10.1186/2046-4053-2-76
90. The Cochrane Collaboration. *Methodological Expectations of Cochrane Intervention Reviews (MECIR): Standards for the Conduct and Reporting of New Cochrane Intervention Reviews 2012*. Cochrane; 2012. [https://community.cochrane.org/sites/default/files/uploads/1MECIR standards booklet.pdf](https://community.cochrane.org/sites/default/files/uploads/1MECIR%20standards%20booklet.pdf)
91. Altman DG, Simera I. A history of the evolution of guidelines for reporting medical research: the long road to the EQUATOR Network. *J R Soc Med*. 2016;109(2):67-77. doi:10.1177/0141076815625599
92. Institute of Medicine (US) Committee on Standards for Systematic Reviews of Comparative Effectiveness Research. *Finding What Works in Health Care: Standards for Systematic Reviews*. (Eden J, Levit L, Berg A, Morton S, eds.). National Academies Press; 2011. doi:10.17226/13059
93. Delaney A, Bagshaw SM, Ferland A, Manns B, Laupland KB, Doig CJ. A systematic evaluation of the quality of meta-analyses in the critical care literature. *Crit Care*. 2005;9(5):R575-82. doi:10.1186/cc3803
94. Golder S, Loke Y, McIntosh HM. Poor reporting and inadequate searches were apparent in systematic reviews of adverse effects. *J Clin Epidemiol*. 2008;61(5):440-448. doi:10.1016/j.jclinepi.2007.06.005
95. Begg C, Cho M, Eastwood S, et al. Improving the Quality of Reporting of Randomized Controlled Trials: The CONSORT Statement. *JAMA*. 1996;276(8):637-639. doi:10.1001/jama.1996.03540080059030

96. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Quality of Reporting of Meta-analyses. Lancet (London, England)*. 1999;354(9193):1896-1900. doi:10.1016/S0140-6736(99)04149-5
97. Clarke M. The QUORUM statement. *Lancet (London, England)*. 2000;355(9205):756-757. doi:10.1016/S0140-6736(05)72172-3
98. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med*. 2009;6(7):e1000100. doi:10.1371/journal.pmed.1000100
99. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLoS Med*. 2021;18(3):1-15. doi:10.1371/JOURNAL.PMED.1003583
100. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med*. 2010;7(9):e1000326. doi:10.1371/journal.pmed.1000326
101. Page MJ, Moher D. Mass Production of Systematic Reviews and Meta-analyses: An Exercise in Megascilliness? *Milbank Q*. 2016;94(3):515-519. doi:10.1111/1468-0009.12211
102. Niforatos JD, Weaver M, Johansen ME. Assessment of Publication Trends of Systematic Reviews and Randomized Clinical Trials, 1995 to 2017. *JAMA Intern Med*. 2019;179(11):1593-1594. doi:10.1001/jamainternmed.2019.3013
103. Sun X, Zhou X, Yu Y, Liu H. Exploring reporting quality of systematic reviews and Meta-analyses on nursing interventions in patients with Alzheimer's disease before and after PRISMA introduction. *BMC Med Res Methodol*. 2018;18(1):154. doi:10.1186/s12874-018-0622-7
104. Rice DB, Kloda LA, Shrier I, Thombs BD. Reporting completeness and transparency of meta-analyses of depression screening tool accuracy: A comparison of meta-analyses published before and after the PRISMA statement. *J Psychosom Res*. 2016;87:57-69. doi:https://doi.org/10.1016/j.jpsychores.2016.06.007
105. Munn Z, Stern C, Aromataris E, Lockwood C, Jordan Z. What kind of systematic review should i conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. *BMC Med Res Methodol*. 2018;18(1):1-9. doi:10.1186/s12874-017-0468-4
106. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. *JBIM Evid Synth*. 2020;18(10):2119-2126. doi:10.11124/JBIES-20-00167
107. Pearson A. Balancing the evidence: incorporating the synthesis of qualitative data into systematic reviews. *JBIM Reports*. 2004;2(2):45-64. doi:https://doi.org/10.1111/j.1479-6988.2004.00008.x
108. Gough D, Thomas J, Oliver S. *An Introduction to Systematic Reviews*. SAGE Publications Ltd; 2017.
109. Gough D, Thomas J, Oliver S. Clarifying differences between review designs and methods. *Syst Rev*. 2012;1:28. doi:10.1186/2046-4053-1-28
110. Bahadoran Z, Mirmiran P, Kashfi K, Ghasemi A. The Principles of Biomedical Scientific Writing: Abstract and Keywords. *Int J Endocrinol Metab*. 2020;18(1):e100159. doi:10.5812/ijem.100159
111. Squires BP. Structured abstracts of original research and review articles. *C Can Med Assoc J = J l'Association medicale Can*. 1990;143(7):619-622.
112. Shiely F, Gallagher K, Millar SR. How, and why, science and health researchers read scientific (IMRAD) papers. *PLoS One*. 2024;19(1):e0297034. https://doi.org/10.1371/journal.pone.0297034
113. Tullu MS. Writing the title and abstract for a research paper: Being concise, precise, and meticulous is the key. *Saudi J Anaesth*. 2019;13(Suppl 1):S12-S17. doi:10.4103/sja.SJA_685_18

114. Cals JWL, Kotz D. Effective writing and publishing scientific papers, part II: title and abstract. *J Clin Epidemiol.* 2013;66(6):585. doi:10.1016/j.jclinepi.2013.01.005
115. Adobes Martin M, Santamans Faustino S, Llario Almiñana I, Aiuto R, Rotundo R, Garcovich D. There is still room for improvement in the completeness of abstract reporting according to the PRISMA-A checklist: a cross-sectional study on systematic reviews in periodontology. *BMC Med Res Methodol.* 2021;21(1):1-12. doi:10.1186/s12874-021-01223-y
116. Nascimento DP, Gonzalez GZ, Araujo AC, Moseley A, Maher C, Costa LOP. Factors associated with the reporting quality of low back pain systematic review abstracts in physical therapy: a methodological study. *Braz J Phys Ther.* 2021;25(3):233-241. doi:10.1016/j.bjpt.2020.10.002
117. Maticic K, Krnic Martinic M, Puljak L. Assessment of reporting quality of abstracts of systematic reviews with meta-analysis using PRISMA-A and discordance in assessments between raters without prior experience. *BMC Med Res Methodol.* 2019;19(1):32. doi:10.1186/s12874-019-0675-2
118. Vásquez-Cárdenas J, Zapata-Noreña Ó, Carvajal-Flórez Á, Barbosa-Liz DM, Giannakopoulos NN, Faggion Jr. CM. Systematic reviews in orthodontics: Impact of the PRISMA for Abstracts checklist on completeness of reporting. *Am J Orthod Dentofac Orthop.* 2019;156(4):442-452.e12. doi:10.1016/j.ajodo.2019.05.009
119. O'Donohoe TJ, Dhillon R, Bridson TL, Tee J. Reporting Quality of Systematic Review Abstracts Published in Leading Neurosurgical Journals: A Research on Research Study. *Neurosurgery.* 2019;85(1):1-10. doi:10.1093/neuros/nyy615
120. Jiancheng W, Jinhui T, Lin H, Yuxia M, Juxia Z. Has the Reporting Quality of Systematic Review Abstracts in Nursing Improved Since the Release of PRISMA for Abstracts? A Survey of High-Profile Nursing Journals. *Worldviews evidence-based Nurs.* 2020;17(2):108-117. doi:10.1111/wvn.12414
121. Bigna JJR, Um LN, Nansseu JRN. A comparison of quality of abstracts of systematic reviews including meta-analysis of randomized controlled trials in high-impact general medicine journals before and after the publication of PRISMA extension for abstracts: A systematic review and meta-an. *Syst Rev.* 2016;5(1):1-9. doi:10.1186/s13643-016-0356-8
122. Helbach J, Hoffmann F, Pieper D, Allers K. Reporting according to the preferred reporting items for systematic reviews and meta-analyses for abstracts (PRISMA-A) depends on abstract length. *J Clin Epidemiol.* 2022;154:167-177. doi:10.1016/j.jclinepi.2022.12.019
123. Helbach J, Wandscher K, Pieper D, Hoffmann F. Cochrane review abstracts are getting longer, but this has no large impact on the reporting quality. *J Evid Based Med.* 2023;16(3):294-302. doi:10.1111/jebm.12545
124. Helbach J, Wandscher K, Pieper D, Hoffmann F. Readers' attention to shorter versus longer abstracts of systematic reviews: a randomised controlled trial. *BMJ evidence-based Med.* Published online October 2025. doi:10.1136/bmjebm-2024-113613
125. Soderberg L, Ergun O, Ding M, et al. Percutaneous nephrolithotomy versus retrograde intrarenal surgery for treatment of renal stones in adults. *Cochrane database Syst Rev.* 2023;11(11):CD013445. doi:10.1002/14651858.CD013445.pub2
126. Soderberg L, Ergun O, Ding M, et al. Percutaneous nephrolithotomy vs retrograde intrarenal surgery for renal stones: a Cochrane Review. *BJU Int.* 2024;133(2):132-140. doi:10.1111/bju.16220
127. Hopewell S, Clarke M, Moher D, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet.* 2008;371(9609):281-283. doi:10.1016/s0140-6736(07)61835-2
128. Tsou AY, Treadwell JR. Quality and clarity in systematic review abstracts: an empirical study. *Res Synth Methods.* 2016;7(4):447-458. doi:10.1002/JRSM.1221

129. Benchimol EI, Smeeth L, Guttman A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLOS Med.* 2015;12(10):e1001885. <https://doi.org/10.1371/journal.pmed.1001885>
130. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ.* 2007;335(7624):806-808. doi:10.1136/bmj.39335.541782.AD
131. Wandscher K, Helbach J, Pieper D, Hoffmann F. “We used standard Cochrane methods” – observational study on reporting according to PRISMA-A in Cochrane review abstracts between 2016 and 2023. *J Clin Epidemiol.* 2025;181(111713). doi:<https://doi.org/10.1016/j.jclinepi.2025.111713>
132. 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(1):e20. doi:10.1371/journal.pmed.0050020
133. Germini F, Marcucci M, Fedele M, et al. Quality of reporting in abstracts of RCTs published in emergency medicine journals: a systematic survey of the literature suggests we can do better. *Emerg Med J.* Published online November 2019:660-665. doi:10.1136/emered-2019-208629
134. Gopinath VK, Shetty RM, Renugalakshmi A, Dharmarajan L, Prakash PSG, Jayaraman J. Reporting Quality of the Abstracts for Randomized Controlled Trials in Pediatric Dentistry. *Eur J Dent.* 2023;18(1):341-348. doi:10.1055/s-0043-1770912
135. Khan MS, Shaikh A, Ochani RK, et al. Assessing the Quality of Abstracts in Randomized Controlled Trials Published in High Impact Cardiovascular Journals. *Circ Cardiovasc Qual Outcomes.* 2019;12(5):e005260. doi:10.1161/CIRCOUTCOMES.118.005260
136. Villa M, Le Pera M, Cassina T, Bottega M. Reporting quality of abstracts from randomised controlled trials published in leading critical care nursing journals: a methodological quality review. *BMJ Open.* 2023;13(3):e070639. doi:10.1136/bmjopen-2022-070639
137. Vrebalov Cindro P, Bukic J, Pranić S, et al. Did an introduction of CONSORT for abstracts guidelines improve reporting quality of randomised controlled trials’ abstracts on *Helicobacter pylori* infection? Observational study. *BMJ Open.* 2022;12(3):e054978. doi:10.1136/bmjopen-2021-054978
138. Higgins J, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.4 (Updated August 2023)*. Cochrane; 2023. www.training.cochrane.org/handbook
139. The Cochrane Collaboration. Cochrane Style Manual (Version 6). 2024. <https://community.cochrane.org/style-manual>
140. Pieper D, Hoffmann F. Retrieving Cochrane reviews is sometimes challenging and their reporting is not always optimal. *Res Synth Methods.* Published online 2022:1-4. doi:10.1002/JRSM.1564
141. Goldkuhle M, Narayan VM, Weigl A, Dahm P, Skoetz N. A systematic assessment of Cochrane reviews and systematic reviews published in high-impact medical journals related to cancer. *BMJ Open.* 2018;8(3):e020869. doi:10.1136/bmjopen-2017-020869
142. Rosengaard LO, Andersen MZ, Rosenberg J, Fonnes S. Citation patterns of Cochrane Reviews and other systematic reviews: a bibliometric analysis. *Curr Med Res Opin.* Published online January 2025:1-9. doi:10.1080/03007995.2024.2442045
143. Cumpston M, Chandler J. Chapter II: Planning a Cochrane Review. In: Higgins J, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.3 (Updated February 2022)*. Cochrane; 2022.
144. Cumpston M, Lasserson T, Chandler J, Page MJ. Chapter III: Reporting the review. In: Higgins J, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.3*

(Updated February 2022). Cochrane; 2022.

145. Page MJ, Cumpston M, Chandler J, Lasserson T. Chapter III: Reporting the review. In: Higgins J, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.0*. Cochrane; 2019.
146. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64(4):383-394. doi:10.1016/j.jclinepi.2010.04.026
147. Lawrenson JG, Huntjens B, Virgili G, et al. Interventions for myopia control in children: a living systematic review and network meta-analysis. *Cochrane database Syst Rev*. 2025;2(2):CD014758. doi:10.1002/14651858.CD014758.pub3
148. Butler AR, Lindson N, Livingstone-Banks J, et al. Interventions for quitting vaping. *Cochrane database Syst Rev*. 2025;1(1):CD016058. doi:10.1002/14651858.CD016058.pub2
149. Yunas I, Gallos ID, Devall AJ, et al. Tests for diagnosis of postpartum haemorrhage at vaginal birth. *Cochrane database Syst Rev*. 2025;1(1):CD016134. doi:10.1002/14651858.CD016134
150. Moseley AM, Elkins MR, Herbert RD, Maher CG, Sherrington C. Cochrane reviews used more rigorous methods than non-Cochrane reviews: survey of systematic reviews in physiotherapy. *J Clin Epidemiol*. 2009;62(10):1021-1030. doi:10.1016/j.jclinepi.2008.09.018
151. Collier A, Heilig L, Schilling L, Williams H, Dellavalle RP. Cochrane Skin Group systematic reviews are more methodologically rigorous than other systematic reviews in dermatology. *Br J Dermatol*. 2006;155(6):1230-1235. doi:10.1111/j.1365-2133.2006.07496.x
152. Jones AP, Remington T, Williamson PR, Ashby D, Smyth RL. High prevalence but low impact of data extraction and reporting errors were found in Cochrane systematic reviews. *J Clin Epidemiol*. 2005;58(7):741-742. doi:10.1016/j.jclinepi.2004.11.024
153. Ji Z, Zhang J, Menniti-Ippolito F, et al. The quality of Cochrane systematic reviews of acupuncture: an overview. *BMC Complement Med Ther*. 2020;20(1):307. doi:10.1186/s12906-020-03099-9
154. Jia PL, Xu B, Cheng JM, et al. Assessment of the abstract reporting of systematic reviews of dose-response meta-analysis: A literature survey. *BMC Med Res Methodol*. 2019;19(1):1-8. doi:10.1186/s12874-019-0798-5
155. Li T, Hua F, Dan S, Zhong Y, Levey C, Song Y. Reporting quality of systematic review abstracts in operative dentistry: An assessment using the PRISMA for Abstracts guidelines. *J Dent*. 2020;102:103471. doi:10.1016/J.JDENT.2020.103471
156. Karjalainen T V, Lusa V, Page MJ, O'Connor D, Massy-Westropp N, Peters SE. Splinting for carpal tunnel syndrome. *Cochrane database Syst Rev*. 2023;2(2):CD010003. doi:10.1002/14651858.CD010003.pub2
157. Pasquali S, Hadjinicolaou A V, Chiarion Sileni V, Rossi CR, Mocellin S. Systemic treatments for metastatic cutaneous melanoma. *Cochrane database Syst Rev*. 2018;2(2):CD011123. doi:10.1002/14651858.CD011123.pub2
158. The Cochrane Collaboration. Cochrane's sustainable path to open access. 2024. <https://www.cochrane.org/about-us/news/cochranes-sustainable-path-open-access>
159. 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(1):e20. doi:10.1371/journal
160. Eungi K, Yong-Gu L. Comparison of length limits and the actual length of abstracts in pharmacology, oncology, and neurology journals listed in PubMed. *Sci Ed*. 2021;8(2):39-46.

161. Silverberg O, Ray JG. Variations in Instructed vs. Published Word Counts in Top Five Medical Journals. *J Gen Intern Med*. 2018;33(1):16-18. doi:10.1007/s11606-017-4196-6

Anhang

Anhang 1 Erklärung

Anhang 2 Publikationen

Anhang 1. Erklärung

Hiermit versichere ich, dass ich Jasmin Helbach die vorliegende Dissertation selbständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe. Die den verwendeten Werken inhaltlich oder wörtlich entnommenen Stellen sind als solche kenntlich gemacht und alle Quellen sind im Literaturverzeichnis angegeben.

Als Erstautorin war ich bei allen Einzelarbeiten, die Bestandteil dieser kumulativen Dissertation sind, maßgeblich für die Konzeption der Studie, die Datenerhebung und -auswertung (in Zusammenarbeit mit den Ko-Autor*innen), die Erstellung des Manuskripts, die Abstimmung mit den Ko-Autor*innen sowie die Einreichung und die Überarbeitung der Artikel im Peer-Review-Verfahren verantwortlich. Für die Publikationen, bei denen ich Ko-Autorin bin und die in dieser Dissertation zitiert werden, war ich substantiell an der Konzeptualisierung, der Datenerhebung und/oder -auswertung sowie der Abstimmung mit dem Hauptautor/der Hauptautorin und den anderen Ko- Autor*innen über die Erstellung und Überarbeitung des Manuskripts beteiligt.

Oldenburg, 03.11.2025

Jasmin Helbach

Anhang 2. Eingeschlossene Einzelarbeiten

Helbach J, Hoffmann F, Pieper D, Allers K. Reporting according to the preferred reporting items for systematic reviews and meta-analyses for abstracts (PRISMA-A) depends on abstract length. *J Clin Epidemiol.* 2022;154:167-177. doi:10.1016/j.jclinepi.2022.12.019

Helbach J, Wandscher K, Pieper D, Hoffmann F. Cochrane review abstracts are getting longer, but this has no large impact on the reporting quality. *J Evid Based Med.* 2023;16(3):294-302. doi:10.1111/jebm.12545

Helbach J, Wandscher K, Pieper D, Hoffmann F. Readers' attention to shorter versus longer abstracts of systematic reviews: a randomised controlled trial. *BMJ Evid Based Med.* 2025 Oct 17:bmjebm-2024-113613. doi: 10.1136/bmjebm-2024-113613.

Hinweis: Die publizierten Originalarbeiten werden aus urheberrechtlichen Gründen nicht veröffentlicht. Alle Arbeiten können unter den eingefügten DOI online eingesehen werden.