

# Multidisziplinärer anonymer Survey zur Vertrautheit von Ärztinnen und Ärzten eines Universitätsklinikums mit gängigen Termini und Kenngrößen der evidenzbasierten Medizin

## Multi-disciplinary anonymous survey on a university hospital's medical doctors' knowledge of established terms and parameters in evidence-based medicine



### Autoren

Christine Baulig, Sünje-Theresa Ahrens, Sabrina Tulka, Frank Krummenauer



### Institut

Institut für Medizinische Biometrie und Epidemiologie der Universität Witten/Herdecke

### Schlüsselwörter

akademische Nachwuchsförderung in der Medizin, evidenzbasierte Medizin, Weiterbildungsbedarf

### Key words

medical junior researchers education, evidence-based medicine, demand for education

### Bibliografie

DOI <https://doi.org/10.1055/a-0961-7073>

Online-Publikation: 27.8.2019

Dtsch Med Wochenschr 2019; 144: e138–e144

© Georg Thieme Verlag KG, Stuttgart · New York

ISSN 0012-0472

### Korrespondenzadresse

Prof. Dr. Christine Baulig MSc.

Institut für Medizinische Biometrie und Epidemiologie (IMBE)

Department für Humanmedizin, Fakultät für Gesundheit der

Universität Witten/Herdecke, Alfred-Herrhausen-Straße 50,

D-58448 Witten

Christine.Baulig@uni-wh.de

### ZUSAMMENFASSUNG

**Hintergrund** Eine Patientenversorgung nach den Grundsätzen der evidenzbasierten Medizin (EbM) ist im Sozialgesetzbuch (u. a. § 137f SGB V) festgelegt. Im Rahmen der Qualitätssicherung in der Medizin wird deshalb immer wieder eine Pflicht-Weiterbildung für Ärztinnen und Ärzte im Bereich EbM gefordert. Es stellt sich die Frage, inwieweit aktuell klinisch tätige Ärztinnen und Ärzten bereits mit gängigen Termini und Kenngrößen der EbM vertraut sind. Ebenso ist die optimale zeitliche Ansiedlung einer möglichen Weiterbildung während der ärztlichen Ausbildung unklar.

**Material und Methoden** Im Februar 2017 wurde an 4 Kliniken der Universität Witten/Herdecke ein anonymer Survey

(Papier-Fragebogen) durchgeführt. Der Fragebogen bestand aus 12 ausformulierten Aufgaben zu klassischen Methoden der EbM. Alle Aufgaben waren mit „richtig“ oder „falsch“ zu beantworten; das Nichtbeantworten einzelner Aufgaben war gestattet. Primärer Endpunkt der Untersuchung war der individuelle Anteil korrekt beantworteter unter den gegebenen Antworten. Die primäre Fragestellung des Surveys bestand in der Schätzung des medianen Anteils korrekter unter den vorliegenden Antworten der Teilnehmenden mittels eines 2-seitigen 95%-Konfidenzintervalls.

**Ergebnisse** Aus den 4 Kliniken resultierten 70 auswertbare Fragebögen (15–19 pro Klinik). Der mediane Anteil korrekter unter den gegebenen Antworten lag bei 47 % (95 %-KI [40 %; 50 %]). Die 4 Kliniken unterschieden sich zum Niveau 5 % bezüglich des Anteils korrekter Antworten nicht signifikant ( $p = 0,388$ ), konkret zeigten sich mediane Anteile zwischen 40 % und 50 % korrekter unter den gegebenen Antworten in den 4 Kliniken.

**Diskussion** Dem klinisch tätigen Personal an Universitätskliniken kann eine gewisse Vertrautheit mit gängigen Methoden der klinischen Forschung unterstellt werden. Der Anteil korrekter Antworten lag jedoch unter 50 %, sodass eine Weiterbildung zu EbM-spezifischen Methoden indiziert erscheint. Hinweise zum optimalen Zeitpunkt solcher Angebote während der ärztlichen Ausbildung konnten nicht aus dem Survey abgeleitet werden.

### ABSTRACT

**Background** Patient care according to the principles of evidence-based medicine (EbM) is defined by the Social Code (SGB, inter alia § 137f SGB V). As part of quality assurance in medicine, a mandatory further education for doctors is demanded. The aim of this investigation was to evaluate how familiar medical practitioners are with common terms and parameters of EbM. Another question was, at which point of the medical training a possible postgraduate training on EbM methods should take place.

**Material and Methods** In February 2017, an anonymous survey (paper questionnaire) was conducted at four clinics of the Witten/Herdecke University. The questionnaire offered 12 application scenarios for EbM methods; all items could be clearly

answered with “right” or “wrong”; the refusal of an answer to single items was allowed. The investigation’s primary endpoint was the individual proportion [%] of correctly among all available answers for each survey participant. The primary analysis sought to estimate this proportion of correct answers by means of a two-sided 95 % confidence interval.

Results A total of 70 evaluable questionnaires could be evaluated (15–19 per clinic). The median individual proportion of correct answers was 47 % (95 % CI 40 %; 50 %). The four clinics did not differ significantly with respect to this proportion of

correct answers (Kruskal/Wallis  $p=0.388$ ) and showed median proportions between 40 % and 50 %.

**Discussion** Medical practitioners at university hospitals can be expected to have certain knowledge on common methods of clinical research. However, less than 50 % of the answers were correct and a need for further training in EbM-specific methods appears to be indicated. The survey did not provide information on the optimal timing of such offers during medical training.

## Einleitung

Das Gesundheitssystem in Deutschland hat seit den 1990er-Jahren unter dem Blickwinkel verschiedenster Entwicklungen einen stetigen Wandel vollzogen, womit auch die Forderung nach einer evidenzbasierten Medizin (EbM) aufwuchs. Sackett et al. definierten 1997 die EbM als den „gewissenhaften, ausdrücklichen und umsichtigen Gebrauch der gegenwärtig besten externen, wissenschaftlichen Evidenz für Entscheidungen in der medizinischen Versorgung individueller Patienten. Die Praxis der EbM bedeutet die Integration individueller klinischer Expertise mit der bestmöglichen externen Evidenz aus systematischer Forschung“ [1]. Dabei ist es der Anspruch an den Arzt im Sinne der EbM, auf eine spezifische Fragestellung, mit welcher sich im klassischen Fall der Patient an ihn richtet, den bestmöglich fundierten Behandlungsvorschlag, also die bestmögliche Antwort im Sinne der EbM zu machen. Dies impliziert die Fähigkeit, sachgerechte Literatur aufzufinden, die in identifizierten Publikationen enthaltenen Aussagen „richtig zu lesen“ und dabei die Belastbarkeit der veröffentlichten Daten kritisch bewerten zu können. Dieses „Lesen und Bewerten“ erfordert aber auch grundlegendes Methodenverständnis und setzt Kenntnisse der Techniken und Methoden der Epidemiologie, medizinischen Biometrie sowie der medizinischen Informatik voraus [2], deren „klassischen“ Methoden letztlich die „neueren“ Methoden der EbM entsprungen sind.

Fest steht, dass die EbM in Form von Leitlinien und Health Technology Assessments (HTA) zwar längst ihren Einzug in die Klinik gefunden hat und auch die Nutzung der „Cochrane-Library“ eine etablierte diagnostische und therapeutische Handlungsbasis darstellt, Patienten trotzdem aber nicht stets eine evidenzbasierte Versorgung erhalten. Als eine mögliche Erklärung, warum EbM im Alltag nach wie vor eher verhalten genutzt wird, wird u. a. postuliert, dass es nur ungenügend Möglichkeiten gäbe, sich die wissenschaftliche Evidenz zu beschaffen [3]. Eine Umfrage unter deutschen Weiterbildungsassistenten ergab 2011 als Hauptgrund eine weit verbreitete Unsicherheit zum Umgang mit EbM-Kenngrößen [4]. Grund hierfür könnte wiederum sein, dass – obwohl die EbM längst in der deutschen Sozialgesetzgebung verankert ist und nach § 137 SGB V „eine Patientenversorgung nach den Grundsätzen der EbM“ fordert – die evidenzbasierte Medizin kein Bestandteil der (Muster-) Weiterbildungsordnung der Bundesärztekammer ist. International konnten Studien aus den USA [5–7], Israel [8], dem Iran [9] und Dänemark [10] belegen, dass hier bei

„Weiterbildungsassistenten“ und bei Fachärzten nur ein mäßiger Kenntnisstand bezüglich gängiger Größen der evidenzbasierten Medizin nachgewiesen werden konnte.

Seit längerem zeigt sich eine Diskussion um eventuelle Pflicht-Weiterbildungen zu Inhalten der evidenzbasierten Medizin (EbM), z. B. für Ärztinnen und Ärzte auf dem Weg zum Facharzt-Status. Unklar ist der bestmögliche Zeitpunkt für solche Weiterbildungen. Während eine Ansiedlung entsprechender Methoden-Schulungen schon im frühen klinischen Studium der Humanmedizin einleuchten kann, wird dort bisher nur bedingt zeitlicher Raum geschaffen neben den schon bestehenden Methoden-Fächern. Faktisch ist eine solche Erweiterung des studentischen Unterrichts noch weitgehend der Ausrichtung lokaler Fachvertretungen, z. B. für medizinische Biometrie und (klinische) Epidemiologie oder auch für Allgemeinmedizin, überlassen. In einer Stellungnahme der AWMF zum „Masterplan Medizinstudium 2020“ wird ausdrücklich unter der Überschrift „Wissenschaftskompetenz ist eine Schlüsselfunktion für jede ärztliche Tätigkeit“ eine EbM-Implementierung in das Medizinstudium gefordert [11]. Als Hauptbarriere der Implementierung von eigenständigen EbM-Curricula in das Medizinstudium erwies sich der Zeitmangel [12], da das Medizinstudium mit einem dichten Lehrplan stark ausgelastet ist.

In zunehmendem Umfang verfügbare postgraduelle Weiterbildungsangebote für klinisch tätiges Personal leiden unter der „Konkurrenz“ mit Fachdisziplinen-spezifischen Weiterbildungsmöglichkeiten und -pflichten, sodass nur bei ausreichend hoher Eigenmotivation von ärztlichem Personal und Abteilungsleitungen Weiterqualifikationen zur EbM wahrgenommen werden. Dabei konnten Windish et al. 2007 in einer Untersuchung an 277 Assistenzärzten demonstrieren, dass biostatistisches Training während der Studienzeit das statistische Wissen sowie die Fähigkeiten der Interpretation von evidenzbasierten Inhalten erhöhte [5].

Neben der zeitlichen Ansiedlung ist auch deren Umfang und Ausrichtung unklar: Je nach Standort bieten Methoden-Abteilungen postgraduelle Weiterbildungen von mehreren Stunden oder auch mehreren Wochen an; speziell im letzteren Kontext werden Pflicht-Weiterbildungen sicherlich umso kritischer diskutiert. Reed und Rao publizierten in den 2000er-Jahren notwendige Inhalte zu solchen Weiterbildungen, in Anlehnung an die gängigsten statistischen Konzepte und Methoden in medizinischen Zeitschriften [13, 14]. Entscheidend für eine Bemessung des möglichen Volumens solcher Weiterbildungsangebote kann jedoch

letztlich nur eine Messung des bestehenden Wissens – und damit des Weiterbildungsbedarfs – sein. Vor diesem Hintergrund sollte ein multidisziplinärer Survey unter klinisch tätigen Ärztinnen und Ärzten eines Universitätsklinikums zur Messung der Vertrautheit mit gängigen Termini und Kenngrößen sowie grafischen Methoden der EbM durchgeführt werden.

## Material und Methoden

### Erhebungsinstrument

Als Erhebungsinstrument wurde ein Survey-Fragebogen konzipiert (► **Anhang**): Dieser Fragebogen bot 12 ausformulierte Aussagen zu „klassischen“ Methoden der EbM und klinischen Epidemiologie (Interpretation und Anwendung von z. B. negativem Vorhersagewert, p-Wert oder relativem Risiko) sowie zu „neueren“ Methoden (Interpretation von Forest-Plots und Bias-Arten). Sämtliche Aussagen waren eindeutig mit „richtig“ oder „falsch“ zu bewerten. Um die zur Bearbeitung notwendige Zeit einschätzen zu können sowie die Verständlichkeit der einzelnen Fragen des Fragebogens zu testen, wurden im Rahmen der Konstruktion des Instruments mehrere Pretests durchgeführt. Rekrutiert wurden hierfür 10 freiwillige Teilnehmer aus verschiedenen Stadien der ärztlichen Ausbildung (Studenten, Ärzte direkt nach der dritten ärztlichen Prüfung und Fachärzte nach 2 Jahren Weiterbildung). Diese Teilnehmer waren von der späteren Erhebung ausgeschlossen. Auf Basis der 10 Pretests wurde die Bearbeitungszeit ermittelt und der Fragebogen hinsichtlich Verständlichkeit und Aufgabenstellungen angepasst. Einleitend enthielt der Fragebogen freiwillige Angaben zu Klinikzugehörigkeit, Weiterbildungsstatus (Fach- oder Assistenzarzt) und zum Geschlecht der Teilnehmenden.

### Erhebung

In 4 Kliniken der Universität Witten/Herdecke (darunter vertreten 2 konservative und 2 operative Disziplinen) wurde im Februar 2017 der anonyme Survey auf Basis eines Papier-Fragebogens unter allen Anwesenden einer Dienstbesprechung durchgeführt. Zur Bearbeitung stand nach einer kurzen Erläuterung zum Ziel des Surveys ein Zeitfenster von 10 Minuten zur Verfügung. Es sollten nur die Aufgaben beantwortet werden, welche eindeutig mit „richtig“ oder „falsch“ beantwortet werden konnten; das „Raten“ von Antworten war explizit nicht erwünscht, das Auslassen einzelner Aufgaben war somit möglich. Eine datenschutzrechtliche Abklärung des Vorgehens erfolgte in Abstimmung mit dem Datenschutzbeauftragten der Universität Witten/Herdecke; ein Unbedenklichkeitsschreiben zum Datenschutz vom 31.10.2016 liegt vor.

### Primärer Endpunkt und Auswertung

Primärer Endpunkt des Surveys war der individuelle Anteil [%] korrekt beantworteter unter allen gegebenen Antworten der Teilnehmenden: Wurden seitens der Teilnehmer von den 12 Aussagen nur 8 beantwortet und davon 6 korrekt, so ergab sich hieraus ein individueller Anteil von  $6/8 = 75\%$ . Die primäre Frage-

stellung des Surveys bestand in der Schätzung des medianen Anteils korrekter Antworten mittels eines 2-seitigen 95 %-Konfidenzintervalls über alle Teilnehmenden der 4 Kliniken hinweg. In einem weiteren Schritt wurde dieser mediane Anteil stratifiziert nach Kliniken und Weiterbildungsstatus der Teilnehmenden geschätzt. Ferner wurden die Verteilungen des individuellen Anteils korrekter unter den gegebenen Antworten zwischen den Kliniken mittels eines Kruskal/Wallis-Tests zum globalen Signifikanzniveau 5 % verglichen.

### Fallzahlplanung

Zur Planung der Fallzahl wurde angenommen, dass, um von einer hinreichend „guten“ Vertrautheit mit den im Fragebogen adressierten Methoden der EbM ausgehen zu können (entsprechend dem Prädikat „gut“ für einen Anteil 80 % korrekter Antworten in einer standardisierten Examens-Prüfung), im Median 80 % der „aktiv“ beantworteten Fragen korrekt zu beantworten sind. Um nun diesen Median von 80 % mit einem 95 %-Konfidenzintervall der maximalen Breite von  $\pm 7,5\%$  schätzen zu können, waren mindestens 62 auswertbare Fragebögen einzubringen (NQuery® Version 4.0 für Windows®).

## Ergebnisse

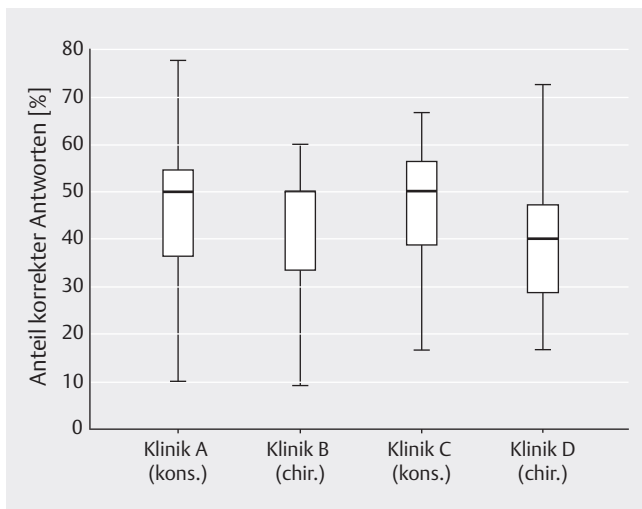
Aus den 4 teilnehmenden Kliniken wurden insgesamt 70 auswertbare Fragebögen bereitgestellt ( $n = 15 + 17 + 19 + 19$ ). Unter den Teilnehmenden befanden sich 30 Fachärzte und 36 Assistenzärzte; 4 Teilnehmende machten keine Angabe zum Weiterbildungsstatus.

Im Median beantworteten die Teilnehmenden 10 der 12 Fragen. Unter den 70 ausgewerteten Fragebögen zeigte sich über alle 12 Aufgaben hinweg ein medianer Anteil korrekter Antworten (unter den gegebenen Antworten) von 47 % (95 %-KI [40 %; 50 %]). Über die 12 Aufgaben hinweg und bezüglich des Anteils korrekter unter den gegebenen Antworten unterschieden sich die 4 teilnehmenden Kliniken nicht statistisch signifikant (Kruskal/Wallis  $p = 0,388$ ) zum Niveau 5 %. Konkret zeigten sich mediane Anteile korrekter Antworten zwischen 40 % und 50 % in den 4 Kliniken (► **Abb. 1**).

Der Anteil korrekter Antworten (unter den gegebenen) lag bei den Teilnehmenden mit Facharztstatus (30 Fragebögen) bei 42 % (95 %-KI [36 %; 50 %]) und bei den Teilnehmenden mit Assistenzarztstatus (36 Fragebögen) bei 50 % (95 %-KI [38 %; 50 %]).

Die Anteile „aktiv gegebener“ Antworten rangierten unter den 12 Aufgaben von 100 % (Aufgabe zur Spezifität) bis 61 % (Aufgabe zum Meta-Effektmaß im Forest-Plot). Es zeigte sich bei den 6 Fragen zu „klassischen“ Methoden der klinischen Epidemiologie ein hoher Anteil an „aktiv gegebenen“ Antworten (64–100 %), der Anteil korrekter unter diesen „aktiv gegebenen“ Antworten lag zwischen 25 % (Aufgabe zum relativen Risiko) und 84 % (Aufgabe zur Number Needed to Treat).

Die Anteile „aktiv gegebener“ Antworten rangierten hingegen bei den 6 Aufgaben zu „neueren“ Methoden der EbM zwischen 61 % und 71 %, darunter Anteile von 19–72 % korrekten unter diesen „aktiv gegebenen“ Antworten. Für die 4 Teilaufgaben zur Interpretation eines Forest-Plots lag der Anteil korrekter unter den



► **Abb. 1** Box-Whisker-Plots zur Verteilung der individuellen Rate [%] korrekt unter maximal 12 beantworteten Aufgaben für 70 Teilnehmende, stratifiziert nach deren Kliniken A–D (Horizontalen kennzeichnen Median und Quartile, Vertikalen Maximum und Minimum).

„aktiv gegebenen“ Antworten dabei zwischen 19 % und 42 % (siehe ► **Tab. 1**).

## Diskussion

In dem vorliegenden Survey zur Vertrautheit mit Methoden der EbM ergab sich ein medianer Anteil korrekter Antworten von 47 %. Somit beantworteten die Ärzte weniger als die Hälfte der Fragen richtig. Dieser Anteil lag damit merklich unter dem erwarteten medianen Durchschnitt von 80 %. Das Ergebnis lässt dennoch die Einschätzung zu, dass trotz des Zurückbleibens des Ergebnisses hinter der erwarteten Spanne eine gewisse Vertrautheit mit den abgefragten Inhalten vorhanden sein muss. Dies kann u. a. damit begründet werden, dass die Teilnehmer ausdrücklich darauf hingewiesen wurden, dass „Raten“ nicht erwünscht sei; der recht hohe Anteil (je nach Frage zwischen 0 % und 39 %) „nicht aktiv gegebener“ Antworten lässt den Schluss zu, dass sich die Teilnehmer im Falle „aktiver“ Antworten zuge- traut haben, die Fragen richtig beantworten zu können. Eine weitere Differenzierung ist bei Betrachtung der „aktiven“ Antwort- raster zwischen den Aufgaben zu „neueren“ EbM-Methoden und denen zu eher „klassischen“ Methoden der klinischen Epidemiologie möglich: Unter den 6 Aufgaben zu „klassischen“ Methoden der klinischen Epidemiologie rangierten die Anteile der aktiv ge- gebenen Antworten zwischen 64 % und 100 %, unter denen zu „neueren“ Methoden der EbM zwischen 61 % und 71 %. Die „klas- sischen“ Methoden scheinen den ärztlichen Teilnehmenden aus den 4 Kliniken vertrauter zu sein als die „neueren“. Denkbar ist, dass die etwas höhere Vertrautheit mit „klassischen“ Methoden auch von deren aktiver Nutzung in einer eigenen Dissertation oder aktuellen Forschungsvorhaben herrührte, wo sicherlich öfter solche „klassischen“ Methoden wie Signifikanzprüfungen und Risikomaße eingehen als „neuere“ Methoden der EbM. Eine weitere Begründung könnte sein, dass das fest im Studienplan der

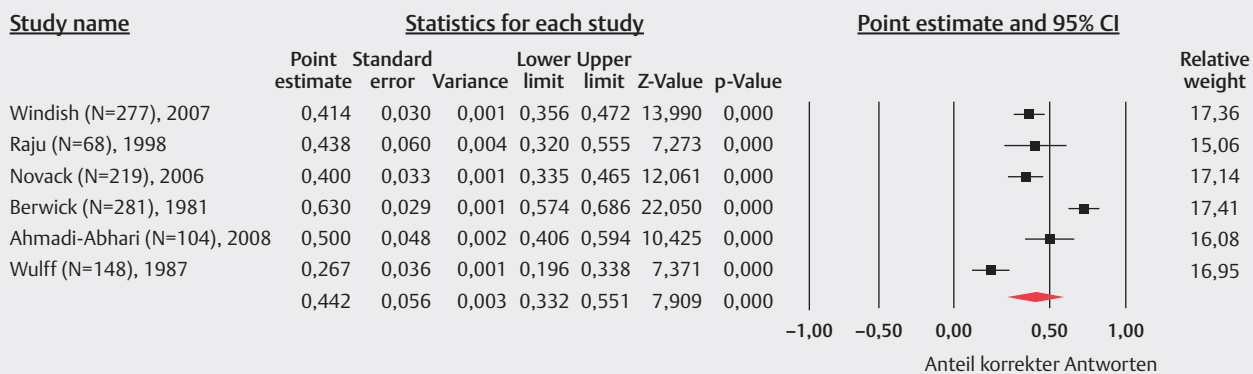
► **Tab. 1** Relative Häufigkeiten korrekter unter den „aktiv“ gegebenen Antworten für jede von 12 Aufgaben eines anonymen Surveys zu Methoden der evidenzbasierten Medizin, stratifiziert nach dem freiwillig „selbstberichteten“ Facharztstatus von 70 ärztlichen Mitarbeiterinnen und Mitarbeitern aus 4 Kliniken der Universität Witten/Herdecke.

	Fachärztin/ Facharzt: Anteil korrekter Antworten [%]	Assistenzärztin/ Assistenzarzt: Anteil korrekter Antworten [%]
Number Needed to Treat	73 %	89 %
relatives Risiko	33 %	17 %
Interpretation des p-Wertes	43 %	42 %
Anforderungen an Randomisierungen	23 %	36 %
Interpretation des Box-Whisker-Plots	47 %	33 %
diagnostische Kenngrößen	40 %	41 %
Lead-Time-Bias	13 %	39 %
Length-Time-Bias	50 %	58 %
Forest-Plot: studien- weise Angaben	17 %	8 %
Forest-Plot: Darstellung Meta-Effekt	20 %	13 %
Forest-Plot: Meta-Signi- fikanz	30 %	8 %
Forest-Plot: studien- weise Signifikanz	30 %	22 %

Humanmedizin verankerte Pflicht-Curriculum zur medizinischen Informatik, Biometrie und Epidemiologie (Querschnittsfach QB 1 laut Ärztlicher Approbationsordnung) Kenntnisse vermittelt hat, welche den Teilnehmenden des Surveys nach eigenen Klausuren im Studium noch in gewissem Maße nachhaltig verfügbar sind.

Es bleibt zu diskutieren, ob eine frühe und obligate Vermittlung von „neueren“ EbM-Methoden im Studium der Humanmedi- zin eine ähnlich hohe Vertrautheit und Sicherheit in deren Nut- zung schulen könnte wie hier für „klassische“ Methoden erkannt. Sicherlich könnte dann auch die spätere Bereitschaft gefördert werden, das aus dem Studium bereits Bekannte in das eigene Handeln zu integrieren [15–17].

Zu erwägen ist schließlich, ob die abgefragten Inhalte von den ärztlichen Teilnehmenden des Surveys nicht befriedigend verstan- den wurden. Inwieweit der intrinsische Schwierigkeitsgrad des Surveys das Ergebnis beeinflusst hat, bleibt offen. Rückmeldun- gen seitens Teilnehmender rangierten hierzu zwischen „motivie- rend“ und „anspruchsvoll“. Vergleicht man jedoch den hier ver- wendeten Survey mit Fragebögen, welche bei anderen Untersuchungen zum Einsatz kamen [5–10], muss festgestellt



Gepoolter Effektschätzer; Anteil korrekter Antworten

► **Abb. 2** Forest-Plot zur grafischen Darstellung eines gepoolten Effektschätzers nach dem Random-Effects-Modell aus der gefundenen internationalen Literatur: 6 Publikationen (1097 Teilnehmer) mit dokumentiertem Outcome „Anteil an korrekten Fragen eines EbM-Fragebogens“ konnten in die Analyse eingeschlossen werden. Dargestellt werden die studienweise Gewichtung, das zugehörige 95%ige Konfidenzintervall sowie der errechnete gepoolte Effektschätzer. Die Berechnung und die grafische Darstellung erfolgte mit der Software „Comprehensive Meta Analysis (CMA®) release 2.2.064, Biostat, Englewood 2011“.

werden, dass unsere Fragen im internationalen Vergleich eher als „leicht“ einzustufen waren.

## Vergleiche mit der Literatur

Zur besseren Übersicht der aktuellen Literatur wurde zu den gefundenen Studien, welche wie unsere Untersuchung aktiv Wissen bei Ärzten mittels eines Fragebogens abfragten, ein gepoolter Effektschätzer mittels Random-Effects-Modell (in Anlehnung an die Methode einer Metaanalyse) berechnet (► **Abb. 2**). Insgesamt konnten 6 Untersuchungen gefunden werden, mit Fallzahlen zwischen 68 [6] und 277 [5] Teilnehmern. Die jeweiligen Fragebögen bestanden aus 6 Fragen [9] bis hin zu 41 Fragen [7]. Das „beste“ Ergebnis publizierten Berwick et al. 1981 in den Vereinigten Staaten unter dem Titel „When Doctors Meet Numbers“: 281 Mediziner unterschiedlicher Ausbildungsstufen (vom Studenten bis zum auf Forschung spezialisierten Arzt) erhielten einen „Statistical Skills Self-Assessment Questionnaire“ mit 36 Multiple-Choice-Fragen. Es wurden im Median 63% der Fragen korrekt beantwortet [7]. Wulff et al. dagegen hatte bei 148 dänischen Ärzten im Median bei 9 Fragen nur 26,7% korrekter Antworten zu verzeichnen [10]. Berechnet man für die gefundenen 6 internationalen Studien einen gemeinsamen „Effektschätzer“, so lag der gepoolte Wissensscore von n = 1097 Medizinern bei 44,2% (95%-KI [33,2%; 55,1%]). Somit zeigte sich, dass weniger als die Hälfte aller gestellten Fragen zu Inhalten der evidenzbasierten Medizin beantwortet werden konnten. Dies deckt sich mit den Ergebnissen unserer Untersuchung, wo der mediane Anteil an korrekten Antworten bei 47% gelegen hat. Dabei muss allerdings beachtet werden, dass die zugrunde liegenden Fragebögen, welche international zum Einsatz gekommen waren, mehr als nur 2 Antwortmöglichkeiten beinhalteten (4 bis 5 Auswahlmöglichkeiten, teilweise Multiple Choice), einen deutlich größeren Umfang aufwiesen und auch inhaltlich wesentlich detaillierteres (noch) spezielleres Wissen um evidenzbasierte Methoden abfragte.

## Methodische Limitationen

Der Survey wurde zur Validierung „vorgetestet“. Insgesamt 10 Probanden, bestehend aus Studierenden und Ärzten, beantworteten den Survey und evaluierten ihn anschließend hinsichtlich der einzelnen Fragestellungen. Daraufhin wurden 2 Fragen entnommen, 1 Frage überarbeitet (umformuliert, gekürzt) und 1 ersetzt, sodass die Finalversion schlussendlich 12 Fragen bot. Sicherlich kann der Survey aber nach dieser eingeschränkten Vorlaufphase noch nicht bezüglich psychometrischer Kenngrößen als valide, reliabel und objektiv demonstriert werden; hierzu werden eigene Untersuchungen folgen müssen, nachdem ggf. auch das Instrument nochmals revidiert wurde.

Ziel des Surveys war es, die Vertrautheit mit Termini und gängigen Kenngrößen der EbM unter Ärzten zu ermitteln. Ob hierfür 12 Aufgaben mit binärer Antwortvorgabe ausreichen, zumal 4 dieser 6 Fragen das gleiche Fragen-Thema abdecken (Forest-Plot), ist zu hinterfragen. Eine größere Abdeckung von Themen zur EbM in einem erweiterten Folge-Survey wäre denkbar, muss aber zugleich auch in Relation zur zeitlichen Durchführbarkeit gesehen werden: Für die Durchführung des vorliegenden Surveys war ein Zeitraum von maximal 10 Minuten angesetzt (und aus den Vor-testungen auch als realistisch erkannt), sodass eine Durchführbarkeit während der Früh- oder Mittagsbesprechungen einer Klinik gewährleistet werden konnte. Umfangreichere Instrumente – wie sie international zum Einsatz kamen – könnten nominell einen höheren Informationsgehalt liefern, welcher aber durch sinkende Compliance und höheren Zeitdruck weniger belastbare Angaben in sich bergen kann. Teilweise waren für die Befragungen in anderen Untersuchungen Zeitfenster bis 25 Minuten veranschlagt [5]. Andere Untersuchungen verschickten die Fragebögen, wobei hier die Rücklaufquoten der Fragebögen stark variierten zwischen 59,2% [10] und 84,2% [8].

Trotz der ausdrücklichen Bitte konnte ferner ein „Raten“ bei einzelnen Aufgaben nicht ausgeschlossen werden. Hierbei bestand



eine 50 %-Erfolgswahrscheinlichkeit durch das einfache Antwort-Format „richtig“/„falsch“. Dieses Antwort-Format wurde in Anbetracht der begrenzten Zeitvorgabe (maximal 10 Minuten) gewählt, um zumindest eine repräsentative Auswahl an medizinstatistischen Themen abbilden zu können. Alternative Formate wie Kurzantworten, Ordnungs-Antwortaufgaben, Auswahlantworten, Ergänzungsaufgaben etc. hätten vermutlich mehr Zeit zum Überlegen und Ausführen erfordert, sodass eine vollständige Bearbeitung des Surveys unter der obigen Zeitvorgabe nicht möglich gewesen wäre bzw. alternativ zu Ergebnisverzerrung hätte führen können. Das Ausmaß, welches diese Verzerrung haben kann, ist nicht bekannt und auch schwerlich abzuschätzen. Einige Autoren anderer Untersuchungen ließen bewusst die Antwortmöglichkeit „ich weiß es nicht“ zu, um auch hier ein „Raten“ von Antworten zu verhindern. In einer zukünftigen Untersuchung sollten unbedingt mehr als 2 Antwortmöglichkeiten gewählt werden, um eine Verzerrung der Ergebnisse durch mögliches Raten auszuschließen.

Eine weitere Limitation der Aussagekraft obiger Ergebnisse resultiert aus der Anzahl der anwesenden ärztlichen Survey-Teilnehmenden. Zwar wurde insgesamt die Anzahl an Fragebögen gemäß vorheriger Fallzahlplanung erfüllt, jedoch war diese Fallzahlplanung von einer deutlich optimistischeren Annahme des Survey-Ergebnisses ausgegangen (siehe auch die höhere Breite des Konfidenzintervalls zum primären Endpunkt gegenüber den Vorgaben der Fallzahlplanung). Alle Anwesenden der 4 Kliniken nahmen am Survey teil; keine/r der Ärztinnen und Ärzte verweigerte die Teilnahme. Es ist jedoch nicht auszuschließen, dass am Tag der Erhebung zufällig solche Ärztinnen und Ärzte im Dienst waren, die entweder mehr oder weniger als ihre abwesenden Kolleginnen und Kollegen mit Methoden und Inhalten der EbM vertraut waren, wodurch das Ergebnis einer Verzerrung mit positiver oder negativer Ergebnistendenz unterlegen sein könnte.

Weiterhin war der vorliegende Survey tendenziell monozentrisch und befragte nur ärztliche Fachkräfte aus der Universität Witten/Herdecke angegliederten Kliniken. Zu diesen Kliniken muss erwähnt werden, dass deren Träger (Städtische Kliniken Köln respektive Helios Wuppertal-Barmen) nicht denen klassischer Universitätskliniken entsprechen. Inwieweit das Personal an Universitätskliniken eines privaten, respektive städtischen Trägers bezüglich akademischer Expertise und Erfahrung vergleichbar ist mit dem ärztlichen Personal staatlicher Unikliniken, kann nur gemutmaßt werden. Wie weit sich also obige Ergebnisse in entsprechenden Kliniken staatlicher Universitätskliniken reproduzieren lassen, müssen multizentrische multidisziplinäre Folgeerhebungen zeigen. Für solche Folge-Untersuchungen wäre es zudem sinnvoll zu evaluieren, an welchen der jeweiligen Standorte EbM-Anteile im Studium oder verpflichtend im Rahmen der Facharzt-ausbildungen postgraduell erfolgen und in welchem Umfang.

#### KERNAUSSAGEN

- Klinisch tätige Ärztinnen und Ärzte zeigten sich in diesem Survey hinreichend vertraut mit klassischen Methoden der klinischen Epidemiologie (z. B. relatives Risiko).
- Aktuellere Methoden der evidenzbasierten Medizin (z. B. Forest-Plot) zeigten eine eingeschränkte Vertrautheit.

- Es besteht ein Bedarf an zielgerichteten Angeboten zur Weiterqualifikation für klinisch tätige Ärztinnen und Ärzte in Methoden der evidenzbasierten Medizin.

#### Autorenschafts-Legitimation

STA und FK haben die Untersuchung konzipiert und das Erhebungsinstrument entwickelt; STA und ST haben die Untersuchung durchgeführt; FK und ST haben die erhaltenen Daten statistisch analysiert; STA, ST, CB und FK haben die Ergebnisse bewertet und aggregiert; CB führte die in der Diskussion enthaltene Analyse des gepoolten Effektschätzers durch, CB und FK haben den Entwurf dieser Publikation erstellt, STA und ST haben die Publikation intensiv revidiert.

#### Interessenkonflikt

Die Autoren geben an, dass keine Interessenkonflikte finanzieller oder politischer Art bezüglich den in diesem Manuskript formulierten Inhalten bestehen. Die Ergebnisse sind enthalten in der Dissertationsschrift von Frau cand. med. Sünje-Theresa Ahrens, der Fakultät für Gesundheit der Universität Witten/Herdecke im Jahr 2019 vorgelegt zur Erlangung des Grades „Dr. med.“; die in diesem Manuskript enthaltenen Ergebnisse wurden zudem als wissenschaftliches Poster (ohne publiziertes Abstract) präsentiert im Rahmen der Jahrestagung der Deutschen Region der Internationalen Biometrischen Gesellschaft (März 2018, Frankfurt/Main).

#### Danksagung

Die Autoren möchten sich ausdrücklich bei den Direktoren der 4 teilnehmenden Kliniken der Universität Witten/Herdecke für die Ermöglichung der Durchführung dieses Surveys bedanken: Prof. Dr. med. Stephan Roth, Direktor der Klinik für Urologie und Kinderurologie, Prostatazentrum Wuppertal und Bergisch Land, Helios-Universitätsklinikum Wuppertal Prof. Dr. med. Melchior Seyfarth, Direktor der Klinik für Kardiologie, Helios-Universitätsklinikum Wuppertal Prof. Dr. med. Erich Stoelben, Chefarzt der Thoraxchirurgie, Kliniken der Stadt Köln, Krankenhaus Merheim, Lungenklinik am Zentrum für Thoraxchirurgie, Pneumologie/Onkologie, Schlaf- und Beatmungsmedizin Prof. Dr. med. Stefan Wirth, Chefarzt des Zentrums für Kinder- und Jugendmedizin, Helios-Universitätsklinikum Wuppertal Weiterhin gilt unser Dank den Ärztinnen und Ärzten dieser 4 Kliniken für deren uneingeschränkte Bereitschaft zur Teilnahme am Survey und die offenen Rückmeldungen zu unserer Erhebung.

#### Literatur

- [1] Sackett DL, Rosenberg WM, Gray JA et al. Evidence based medicine: what it is and what it isn't. *Bmj* 1996; 312: 71–72
- [2] Hilgers R, Feldmann U, Jöckel K et al. Empfehlungen zur Umsetzung der Approbationsordnung für Ärzte vom 27.06.2002 in den Fächern Epidemiologie, Medizinische Biometrie und Medizinische Informatik. *GMS Med Inform Biom Epidemiol* 2005; 1: Doc05 (20050407)
- [3] Burnand B. Warum Patienten zu wenig evidenzbasiert behandelt werden. In 2014. Im Internet: [https://swiss.cochrane.org/sites/swiss.cochrane.org/files/public/uploads/CochraneCH\\_MedicalTribune\\_2014.pdf](https://swiss.cochrane.org/sites/swiss.cochrane.org/files/public/uploads/CochraneCH_MedicalTribune_2014.pdf); Stand: 06.08.2018
- [4] Timmer A. Weiterbildung: Plädoyer für die evidenzbasierte Medizin. *Dtsch Arztebl* 2012; 109: A-1418/B-1227/C-1205

- [5] Windish DM, Huot SJ, Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. *Jama* 2007; 298: 1010–1022. doi:10.1001/jama.298.9.1010
- [6] Raju TN, Langenberg PW, Vidyasagar D et al. A biostatistical survey questionnaire. *The Journal of pediatrics* 1988; 112: 859–863
- [7] Berwick DM, Fineberg HV, Weinstein MC. When doctors meet numbers. *The American journal of medicine* 1981; 71: 991–998
- [8] Novack L, Jotkowitz A, Knyazer B et al. Evidence-based medicine: assessment of knowledge of basic epidemiological and research methods among medical doctors. *Postgraduate medical journal* 2006; 82: 817–822. doi:10.1136/pgmj.2006.049262
- [9] Ahmadi-Abhari S, Soltani A, Hosseinpanah F. Knowledge and attitudes of trainee physicians regarding evidence-based medicine: a questionnaire survey in Tehran, Iran. *Journal of Evaluation in Clinical Practice* 2008; 14: 775–779
- [10] Wulff HR, Andersen B, Brandenhoff P et al. What do doctors know about statistics? *Statistics in medicine* 1987; 6: 3–10
- [11] Trede R. Stellungnahme der AWMF zum „Masterplan Medizinstudium 2020“: Wissenschaftskompetenz ist eine Schlüsselqualifikation für jede ärztliche Tätigkeit. In 2016. Im Internet: [https://www.awmf.org/fileadmin/user\\_upload/Stellungnahmen/Aus-\\_und\\_Weiterbildung/AWMF-Stellungnahme%20Masterplan%20Medizinstudium%202020\\_2016-11-05.pdf](https://www.awmf.org/fileadmin/user_upload/Stellungnahmen/Aus-_und_Weiterbildung/AWMF-Stellungnahme%20Masterplan%20Medizinstudium%202020_2016-11-05.pdf); Stand: 06.08.2018
- [12] Blanco MA, Capello CF, Dorsch JL et al. A survey study of evidence-based medicine training in US and Canadian medical schools. *J Med Libr Assoc* 2014; 102: 160–168. doi:10.3163/1536-5050.102.3.005
- [13] Rao G. Physician numeracy: essential skills for practicing evidence-based medicine. *Family medicine* 2008; 40: 354–358
- [14] Reed JF 3rd, Salen P, Bagher P. Methodological and statistical techniques: what do residents really need to know about statistics? *Journal of medical systems* 2003; 27: 233–238
- [15] Ilic D, Forbes K. Undergraduate medical student perceptions and use of Evidence Based Medicine: a qualitative study. *BMC Med Educ* 2010; 10: 58. doi:10.1186/1472-6920-10-58
- [16] Kotur PF. Introduction of evidence-based medicine in undergraduate medical curriculum for development of professional competencies in medical students. *Curr Opin Anaesthesiol* 2012; 25: 719–723. doi:10.1097/ACO.0b013e32835a1112
- [17] Taheri H, Mirmohamadsadeghi M, Adibi I et al. Evidence-based medicine (EBM) for undergraduate medical students. *Ann Acad Med Singapore* 2008; 37: 764–768