

**Machbarkeitsstudie zur praktischen Anwendung
von Evidenzbasierter Medizin im Rahmen des
humanmedizinischen Studiums im Fachbereich für
Allgemeinmedizin**

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Hohen Medizinischen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

Miriam Esser
aus Mechernich
2020

Angefertigt mit der Genehmigung
der Medizinischen Fakultät der Universität Bonn

1. Gutachter: Prof. Dr. med. Klaus Weckbecker
2. Gutachterin: PD Dr. med. Mignon-Denise Keyver-Paik

Tag der Mündlichen Prüfung: 28.01.2020

Aus dem Institut für Hausarztmedizin
Direktorin: Prof. Dr. med. Birgitta Weltermann

Inhaltsverzeichnis

	Abkürzungsverzeichnis	6
1.	Einleitung	7
1.1	Hintergrund	9
1.1.1	Einführung: Der 5-A-Schritteprozess	9
1.1.2	Einführung: Das PICO-Schema	10
1.1.3	Einführung: Studientypen	10
1.1.4	Einführung: Recherchequellen	12
1.1.5	Einführung: PubMed	14
1.1.6	Einführung: Evidenzstufen	15
1.1.7	Einführung: Valide EbM-Evaluationsinstrumente	16
1.2	Fragestellung	17
2.	Material und Methoden	18
2.1	Entwicklung des Lehrkonzepts	18
2.2	Das Online-Lernmodul	19
2.2.1	Lernmodul: Einführung in die EbM	20
2.2.2	Lernmodul: Der 5-A-Schritteprozess	21
2.2.3	Lernmodul: Ask	21
2.2.4	Lernmodul: Acquire – Studientypen	22
2.2.5	Lernmodul: Acquire – Recherchequellen	22
2.2.6	Lernmodul: Acquire – PubMed	23
2.2.7	Lernmodul: Appraise – Validität	24
2.2.8	Lernmodul: Appraise – Größe und Präzision	24
2.2.9	Lernmodul: Appraise – Übertragbarkeit und Anwendbarkeit	25
2.2.10	Lernmodul: Apply	26
2.2.11	Lernmodul: Assess	26
2.3	Der Arbeitsauftrag	26
2.4	Die Seminare	27
2.4.1	Der erste Seminartermin	28
2.4.2	Der zweite Seminartermin	29

2.5	Der Bonn Test	29
2.5.1	Valide EbM-Evaluationsinstrumente	29
2.5.2	Adaptationen: Vom Fresno Test zum Bonn Test	30
2.6	Überprüfung der Fragestellungen	34
3.	Ergebnisse	36
3.1	Ergebnisse der fiktiven Fallbeispiele A, B und C	36
3.2	Ergebnisse der eigenen Fallbeispiele	46
3.3	Ergebnisse der Eigenevaluation	53
3.3.1	Ergebnisse der Evaluation: Das Lernmodul	53
3.3.2	Ergebnisse der Evaluation: Die Seminare	57
3.3.3	Ergebnisse der Evaluation: Das Blockpraktikum	59
3.3.4	Ergebnisse der Evaluation: Gesamtnote	62
4.	Diskussion	63
4.1	Überprüfung der Fragestellungen	63
4.2	Grundlagenvermittlung	64
4.2.1	Inhalt des Lernmoduls	64
4.2.2	Format des Lernmoduls	65
4.2.3	Organisatorischer Rahmen des Lehrkonzepts	65
4.3	Evaluierbarkeit	66
4.3.1	Bonn Test – Theorie	66
4.3.2	Bonn Test – Theorie: Ask	67
4.3.3	Bonn Test – Theorie: Acquire – Suchstrategie	69
4.3.4	Bonn Test – Theorie: Acquire – Abstract	69
4.3.5	Bonn Test – Theorie: Appraise	70
4.3.6	Bonn Test – Theorie: Apply	72
4.3.7	Bonn Test – Theorie: Assess	73
4.3.8	Bonn Test: Verbesserungsmöglichkeiten	74
4.4	Universelle Gültigkeit (Bonn Test – Praxis)	75
4.5	Akzeptanz	77
4.5.1	Studentische Zufriedenheit – Lernmodul	77
4.5.2	Studentische Zufriedenheit – Seminare	78

4.5.3	Studentische Zufriedenheit – Blockpraktikum	79
4.6	Stärken und Schwächen	80
5.	Zusammenfassung	82
6.	Abbildungsverzeichnis	83
7.	Tabellenverzeichnis	86
8.	Literaturverzeichnis	87
9.	Anlagen	91
9.1	Anlage I: Druckversion des Online Lernmoduls	91
9.2	Anlage II: Aufgabenstellung der Fallbeispiele A, B und C	112
9.3	Anlage III: Erwartungshorizont des Bonn Tests	113

Abkürzungsverzeichnis

5ASP	5-A-Schritteprozess
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaft e.V.
BP	Blockpraktikum
CDSR	Cochrane Database of Systematic Reviews
DARE	Database of Abstracts of Reviews of Effects
DEGAM	Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin e.V.
DNEbM	Deutsches Netzwerk für Evidenzbasierte Medizin e.V.
EbM	Evidenzbasierte Medizin
G-I-N	Guidelines International Network
HTA Database	Cochrane Health Technology Assessment Database
MeSH	Medical Subject Headings
NGC	National Guideline Clearinghouse
NICE	National Institute for Health & Care Excellence
OCEBM	Oxford Centre for Evidence Based Medicine
PICO	Patient Problem, Intervention, Comparison und Outcome
RCT	Randomized Controlled Trial
SIGN	Scottish Intercollegiate Guidelines Network

1. Einleitung

Die moderne Medizin unterliegt einem stetigen Wandel. Wissenschaftliche Ansätze und Methoden revolutionieren tagtäglich den Forschungsstand. Bestehende Leitlinien werden regelmäßig angepasst und dienen der zeitnahen Implementation neugewonnener Erkenntnisse in den klinischen Alltag. Eine Metaanalyse zeigt jedoch, dass solche Neuerungen selten bzw. zeitverzögert Anwendung in der Praxis finden (Francke et al., 2008). Damit die Patienten von der bestmöglichen Therapie profitieren können, muss sich der behandelnde Arzt daher kritisch mit der entsprechenden Fachliteratur auseinandersetzen können (General Medical Council, 2015). Um zukünftige Ärzte frühzeitig mit diesem Aufgabenbereich vertraut zu machen, sollte die Evidenzbasierte Medizin (EbM) fester Bestandteil des Studiums werden. Dabei handelt es sich um die „gewissenhafte, explizite und umsichtige Anwendung der besten verfügbaren Evidenz um Behandlungsentscheidungen zum Wohle des individuellen Patienten zu treffen“ (“Evidence based medicine ist the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients“) (Sackett et al., 1996). Der Begriff Evidenz (lat. *evidentia* = Augenscheinlichkeit) bezeichnet eigentlich eine Offensichtlichkeit. Im Zusammenhang mit der Evidenzbasierten Medizin leitet er sich hingegen von dem englischen Wort „evidence“ ab und steht für einen Beweis oder einen Beleg (DNEbM e.V., 2011). Diese Art der belegenden Evidenz kann jeweils in eine interne und eine externe Form unterschieden werden. Bei der internen Evidenz handelt es sich um durch das Medizinstudium erworbenes, allgemein anerkanntes fachliches Grundwissen, auf das der Arzt zugreifen kann. Die externe Evidenz geht aus der gezielten Literaturrecherche hervor und muss noch vor der Anwendung evaluiert werden (Beyer M et al., 2005). Um die bestmögliche Behandlungsentscheidung treffen zu können müssen beide Formen miteinander kombiniert werden. Ohne die durch Erfahrung gewonnene Expertise des Mediziners kann auch die beste Evidenz auf den einzelnen Patienten fehlerhaft angewendet werden, während eine Behandlung auf Grundlage der besten Expertise ohne entsprechende Evidenz bestenfalls überholt ist (Sackett et al., 1996). Die Initiative der Evidenzbasierten Medizin wurde 1992 von einer Arbeitsgruppe der McMaster University in Kanada ins Leben gerufen und markiert ein Umdenken bei der Behandlungsentscheidung im klinischen Alltag. Während der Fokus zuvor auf dem

pathophysiologischen Verständnis von Krankheitsmechanismen und einzelnen klinischen Erfahrungen bzw. Expertenmeinungen lag, rückte sie die kritische Auseinandersetzung mit den vorliegenden Forschungsergebnissen in den Vordergrund (Guyatt G et al., 1992). Ziel dieser Umorientierung war, die Patientenversorgung durch eine wissenschaftlich und empirisch fundierte Anpassung insgesamt sicherer, unabhängig vom Behandler und kosteneffektiver zu gestalten (Pope C, 2003). Als Folge der Initiative wurden zahlreiche Institutionen wie die Cochrane Collaboration gegründet, um EbM einfacher im Alltag anwenden zu können. In diesem internationalen Netzwerk schließen sich Wissenschaftler zur Erstellung von systematischen Übersichtsarbeiten zusammen. Experten prüfen die Studienlage zu häufig gestellten klinischen Fragestellungen und evaluieren sie vorab (Bero, 1995). Bei der Entwicklung und Aktualisierung von Leitlinien werden ebenfalls Techniken der Evidenzbasierten Medizin genutzt um die Therapie der Patienten an die neusten wissenschaftlichen Erkenntnisse anzupassen (Hill et al., 2011). Auch im Rahmen der Lehre konnten medizinische Ausbildungsstätten weltweit EbM bereits erfolgreich in das Curriculum integrieren. Obwohl die Evidenzbasierte Medizin bereits in zahlreichen wichtigen Bereichen etabliert werden konnte, zeigt die aktuelle Studienlage, dass sie bisher noch zu selten Anwendung im klinischen Alltag findet. Eine mögliche Erklärung hierfür sehen die Autoren systematischer Übersichtsarbeiten auf diesem Gebiet darin, dass der Schwerpunkt aktueller Lehransätze stets auf dem theoretischen Verständnis für EbM liegt und dadurch die praktische Anwendung dieser Fähigkeiten vernachlässigt wird (Schilling et al., 2006; Thomas und Kreptul, 2015). Auch an der Universität Bonn wurde die Thematik EbM bisher nur theoretisch und eher beiläufig im Rahmen anderer Fächer behandelt. Damit die Evidenzbasierte Medizin zukünftig besser ausgeschöpft werden kann, überprüft die vorliegende Arbeit den Effekt eines Lehrkonzepts, das den Studenten ein systematisches Selbststudium der Grundlagen Evidenzbasierter Medizin mit einem Online-Lernmodul ermöglicht und sie bei der praktischen Anwendung der erlernten Fähigkeiten zunächst im Rahmen von theoretischen Fallbeispielen und später im realen Patientenkontakt während ihres zweiwöchigen Blockpraktikums für Allgemeinmedizin in einer Hausarztpraxis unterstützt. Die Leistung der Studenten in den wichtigsten EbM-Kategorien wird jeweils durch den eigens dafür konzipierten „Bonn Test“ erfasst und kann mit Hilfe eines universellen Erwartungshorizonts sowohl für die theoretischen als auch für die praktischen Fallbeispiele ausgewertet werden. Der organisatorische Rahmen richtet

sich dabei nach den Empfehlungen des „Curriculum Evidenzbasierte Medizin im Studium“, welches 2005 gemeinsam von dem Deutschen Netzwerk für Evidenzbasierte Medizin (DNEbM) e.V. und der Bundesärztekammer veröffentlicht wurde (Beyer M et al., 2005). So werden zusätzlich zu dem Selbststudium zwei Seminare in interaktiven Kleingruppen abgehalten, die selbstständig, aber tutorengestützt die Lernziele erarbeiten.

1.1 Hintergrund

Für ein besseres Verständnis der Inhalte der Lehreinheit werden im Folgenden wichtige Konzepte der Evidenzbasierten Medizin vorab erläutert.

1.1.1 Einführung: Der 5-A-Schritteprozess

Der 5-A-Schritteprozess (5ASP) ist eine Anleitung für die systematische Anwendung von Evidenzbasierter Medizin und wurde bereits während der Anfänge der EbM-Bewegung veröffentlicht (Cook et al., 1992, **Abb. 1**). Er ist als ein Kreislauf von Abläufen zu verstehen. Die 5 As stehen für die Punkte Ask, Acquire, Appraise, Apply und Assess. Am Anfang des 5-A-Schritteprozesses steht der Patient mit seiner individuellen Erkrankung, seinen Eigenschaften, sowie Bedürfnissen und Wünschen. Aus der Gesamtheit dieses Szenarios wird eine wissenschaftliche Fragestellung formuliert (Ask). Anschließend wird eine Literaturrecherche zur Beantwortung der Fragestellung durchgeführt (Acquire) und die Rechercheergebnisse werden kritisch evaluiert (Appraise). Schließlich wird aufgrund der gefundenen Evidenz, der klinischen Erfahrung des behandelnden Arztes, sowie der Bedürfnisse des Patienten eine Behandlung ausgewählt (Apply). In der Nachverfolgung wird die gewählte Behandlung evaluiert (Assess). Im Falle eines nicht zufriedenstellenden Ergebnisses wird die Fragestellung optimiert und der Ablauf wiederholt.

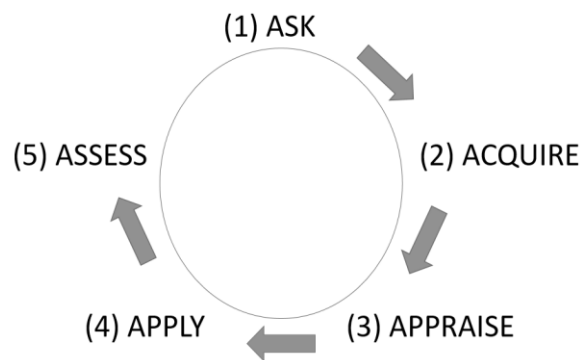


Abb. 1: Darstellung des 5-A-Schritteprozesses.

1.1.2 Einführung: Das PICO-Schema

Das PICO-Schema ist ein valides Modell bei der Entwicklung wissenschaftlicher Fragestellungen (Duke University Medical Center Library und Health Sciences Library at the University of North Carolina at Chapel Hill, 2014). Die Buchstaben P, I, C und O stehen jeweils für die wichtigsten Aspekte, die bei der Frage nach der bestmöglichen Therapie berücksichtigt werden sollten:

- P = „Patient Problem“: Identifizierung der wichtigsten charakteristischen Merkmale wie der Haupterkrankung und Begleiterkrankungen aber auch allgemeinen Aspekten wie Alter, Geschlecht oder ethnische Herkunft
- I = „Intervention“: Benennung der angestrebten Therapieform wie Medikamenten oder einer Operation
- C = „Comparison“: Bestimmung einer naheliegenden alternativen Behandlung
- O = „Outcome“: Definition der erhofften Wirkung der Intervention.

1.1.3 Einführung: Studientypen

Am Anfang einer systematischen Literaturrecherche steht die Festlegung geeigneter Studientypen und Recherchequellen (Kapitel 2.2.4 Lernmodul: Acquire – Studientypen). Ein Überblick über die wichtigsten Studientypen gehört daher zu dem grundlegenden Basiswissen bei der Anwendung von EbM.

- Querschnittsuntersuchungen befassen sich mit der Beziehung zwischen einer Krankheit und anderen Faktoren zu einem bestimmten Zeitpunkt. Sie enthalten ausschließlich prävalente Fälle und sind daher gefragte Instrumente für Diagnostik-Tests, aber ungeeignet für vergleichende Studien beispielsweise zu Therapieformen.
- Case Reports und Case Series sind eine deskriptive Darstellung einzelner Fallberichte bzw. Sammlungen ähnlicher Fälle. Sie haben aufgrund dieser Individualität nur eine geringe statistische Validität.
- In Fall-Kontroll-Studien werden Betroffene und Nicht-Betroffene retrospektiv untersucht, um mögliche Auslöser für eine Krankheit zu identifizieren. Hierbei ist es schwer, einzelne Faktoren zu identifizieren, da häufig eine große Vielfalt an Variablen Einfluss auf eine Krankheit und deren Verlauf nehmen kann und statistische Zusammenhänge allein noch keine Kausalität beweisen.
- In einer Kohorten Studie wird eine Gruppe, die einem bestimmten Reiz ausgesetzt ist oder eine Therapie einhält, prospektiv mit einer Vergleichsgruppe abgeglichen, die dem Faktor nicht ausgesetzt war. Fehlerhafte oder fehlende Erinnerungen an Einflussfaktoren haben auf Kohorten Studien zwar weniger Einfluss als auf Fall-Kontroll-Studien, dennoch können auch hier andere Faktoren als die gesuchte Variable unterschiedliche Ausgänge bei Interventions- und Kontrollgruppe provozieren.
- Crossover-Studien umgehen diesen Störfaktor, indem jeder Patient als seine eigene Kontrolle dient, da er zu verschiedenen Zeitpunkten jeweils ein Medikament oder Placebo erhält. Möglicher Störfaktor ist bei dieser Studie eine zu kurz gewählte „Auswaschphase“ bis zum Wechsel des Präparats, sodass ein Überhang-Effekt („Carryover“) entsteht.
- Eine randomisierte kontrollierte Studie (RCT) ist eine sorgfältig geplante Studie, bei der die Interventionsgruppe mit einer Kontrollgruppe verglichen wird und systematische Fehler durch Randomisierung und Verblindung minimiert werden. Bei einer Randomisierung werden die Patienten zufällig und nicht vorhersagbar auf Behandlungs- und Kontrollgruppe aufgeteilt, sodass eine strukturell identische Verteilung von Störgrößen gewährleistet ist. RCTs gelten allgemein als der Goldstandard zur Überprüfung der Wirksamkeit und Sicherheit einer Therapie,

aber die externe Validität sollte für jeden Fall hinterfragt werden, da in der Realität keine standardisierten und kontrollierten Studienbedingungen existieren und andere Faktoren den Ausgang deutlich beeinflussen können (Kabisch et al., 2011).

- Systematic Reviews und Metaanalysen sind Zusammenfassungen von Studienergebnissen. Sie bieten sich besonders dann an, wenn geringe Fallzahlen ausgeglichen werden sollen oder Einzelstudien unübersichtlich oder inkonsistent sind (Ressing et al., 2009). In Systematic Reviews werden die Ergebnisse einer ausführlichen Literaturrecherche zu einer bestimmten medizinischen Fragestellung zusammengefasst. Sie stützen sich auf zuvor evaluierte, methodisch hochwertige Studien.
- Bei Metaanalysen werden die Ergebnisse verschiedener valider Studien statistisch korrekt kombiniert, als handle es sich dabei um eine große Studie (Duke University Medical Center Library und Health Sciences Library at the University of North Carolina at Chapel Hill, 2014). Metaanalysen der publizierten Daten aus kontrollierten klinischen Studien sind gefolgt von den Studienergebnissen randomisierter, kontrollierter Untersuchungen die Studientypen mit den höchsten Evidenzgraden (Du Prel JB et al., 2009). Dennoch sollte ihre Qualität individuell beurteilt und beispielsweise auf die mögliche Heterogenität der zusammengefassten Studien oder die Berücksichtigung anderer Limitationen hin geprüft werden (Ressing et al., 2009).

1.1.4 Einführung: Recherchequellen

Für die Anwendung von EbM sollten gängige Recherchequellen bekannt sein und entsprechend ihrer Relevanz und Validität eingeordnet werden können (Kapitel 5.2.5 Acquire – Recherchequellen).

- Eine wichtige Informationsquelle im Internet sind Literaturdatenbanken. Die beiden größten Datenbanken für biomedizinische Literatur sind Embase und MEDLINE. MEDLINE liefert neben bibliografischen Daten, Abstracts und inhaltsabbildenden Schlagwörtern zum Teil auch Links zu Volltexten und ist über Portale wie PubMed oder DIMDI kostenlos zugänglich. Der Schwerpunkt liegt bei MEDLINE auf U.S.-

amerikanischer Literatur während sich Embase stärker auf europäische Publikationen konzentriert, jedoch nur kostenpflichtig nutzbar ist (PUBMED.DE; Universitätsklinikum Freiburg / Deutsches Cochrane Zentrum, 2013).

- Andere Literaturdatenbanken wie die Cochrane Database of Systematic Reviews (CDSR), die Database of Abstracts of Reviews of Effects (DARE) oder die Cochrane Health Technology Assessment Database (HTA Database) haben sich auf die Suche nach Systematic Reviews spezialisiert und sind über die Cochrane Library kostenlos zugänglich (Universitätsklinikum Freiburg / Deutsches Cochrane Zentrum, 2013). Systematic Reviews oder auch systematische Übersichtsarbeiten zeichnen sich im Vergleich mit traditionell narrativen Übersichtsarbeiten auf der Basis von selektiver Literaturlauswahl dadurch aus, dass sie sich hauptsächlich mit den Studienergebnissen aus randomisierten, kontrollierten Studien auseinandersetzen und als oberstes Ziel die Reduktion von Verzerrungen verfolgen (Blümle et al., 2009). Cochrane ist ein globales, unabhängiges Netzwerk in mehr als 130 Ländern. Es schafft durch die regelmäßige Überarbeitung der Cochrane Reviews sogenannte „Living Documents“, die sich durch einen hohen Grad der Standardisierung auszeichnen und eine objektive Interpretation der Ergebnisse anbieten (Cochrane Deutschland, 2016).
- EBP-Quellen werden auch „Pre-appraised Resources“ genannt, da sie wie die Systematic Reviews der Cochrane Library aktuelle Berichte zu klinisch hochwertigen Studien zusammenfassen, vorab evaluieren und zum Teil durch Experten kommentieren. Hierzu zählen zum Beispiel ACP Journal Club, Clinical Evidence, Dynamed, Essential Evidence oder UpToDate (Duke University Medical Center Library und Health Sciences Library at the University of North Carolina at Chapel Hill, 2014). Im direkten Vergleich mit MEDLINE können klinische Fragestellungen durch Pre-appraised Resources schneller beantwortet werden, da die Studien bereits vorab zusammengefasst und evaluiert werden. Dennoch bleiben mehr Fragen unbeantwortet als mit MEDLINE, welches ein deutlich breiteres Spektrum an Informationen bietet, das der Nutzer selbstständig evaluieren muss. Während die Nutzung von MEDLINE kostenlos ist, kann der Preis für ein Abonnement von Pre-appraised Resources zwischen 200 \$ für Einzelpersonen und mehr als 40.000 \$ für Institute liegen (Patel et al., 2006).

- Speziell für Fragen zur Therapie können Leitliniendatenbanken wie AWMF, G-I-N (Guidelines International Network) oder NGC (National Guideline Clearinghouse) bzw. die Websites von Leitlinien-Erstellern selber wie SIGN (Scottish Intercollegiate Guidelines Network) oder NICE (National Institute for Health & Care Excellence) nützliche Informationen zu empfohlenen Behandlungsmaßnahmen liefern (Universitätsklinikum Freiburg / Deutsches Cochrane Zentrum, 2013).

1.1.5 Einführung: PubMed

Da die Literaturrecherche im Lernmodul anhand von PubMed demonstriert wird (Kapitel 2.2.6 Lernmodul: Acquire – PubMed), folgt eine ausführlichere Vorstellung dieser Recherchequelle. PubMed stellt mehr als 25 Millionen Verweise auf biomedizinische Literatur von MEDLINE, Journals und online Büchern zur Verfügung und ist eine häufig genutzte Quelle für die Literaturrecherche (Bethesda M D, 2005). Bei der Umwandlung einer PICO-Frage in eine Suchanfrage ist es notwendig abzuwägen welche Informationen übernommen werden sollen, damit die Suche zwar präzise, aber nicht zu spezifisch ist. Die Suchsprache bei PubMed ist wie in den meisten anderen online Recherchequellen auch Englisch, sodass die Begriffe übersetzt werden müssen. In PubMed kann nach Stichworten (engl.: keyword) oder Schlagwörtern (engl.: descriptor) gesucht werden. Stichwörter kommen in Titel oder Abstract vor und sind besonders relevant für die Suche nach aktuellen Artikeln, da z. B. 2000 – 4000 neue Einträge täglich allein bei MEDLINE veröffentlicht werden und diese in der Regel noch nicht verschlagwortet sind. Schlagwörter fassen den Inhalt eines Artikels zusammen und werden automatisch mit Synonymen verknüpft, sodass unterschiedliche Wörter mit derselben Bedeutung auch zum selben Suchergebnis führen (PUBMED.DE; Universitätsklinikum Freiburg / Deutsches Cochrane Zentrum, 2013). Medical Subject Headings (MeSH) ist ein von der U.S.-amerikanischen National Library of Medicine entwickeltes und weltweit verwendetes Schlagwortvokabular, das durch die Einteilung in 16 Hauptkategorien und untergeordnete Begriffe einen hierarchischen Aufbau (MeSH Tree Structure) aufweist. Die Artikel werden von Experten untersucht und mit meist 5 – 15 MeSH-Begriffen etikettiert, die am stärksten den Inhalt des Artikels wiedergeben oder Eigenschaften des Studienobjektes bzw. der Publikationsform reflektieren. Zusätzlich zu „MeSH-Headings“ können „Subheadings“ vergeben werden, um den Aspekt näher zu beschreiben. Medical Subject Headings

gewährleisten einheitliches und konsistentes Indexing von biomedizinischer Literatur und werden jährlich aktualisiert. Die MeSH-Datenbank kann über PubMed eingesehen werden und liefert bei Eingabe von Suchbegriffen, deren Definition und gegebenenfalls Vorschläge für bessere Synonyme (Universitätsklinikum Freiburg / Deutsches Cochrane Zentrum, 2013; Bethesda M D, 2005; U.S. National Library of Medicine, 01.03.2001). Erzielt die Suchanfrage zu wenige Treffer, kann die Suchstrategie erweitert werden, indem alternative, allgemeinere Suchbegriffe verwendet werden oder die Option „Similar Articles See All“ ausgewählt wird, um ähnliche PubMed-Zitate zu finden. Im Kontrast dazu kann die Suchstrategie bei zu vielen Ergebnissen eingeschränkt werden. Dazu werden Suchbegriffe durch spezifischere Alternativen ersetzt und zusätzliche Begriffe oder Operatoren und andere Filter verwendet. Filter können über die Filterleiste ausgewählt werden und Ergebnisse in verschiedenen Kategorien wie Sprache oder Verfügbarkeit einschränken (Bethesda M D, 2005). Operatoren verknüpfen die Suchbegriffe von links nach rechts, es sei denn diese automatische Reihenfolge wird durch das Setzen von Klammern verändert (U.S. National Library of Medicine, 01.03.2001). Die wichtigsten Operatoren sind:

- AND (Suche nach beiden Begriffen)
- OR (Suche nach mindestens einem der Begriffe)
- NOT (Schließt den zweiten Begriff aus der Suche aus) (Bethesda M D, 2005; Universitätsklinikum Freiburg / Deutsches Cochrane Zentrum, 2013).

Nach Festlegung einer geeigneten Suchstrategie werden die Überschriften der entsprechenden Suchergebnisse überflogen und das beste passende Abstract ausgewählt.

1.1.6 Einführung: Evidenzstufen

Damit die Rechercheergebnisse angemessen evaluiert werden können, sollte die hierarchische Einordnung externer Evidenz nach Validitätskriterien bekannt sein. Diese wurden von Cochrane Deutschland veröffentlicht und sind in absteigender Reihenfolge:

- Ia) mindestens ein Systematic Review auf Basis von randomisierten kontrollierten Studien (RCTs)
- Ib) mindestens ein RCT
- IIa) mindestens eine hochwertige Studie ohne Randomisierung (z. B. Crossover-Studie)
- IIb) mindestens eine hochwertige Studie eines anderen Typs quasi-experimenteller Studien (z. B. Kohorten-Studie)
- III) mehr als eine methodisch hochwertige nichtexperimentelle Studie (z. B. Fall-Kontroll-Studie)
- IV) Expertenmeinung oder beschreibende Studien (z. B. Case Report/ Case Series) (Cochrane Deutschland).

Weitere konkrete Ausführungen zur Evaluation der Rechercheergebnisse (Appraise) liefern unter anderem das „2011 OCEBM Levels of Evidence“-Konzept des Oxford Centre for Evidence Based Medicine (Howick J et al.), eine 22-teilige, im deutschen Ärzteblatt erschienene Serie zur Bewertung wissenschaftlicher Publikationen (Du Prel et al., 2009), die Erläuterungen zu Klassifikationssystemen der Evaluation auf der Website der Cochrane Library (Cochrane Deutschland), sowie eine sechsteilige Serie des Journals der American Osteopathic Association (Virgilio et al., 2007).

1.1.7 Einführung: Valide EbM-Evaluationsinstrumente

Als anerkannte Instrumente zur Überprüfung von Fähigkeiten in der Evidenzbasierten Medizin dienen Fresno Test und Berlin Questionnaire (Ramos KD et al., 2003; Fritsche, 2002). Bei dem Fresno Test handelt es sich um 12 größtenteils offen gestellte, meist fallbasierte Fragen. Sie beziehen sich auf die vier wichtigen Methodik Schwerpunkte Ask, Access, Appraise und Apply, sowie auf statistische Basiskenntnisse. Im Vergleich dazu besteht der Berlin Questionnaire aus 15 Szenario-basierten Multiple Choice-Fragen und deckt neben den Schwerpunkten Appraise und Apply noch tiefergehende mathematische Kenntnisse ab. Obwohl beide Instrumente für sich allein als valide Instrumente anerkannt sind, zeigte eine Studie von Lai et al., dass die beiden Tests nur geringfügig korrelierten. Hierbei wurden die beiden Formate jeweils leicht adaptiert und hintereinander von Medizinstudenten, die EbM-Training erhielten, bearbeitet (Lai NM et al., 2012). Im direkten

Vergleich weist der Fresno Test nicht nur die beste dokumentierte Validität und Reliabilität auf, sondern ist auch besonders geeignet um eine Änderung des Wissensstandes, zum Beispiel nach einer Lerneinheit in EbM zu testen. Daher habe ich den Fresno Test als Vorlage für meinen eigenen Test zur Evaluation der praktischen EbM-Fähigkeiten gewählt. Eine genaue Übersicht der hierfür notwendigen Adaptationen folgt im Kapitel 3.5.2 Adaptationen: Vom Fresno Test zum Bonn Test. Der Fresno-Test ist gegliedert in sieben Kurzfragen, zwei mathematische Aufgaben und 3 Lückentexte und nimmt etwa 30 Minuten Bearbeitungszeit in Anspruch. Für jede Frage legt ein Erwartungshorizont fest, welche Inhalte in der Antwort wie ausführlich abgedeckt sein müssen, um eine bestimmte Punktzahl zu erzielen. Insgesamt können 212 Punkte erreicht werden. Eine Frage gilt ab 50 % der maximal erreichbaren Punkte als bestanden. (Ramos KD et al., 2003; Thomas und Kreptul, 2015).

1.2 Fragestellung

Das Ziel dieser Studie ist es die Machbarkeit der praktischen Anwendung von Evidenzbasierter Medizin im Rahmen des humanmedizinischen Studiums im Fachbereich für Allgemeinmedizin unter Berücksichtigung der folgenden Fragestellungen zu prüfen:

1. Ist die Implementierung einer Lerneinheit zum Thema EbM mit praktischer Anwendung von EbM in das Blockpraktikum Allgemeinmedizin machbar?
2. Ist die Erfassung von Leistungen der Studenten in Bezug auf die unterschiedlichen EbM-Fähigkeiten durch den Bonn Test und seinen Erwartungshorizont machbar?
3. Wie beurteilen die Studierenden die Implementierung einer Lerneinheit zum Thema EbM in das Blockpraktikum Allgemeinmedizin?

2. Material und Methoden

2.1 Entwicklung des Lehrkonzepts

Für den organisatorischen Rahmen bei der Vermittlung von EbM-Grundlagen konnten sich in der Vergangenheit interaktiv gestaltete Schulungen im Blended Learning Stil als besonders erfolgreich erweisen (Ilic et al., 2015). Auf dieser Basis entwickelte ich ein interaktives Lernmodul, auf das die Studenten über einen online Zugang von zu Hause zugreifen können. Bei der Ausarbeitung des Lehrkonzepts standen Inhalt, praktische Umsetzung und Evaluationsmöglichkeiten des Projekts im Vordergrund. Nach einer Schulung des Hochschulrechenzentrums zur Programmierung von Lernmodulen entwickelte ich ein Modul auf der online Plattform E-Campus, die von zahlreichen Instituten zur Bereitstellung von Lehrmaterial für die Studenten genutzt wird. Im Rahmen eines Pretests bearbeiteten vier Kommilitonen das interaktive Lernmodul selbstständig und gaben unmittelbar danach Rückmeldung bezüglich der didaktischen, inhaltlichen und technischen Umsetzung. Ihre Verbesserungsvorschläge konnten die benutzerfreundliche Anwendung weiterhin optimieren. Inhaltlich wird der häufig in EbM-Schulungen verwendete 5-A-Schritteprozess untersucht. Die 5 As werden jeweils anschaulich dargestellt und durch weitere Modelle wie das PICO-Schema zur Entwicklung einer wissenschaftlichen Fragestellung ergänzt. Insgesamt erhalten die Studenten dadurch eine Anleitung zur strukturierten Anwendung von EbM. Das Selbststudium mit dem Lernmodul wird zusätzlich ergänzt durch zwei Seminartermine und die praktische Anwendung der erworbenen Fähigkeiten in den Lehrarztpraxen. Um die Leistung der Studenten dabei beurteilen zu können habe ich den „Bonn Test“ mit Erwartungshorizont konstruiert. Dieser ist angelehnt an den Fresno Test, kann aber praktische Fertigkeiten anstelle der theoretischen Grundlagen bewerten.

Tab. 1: Ablauf des EbM-Lehrkonzepts.

1 Woche vor Beginn des BPs	1. Woche BP	2. Woche BP
Online Modul zur Bearbeitung freigeschaltet	Bis einschließlich Montag Vormittag: Bearbeitungszeit für online Lernmodul und fiktive Fallbeispiele	Bis einschließlich Montag Vormittag: Anwesenheit im BP/ Bearbeitungszeit für reale Fallbeispiele
Hausaufgabe: Fiktive Fallbeispiele A, B und C	Montag Nachmittag: 1. Seminartermin (Besprechung der fiktiven Fallbeispiele/ Arbeitsauftrag: reale Fallbeispiele im BP)	Montag Nachmittag: 2. Seminartermin (Besprechung der realen Fallbeispiele/ gemeinsame Bearbeitung eines Fallbeispiels)
	Einsammeln der bearbeiteten fiktiven Fallbeispiele und Auswertung mit Hilfe des Erwartungshorizonts	Einsammeln der bearbeiteten realen Fallbeispiele und Auswertung mit Hilfe des Erwartungshorizonts

2.2 Das Online-Lernmodul

Das Lernmodul ist inhaltlich in eine allgemeine Einführung zu Evidenzbasierter Medizin und Kapitel zu jedem Unterpunkt des 5-A-Schritteprozesses unterteilt. Um das Selbststudium mit dem Lernmodul möglichst interaktiv zu gestalten, sind den einzelnen Kapiteln jeweils drei bis vier Wiederholungsfragen zwischengeschaltet. Sobald die Studenten ein Kapitel abgeschlossen haben, können sie mit Hilfe der Multiple Choice- und Zuordnungsfragen ihr Verständnis für die wichtigsten Inhalte umgehend überprüfen. Das Format auf E-Campus erlaubt dabei die unmittelbare Rückmeldung für die Studenten, ob ihre Antwort richtig ist bzw. welche Wahl die richtige gewesen wäre. Anschließend haben die Blockpraktikanten Zugriff auf die Fallbeispiele A, B und C (Kapitel 2.3: Der Arbeitsauftrag), sowie auf den Bonn Test (Kapitel 2.5: Der Bonn Test). Eine Druckversion

des online Lernmoduls kann in Anlage I eingesehen werden. Anlage II umfasst die Aufgabenstellungen der drei Fallbeispiele A, B und C und Anlage III den Erwartungshorizont des Bonn Tests.

2.2.1 Lernmodul: Einführung in die EbM

Einen Einstieg in die Thematik bietet die EbM-Definition nach Sackett, der als Gründungsdirektor des Centre for Evidence-Based Medicine des National Health Service an der Universität Oxford dieses Konzept maßgeblich prägte (Heneghan C, 2015). Danach ist EbM die „gewissenhafte, explizite und umsichtige Anwendung der besten verfügbaren Evidenz um Behandlungsentscheidungen zum Wohle des individuellen Patienten zu treffen“ (“Evidence based medicine ist the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients“) (Sackett et al., 1996). Diese Definition wird ergänzt durch das Bild der „drei Pfeiler“, von denen die EbM gestützt wird (**Abb. 2**):

1. „Attitude“ ist die Bereitschaft des behandelnden Arztes sich ein Leben lang fortzubilden. Er sollte fähig sein, wissenschaftliche Arbeiten kritisch zu evaluieren, damit er stets gut fundierte Entscheidungen treffen kann, ohne den Fokus von den individuellen Bedürfnissen und Wünschen des Patienten abzuwenden.
2. „Skills“ sind die Fertigkeiten zur Anwendung von EbM. Sie werden im weiteren Verlauf des Moduls näher besprochen.
3. „Knowledge“ ist die Verbindung aus interner und externer Evidenz. Bei der internen Evidenz handelt es sich um durch das Medizinstudium erworbenes, allgemein anerkanntes fachliches Grundwissen, auf das der Arzt zugreifen kann. Die externe Evidenz geht aus der gezielten Literaturrecherche hervor und muss noch vor der Anwendung evaluiert werden (Beyer M et al., 2005).

Hervorgehoben wird hier auch, dass neben der Herausarbeitung der Patientenbedürfnisse als höchster Priorität das Zusammenspiel zwischen klinisch externer Evidenz und klinischer Expertise wichtig ist. Ohne die durch Erfahrung gewonnene Expertise des Mediziners kann auch die beste Evidenz auf den einzelnen Patienten fehlerhaft angewendet werden, während eine Behandlung auf Grundlage der

besten Expertise ohne entsprechende Evidenz bestenfalls überholt ist (Sackett et al., 1996).

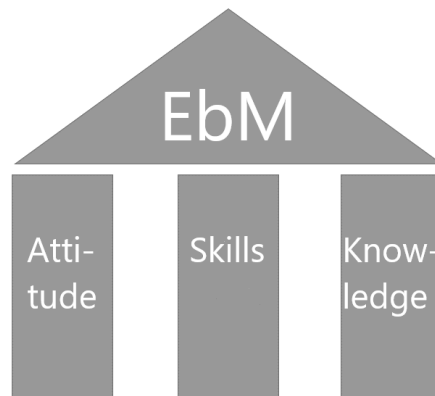


Abb. 2: Darstellung der „3 Pfeiler“ Attitude, Skills und Knowledge, von denen die EbM gestützt wird.

2.2.2 Lernmodul: Der 5-A-Schritteprozess

Nach dieser allgemeinen Einführung des Lernmoduls in die Evidenzbasierte Medizin erfolgt die inhaltliche Präsentation des 5-A-Schritte-Prozesses nach Cook et al.:

1. Ask: Formulierung einer wissenschaftlichen Frage
2. Acquire: Literaturrecherche zur Beantwortung der Fragestellung
3. Appraise: Kritische Evaluation der Rechercheergebnisse
4. Apply: Auswahl der Behandlung aufgrund der gefundenen Evidenz, klinischer Erfahrung des behandelnden Arztes, sowie der Bedürfnisse des Patienten
5. Assess: Evaluation der Behandlung (Cook et al., 1992)

Es folgt der erste interaktive Abschnitt mit einer Multiple Choice Frage zu den „drei Pfeilern“ der EbM und zwei Zuordnungsaufgaben der einzelnen Punkte des 5ASP.

2.2.3 Lernmodul: Ask

Die Formulierung einer wissenschaftlichen Fragestellung sollte nach dem PICO-Schema erfolgen. Das Fallbeispiel einer 42-jährigen Patientin mit chronischen Schmerzen im Lendenwirbelbereich, die aus Zeitgründen eine Therapie mit Akupunktur einem regelmäßigen Sportprogramm vorziehen würde, soll das PICO-Schema veranschaulichen. Als praktische Übung sollen die Studenten im Anschluss jeweils die Bedeutung der Buchstabenabkürzungen reproduzieren.

2.2.4 Lernmodul: Acquire - Studientypen

Eine systematische Literaturrecherche sollte entsprechend der folgenden Schritte ablaufen:

1. Auswahl von Studientypen und Recherchequellen
2. Festlegung von Suchvokabular und -strategie, Durchführung der ersten Suche
3. Sichtung der Ergebnisse (und ggf. Durchführung einer angepassten zweiten Suche)
4. Dokumentation der Recherche (Universitätsklinikum Freiburg / Deutsches Cochrane Zentrum, 2013).

Die Ergebnisqualität der Suche ist stark abhängig von der Entwicklung der Suchstrategie. Publikationsbias und andere Verzerrungen können durch eine umfassende und systematische Auseinandersetzung mit einer Vielzahl von Quellen minimiert werden (Universitätsklinikum Freiburg / Deutsches Cochrane Zentrum, 2013). Die Literaturrecherche konzentriert sich zunächst auf die erfolgversprechendsten Studientypen und wird auf den nächstbesten Studientypen ausgeweitet, wenn sich keine Evidenz findet (Howick J et al.). Um die Forschungsergebnisse für den Patienten richtig beurteilen zu können, ist ein Grundverständnis für Planung, Auswertung und Aussagekraft der unterschiedlichen Studientypen notwendig (Du Prel JB et al., 2009). Es folgt daher in unserem Lernmodul eine Vorstellung der wichtigsten Kategorien von Studien (Kapitel 1.1.3 Einführung: Studientypen) und ihre Anordnung zu einer Pyramide. In aufsteigender Reihenfolge nimmt ihre Empfindlichkeit für systematische Fehler und Verzerrungen ab (Duke University Medical Center Library und Health Sciences Library at the University of North Carolina at Chapel Hill, 2014).

2.2.5 Lernmodul: Acquire – Recherchequellen

Es folgt eine kurze Auflistung wichtiger Recherchequellen in Bezug auf Relevanz und Validität (Kapitel 1.1.4: Einführung: Recherchequellen). Im Anschluss wird das Verständnis der Studenten für die Validität überprüft. Aus einer Liste von fünf Quellen sollen sie diejenigen bestimmen, die sich auf die Bereitstellung von vorab evaluierten Informationen spezialisiert haben. Drei weitere Fragen befassen sich mit den unterschiedlichen Studientypen.

2.2.6 Lernmodul: Acquire – PubMed

Nach der Auseinandersetzung mit verschiedenen Recherchequellen und Studientypen werden im nächsten Schritt wichtige Suchstrategien am Beispiel von PubMed erläutert. Auch wenn in der Praxis häufig vorab evaluierte Recherchequellen zur Beantwortung klinischer Fragestellungen herangezogen werden, habe ich für die Suchanfrage im Lernmodul PubMed gewählt. Auf diese Weise haben die Studenten die Möglichkeit den 5-A-Schritteprozess systematisch zu durchlaufen und die Gelegenheit ihre Rechercheergebnisse kritisch zu evaluieren. Bei einer Suchstrategie liegt der Schwerpunkt entweder auf der Genauigkeit oder auf der Vollständigkeit der Ergebnisse (Universitätsklinikum Freiburg / Deutsches Cochrane Zentrum, 2013). Für eine kurzfristige Recherche, wie sie im Rahmen des Seminars stattfindet, ist die Genauigkeit wichtiger und sollte anhand von 3 Schritten erfolgen:

1. Umwandlung der PICO-Elemente in eine Suchanfrage
2. Berücksichtigung von „Medical Subject Headings“ (Schlagwortvokabular) und ggf. Verwendung von Filtern und Operatoren
3. Limitierung auf ein geeignetes Studiendesign (Duke University Medical Center Library und Health Sciences Library at the University of North Carolina at Chapel Hill, 2014).

An dieser Stelle wird noch einmal das Fallbeispiel der 42-jährigen Sekretärin mit chronischen Schmerzen im Lendenwirbelbereich aufgegriffen, das bereits in dem Kapitel zum PICO-Schema Anwendung gefunden hat. Um die Notwendigkeit einer präzisen, aber nicht zu spezifischen Suchanfrage zu demonstrieren, werden zwei beispielhafte Suchstrategien dokumentiert. Die allgemeine Formulierung „chronic back pain treatment“ erzielte dabei 9314 Treffer auf PubMed. Die spezifischere Suchanfrage „chronic back pain acupuncture“ grenzte die Trefferanzahl bereits auf 416 ein. Bei der Suchanfrage sollten möglichst Schlagwörter anstelle von Stichwörtern verwendet werden (Kapitel 1.1.5 Einführung: PubMed). Mit gleichzeitiger Filterauswahl für Reviews und frei verfügbare Artikel reduziert sich die Anzahl der Suchergebnisse für die Anfrage „chronic back pain acupuncture“ nochmals auf 27. Ist die Trefferanzahl nach diesen Schritten immer noch zu hoch, können zusätzlich Operatoren verwendet werden, um die Ergebnisse

einzu­schränken (Kapitel 1.1.5 Einführung PubMed). Ich habe den inhaltlich passendsten Review dieser 27 Beiträge ausgewählt und in Zitat­form dokumentiert. Somit werden den Studenten im Rahmen des Lernmoduls unterschiedliche Möglichkeiten vorgestellt die Suchstrategie auf PubMed zu gestalten. Ein Link zu einem PubMed-internen Schulungsinstrument liefert optional weitere Informationen zum Umgang mit der Literaturdatenbank. Nach Abschluss des Kapitels Acquire besteht wieder die Möglichkeit zur Selbstkontrolle mit drei Multiple Choice Fragen zu diesem Thema.

2.2.7 Lernmodul: Appraise – Validität

Der Cochrane Deutschland Vereinigung zu folge sollte die Evaluation einer Studie systematisch nach 3 Aspekten erfolgen:

1. Validität
2. Größe und Präzision
3. Übertragbarkeit und Anwendbarkeit (Cochrane Deutschland).

Um die Gültigkeit einer Studie kritisch beurteilen zu können, wird den Studenten die hierarchische Ordnung externer Evidenz nach Validitätskriterien vermittelt (Kapitel 1.1.6 Einführung: Evidenzstufen). Sie sollen dafür sensibilisiert werden, dass die Nutzer solcher Hierarchien von Evidenz häufig die Tendenz entwickeln, das eigene Urteilsvermögen zu vernachlässigen. Die einzelnen Studienarten können im Evidenz-Ranking je nach Inhalt ein „Up“- beziehungsweise „Downgrade“ erhalten (Oxford Centre for Evidence-Based Medicine). RCTs sind besonders aussagekräftig bei Studien zu möglichen negativen Wirkungen, da die Ergebnisse von Studien- und Kontrollgruppe deutlich differenzierbar sind (Cardarelli und Seater, 2007). Kohorten Studien sind hingegen besser geeignet um seltene Nebenwirkungen aufzudecken (Kabisch et al., 2011).

2.2.8 Lernmodul: Appraise – Größe und Präzision

Die Beurteilung wissenschaftlicher Veröffentlichungen kann jeweils in drei Kategorien erfolgen:

1. Design
2. Durchführung
3. Auswertung

Das Studiendesign umfasst nicht nur die Wahl eines angemessenen Studientyps, sondern auch das Gesamtkonzept der Vorgehensweise einer Studie. Relevante Aspekte sind hierbei Fragestellung, Studienpopulation, Messverfahren und Beobachtungseinheit sowie Fallzahlplanung. Die Fragestellung legt fest, ob die Studie deskriptiv, explorativ oder konfirmatorisch auszuwerten ist und sollte innovativ und medizinisch relevant sein (Du Prel et al., 2009; Röhrig et al., 2009). Sie ist in einem Großteil der Studien „disease-oriented“ und macht den Erfolg oder Nicht-Erfolg einer Behandlung an einer Verbesserung von Parametern wie Blutdruck oder Puls fest. Tatsächlich wird der Erfolg in der Praxis häufiger mit Hilfe von „patient-oriented“ Kriterien bestimmt. Dabei stehen dann Faktoren wie die Mortalitätsrate oder ein Schmerzrückgang eher im Vordergrund (Cardarelli et al., 2007). Die Wahl eines geeigneten statistischen Testverfahrens ist unter anderem abhängig von der Zielgröße und der ihr zugrundeliegenden Verteilung (Du Prel et al., 2010). Die Fallzahl muss mindestens so groß sein, dass die Aussagekraft gewährleistet ist. Idealerweise sollte sie noch größer gewählt werden, um das möglich Ausscheiden von einzelnen Teilnehmern aus der Studie kompensieren zu können (Röhrig et al., 2010). Eine multizentrische, das heißt an mehreren Instituten durchgeführte Studie, kann die nötige Fallzahl im Vergleich zu einer unizentrischen Studie leichter erreichen (Röhrig et al., 2009). Für die Beurteilung der Durchführung einer Studie sollte eine möglichst detaillierte Beschreibung des Studienablaufs vorliegen, die es ermöglicht, die Zuverlässigkeit und Gültigkeit der Studie zu bewerten (Du Prel et al., 2009).

2.2.9 Lernmodul: Appraise – Übertragbarkeit und Anwendbarkeit

Neben den Methoden, mit denen die Studienergebnisse erzielt werden, sollten auch die aus ihnen gezogenen Schlüsse auf ihre Gültigkeit hin überprüft werden (Schranz und Dunn, 2007). Außer der internen Validität sollte auch die externe Validität der Studie, also die Übertragbarkeit der Ergebnisse von der Studienpopulation auf die Zielpopulation, kritisch hinterfragt werden (Röhrig et al., 2009). Zuletzt ist die Relation zwischen Kosten und Nutzen für die Behandlung besonders in Hinblick auf die individuellen Endpunkte des Patienten abzuwägen (Cardarelli et al., 2007; Cochrane Deutschland). Weitere Kriterien zur Beurteilung wissenschaftlicher Veröffentlichungen sind unter anderem die korrekte Verwendung von Literaturangaben oder etwaigen Interessenkonflikte der Autoren (Du Prel et al., 2009).

Als Beispiel wird an dieser Stelle wieder das Abstract aus dem vorherigen Kapitel gewählt, das bei der Sichtung der Überschriften für die Bearbeitung des klinischen Falles zu Rückenschmerzen und Akupunktur in Betracht kam. Die relevanten Informationen zu Validität, Größe und Präzision, sowie Übertragbarkeit und Anwendbarkeit, die dem Abstract entnommen werden können, sind jeweils farblich hervorgehoben. Im nächsten interaktiven Abschnitt des online Lernmoduls haben die Studenten Gelegenheit die Evidenzstufen den entsprechenden Definitionen zuzuordnen. Zwei weitere Multiple Choice Fragen wiederholen die Kriterien zur Beurteilung der Größe und Präzision, sowie der Übertragbarkeit und Anwendbarkeit einer Studie.

2.2.10 Lernmodul: Apply

Die Anwendung der Rechercheergebnisse ist ein Prozess, der sich auf die vorherigen Punkte Ask, Acquire und Appraise stützt. Die Therapieempfehlung sollte sorgfältig evaluiert werden und zu den individuellen Bedürfnissen des Patienten passen.

2.2.11 Lernmodul: Assess

Die Evaluation der Behandlung dient der Beurteilung des gesamten 5-A-Schritteprozesses. Die Patientenproblematik ist dabei jeweils Start- und Endpunkt des Prozesses. Bleibt der Erfolg der Therapie aus, wird der 5ASP erneut mit einer angepassten Fragestellung durchlaufen.

2.3 Der Arbeitsauftrag

Im Anschluss an das Lernmodul sollen die Blockpraktikanten den 5-A-Schritteprozess praktisch anwenden. Hierfür habe ich 3 Fallbeispiele entwickelt (Kapitel 9.2 Anlage II: Aufgabenstellung der Fallbeispiele A, B und C). Welches der drei Fallbeispiele sie bearbeiten sollen, wurde ihnen zuvor per E-Mail mitgeteilt. Anhand von 10 vorgegebenen Fragen sollten die Studenten ihr erlerntes Grundwissen erstmalig anwenden. Da die Blockpraktikanten diese Aufgabe im Rahmen des Selbststudiums durchführen sollen, hilft ein weiteres interaktives Fallbeispiel die einzelnen Schritte nochmals zu durchlaufen, um mögliche Fragen vor der Bearbeitung der Fallbeispiele A bis C auszuräumen. Der Patient in diesem interaktiven Fallbeispiel ist ein chronischer Gichtpatient mit Allopurinol-unverträglichkeit, der den Wunsch nach einer alternativen Therapie äußert. Die drei

Fallbeispiele A, B und C stammen ebenfalls aus dem Themenfeld Gicht. Dazu haben wir uns entschieden, weil die Studenten im Rahmen einer zweiten Studie des Instituts für Hausarztmedizin während des Blockpraktikums in den Lehrarztpraxen mit Gichtpatienten Fragebögen ausfüllen sollen. Auf diese Weise frischen die Studenten ihr Wissen zu dem Krankheitsbild Gicht auf und können es um die aktuelle Evidenz ergänzen. Die konstruierten Patientenbeispiele orientieren sich jeweils an der DEGAM-Leitlinie zur Gichttherapie und gehen inhaltlich auf unterschiedliche Komplikationen bei der Behandlung ein (Engel B und Prautzsch H, 2013). So stellt sich in Fall A die Frage nach der Gichttherapie bei Patienten mit chronischer Niereninsuffizienz. Fallbeispiel B beschäftigt sich mit diätetischen Maßnahmen zur Vermeidung von Gichtanfällen und Patient C will wissen, ob erhöhte Harnsäurewerte einen Risikofaktor für kardiovaskuläre Erkrankungen darstellen. Am Ende des Lernmoduls sollte es den Studenten möglich sein aus der geschilderten Situation eine Fragestellung zu gewinnen (Ask), eine Literaturrecherche zu diesem Thema durchzuführen (Acquire), die gefundene Evidenz zu evaluieren (Appraise), daraus die bestmögliche Behandlungsoption für den Patienten abzuleiten (Apply) und Maßnahmen für die Evaluation der Behandlung festzulegen (Assess). Das Ergebnis ihrer Recherche bringen die Studenten mit zu dem ersten Seminartermin.

2.4 Die Seminare

Die Studenten nehmen jeweils am Montag der beiden Blockpraktikumsgruppen an einem Seminar im Konferenzraum des Instituts für Hausarztmedizin teil. Ziel des Seminars ist die Vertiefung der erlernten Inhalte des online-Lernmoduls. Bei den drei Testdurchläufen leitete meine Betreuerin Dr. med. Bettina Engel die Seminare. Ich war ebenfalls bei allen Terminen anwesend, um die direkte Rückmeldung der Studenten einzuholen und auf mögliche auftretende Probleme unmittelbar mit Anpassungen des Lehrkonzepts reagieren zu können. Auch in diesem Bereich verfolgten wir den Blended Learning Ansatz anstelle des klassischen Frontalunterrichts. So wurden die Ergebnisse des Bonn Tests in Kleingruppen gemeinsam besprochen und die Inhalte des Lernmoduls durch weitere interaktive Aufgaben aufgearbeitet.

2.4.1 Der erste Seminartermin

Das erste Seminar beginnt jeweils mit einer kurzen Vorstellungsrunde, bei der die Studenten auf ihre Zufriedenheit mit dem Lernmodul, eventuelle Verständnisprobleme, sowie Erwartungen und Wünsche bezüglich des Blockpraktikums eingehen sollen. Anschließend stellt der Dozent das Thema EbM nochmals in einer kompakten PowerPoint Präsentation vor. Für die Präsentation habe ich hauptsächlich die wichtigsten Graphiken und Kernaussagen des Lernmoduls übernommen. Bettina Engel ergänzte sie um weitere Folien zur Veranschaulichung der Inhalte. Dieser Vortrag dient dabei lediglich dem Auffrischen des Erlernten und ersetzt nicht das Selbststudium mit dem Lernmodul. Es folgt eine interaktive Einheit, bei der die Studenten an einer Pinnwand Karten mit verschiedenen Studientypen den richtigen Definitionen zuordnen sollen. Anschließend müssen die Studientypen entsprechend ihrer Evidenzstufen angeordnet werden. Danach besprechen die Studenten mit dem Dozenten unterschiedliche Recherchequellen. Obwohl wir für unser Lernmodul PubMed als Medium ausgewählt haben, da sich hierbei die einzelnen Schritte besser nachvollziehen lassen, sollen die Studenten sich auch mit häufig verwendeten Quellen wie der Cochrane Library oder UptoDate vertraut machen. Zuletzt stellen die Studenten ihre eigenen Suchstrategien und Rechercheergebnisse vor. Diese vergleichen sie zunächst in Kleingruppen mit Kommilitonen, die dasselbe Fallbeispiel A, B beziehungsweise C bearbeitet haben. Anschließend stellt jede Gruppe einen bearbeiteten Bonn Test vor. Die Vorgehensweise wird dann gemeinsam in der großen Gruppe diskutiert. Am Ende des Seminars erhalten die Studenten ihren Auftrag für das Blockpraktikum. Bis zum nächsten Seminartermin am darauffolgenden Montag sollen sie während ihres Aufenthaltes in den Lehrarztpraxen eine eigene wissenschaftliche Fragestellung entwickeln und mit Hilfe des Bonn Tests bearbeiten. Dabei ist ihnen die Wahl des Themas freigestellt. Möglich wäre beispielsweise ein konkreter Patientenfall aus der Lehrarztpraxis oder der neuste Stand bestimmter Therapieoptionen, für deren Recherche der Lehrarzt keine Zeit hat. Die ausgefüllten Unterlagen zu den fiktiven Fallbeispielen A, B und C werden am Ende des ersten Seminartermins eingesammelt und anschließend ausgewertet. Wenn die Antwort eines Studenten noch nicht in dem Erwartungshorizont berücksichtigt wird, kann die Qualität der Antwort in den bereits bestehenden Abstufungen eingeordnet und eine Punktzahl

dafür festgelegt werden. Auf diese Weise kann der Bonn Test durch die vermehrte Anwendung kontinuierlich weiterentwickelt werden.

2.4.2 Der zweite Seminartermin

Im Rahmen des zweiten Seminartermins tauschen sich die Studenten über die bisherigen Erfahrungen in den Lehrarztpraxen aus. Hierbei besteht die Möglichkeit auf Probleme und Unklarheiten einzugehen. Die Studenten stellen ihre selbstentwickelten Fragestellungen zu klinischen Fällen aus den Lehrarztpraxen sowie ihre Rechercheergebnisse und Therapieempfehlungen vor. Anschließend wird eine der Fragestellungen gemeinsam mit Hilfe des online Zugangs vor Ort recherchiert, um die praktische Anwendung von EbM nachvollziehbar auszuführen. Am Ende des Kurses werden die ausgefüllten Unterlagen zu den eigenen klinischen Fragestellungen eingesammelt. Mit Hilfe des Erwartungshorizonts des Bonn Tests können sie wie auch schon zuvor die konstruierten Fallbeispiele ausgewertet werden. Die Leistung der Studenten wird dabei sowohl untereinander als auch im Vergleich der theoretischen mit der praktischen Anwendung beurteilt.

2.5 Der Bonn Test

2.5.1 Valide EbM-Evaluationsinstrumente

Um die Leistungen in den unterschiedlichen Kategorien des 5-A-Schritteprozesses differenzieren und Stärken und Schwächen der Studenten untereinander vergleichen zu können, habe ich den Bonn Test entwickelt. Dabei handelt es sich um 10 offen formulierte Fragen, die an die Inhalte des Lernmoduls angepasst sind, sodass die Reihenfolge der 5 As beibehalten wird. Bei der Entwicklung des Bonn Tests habe ich mich an den etablierten Tests zur Überprüfung von EbM-Schulungen orientiert. Nachdem eine Studie Von Lai et al. zeigte, dass die beiden anerkannten Evaluationsinstrumente Fresno Test und Berlin Questionnaire wenig korrelieren (Lai NM et al., 2012), findet die Evaluation des Bonner Lernmoduls nicht aus einer Kombination beider Instrumente statt. Da es sich bei dem Berlin Questionnaire um Multiple Choice-Fragen handelt, eignet sich der Fresno Test aufgrund seiner größtenteils offen gestellten, meist fallbasierten Fragen eher als Vorlage für den Bonn Test, der in der Praxis auf jede beliebige Fragestellung anwendbar sein soll.

Die vier wichtigen Methodik Schwerpunkte des Fresno Tests Ask, Access, Appraise und Apply habe ich um die restlichen Schritte Acquire und Assess ergänzt, während die Überprüfung statistischer Basiskenntnisse im Bonn Test wegfällt. Den entsprechenden Erwartungshorizont habe ich als Ausgangspunkt genutzt und die nötigen Anpassungen durchgeführt, damit die EbM-Fähigkeiten nicht nur in der Theorie, sondern auch in der praktischen Anwendung geprüft werden können. Zusätzlich ist mein Erwartungshorizont auf die im Lernmodul vermittelten Inhalte abgestimmt (Kapitel 9.3 Anlage III: Erwartungshorizont des Bonn Tests). Bei der Festlegung der Punkte habe ich versucht so nah wie möglich an dem bereits existierenden Schema des Fresno Tests zu bleiben. Für jede geprüfte Fähigkeit des 5-A-Schritteprozesses sind jeweils 24 Punkte erreichbar. Acquire stellt die Ausnahme dar, da hier noch einmal in die Formulierung der Suchstrategie und die Auswahl des Abstracts unterschieden wird und somit zweimal 24 Punkte erreichbar sind. In festgelegten Abstufungen der Ausführlichkeit, werden die Antworten als „hervorragend“, „stark“, „limitiert“ bzw. als „ohne Evidenz“ bewertet und mit einer angepassten Punktzahl bewertet.

2.5.2 Adaptationen: Vom Fresno Test zum Bonn Test

Die erste Aufgabe des Fresno Tests befasst sich mit der Formulierung einer klinischen Fragestellung. Es werden jeweils zwei Fallbeispiele vorgegeben, aus denen alle notwendigen Informationen für eine Suchanfrage herauszufiltern sind. Der Erwartungshorizont orientiert sich dabei an dem PICO-Schema. Zum Erreichen der vollen Punktzahl müssen mehrere relevante Eigenschaften des Patienten berücksichtigt, eine Intervention und alternative Behandlung spezifisch benannt und ein objektiver und für den Patienten bedeutsamer Ausgang der Intervention festgelegt werden. In den einzelnen Kategorien Patient, Intervention, Comparison und Outcome des PICO-Schemas können maximal 3 Punkte also pro Fallbeispiel 12 Punkte und für die gesamte erste Aufgabe 24 Punkte erreicht werden. Sind die Angaben in den einzelnen Kategorien zu unspezifisch oder nicht zielführend für die Recherche, werden die Punkte dem Erwartungshorizont entsprechend abgestuft. Da die erste Aufgabe des Fresno Tests den ersten Schritt des 5-A-Schritteprozesses (Ask) abdeckt und sich dabei, wie im Lernmodul empfohlen, am PICO-Schema orientiert, kann sie in dieser Form für den Bonn Test übernommen werden. Dadurch, dass die Blockpraktikanten jedoch nur ein Fallbeispiel am Ende des Lernmoduls

bearbeiten sollen, werden die im Erwartungshorizont festgelegten Fähigkeiten jeweils mit der doppelten Punktzahl bewertet, sodass insgesamt wieder 24 Punkte erreicht werden können. Die tatsächliche Aufgabenstellung des Bonn Tests leitet sich von dem angepassten Erwartungshorizont ab.

- In der ersten Aufgabe des Bonn Tests sind die wichtigen Informationen des Falles in kurzen Stichpunkten zusammenzufassen. Diese Aufgabe fließt nicht in die Bewertung des Erwartungshorizonts ein. Dennoch habe ich sie als erste Aufgabe festgelegt, damit der Ausgangspunkt des Studenten für die Korrektur des Bonn Tests dokumentiert ist. Treten beispielsweise Probleme bei der Anwendung des PICO-Schemas auf, so kann mit Hilfe der Stichworte nachvollzogen werden, ob die relevanten Informationen nicht aus der Fallbeschreibung herausgearbeitet wurden, oder ob die Translation in das PICO-Schema nicht funktioniert hat.
- In der zweiten Aufgabe soll aus dem Fall eine wissenschaftliche Fragestellung im PICO-Schema abgeleitet werden (Ask).

Aufgabe zwei bis vier des Fresno-Tests prüfen die Fähigkeit der Literaturrecherche. Dabei sollen in Aufgabe zwei unterschiedliche Recherchequellen benannt und in Bezug auf ihre Zweckmäßigkeit, klinische Relevanz und Validität diskutiert werden. In Aufgabe drei soll anschließend eine konkrete Suchstrategie bei MEDLINE (PubMed) ausformuliert werden. Neben einer angemessenen Auswahl von Suchbegriffen, die die relevanten Informationen des PICO-Schemas berücksichtigen, ist das Verständnis für die PubMed eigenen Index-Kategorien („tags“) gefordert. Diese ermöglichen es zum Beispiel explizit nach einem Titel, Autor, Journal oder Schlag- bzw. Stichwörtern zu suchen. Zusätzlich werden Punkte für das Einschränken der Suchergebnisse vergeben, beispielsweise durch die Verwendung von Operatoren oder der Limitation auf ein Studiendesign. In Aufgabe vier ist die begründete Festlegung auf ein favorisiertes Studiendesign gefordert. In Aufgabe zwei bis vier des Fresno Tests sind jeweils maximal 24 Punkte zu erreichen. Gemeinsam repräsentieren sie den zweiten Punkt des 5-A-Schritteprozesses, Acquire. Der Erwartungshorizont des Fresno Tests wird für den Bonn Test hinsichtlich der im Lernmodul vermittelten Inhalte angepasst. Somit wird die Suchstrategie in Bezug auf die Wortwahl, sowie die Verwendung von Schlagwörtern (MeSH) und präzisierenden Optionen bewertet. Bei der Auswahl des Abstracts gibt es jeweils für den Inhalt und die

Qualität des zugrundeliegenden Designs Punkte. Sowohl mit der Suchstrategie, als auch mit der Auswahl des Abstracts können maximal 24 Punkte erreicht werden.

- Die Aufgabenstellungen drei bis fünf des Bonn Tests decken die beiden Teilbereiche Suchstrategie und Abstractauswahl des Schrittes Acquire ab. In Aufgabe drei soll zunächst die Suchanfrage dokumentiert werden.
- Aufgabe vier erfragt zusätzliche Optionen, die bei der Arbeit mit PubMed genutzt werden können, um die Suchanfrage möglichst präzise zu gestalten.
- In Aufgabe fünf ist ein Abstract auszuwählen, das am besten geeignet ist, um die zuvor festgelegte Fragestellung zu beantworten.

In Aufgabe fünf, sechs und sieben des Fresno Tests soll das ausgewählte Abstract jeweils in Bezug auf seine Relevanz, Validität, sowie die Aussagekraft und statistische Signifikanz bewertet werden. Abhängig von der Ausführlichkeit, mit der die einzelnen Kategorien mit Beispielen besprochen werden, können je Kategorie wieder maximal 24 Punkte erreicht werden. Gemeinsam repräsentieren die Aufgaben fünf bis sechs die Evaluation der gewählten Studie. Da diese Fähigkeit als Appraise im Lernmodul vergleichsweise komprimiert vermittelt wird, können entsprechend des adaptierten Erwartungshorizonts in den Kategorien Validität, Größe und Präzision sowie Übertragbarkeit und Anwendbarkeit jeweils acht und somit für den gesamten Schritt Appraise 24 Punkte erreicht werden.

- Die Aufgaben sechs, sieben und acht des Bonn Tests decken den Schritt Appraise ab. In Aufgabe sechs soll für das ausgewählte Abstract die richtige Evidenzstufe festgelegt werden.
- In Aufgabe sieben sollen Design, Durchführung und Auswertung beurteilt werden.
- In Aufgabe acht soll die Übertragbarkeit auf den eigenen Patienten beurteilt werden.

Aufgabe acht bis zehn des Fresno Tests befassen sich jeweils mit statistischen Prinzipien, die im Lernmodul nicht berücksichtigt werden und somit auch für den Erwartungshorizont des Bonn Tests entfallen. In den letzten beiden Aufgaben, elf und zwölf, wird das Verständnis für die Auswahl eines geeigneten Studiendesigns in Abhängigkeit der entsprechenden Fragestellung evaluiert. In dem adaptierten Erwartungshorizont wird diese Fähigkeit nicht gesondert bewertet, sondern ist bereits in der Kategorie Acquire

(Auswahl des Abstracts) integriert. Um den 5-A-Schritteprozess zu vervollständigen wird der Erwartungshorizont des Fresno Tests für den Bonn Test um die beiden Kategorien Apply und Assess erweitert. Die gezogenen Schlüsse für die Therapieempfehlung (Apply) aus den evaluierten Rechercheergebnissen werden jeweils nach ihrer internen beziehungsweise externen Validität bewertet. Gemeinsam können mit der internen und externen Validität im Bereich Apply 24 Punkte erreicht werden.

- In Aufgabe neun des Bonn Tests wird die Fähigkeit Apply überprüft, indem aufgrund der evaluierten Evidenz eine Behandlung ausgewählt werden soll.

Im letzten Schritt, Assess, soll eine Evaluationsmaßnahme für die gewählte Therapieoption empfohlen werden, bei der die Evaluation nach einem festgelegten Zeitintervall erfolgt und die Möglichkeit bietet, Schäden für den Patienten frühzeitig zu erkennen und abzuwenden. Auch hierfür werden maximal 24 Punkte vergeben.

- Der entsprechende Arbeitsauftrag in Aufgabe zehn des Bonn Tests sieht vor, dass eine sinnvolle Maßnahme festgelegt werden soll, um die zuvor gewählte Behandlung evaluieren zu können.

Der adaptierte Erwartungshorizont deckt somit alle Schritte des 5-A-Schritteprozesses ab. Insgesamt können im Rahmen des Bonn Tests 144 Punkte erreicht werden. Im Folgenden sind die 10 Fragen des Bonn Tests aufgeführt.

1. Fassen Sie den Fall in kurzen Stichpunkten zusammen.
2. Welche Fragestellung leiten Sie aus dem vorliegenden Fall ab?
3. Wie lautet ihre Suchanfrage bei PubMed?
4. Welche zusätzlichen Optionen bei PubMed nutzen Sie, um die Suchanfrage möglichst präzise zu gestalten?
5. Welches Abstract deckt Ihre zuvor formulierte Fragestellung am besten ab? (Geben Sie die Literaturangabe an und bringen Sie ihr gewähltes Abstract ausgedruckt mit zum nächsten Seminartermin.)
6. Welche Evidenzstufe liegt bei Ihrem Abstract vor?
7. Welche Information zu Design, Durchführung & Auswertung gewinnen Sie aus dem Abstract?

8. Sind die Ergebnisse aus der Studie auf Ihren Patienten übertragbar? Notieren Sie mögliche Differenzen.
9. Für welche Behandlung würden Sie sich aufgrund der evaluierten Evidenz entscheiden?
10. Mit welcher Maßnahme würden Sie Ihre gewählte Behandlung im Verlauf evaluieren?

2.6 Überprüfung der Fragestellungen

Insgesamt werden drei Datensätze generiert.

1. Das Ergebnis des Bonn Tests liegt jeweils für die fiktiven Fallbeispiele A, B und C vor. Mit Hilfe dieser Daten wird zunächst die allgemeine Implementierung der Lehreinheit zum Thema EbM unabhängig vom Kontext der praktischen Anwendung überprüft. Dazu habe ich die folgenden Kriterien definiert, die das Lehrkonzept erfüllen sollte:
 - Grundlagenvermittlung: Die Studenten sollten eine ausreichende Basisausbildung auf dem Gebiet der EbM erhalten, sodass sie das Konzept der evidenzbasierten Medizin verstehen. Insbesondere sollte hier das Wissensmanagement geschult werden.
 - Evaluierbarkeit: Um das Lehrkonzept bzw. die dadurch gewonnenen Kompetenzen evaluieren zu können, muss der Bonn Test als Evaluationsinstrument alle Fähigkeiten des 5-A-Schritteprozesses angemessen erfassen und bewerten. Die differenzierte Leistungsdarstellung in den einzelnen Kategorien mit Hilfe des Erwartungshorizonts kann dabei spezifische Schwierigkeiten aufdecken, sodass das Lehrkonzept fortlaufend weiterentwickelt und verbessert werden kann.
2. Die selbstentwickelten Fragestellungen der Studenten werden wieder mit Hilfe des Bonn Tests bearbeitet und die Qualität der Lösungsansätze entsprechend des Erwartungshorizonts bewertet. Dieser Datensatz prüft die Implementierung der Lehreinheit zum Thema EbM im Kontext der praktischen Anwendung. Damit die

Translation der vermittelten Inhalte in die Praxis gelingen kann muss ein zusätzliches Kriterium erfüllt sein:

- **Universelle Gültigkeit:** Das vermittelte Wissen sollte generalisierbar und damit für die unterschiedlichsten Fragestellungen im klinischen Alltag anwendbar sein. Ebenso notwendig ist die Anwendbarkeit von Bonn Test und Erwartungshorizont auf die realen Patientenfälle.

3. Bei dem dritten Datensatz handelt es sich um die Evaluationsbögen, die die Zufriedenheit der Studenten mit der Implementierung einer Lehreinheit zum Thema EbM in das Blockpraktikum Allgemeinmedizin erfassen sollen. Die wichtigsten Aspekte rund um das Lehrkonzept werden jeweils mit Hilfe einer numerischen Skala von 0 („trifft nicht zu“) bis 10 („trifft voll zu“) bewertet, sodass die Blockpraktikanten auch durch ihre direkte Rückmeldung Einfluss auf die Weiterentwicklung des Lehrkonzepts nehmen können. Dabei steht das folgende Kriterium im Vordergrund:

- **Studentische Akzeptanz:** Das Lehrkonzept kann nur erfolgreich implementiert werden, wenn auch die Perspektive der Studenten in die Entwicklung des Konzeptes mit einfließen.

In Zusammenschau der ersten beiden Datensätze können also die vorab formulierten Fragestellungen 1 und 2 beantwortet werden:

1. Ist die Implementierung einer Lehreinheit zum Thema EbM mit praktischer Anwendung von EbM in das Blockpraktikum Allgemeinmedizin machbar?
2. Ist die Erfassung von Leistungen der Studenten in Bezug auf die unterschiedlichen EbM-Fähigkeiten durch den Bonn Test und seinen Erwartungshorizont machbar?

Der dritte Datensatz dient der Beantwortung der dritten Fragestellung:

3. Wie beurteilen die Studierenden die Implementierung einer Lehreinheit zum Thema EbM in das Blockpraktikum Allgemeinmedizin?

3. Ergebnisse

3.1 Ergebnisse der fiktiven Fallbeispiele A, B und C

Von insgesamt 38 Studenten, die aufgeteilt auf drei Gruppen für die Probedurchläufe des erweiterten Blockpraktikums vorgesehen waren, erschienen 35 zum ersten Seminar. 91,4 % der Seminarteilnehmer reichten ihr bearbeitetes fiktives Fallbeispiel ein. Davon waren 84,4 % der Beiträge vollständig bearbeitet und somit auswertbar (**Abb. 3**). Einer der drei abwesenden Studenten hatte seine E-Mails mit allen organisatorischen Informationen nicht gelesen und war am ersten Montag des Blockpraktikums in die Arztpraxis gefahren. Ein weiterer Student hatte sich im Vorfeld abgemeldet, da er das Blockpraktikum im Fachgebiet Allgemeinmedizin bereits in den Semesterferien vorgezogen hatte und der dritte Student nahm einen Fehltermin in Anspruch. Von den eingereichten Bonn Tests sind insgesamt vier Beiträge nicht auswertbar, da kein Abstract oder eine entsprechende Literaturangabe vorliegt. Ein weiterer Bonn Test wurde laut des Studenten „aus Zeitgründen“ nach der Hälfte abgebrochen und kann somit ebenfalls nicht ausgewertet werden.

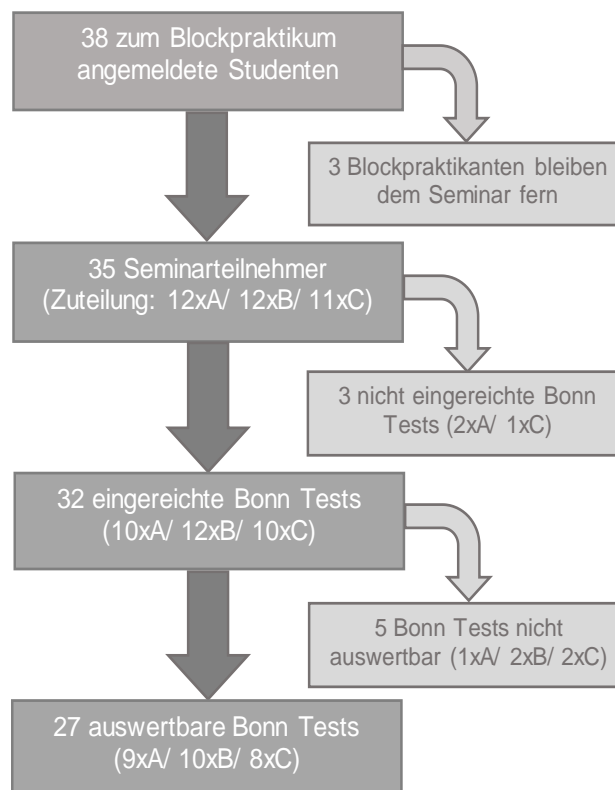


Abb. 3: Darstellung der Rekrutierung von auswertbaren Bonn Tests zu fiktiven Fallbeispielen.

Alle der 27 auswertbaren Beiträge erreichen mindestens die zum Bestehen notwendige Hälfte der möglichen Punktzahl. Im Durchschnitt werden 108,4 Punkte mit einer Standardabweichung von 15,4 Punkten erzielt. Im direkten Vergleich schneiden die Studenten in den Fallbeispielen A ($\bar{x}=106,2$; $s=12,7$, $n=9$), B ($\bar{x}=107,9$; $s=10,9$; $n=10$) und C ($\bar{x}=111,6$; $s=21,3$; $n=8$) ähnlich gut ab, allerdings fällt bei Fallbeispiel C eine höhere Abweichung der Ergebnisse vom Mittelwert auf (**Abb. 4**).

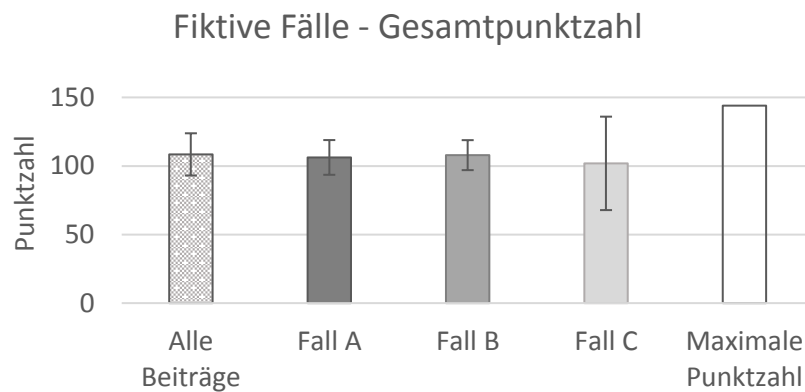


Abb. 4: Darstellung der im Bonn Test erreichten Gesamtpunktzahl für die fiktiven Fälle und direkter Vergleich aller Beiträge mit den fallspezifischen Beiträgen A, B und C sowie der maximal möglichen Punktzahl.

Für die vier Bonn Tests ohne Abstract lässt sich keine Gesamtpunktzahl ermitteln, da die Kategorien des 5-A-Schritteprozesses, mit Ausnahme der Fragestellung (Ask) und der Suchstrategie (Acquire – Suchstrategie), lediglich im Kontext des gewählten Abstracts evaluiert werden können. In diesen beiden Kategorien wird in den unvollständigen Beiträgen mehr als die Hälfte der möglichen Punkte erzielt (**Abb. 5**).

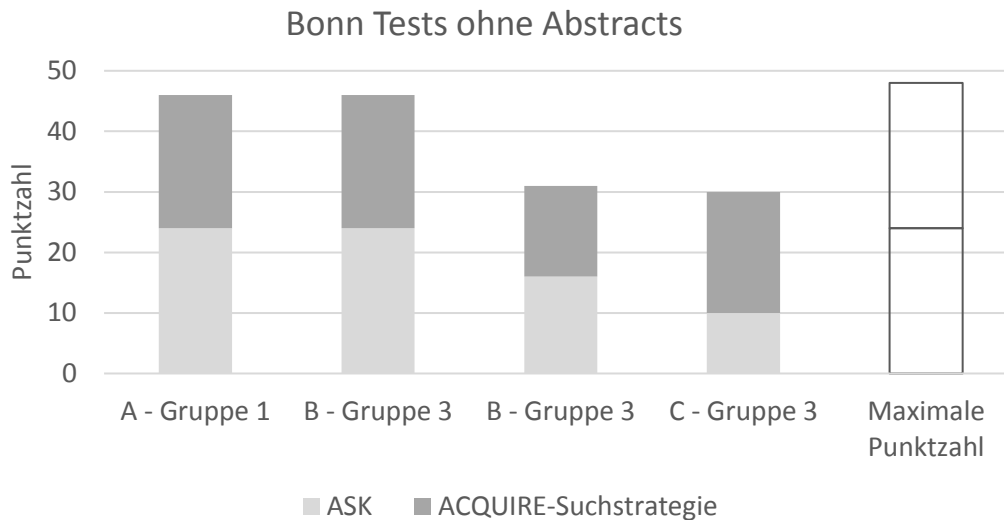


Abb. 5: Darstellung der auswertbaren Kategorien der unvollständig eingereichten Bonn Tests, die aufgrund eines fehlenden Abstracts bzw. einer entsprechenden Literaturangabe nicht vollständig evaluiert werden können.

Der fünfte Beitrag, der „aus Zeitgründen“ unvollständig bearbeitet wurde und somit aus der Wertung fällt, erreicht in den ersten drei auswertbaren Kategorien 24 von 72 möglichen Punkten. Somit liegt er als einziger Beitrag unterhalb der Grenze zum Bestehen. In den vollständig auswertbaren Beiträgen variiert die Leistung der Studenten in den unterschiedlichen Kategorien des 5-A-Schritteprozesses (**Abb. 6**). Am stärksten sind die Blockpraktikanten insgesamt bei der Auswahl des Abstracts mit durchschnittlich 21,3 von 24 möglichen Punkten ($s=3,3$), dicht gefolgt von der Festlegung der Suchstrategie ($\bar{x}=20,9$; $s=2,8$). Beide Fähigkeiten sind dem Schritt Acquire zuzuordnen. In den Kategorien Assess ($\bar{x}=18,6$; $s=4,4$), Appraise ($\bar{x}=18,3$; $s=6,5$) und Ask ($\bar{x}=17,6$; $s=7,1$) erzielten sie immer noch etwa dreiviertel der zu erreichenden Punkte. In der schwächsten Kategorie, Apply, lag der Mittelwert mit 11,9 Punkten ($s=6$) sogar knapp unter der Grenze zum Bestehen dieser Teildisziplin des 5-A-Schritteprozesses von 12 Punkten.

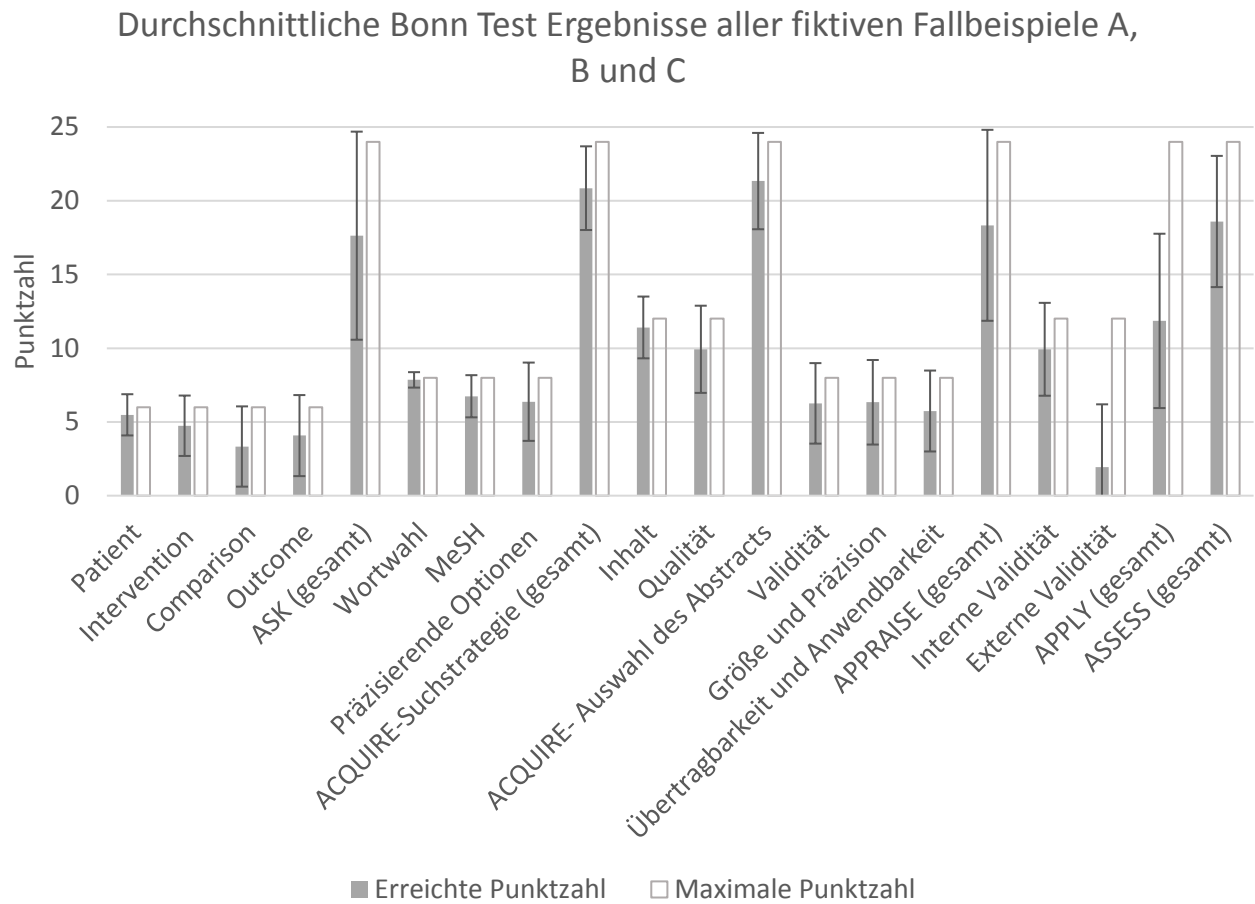


Abb. 6: Darstellung der durchschnittlich erreichten Punktzahl aller fiktiven Fallbeispiele A, B und C in den Kategorien des Erwartungshorizonts.

Betrachtet man die Ergebnisse der 5-As separat für die Fallbeispiele A, B und C, so ergeben sich zum Teil einige Abweichungen für die hierarchische Leistungseinordnung der Studenten (**Abb. 7 A-D**).

Abb. 7 A: Fallübergreifende Leistung

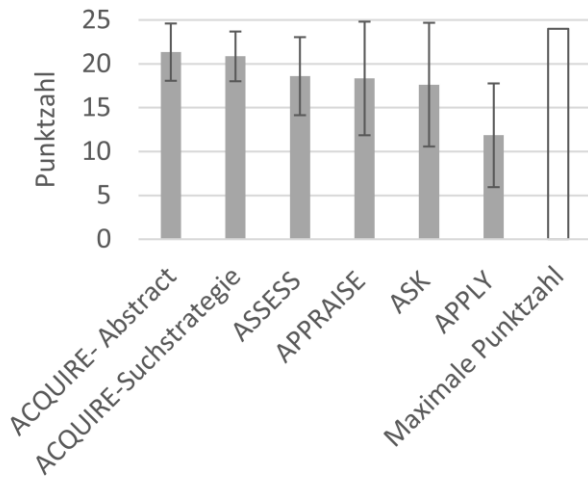


Abb. 7 B: Leistung im Fall A

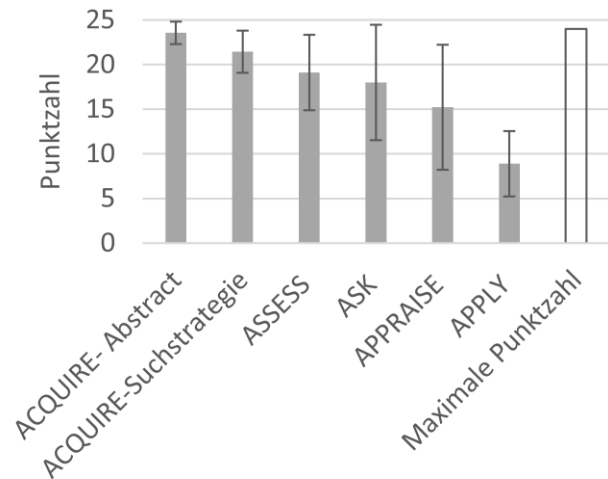


Abb. 7 C: Leistung im Fall B

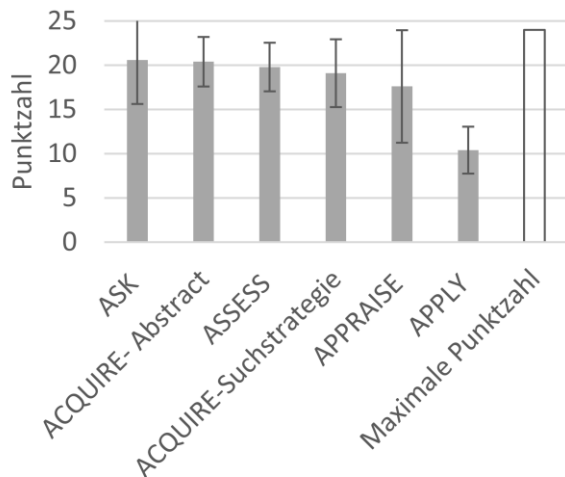


Abb. 7 D: Leistung im Fall C

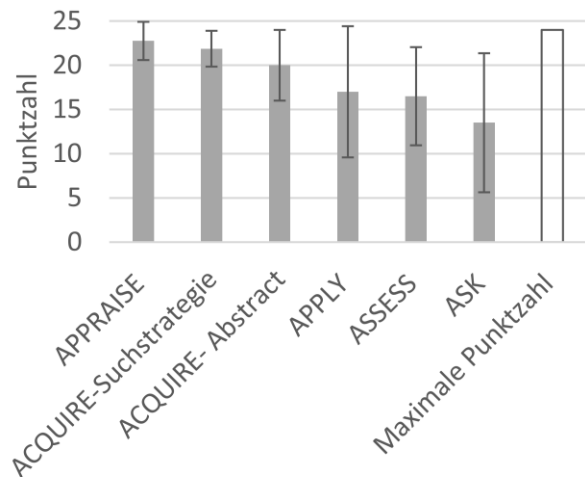


Abb.7 A-D: Hierarchische Darstellung der Leistungen in den Kategorien des 5-A-Schritteprozesses für alle eingereichten fiktiven Fälle (**7 A**), sowie spezifisch für die Fälle A (**7 B**), B (**7 C**) und C (**7 D**).

Bei der Formulierung einer wissenschaftlichen Fragestellung (Ask) wird die Patientenproblematik mit durchschnittlich 5,5 von 6 Punkten ($s=1,4$) am ausführlichsten berücksichtigt, gefolgt von der Benennung der Intervention ($\bar{x}=4,7$; $s=2$) und des Outcomes ($\bar{x}=4,1$; $s=2,70$; **Abb. 8 A**). 67 % der Studenten benennen die spezifische Intervention korrekt und weitere 15 % gehen zumindest auf den Typ der Intervention ein oder nennen die Intervention, aber ordnen sie irrtümlich als Comparison ein. Für das Outcome fehlt in 30 % der Beiträge eine Aussage zu dem gewünschten Therapieziel. Gleichzeitig erreichen 67 % in dieser Unterkategorie die volle Punktzahl, da sie einen objektiven oder für den Patienten bedeutsamen Ausgang der Intervention benennen können (**Abb. 8 B**). Die meisten Probleme haben die Studenten bei der Benennung einer alternativen Intervention (Comparison) ($\bar{x}=3,3$; $s=2,7$). Vergleicht man die Fälle A, B und C miteinander, so schneidet Fallbeispiel B in der Kategorie Ask besonders stark ab mit 20,6 von 24 Punkten. Fallbeispiel C schneidet mit nur 13,5 Punkten am schlechtesten ab (**Abb. 8 A**). 50% dieser Beiträge erreichen weniger als die Hälfte der möglichen Punkte in dieser Kategorie und liegen somit unterhalb der Grenze zum Bestehen (**Abb. 8 C**).

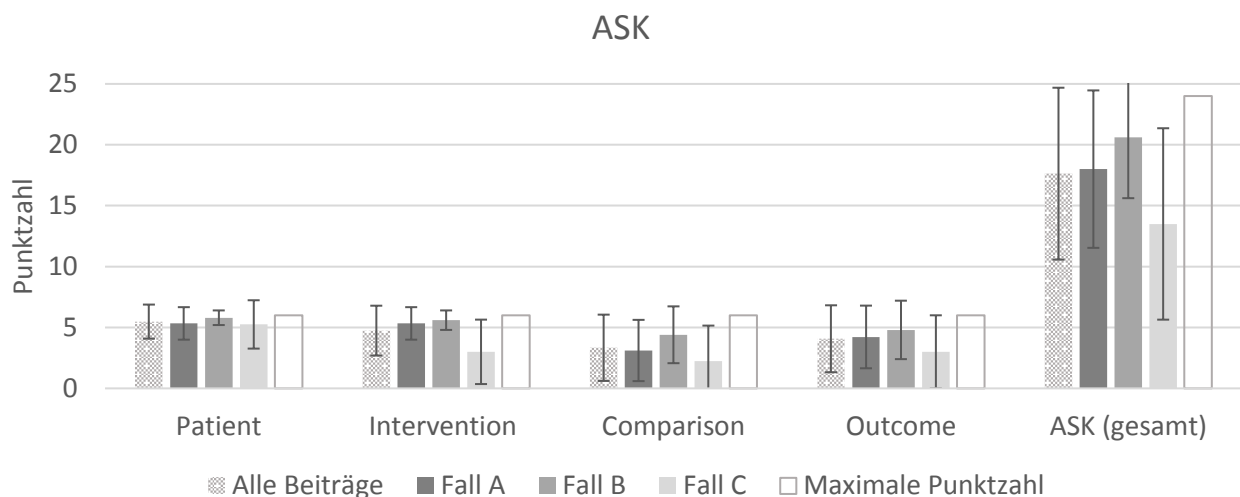


Abb. 8 A: Darstellung der mittleren erreichten Punktzahlen aller fiktiven Fallbeiträge in der Kategorie Ask und den 6 entsprechenden Unterkategorien Patient, Intervention, Comparison und Outcome im Vergleich zur mittleren fallspezifischen Leistung der Studenten.

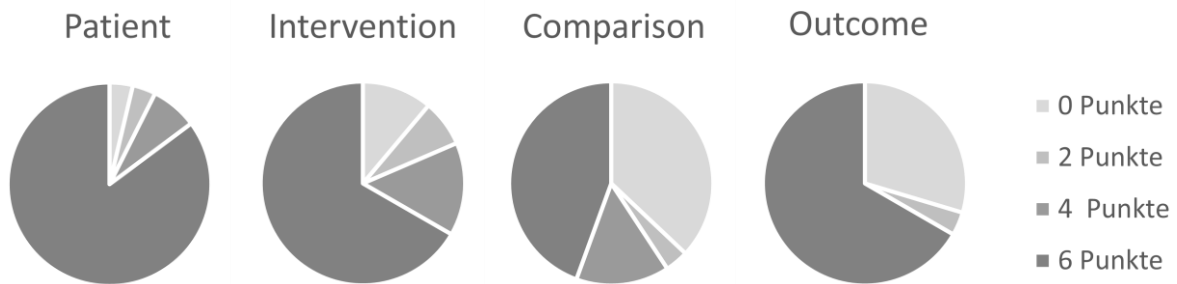


Abb. 8 B: Darstellung der Häufigkeit der jeweils im Bonn Test erreichten Punktzahlen in den Unterkategorien Patient, Intervention, Comparison und Outcome.

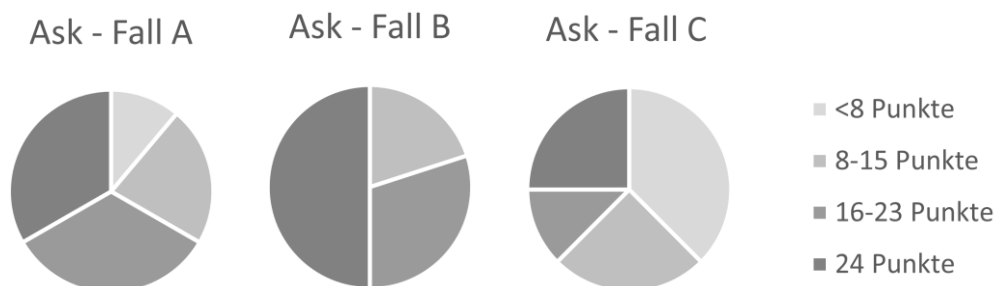


Abb. 8 C: Darstellung der Häufigkeit der jeweils im Bonn Test erreichten Punktzahlen in der Kategorie Ask für die Fallbeispiele A, B und C.

Mit durchschnittlich 20,9 von 24 Punkten ($s=2,8$) schneiden die Blockpraktikanten bei der Formulierung einer Suchstrategie besonders stark ab (**Abb. 9 A**). 93 % der Studenten gelingt es ausschließlich die markantesten PICO-Elemente für die Suchanfrage zu wählen. Damit ist die Wortwahl die stärkste Leistung der Studenten im gesamten Bonn Test. In nur 48 % der Beiträge werden hingegen ausschließliche MeSH verwendet (**Abb. 9 B**). Im direkten Vergleich schneiden die Studenten mit Fallbeispiel B in allen drei Unterkategorien Wortwahl, MeSH und präzisierende Optionen schlechter ab (**Abb. 9 A**). Die präzisierenden Optionen stellen jedoch die größte Abweichung dar. Während allgemein nur 19 % der Studenten in dieser Unterkategorie limitierte oder gar keine Fähigkeiten zeigen, sind es im Fall B schon 40 % (**Abb. 9 B; 9 C**).

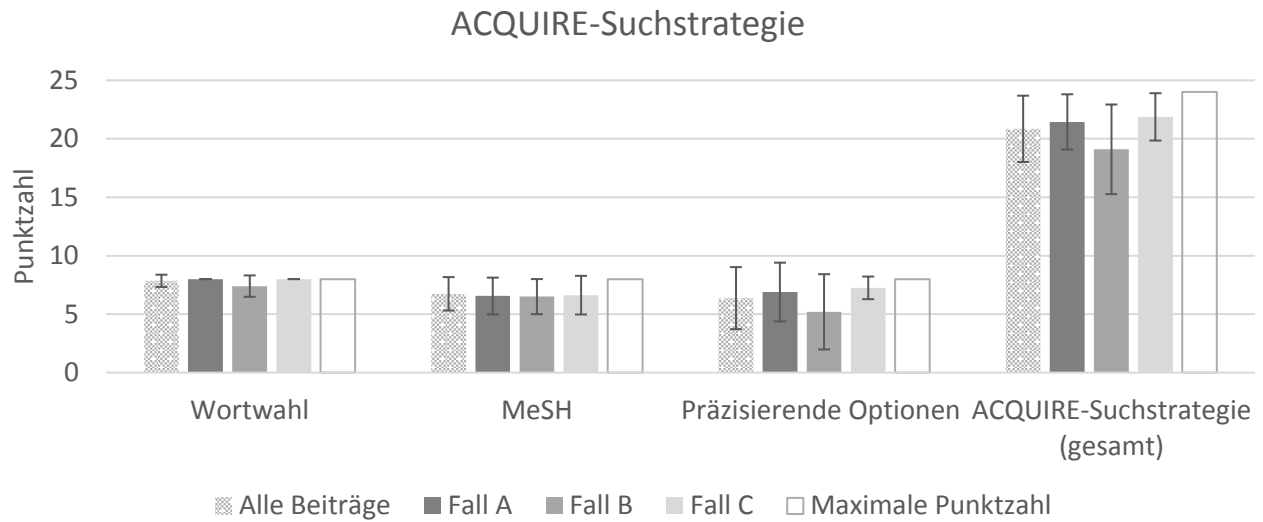


Abb. 9 A: Darstellung der mittleren erreichten Punktzahlen aller fiktiven Fallbeiträge in der Kategorie Acquire - Suchstrategie und den entsprechenden Unterkategorien Wortwahl, MeSH und Präzisierungsoptionen im Vergleich zur mittleren fallspezifischen Leistung der Studenten.

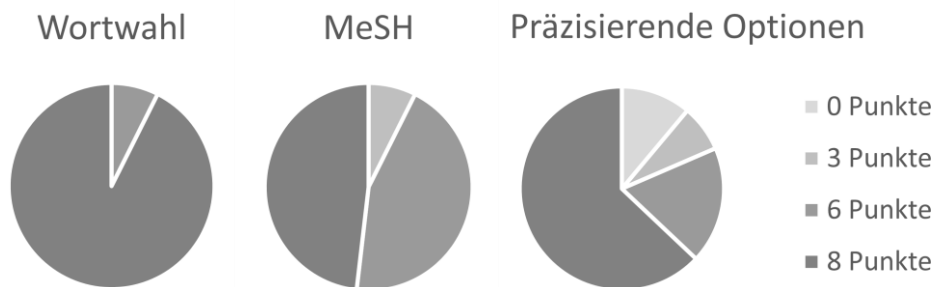


Abb. 9 B: Darstellung der Häufigkeit der jeweils im Bonn Test erreichten Punktzahlen in den Unterkategorien Wortwahl, MeSH und Präzisierungsoptionen.

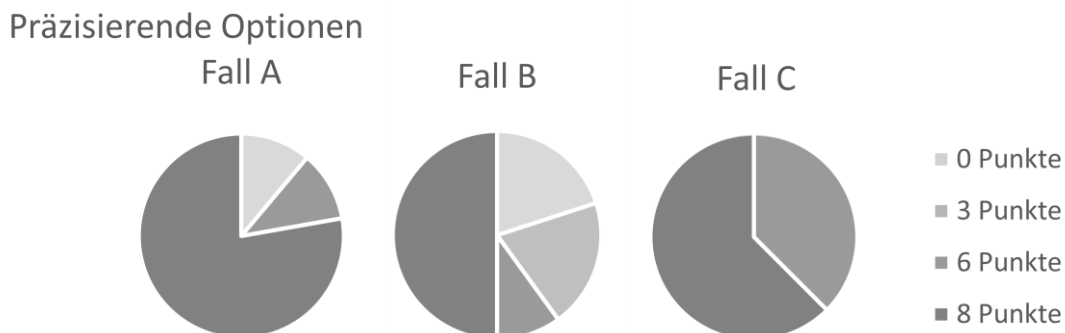


Abb. 9 C: Darstellung der Häufigkeit der jeweils im Bonn Test erreichten Punktzahlen in der Unterkategorie Präzisierungsoptionen für die Fallbeispiele A, B und C.

Die Kategorie Acquire - Auswahl des Abstracts stellt insgesamt die stärkste Fähigkeit der Studenten dar. Es gelingt 93 % der Studenten ein inhaltlich geeignetes Abstract zu wählen, dennoch wird dabei nur in 63 % der Beiträge auf die höchste verfügbare Evidenzstufe zurückgegriffen (**Abb. 10 B**). Besonders hervorzuheben ist das Fallbeispiel A, da die Blockpraktikanten hier zu 89 % die Höchstpunktzahl in beiden Unterkategorien Inhalt und Qualität erzielen. Unter allen Beiträgen der drei Fallbeispiele gelingt dies im Vergleich nur 56 % (**Abb. 10 C**). Während die Studenten mit Fall A und C inhaltlich jeweils die volle Punktzahl erreichen, schneidet Fall B im Mittel mit 10,4 Punkten ($s=3,2$) schlechter ab. Fall C fällt hingegen mit weniger qualitativ hochwertigen Abstracts ($\bar{x}=8$; $s=4$) als Fall A ($\bar{x}=11,6$; $s=1,3$) oder Fall B ($\bar{x}=10$; $s=2$) auf (**Abb 10 A**).

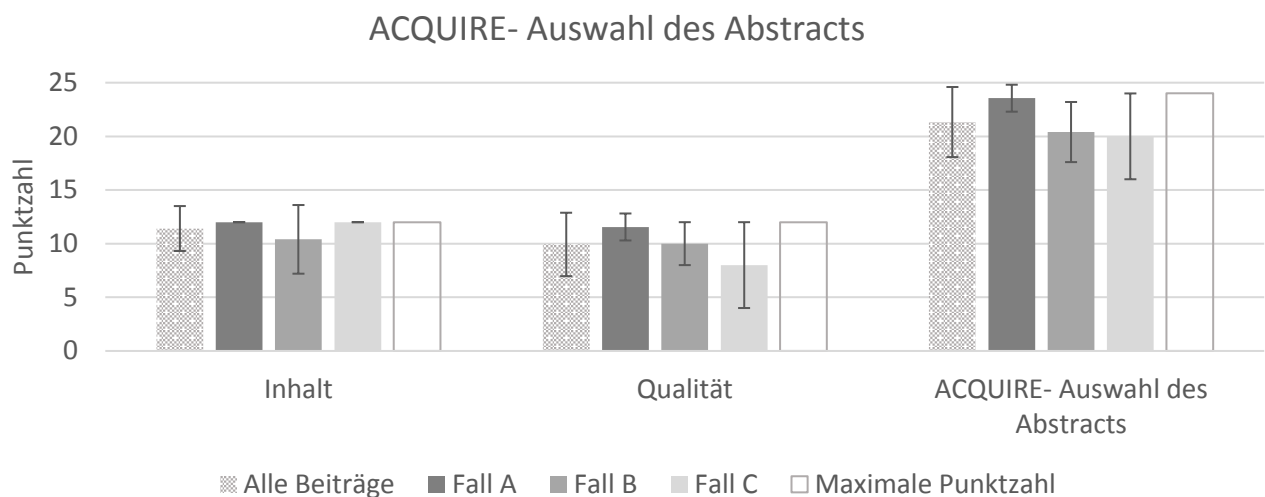


Abb. 10 A: Darstellung der mittleren erreichten Punktzahlen aller fiktiven Fallbeiträge in der Kategorie Acquire – Auswahl des Abstracts und den entsprechenden Unterkategorien Inhalt und Qualität im Vergleich zur mittleren fallspezifischen Leistung der Studenten.

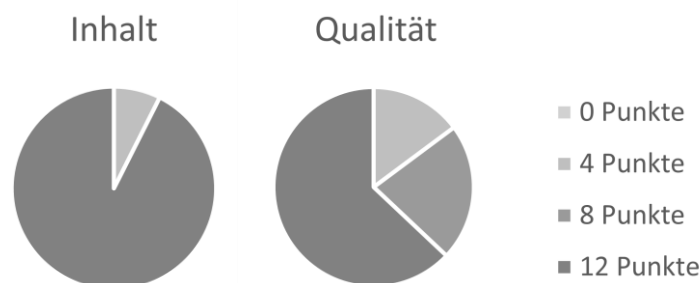


Abb. 10 B: Darstellung der Häufigkeit der jeweils im Bonn Test erreichten Punktzahlen in den Unterkategorien Inhalt und Qualität.

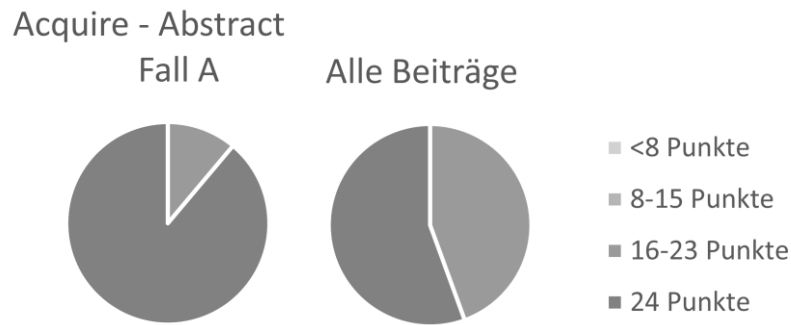


Abb. 10 C: Darstellung der Häufigkeit der jeweils im Bonn Test erreichten Punktzahlen in der Kategorie Acquire – Abstract für das Fallbeispiel A und alle Beiträge.

Die Evaluation der Rechercheergebnisse (Appraise) gelingt in allen drei Unterkategorien Validität, Größe und Präzision, sowie Übertragbarkeit und Anwendbarkeit für das Fallbeispiel C am besten (**Abb. 11 A**). 25 % der Beiträge in dieser Kategorie erzielen starke und 75 % sogar hervorragende Leistungen (**Abb. 11 B**). Fallbeispiel A schneidet am schlechtesten ab. Dabei zeigen nur 56 % der Beiträge starke oder hervorragende Leistungen (**Abb. 11 C**). Insgesamt gelingt nur in 63 % der Beiträge die korrekte Zuordnung der Evidenzstufe. In 15 % wird die Evidenzstufe geringfügig verwechselt oder ein Abstract gewählt, aus dem der Studientyp nicht ersichtlich ist. Bei dem Rest ist die Einschätzung der Validität limitiert oder gar nicht vorhanden. Während 81 % der Blockpraktikanten die meisten oder alle im Abstract genannten Qualitätsmerkmale bezüglich Design, Durchführung und Auswertung auflisten, zeigen sich größere Probleme bei der Untersuchung von Übertragbarkeit und Anwendbarkeit der Rechercheergebnisse. Nur 48 % stellen konkrete Parallelen zum Patientenprofil her oder gehen auf Abweichungen zwischen dem Studienkollektiv und dem Patienten oder auf andere mögliche Einschränkungen für die Übertragbarkeit ein (**Abb. 11 B**).

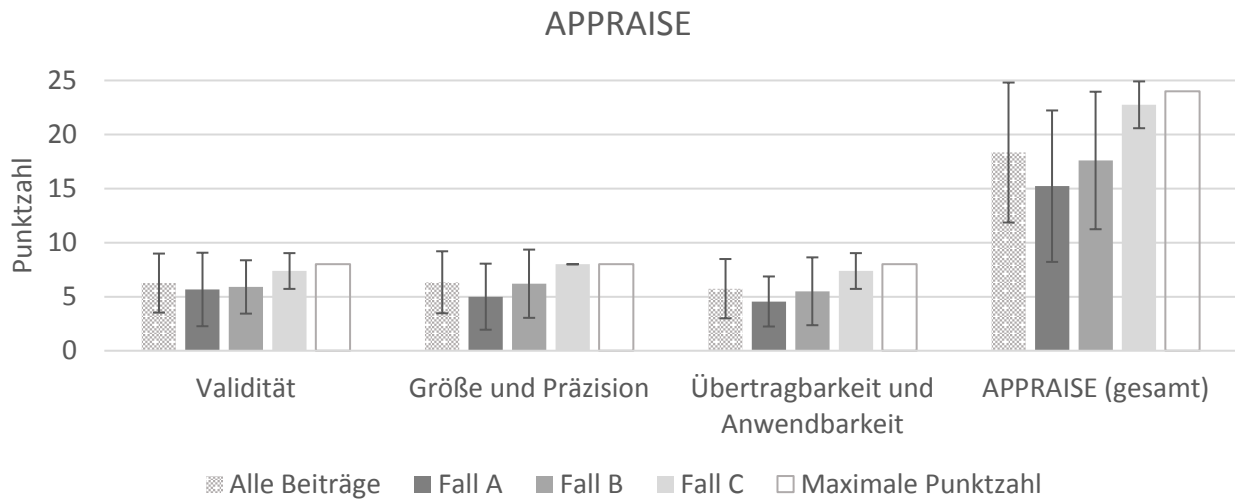


Abb. 11 A: Darstellung der mittleren erreichten Punktzahl aller fiktiven Fallbeiträge in der Kategorie Appraise und den entsprechenden Unterkategorien Validität, Größe und Präzision, sowie Übertragbarkeit und Anwendbarkeit im Vergleich zur mittleren fallspezifischen Punktzahl.

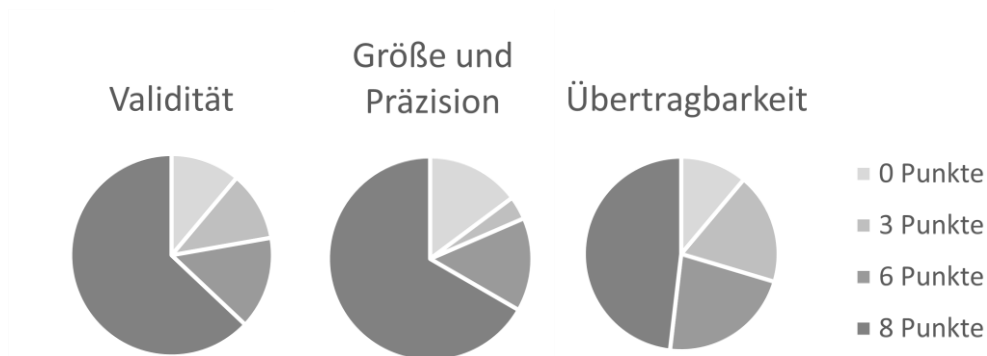


Abb. 11 B: Darstellung der Häufigkeit der jeweils im Bonn Test erreichten Punktzahlen in den Unterkategorien Validität, Größe und Präzision und Übertragbarkeit.

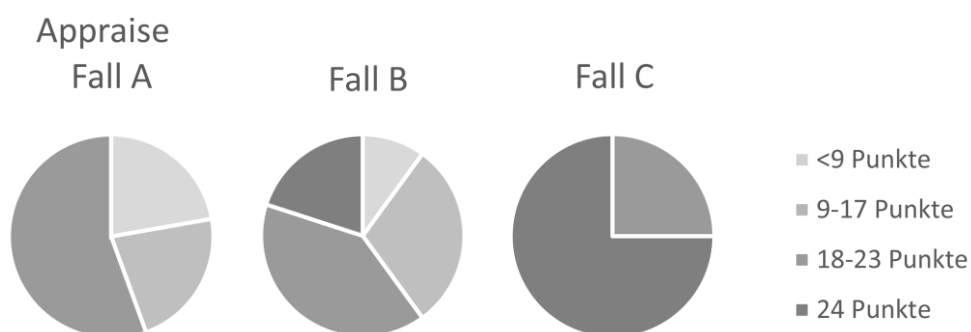


Abb. 11 C: Darstellung der Häufigkeit der jeweils im Bonn Test erreichten Punktzahlen in der Kategorie Appraise für die Fallbeispiele A, B und C.

Die Anwendung der evaluierten Rechercheergebnisse (Apply) (**Abb. 12 A**) stellt im Vergleich zu den anderen As des 5-A-Schritteprozesses die schwächste Kategorie dar. Das liegt in erster Linie daran, dass 81 % der Studenten 0 Punkte in der Unterkategorie externe Validität erzielen. In der anderen Unterkategorie gelingt 67 % der Blockpraktikanten die schlüssige Ableitung einer Therapieempfehlung aus den evaluierten Rechercheergebnissen, was einer hervorragenden internen Validität entspricht. In weiteren 15 % der Fälle ist die ausgewählte Therapieempfehlung auf die Thematik des Abstracts zurückzuführen, wird darin aber nicht eindeutig benannt. In den restlichen Beiträgen ist die ausgewählte Therapieempfehlung zu unspezifisch oder verfehlt das Thema, um die Fragestellung beantworten zu können (**Abb. 12 B**). Die isolierte Betrachtung der drei Fallbeispiele A, B und C spiegelt dieses Verhältnis wider. Lediglich im Fall C wird zu 50 % die Höchstpunktzahl in der Unterkategorie der externen Validität erzielt (**Abb. 12 C**).

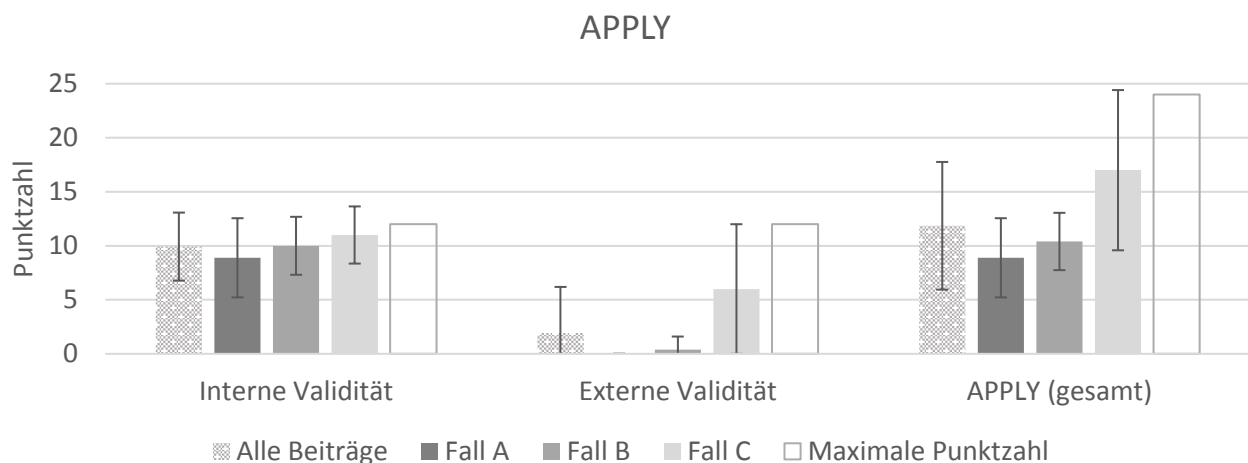


Abb. 12 A: Darstellung der mittleren erreichten Punktzahl aller fiktiven Fallbeiträge in der Kategorie Apply und den entsprechenden Unterkategorien Interne Validität und Externe Validität im Vergleich zur mittleren fallspezifischen Punktzahl.

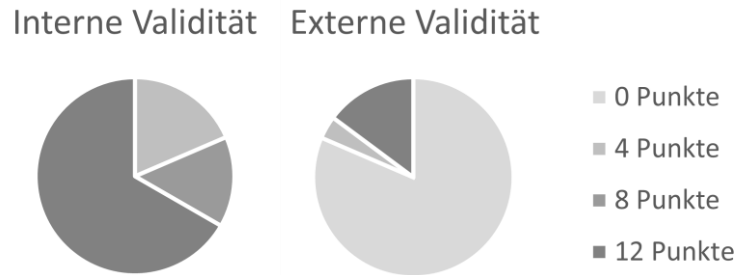


Abb. 12 B: Darstellung der Häufigkeit der jeweils im Bonn Test erreichten Punktzahlen in den Unterkategorien interne Validität und externe Validität.

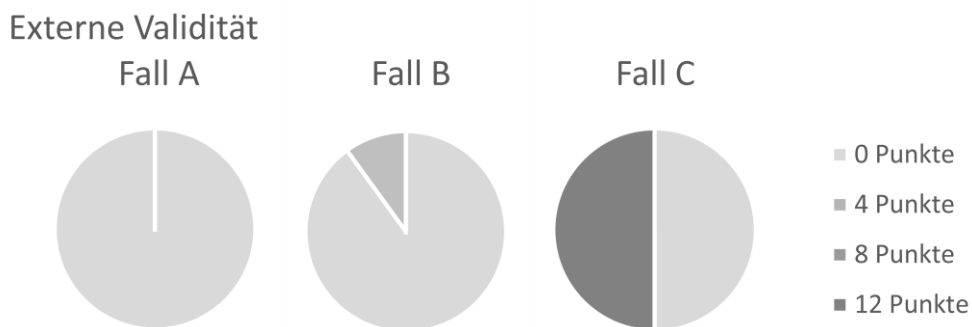


Abb. 12 C: Darstellung der Häufigkeit der jeweils im Bonn Test erreichten Punktzahlen in der Unterkategorie Externe Validität für die Fallbeispiele A, B und C.

Die drittstärkste Kategorie, Assess, fällt den Blockpraktikanten mit dem Fallbeispiel C schwerer als der Gesamtheit (**Abb. 13 A; Abb. 13 B**). Hier weisen 38 % der Beiträge im Vergleich zu insgesamt nur 15 % limitierte Fähigkeiten auf. In 86 % der Beiträge wird eine sinnvolle Evaluationsmaßnahme gewählt, allerdings treten häufig Probleme bei der Festlegung eines Zeitintervalls auf. Nur 30 % der Studenten wählen ein Zeitintervall, das die Möglichkeit bietet Schäden für den Patienten frühzeitig zu erkennen und abzuwenden, während in 56 % der Beiträge entweder kein Zeitintervall genannt wird, oder dieses zu groß gewählt wird und somit nicht gewährleisten kann kurzfristig auftretende Probleme zu erkennen und zu beheben (**Abb. 13 B**).

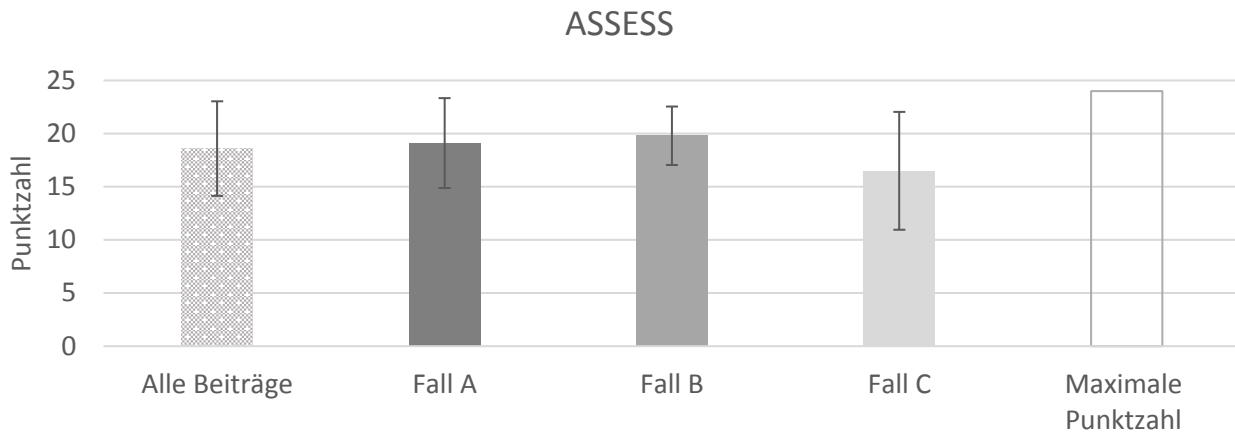


Abb. 13 A: Darstellung der mittleren erreichten Punktzahl aller fiktiven Fallbeiträge in der Kategorie Assess im Vergleich zur mittleren fallspezifischen Punktzahl im Fall A, B und C.

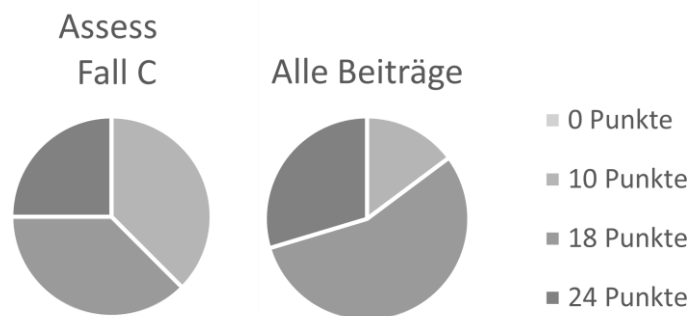


Abb. 13 B: Darstellung der Häufigkeit der jeweils im Bonn Test erreichten Punktzahlen in der Kategorie Assess für das Fallbeispiel C und alle Beiträge.

3.2 Ergebnisse der eigenen Fallbeispiele

Von den 37 angemeldeten Teilnehmern der drei Probedurchläufe liegen am Ende des zweiten Seminartermins 22 ausgefüllte Bonn Tests vor. Während in der zweiten Gruppe 78,6 % der Studenten eine eigene Fragestellung bearbeiten, beträgt die Beteiligung in der dritten Gruppe 47,1 % und in der ersten Gruppe nur 42,9 %. Im Durchschnitt liegt die Gesamtpunktzahl der ausgewerteten Tests mit 99,8 Punkten ($s=20,5$) unter dem Mittelwert für die fiktiven Fallbeispiele A, B und C ($\bar{x}=108,4$; $s=15,4$; **Abb. 14**).

Bonn Test Ergebnisse - Theorie und Praxis

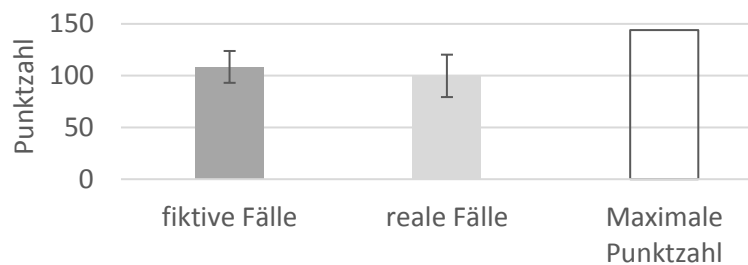


Abb. 14: Vergleich der mittleren erreichten Gesamtpunktzahlen in fiktiven Fällen und realen Fällen.

Wie auch schon bei diesen theoretischen Fällen schneiden die Studenten bei den eigenen Fragestellungen am stärksten bei der Auswahl eines Abstracts (Acquire) ($\bar{x}=20,5$; $s=4,2$) und am schwächsten in der Kategorie Apply ab ($\bar{x}=13,5$; $s=6,9$; **Abb. 15 A-B**).

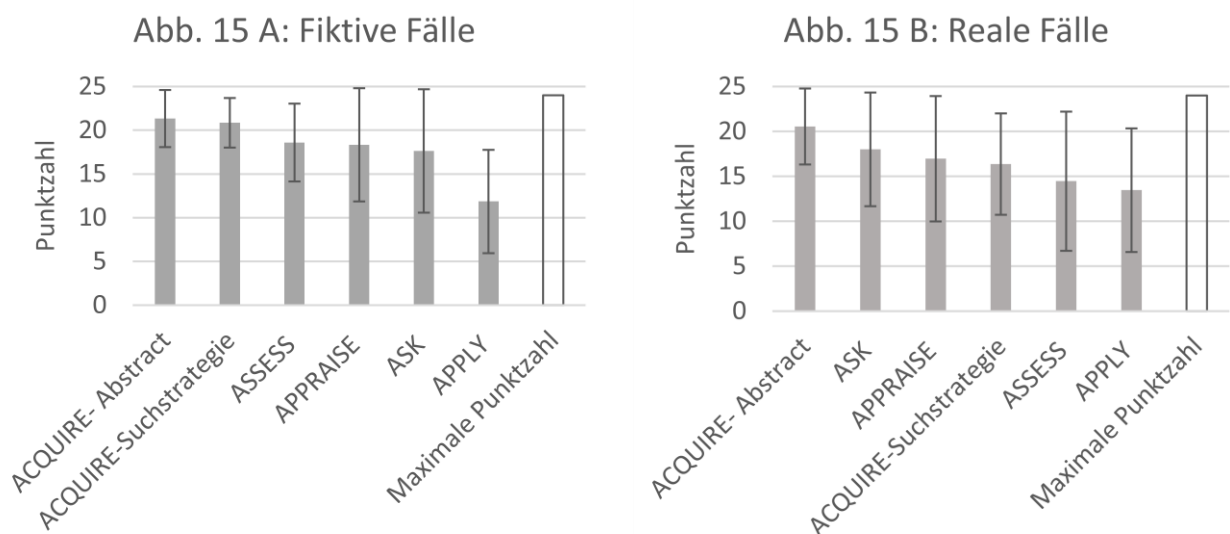


Abb. 15 A-B: Hierarchische Darstellung der mittleren erreichten Punktzahl in den Kategorien des 5-A-Schritteprozesses für die fiktiven Fälle (**15 A**) und die realen Fälle (**15 B**) im Vergleich).

Die restlichen As des 5-A-Schritteprozesses sind im Vergleich zu den vorgegebenen Fallbeispielen bezüglich der erreichten Punktzahlen jedoch neu einzuordnen (**Abb. 16**). Hier belegt Ask den zweiten Platz ($\bar{x}=18$; $s=6,3$), gefolgt von Appraise ($\bar{x}=14,5$; $s=7,7$), Acquire (- Suchstrategie) ($\bar{x}=16,4$; $s=5,6$) und Assess ($\bar{x}=14,5$; $s=7,7$).

Im direkten Vergleich verbessern sich die Studenten in den Kategorien Ask von 17,6 Punkten ($s=7,1$) auf 18 Punkte ($s=6,3$) und Apply von 11,9 Punkten ($s=5,9$) auf 13,5 Punkte ($s=6,9$). Bei allen anderen Schritten des 5-A-Schritteprozesses erreichen sie weniger Punkte. Besonders groß ist diese Differenz bei der Festlegung einer Suchstrategie von 20,9 Punkten ($s=2,8$) auf 16,4 Punkten ($s=5,6$), sowie bei der Evaluation der empfohlenen Therapie von 18,6 Punkten ($s=4,4$) auf 14,5 Punkte ($s=7,7$).

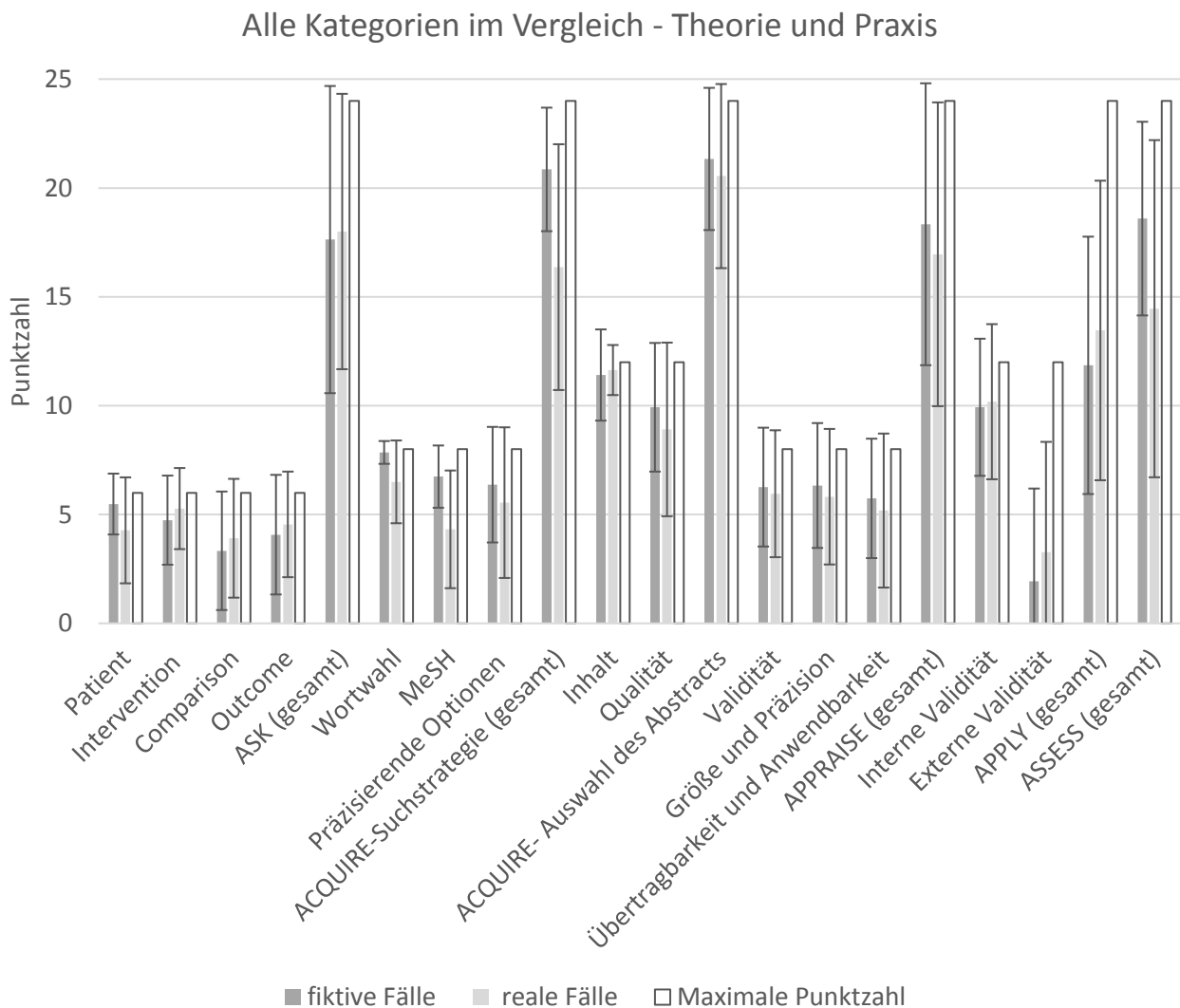


Abb. 16: Darstellung der Ergebnisse in allen Kategorien und Unterkategorien für die fiktiven und realen Fälle im Vergleich.

Die ausgewählten Fragestellungen der Blockpraktikanten werden der Aufgabenstellung gerecht und umfassen ein weites Themenspektrum (**Tab. 2**).

Tab. 2: Auflistung der Themen für die selbstgewählten Fragestellungen.

Thema	Punkte
Medikamenteninteraktion durch Johanniskrauteinnahme bei Depressionen	104
Postexpositionsprophylaxe nach Nadelstichverletzung	70
Diätumstellung oder Lactase-Einnahme bei Lactoseintoleranz	110
ACE-Hemmer bei Migräneprophylaxe	129
Wassereinlagerung bei Amlodipin-Einnahme	113
Phytotherapeutika bei akuter Bronchitis	121
Canephron als Dauertherapie bei rezidivierenden HWI	121
Einfluss von PPI-Einnahme auf Vitamin D und Vitamin B12	52
Notwendigkeit einer Stuhlprobe bei Durchfall nach Auslandsaufenthalt	99
Hyponatriämie als Nebenwirkung von ACE-Hemmern	76
Ideale Trinkmenge bei Altenheimbewohnern	91
Melatonin Supplementierung bei Schlafstörung	109
Einfluss von Vitamin B12-Supplementierung auf die kognitiven Fähigkeiten	100
Langzeitauswirkung von einer Lebendnierenspende	88
Osteoporose Prophylaxe bei gleichzeitiger Therapie mit Aromataseinhibitoren	113
Aktuelle Datenlage zu dem Post-Lyme-Syndrom	65
Risikofaktoren für eine Strahlenpneumonitis	126
Benefit regelmäßiger Teilnahme an Disease Management Programmen für Patienten mit Diabetes Mellitus Typ 2	92
Akupunktur bei Tinnitus	108
Starke Durchfälle als Nebenwirkung einer Interferontherapie	78
Aerosolisierte Vasodilatoren bei pulmonaler Hypertonie	117
Lasea bei Angststörungen	113
Mittelwert	99,8

3.3 Ergebnisse der Eigenevaluation

Von insgesamt 34 Evaluationsbögen sind nur 31 vollständig ausgefüllt und werden im weiteren Verlauf berücksichtigt.

3.3.1 Ergebnisse der Evaluation: Das Lernmodul

Insgesamt stehen die Blockpraktikanten dem gewählten Blended Learning Stil als Ersatz für ein klassisches Lehrkonzept mit Frontalunterricht gemischt gegenüber (**Abb. 17 A.1**). 42 % der Studenten bewerten diesen Lehransatz als gut oder sehr gut, 39 % stehen ihm neutral oder eher neutral gegenüber und 19 % empfinden ihn als negativ (**Abb. 17 A.2**).

Abb. 17 A.1: LM 1

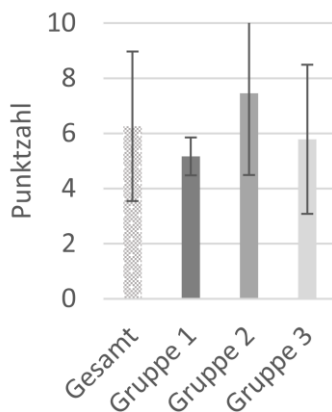


Abb. 17 A.2: Lernmodul 1

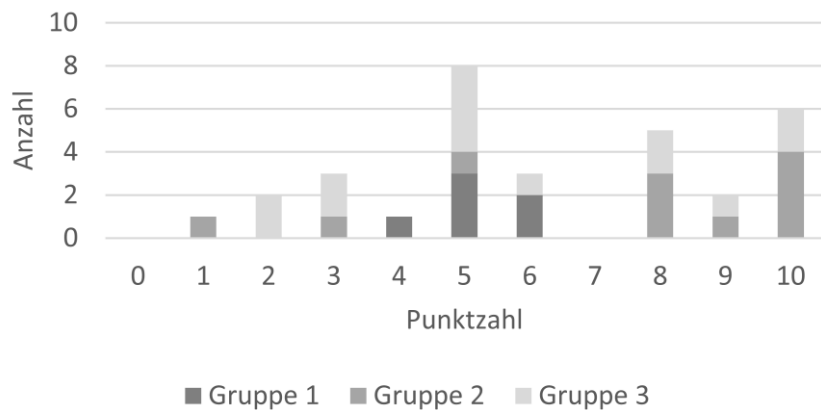


Abb. 17 A: Gruppenspezifische Darstellung der insgesamt vergebenen Punktzahl (**17 A.1**) sowie der Häufigkeit der einzelnen Punkte (**17 A.2**) bei der Evaluation des Lernmoduls für die Aussage „Ich ziehe die selbstständige Arbeit zu Hause mit einem online Lernmodul einem klassischen Frontalunterricht vor.“.

Mit der inhaltlichen, didaktischen und technischen Umsetzung des Lernmoduls sind die Studenten in ihrer Evaluation aber besonders zufrieden (**Abb. 17 B.1; Abb. 17 C.1; Abb. 17 D.1**). 81 % bewerten die inhaltliche, 84 % die didaktische und 77 % die technische Umsetzung mit gut oder sehr gut (**Abb. 17 B.2; Abb. 17 C.2; Abb. 17 D.2**). Im direkten Vergleich der drei Gruppen, die das Lernmodul bearbeitet haben, weichen die Evaluationen der dritten Gruppe zum Teil stark von denen der anderen Studenten ab. Während in den ersten beiden Gruppen Inhalt, Didaktik und Technik ausschließlich als gut oder sehr gut eingestuft werden, sind es in der dritten Gruppe lediglich zwischen 50 % und 64 %.

Abb. 17 B.1: LM 2

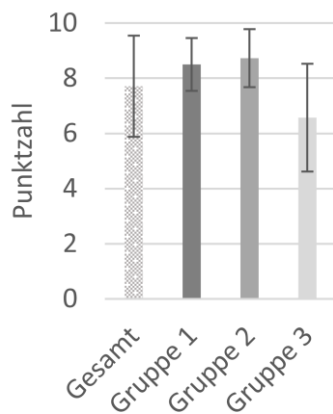


Abb. 17 B.2: Lernmodul 2

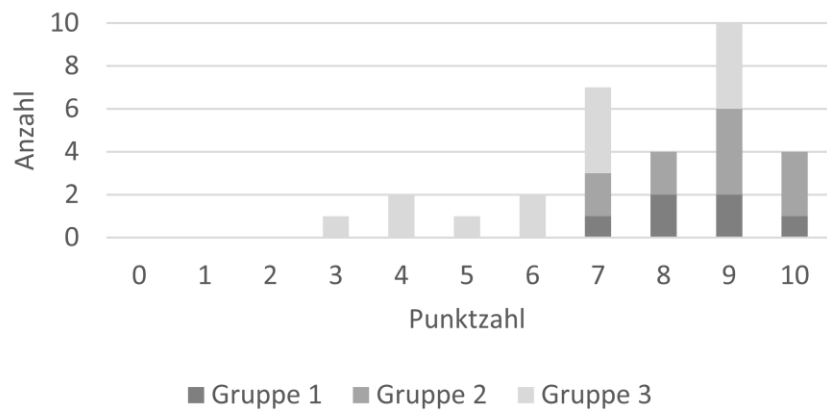


Abb. 17 C.1: LM 3

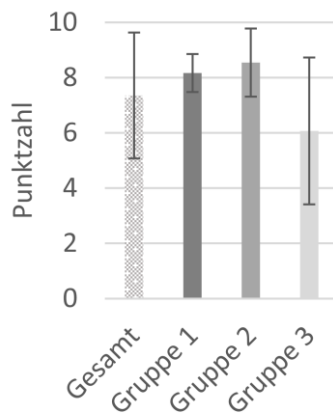


Abb. 17 C.2: Lernmodul 3

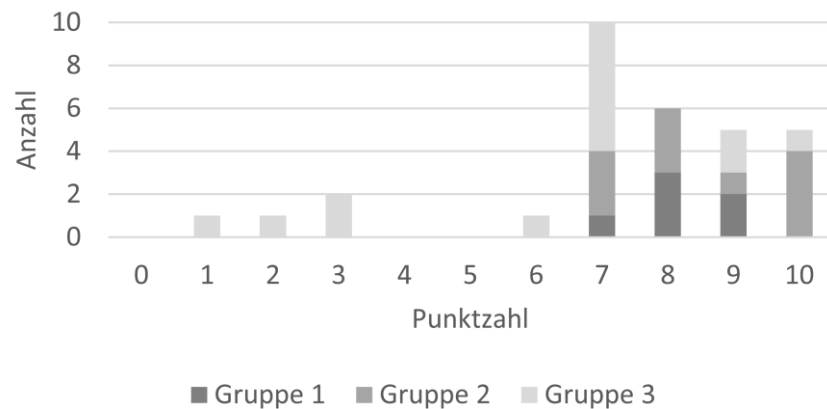


Abb. 17 D.1: LM 4

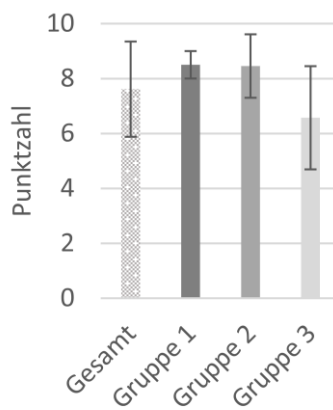


Abb. 17 D.2: Lernmodul 4

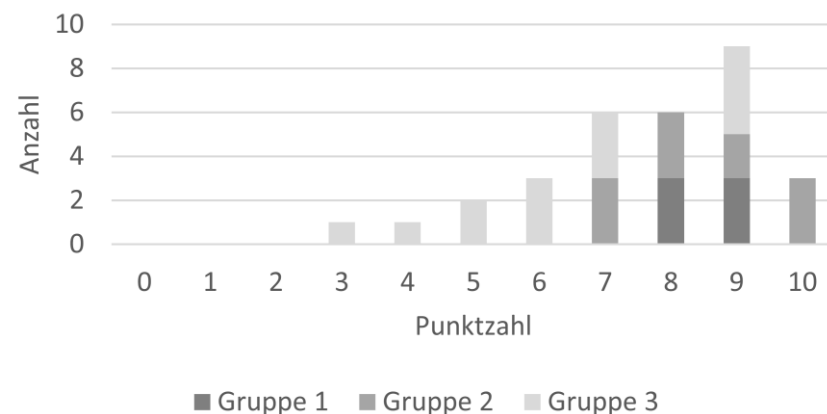


Abb. 17 B-D : Gruppenspezifische Darstellung der insgesamt vergebenen Punktzahl (**17 B.1**; **C.1**; **D.1**) sowie der Häufigkeit der einzelnen Punkte (**17 B.2**; **C.2**; **D.2**) bei der Evaluation des Lernmoduls für die Aussage **17 B**: „Ich bin mit der inhaltlichen Umsetzung des Lernmoduls auf E-campus zufrieden.“, **17 C**: „Ich bin mit der didaktischen Umsetzung des Lernmoduls zufrieden.“, **17 D**: „Ich bin mit der technischen Umsetzung des Lernmoduls zufrieden.“.

Im Laufe der Evaluation zeichnet sich diese Gruppendynamik auch bei anderen Aspekten des erweiterten Blockpraktikums ab (**Abb. 17 E.2; 17 F.2; 17 G.2**). So bestätigen die Blockpraktikanten insgesamt mit 7,4 von 10 Punkten, dass sie ihr Grundlagenwissen zu Inhalt und Technik der evidenzbasierten Medizin verbessern konnten ($s=1,8$; **Abb. 17 E.1**). Während Gruppe eins den Benefit mit 8,5 ($s=1$) und Gruppe zwei mit 7,7 Punkten ($s=1,9$) bewerten, vergibt Gruppe drei hierfür im Schnitt 6,7 Punkte ($s=1,6$).

Abb. 17 E.1: LM 5

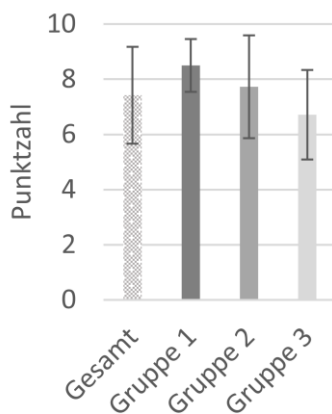


Abb. 17 E.2: Lernmodul 5

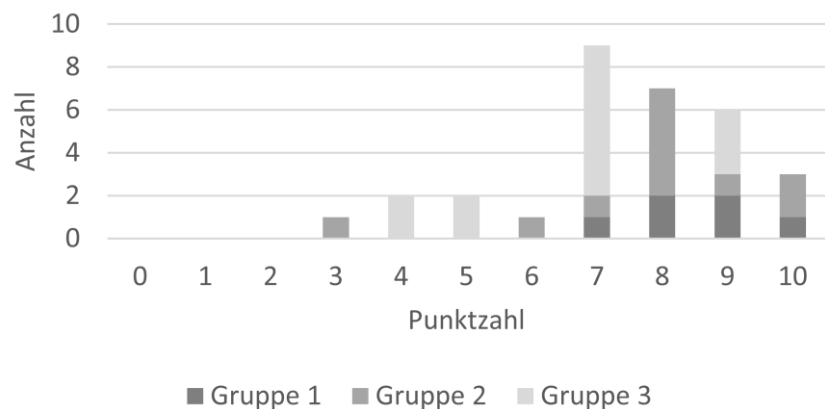


Abb. 17 E: Gruppenspezifische Darstellung der insgesamt vergebenen Punktzahl (**17 E.1**) sowie der Häufigkeit der einzelnen Punkte (**17 E.2**) bei der Evaluation des Lernmoduls für die Aussage „Durch das Lernmodul konnte ich mein Grundlagenwissen zu Inhalt und Technik der Evidenzbasierten Medizin verbessern.“.

Bei der Evaluation ihrer eigenen Fertigkeiten fühlen sich die Blockpraktikanten mit durchschnittlich 7,2 von 10 Punkten ($s=2$; **Abb. 17 F.1**) überwiegend sicher. Mit durchschnittlich 7,4 von 10 Punkten ($s=2,6$) geben die Studenten an, dass die zehn Fragen des Bonn Tests ihnen geholfen haben, den 5-A-Schritteprozess systematisch anzuwenden (**Abb. 17 G.1**). Auch hier liegen die Punkte der dritten Gruppe mit 6,3 Punkten ($s=2,8$) unter den Einschätzungen der ersten Gruppe mit 7,5 Punkten ($s=2,4$) oder der zweiten Gruppe mit 8,8 Punkten ($s=1,5$). Allgemein sind sich alle Blockpraktikanten darin einig, dass die Implementierung einer EbM Schulung zu einem früheren Zeitpunkt im Studium erfolgen sollte.

Abb. 17 F.1: LM 6

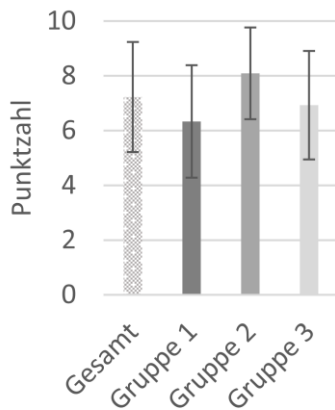


Abb. 17 F.2: Lernmodul 6

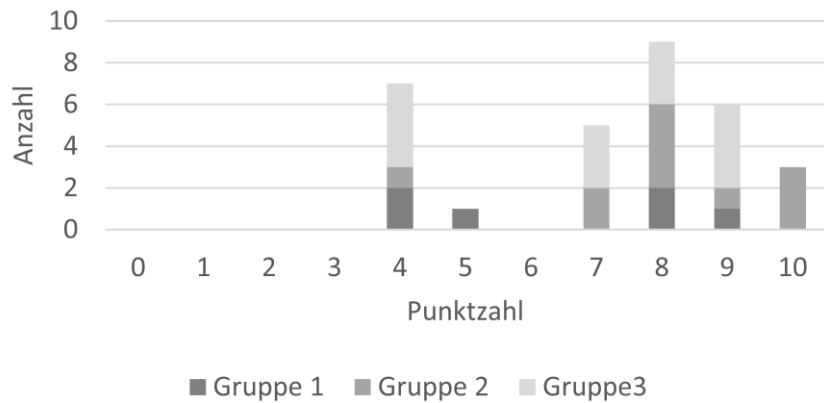


Abb. 17 G.1: LM 7

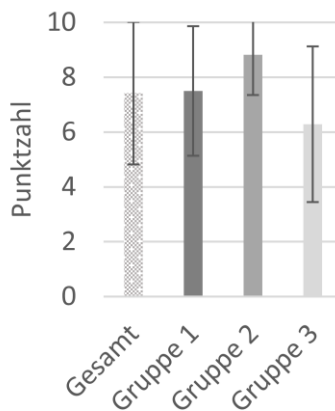


Abb. 17 G.2: Lernmodul 7

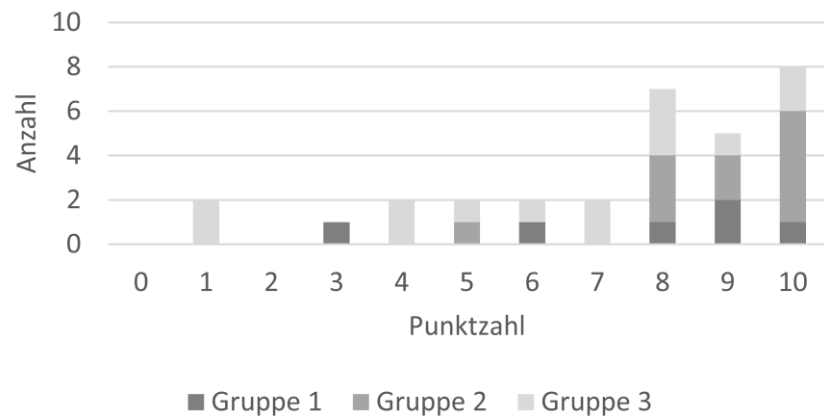


Abb. 17 F-G: Gruppenspezifische Darstellung der insgesamt vergebenen Punktzahl (**17 F.1; G.1**) sowie der Häufigkeit der einzelnen Punkte (**17 F.2; G.2**) bei der Evaluation des Lernmoduls für die Aussage **17 F:** „Die durch das Lernmodul erworbenen EbM-Fähigkeiten konnte ich problemlos auf das zugewiesene Fallbeispiel anwenden.“ und **17 G:** „Das Arbeitsblatt im Anhang des Lernmoduls hat mir geholfen, den 5-A-Schritteprozess systematisch anzuwenden.“.

3.3.2 Ergebnisse der Evaluation: Die Seminare

Insgesamt sind die Blockpraktikanten zufrieden mit der Umsetzung des Seminars als weiteres Kernstück des Blended Learning Ansatzes, schätzen ihren Mehrgewinn durch die Seminare aber geringer ein als durch das Lernmodul. Hier sind nur 65 % der Studenten der Meinung ihr Grundlagenwissen zu Inhalt und Technik von EbM mit Hilfe der Seminare gut oder sehr gut verbessert zu haben (**Abb. 18 A.2**). Dabei vergeben Gruppe eins und zwei im Schnitt wieder einen Punkt mehr als Gruppe drei (**Abb. 18 A.1**).

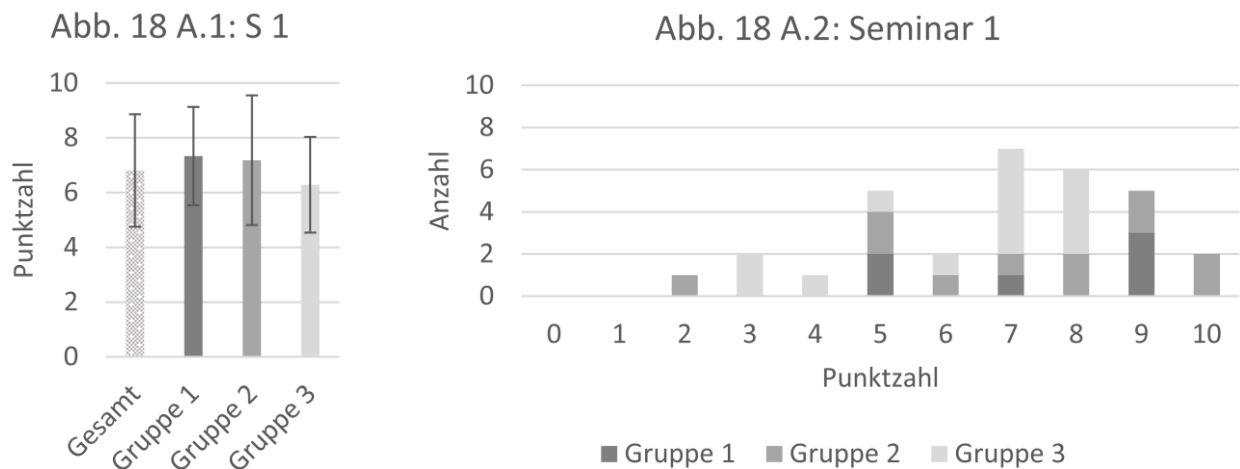


Abb. 18 A: Gruppenspezifische Darstellung der insgesamt vergebenen Punktzahl (**18 A.1**) sowie der Häufigkeit der einzelnen Punkte (**18 A.2**) bei der Evaluation der Seminare für die Aussage „Das Seminar hat mir geholfen mein Grundlagenwissen zu Inhalt und Technik der Evidenzbasierten Medizin zu festigen.“.

Gleichzeitig ist die Zufriedenheit der Studenten mit der Umsetzung der Seminarinhalte durch die Dozenten besonders hoch (**Abb. 18 B.1**). 87 % der Blockpraktikanten bewerten die fachliche Kompetenz und Vorbereitung ihres Dozenten mit gut oder sehr gut (**Abb. 18 B.2**). Bei der Abwägung des Blended Learning Stils gegen konventionelle Unterrichtsmodelle zeigt sich ein gespaltenes Meinungsbild (**Abb. 18 C.1**; **18 C.2**). So zeigten sich die Blockpraktikanten überwiegend zufrieden mit dem tutoren-gestützten Unterricht in Kleingruppen anstelle des klassischen Frontalunterrichts. Die erste und zweite Gruppe honorierte diesen Lehransatz mit durchschnittlich 8,3 bzw. 7,8 von 10 Punkten ($s=1,5$ bzw. $s=2,3$). Lediglich die Teilnehmer der dritten Gruppe standen diesem Format gemischt gegenüber und vergaben im Mittel 5,4 Punkte ($s=3,3$).

Abb. 18 B.1: S 2

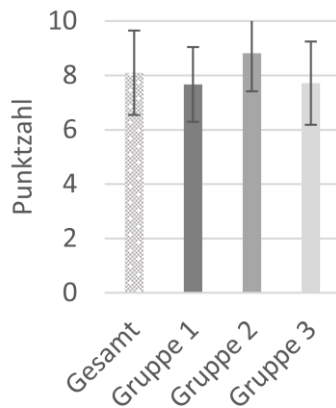


Abb. 18 B.2: Seminar 2

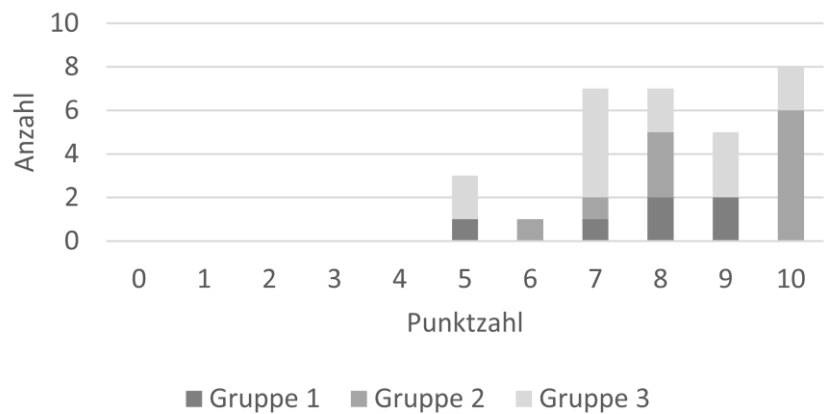


Abb. 18 C.1: S 3

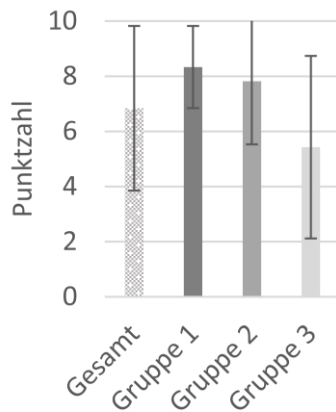


Abb. 18 C.2: Seminar 3

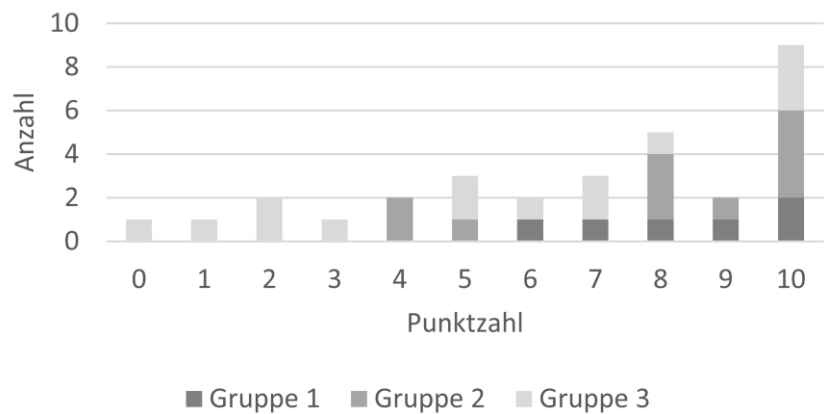


Abb. 18 B-C: Gruppenspezifische Darstellung der insgesamt vergebenen Punktzahl (**18 B.1; C.1**) sowie der Häufigkeit der einzelnen Punkte (**18 B.2; C.2**) bei der Evaluation der Seminare für die Aussage **18 B:** „Die Dozenten haben die Seminarinhalte didaktisch und fachlich kompetent vermittelt und waren gut vorbereitet.“ Und **18 C:** „Ich ziehe den Tutorengestützten Unterricht in Kleingruppen einem klassischen Frontalunterricht vor.“.

3.3.3 Ergebnisse der Evaluation: Das Blockpraktikum

Laut ihrer Selbsteinschätzung ist es 71 % der Studenten gut oder sehr gut möglich im Rahmen des Blockpraktikums ein eigenes Fallbeispiel zu entwickeln (**Abb. 19 A.2**). Dabei gibt ein Viertel der Blockpraktikanten Probleme bei der Entwicklung eines Patientenprofils für das eigene Fallbeispiel an, während dies dem restlichen Anteil der Studenten leichter fällt (**Abb. 19 B.2**). So bewerten sie ihre Fähigkeit zur Konstruktion eines Patientenprofils mit nur durchschnittlich 6,3 Punkten ($s=2,9$; **Abb. 19 A.1**) im Gegensatz zur Entwicklung eines eigenen Fallbeispiels mit 7,5 Punkten ($s=2,5$; **Abb. 19 B.1**).

Abb. 19 A.1: BP 1

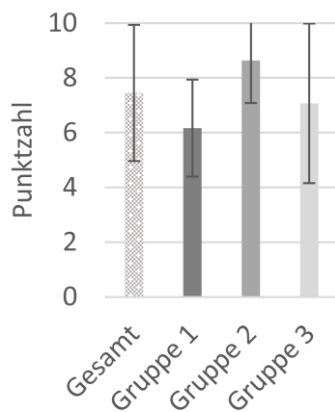


Abb. 19 A.2: Blockpraktikum 1

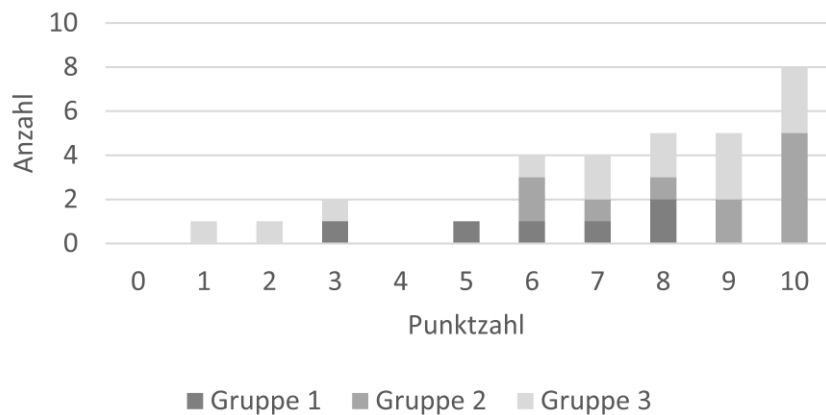


Abb. 19 B.1: BP 2

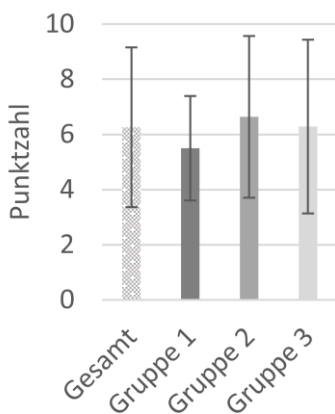


Abb. 19 B.2: Blockpraktikum 2

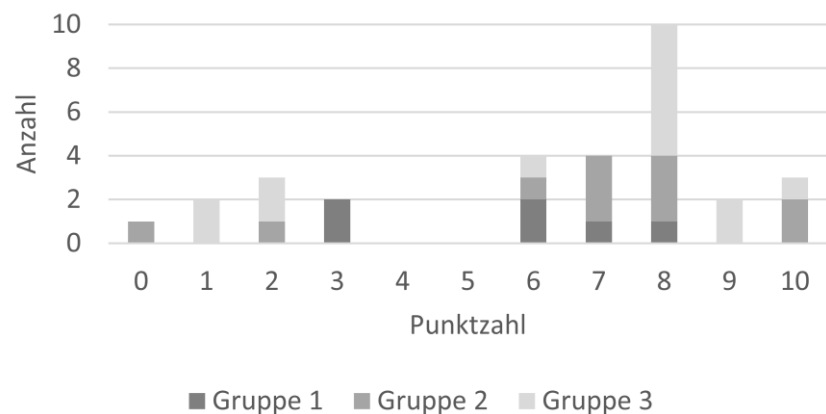


Abb. 19 A-B: Gruppenspezifische Darstellung der insgesamt vergebenen Punktzahl (**19 A.1; B.1**) sowie der Häufigkeit der einzelnen Punkte (**19 A.2; B.2**) bei der Evaluation des Blockpraktikums für die Aussage **19 A:** „Ich konnte problemlos ein eigenes Fallbeispiel entwickeln.“, **19 B:** „Auf Grundlage der Anamnese und Patientendaten konnte ich im Blockpraktikum ein entsprechendes Patientenprofil für mein Fallbeispiel festlegen.“.

Ebenso gemischt schätzen die Studenten die Recherche der notwendigen Patientendaten für die Therapiefindung in der Hausarztpraxis mit durchschnittlich 6,2 Punkten bei einer Standardabweichung von 3 Punkten ein (**Abb. 19 C.1**). In der Literaturrecherche fühlen sie sich mit durchschnittlich 7,5 Punkten ($s=2,1$; **Abb. 19 D.1**) deutlich sicherer. Die gruppenspezifische Aufschlüsselung bestätigt die größere Selbstsicherheit der Studenten bei der Literaturrecherche im Vergleich zur Recherche der Patientendaten (**Abb. 19 C.2**; **19 D.2**).

Abb. 19 C.1: BP 3

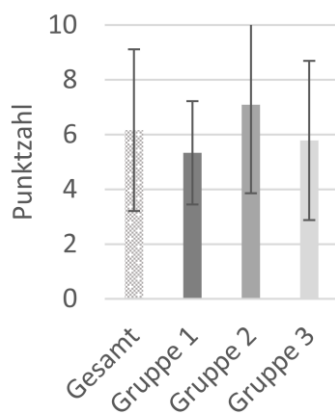


Abb. 19 C.2: Blockpraktikum 3

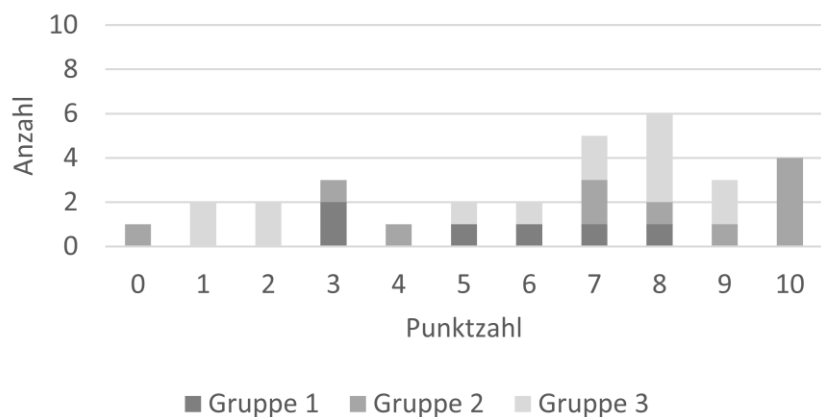


Abb. 19 D.1: BP 4

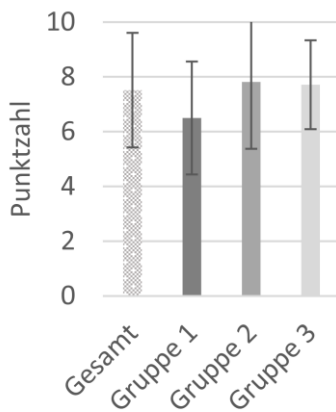


Abb. 19 D.2: Blockpraktikum 4

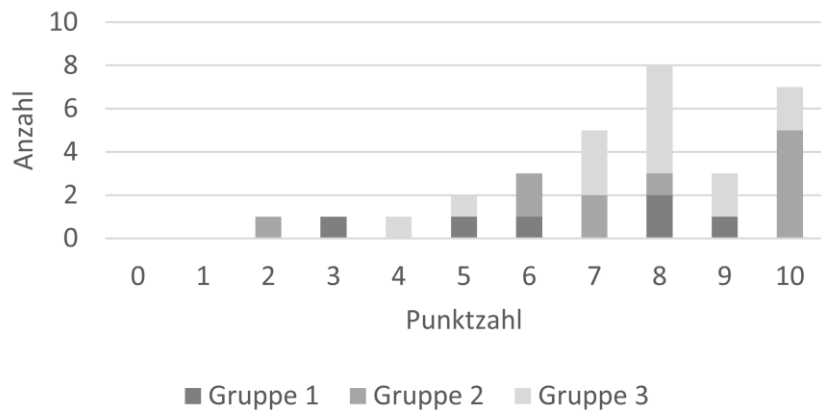


Abb. 19 C-D: Gruppenspezifische Darstellung der insgesamt vergebenen Punktzahl (**19 C.1**; **D.1**) sowie der Häufigkeit der einzelnen Punkte (**19 C.2**; **D.2**) bei der Evaluation des Blockpraktikums für die Aussage **19 C:** „Die für die Therapiefindung notwendigen Patientendaten konnte ich in der Hausarztpraxis problemlos erheben.“, **19 D:** „Mittels Literaturrecherche konnte ich eine Antwort auf meine Fragestellung finden.“.

Mit im Mittel 7,7 Punkten ($s=2,1$; **Abb. 19 E.1**) geben besonders viele Studenten an, Gelegenheit gehabt zu haben die Therapie des Patienten mit dem Lehrarzt zu besprechen. In der zweiten Gruppe zeigen sich sogar mehr als 90 % der Teilnehmer sehr zufrieden mit dem Kontakt zu ihren Lehrärzten (**Abb. 19 E.2**). 74 % der Blockpraktikanten profitierten laut eigenen Angaben von dem Fachwissen und der Erfahrung des Lehrarztes bei der Entwicklung eines eigenen Therapiekonzepts (**Abb. 19 F.2**). Ein kleinerer Anteil gab an, wenig bis gar keine neuen Aspekte aus dem Gespräch mit dem Lehrarzt erhalten zu haben, sodass der Mehrgewinn durch die Zusammenarbeit mit insgesamt 6,9 Punkten bei einer erhöhten Standardabweichung von $s=2,6$ Punkten evaluiert wird (**Abb. 19 F.1**).

Abb. 19 E.1: BP 5

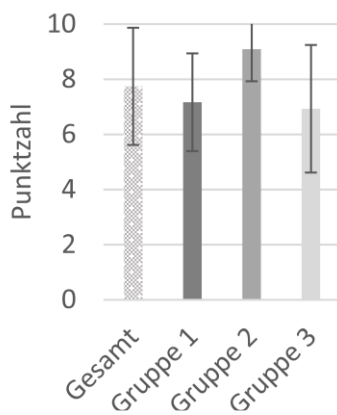


Abb. 19 E.2: Blockpraktikum 5

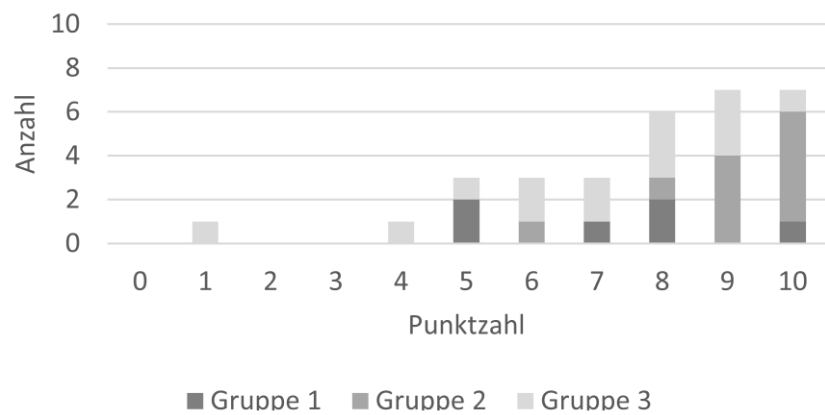


Abb. 19 F.1: BP 6

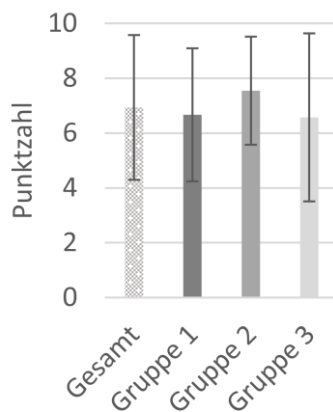


Abb. 19 F.2: Blockpraktikum 6

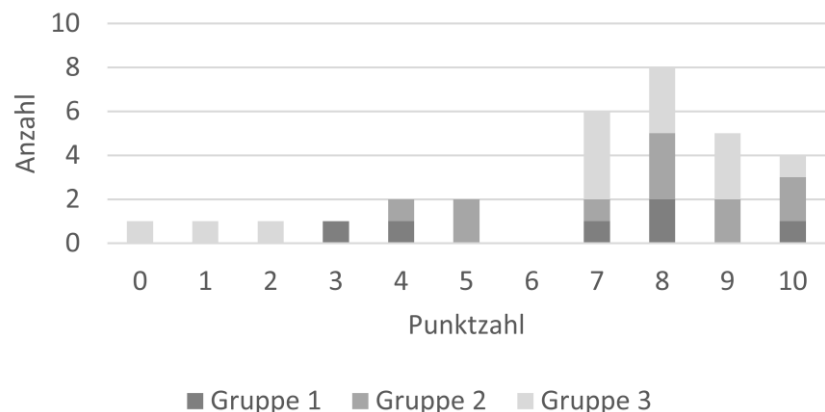


Abb. 19 E-F: Gruppenspezifische Darstellung der insgesamt vergebenen Punktzahl (**19 E.1; F.1**) sowie der Häufigkeit der einzelnen Punkte (**19 E.2; F.2**) bei der Evaluation des Blockpraktikums für die Aussage **19 E:** „Im Rahmen des Blockpraktikums hatte ich Gelegenheit die Therapie des Patienten mit dem Lehrarzt zu besprechen.“ und **19 F:** „Das Fachwissen und die Erfahrung des Lehrarztes konnten mir neue Aspekte zu meinem Fallbeispiel vermitteln, die mich bei der Entwicklung eines Therapiekonzepts beeinflusst haben.“.

3.3.4. Ergebnisse der Evaluation: Gesamtnote

Insgesamt bewerten 74 % der Studenten die Erweiterung des Blockpraktikums mit gut oder sehr gut. In Gruppe eins und zwei sind es sogar 83 % bzw. 82 % (**Abb. 20.2**). Die Dritte Gruppe bewertet die Umstellung wie auch alle vorherigen Aspekte, die mit mehr Zeitaufwand verbunden sind, im Durchschnitt mit 6,6 von 10 Punkten ($s= 2,3$) schlechter als die Blockpraktikanten der ersten beiden Testdurchläufe (**Abb. 20.1**). Die Blockpraktikanten sprechen sich gruppenübergreifend für die Implementierung des Lehrkonzepts zu einem früheren Zeitpunkt im Curriculum aus.

Abb. 20.1: Note

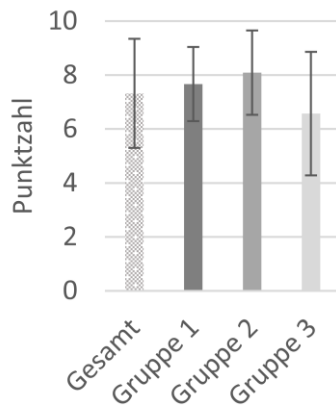


Abb. 20.2: Gesamtnote

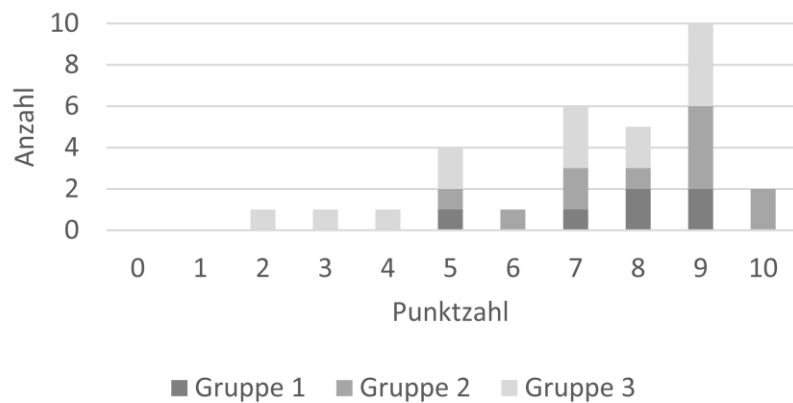


Abb. 20: Gruppenspezifische Darstellung der insgesamt vergebenen Punktzahl (**20.1**) sowie der Häufigkeit der einzelnen Punkte (**20.2**) für die Gesamtnote.

4. Diskussion

4.1 Überprüfung der Fragestellungen

Im Rahmen der Bonn Studie konnte mit Hilfe der drei erhobenen Datensätze die Machbarkeit der praktischen Anwendung von evidenzbasierter Medizin im Rahmen des humanmedizinischen Studiums im Fachbereich für Allgemeinmedizin gezeigt werden.

1. Die Ergebnisse des Bonn Tests für die fiktiven Fallbeispiele A, B und C sprechen dafür, dass die Studenten durch das Lernmodul eine ausreichende Basisausbildung auf dem Gebiet der EbM erhalten. Der Erwartungshorizont des Bonn Tests kann die neu erworbenen Fähigkeiten gezielt erfassen und Stärken bzw. Schwächen in Bezug auf den 5-A-Schritteprozess angemessen differenzieren. Da die Kriterien der Grundlagenvermittlung und Evaluierbarkeit erfüllt sind, ist die allgemeine Implementierung der Lehreinheit zum Thema EbM unabhängig vom Kontext der praktischen Anwendung machbar.
2. Die selbstentwickelten Fragestellungen können ebenfalls mit dem Bonn Test bearbeitet und ausgewertet werden. Dies spricht für die universelle Gültigkeit der im Lernmodul vermittelten Inhalte. Der Erwartungshorizont des Bonn Tests stellt sich wie bei den fiktiven Fallbeispielen als adäquates Evaluationsinstrument für die relevanten EbM-Fähigkeiten dar. Somit ist die Implementierung der Lehreinheit zum Thema EbM auch im Kontext der praktischen Anwendung machbar und die Erfassung von Leistungen der Studenten in Bezug auf die unterschiedlichen EbM-Fähigkeiten durch den Bonn Test und seinen Erwartungshorizont universell möglich.
3. Die Evaluation der studentischen Zufriedenheit mit der Lehreinheit zum Thema EbM im Rahmen des Blockpraktikums für Allgemeinmedizin zeigt eine hohe Akzeptanz, sodass eine zukünftige Implementierung des Lehrkonzepts in das Curriculum möglich ist. Die Studenten sprechen sich jedoch einvernehmlich

dafür aus, dass dies zu einem früheren Zeitpunkt im Studium sinnvoller ist.

In den folgenden Kapiteln diskutiere ich die Ergebnisse der erhobenen Datensätze in Hinblick auf die zuvor festgelegten Kriterien, die erfüllt sein müssen, damit die praktische Anwendung von evidenzbasierter Medizin im Rahmen des Blockpraktikums für Allgemeinmedizin machbar ist (Kapitel 4.6 Überprüfung der Fragestellungen):

- Grundlagenvermittlung
- Evaluierbarkeit
- Universelle Gültigkeit
- Studentische Akzeptanz

4.2 Grundlagenvermittlung

Die erfolgreiche Grundlagenvermittlung durch das Lernmodul ist jeweils abhängig von Inhalt, Format und organisatorischer Einbindung dieses Moduls.

4.2.1 Inhalt des Lernmoduls

Den Inhalten, die im Lernmodul und den Seminaren vermittelt werden, liegt eine ausführliche Recherche unterschiedlicher EbM-Modelle zu Grunde. Die Informationen sind valide, da sie ausschließlich qualitativ hochwertigen Quellen entstammen und bereits in der Vergangenheit erfolgreich in EbM-Schulungen angewendet werden konnten. Die Evidenzbasierte Medizin ist ein großer Themenkomplex, der sich innerhalb eines zweiwöchigen Blockpraktikums in seiner gänzlichen Tiefe nicht erfassen lässt. Der Großteil der EbM-Schulungen im Medizinstudium ist daher auf einen längeren Zeitraum ausgelegt bzw. umfasst deutlich mehr Einzeltermine (Beyer M et al., 2005). Somit besteht die erste Schwierigkeit des Lehrkonzepts darin, die Inhalte so weit zu komprimieren, dass die Studenten innerhalb des zweiwöchigen Blockpraktikums dennoch an den Abläufen in den Hausarztpraxen teilnehmen können. Die Entwicklung des online Lernmoduls ist ein Lösungsansatz für dieses Problem. Hierdurch haben die Blockpraktikanten ab Freischaltung des Lernmoduls auf der Plattform E-Campus eine Woche Zeit das

Lernmodul selbstständig von ihrem eigenen PC aus zu bearbeiten. Solche interaktiven Lerneinheiten als Teil eines Lehrkonzepts im Blended Learning Stil haben sich in der Vergangenheit bewährt und werden daher für EbM-Schulungen empfohlen (Ilic et al., 2015). Insgesamt wird das Thema EbM an der Universität Bonn bisher nur eingeschränkt im Rahmen anderer Fächer wie Biochemie oder medizinische Statistik angeschnitten. Im Vorlesungsverzeichnis der Universität Bonn ist für das Wintersemester 17/18 nur im Studium der Pharmazie ein eigenes Hauptseminar für Evidenzbasierte Medizin aufgeführt (Universität Bonn, 2018). Somit ergänzen die Inhalte des Lernmoduls das Curriculum sinnvoll.

4.2.2 Format des Lernmoduls

Das Format des Lernmoduls ist an die zeitlichen Einschränkungen angepasst und ermöglicht eine selbstständige Arbeitsweise. Kapitelweise werden die Kategorien des 5-A-Schritteprozesses angemessen behandelt und jeweils mit Hilfe eines fortlaufenden Fallbeispiels verdeutlicht. Die zugehörige Grafik, die den aktuellen Arbeitsschritt der 5As hervorhebt, bietet eine zusätzliche Orientierung und hilft den schematischen Ablauf der EbM-Anwendung zu verinnerlichen. Die interaktiven Fragen zur Selbstüberprüfung am Ende jedes Kapitels entsprechen jeweils unterschiedlichen Aufgabentypen und werden somit den interdisziplinären Ansprüchen des Blended Learning Modells gerecht. Das ihnen zugeteilte Fallbeispiel A, B oder C sowie die Seminartermine dienen der Prüfung von Bonn Test und Erwartungshorizont unter kontrollierten Rahmenbedingungen vor dem Einsatz in den Hausarztpraxen.

4.2.3 Organisatorischer Rahmen des Lehrkonzepts

Obwohl das Konzept erstmalig in drei Gruppen getestet wird, halten sich die Studenten an die Arbeitsanweisungen und erscheinen vorbereitet zum Kurs. Lediglich ein Student hatte die drei E-Mails mit Informationen zum Ablauf des erweiterten Blockpraktikums nicht gelesen und war auch am Montag in die Hausarztpraxis gefahren. Generell sollte die Versendung von E-Mails jeweils zu Semesterbeginn, eine Woche vor der Freischaltung des Lernmoduls, sowie am Tag der Freischaltung an die Blockpraktikanten aber

ausreichend sein. Lediglich dem Arbeitsauftrag bezüglich des Ausdrucks von Abstracts bzw. der Angabe eines entsprechenden Literaturvermerks sollte darin noch mehr Nachdruck verliehen werden, da die Studenten dies vermehrt versäumen. Eine semesterübergreifende Durchführung der EbM Schulung würde ebenfalls zur Reduktion fehlender Unterlagen beitragen, da sich die Studenten bei Unklarheiten der Arbeitsaufträge an ihre Kommilitonen wenden können. Langfristig ist auch die Schaffung weiterer Anreize denkbar, um die Quote der eingereichten Beiträge zu erhöhen, z. B. ein Preis für das beste Ergebnis im Bonn Test. Auch ein Anteil des Bonn Tests an der Endnote wäre denkbar. Im Rahmen der Machbarkeitsstudie ist dies nicht durchführbar, da aus ethischen Gründen für die Blockpraktikanten durch die Teilnahme an der Studie kein Nachteil entstehen darf.

4.3 Evaluierbarkeit

Bonn Test und Erwartungshorizont präsentieren sich in Anbetracht der Ergebnisse als adäquates Evaluationsinstrument für die im Lernmodul vermittelten EbM-Fähigkeiten. Dabei kann sowohl der Vergleich der Punktzahlen für die theoretischen Fallbeispiele A, B und C als auch für die einzelnen Kategorien des 5-A-Schritteprozesses die Stärken und Schwächen der Studenten im Umgang mit EbM gezielt herausstellen.

4.3.1 Bonn Test – Theorie

Dass in jeder Kategorie des Bonn Tests mindestens ein Student die volle Punktzahl erreicht, zeigt nochmals, dass das Lernmodul die zum Bestehen notwendigen Fähigkeiten ausreichend vermittelt. In dem ersten Durchlauf des Bonn Test, bei dem die Studenten EbM auf die theoretischen Fallbeispiele A, B bzw. C anwenden, liegen alle auswertbaren Beiträge oberhalb der Grenze zum Bestehen. Dies spricht sowohl für eine ausreichende Vermittlung der Inhalte des Lernmoduls und einen angemessenen Erwartungshorizont des Bonn Tests als auch für ein intaktes Zusammenspiel von Lernmodul und Bonn Test. Gleichzeitig variieren die Leistungen in den unterschiedlichen Kategorien von Student zu Student, sodass davon auszugehen ist, dass die Aufgabenstellung des Bonn Tests durchaus anspruchsvoll ist und die Punkte die Antwortqualität angemessen

widerspiegeln. Durch die Analyse der Punkte in den einzelnen Kategorien können Schwachstellen aufgedeckt und die Lerninhalte des Lernmoduls und der Seminare daran angepasst werden, um die Lehre noch weiter zu verbessern.

4.3.2 Bonn Test Ergebnisse – Theorie: Ask

Die Unterkategorie Patient Problem des PICO-Schemas ist der mit Abstand stärkste Teilbereich der Kategorie Ask, sodass der Patient eine zentrale Stellung bei der Formulierung der Fragestellung bei den Studenten einnimmt. Somit erfüllen sie bereits eine wichtige Voraussetzung für die Anwendung von Evidenzbasierter Medizin, denn der Patient bildet stets den Mittelpunkt des 5-A-Schritteprozesses (Duke University Medical Center Library und Health Sciences Library at the University of North Carolina at Chapel Hill, 2014). Seine Problematik ist Ausgangspunkt und sein Wohlergehen, seine Bedürfnisse und seine individuellen Eigenschaften sind allgegenwärtig bei der Entwicklung einer Behandlungsstrategie. Im Lernmodul wird der Patient nicht mehr hervorgehoben als die anderen Punkte des PICO-Schemas. Die Blockpraktikanten scheinen ihren Schwerpunkt also intuitiv auf ihn zu legen. Die Intervention steht für die Studenten gleich an zweiter Stelle hinter dem Patienten. Die zum Teil unsichere Differenzierung zwischen den Unterkategorien Intervention und Comparison trägt unter anderem dazu bei, dass mehr als ein Drittel der Blockpraktikanten den Unterpunkt Comparison vollkommen außer Acht lässt. Eine weitere Rolle spielt dabei wohl, dass im Studium eher Leitlinien und der Goldstandard vermittelt werden, aber die Lehre weniger auf alternative Behandlungsmethoden eingeht. Dabei können diese aufgrund der individuellen Veranlagung des Patienten im Einzelfall besser geeignet sein. So ist beispielsweise die Gabe von kontraindizierten Medikamenten bei Missachtung von Arzneimittelallergien ein häufiger Grund für Behandlungsfehlervorwürfe (Madea B et al.). Es ist also sinnvoll im Rahmen des Seminars einen Schwerpunkt auf die Unterkategorie Comparison zu legen, um die Studenten für die Suche nach Alternativen zu den herkömmlichen Therapieansätzen zu sensibilisieren. Das Outcome rückt im direkten Vergleich etwas mehr in das Bewusstsein der Studenten. Da in dieser Unterkategorie jedoch entweder die Höchstpunktzahl oder überhaupt keine Punkte erreicht werden, lässt sich darüber mutmaßen, ob der Ausgang der Therapie für diesen geringeren Teil der

Studenten nicht präsent ist oder ob es möglicherweise irritiert, dass das Outcome zwar im PICO-Schema vorgesehen ist, aber häufig nicht mit in die Suchanfrage übernommen wird um die Ergebnisse nicht übermäßig einzuschränken. Dennoch sollte auch dieser Punkt bei der Vermittlung von EbM hervorgehoben werden. Nur durch das gemeinsame Abstecken der Behandlungsziele mit dem Patienten, können seine Bedürfnisse und Wünsche erfasst und so die gesamten folgenden Schritte der EbM maßgeblich beeinflusst werden. Der Patient definiert Lebensqualität möglicherweise anders als der Arzt und nur das Gespräch kann Aufschluss über seine Vorstellungen geben. Daher spielt die Patientenorientierung zunehmend eine wichtige Rolle als Qualitätsmerkmal bei der Patientenversorgung (Lecher et al., 2002). Überwiegend handeln die Blockpraktikanten bei der Entwicklung einer Fragestellung also wie empfohlen nach dem PICO-Schema. Eine Vertiefung zur Suche nach alternativen Therapien sowie der gemeinsamen Zielfindung mit dem Patienten können gegebenenfalls zu einer Verbesserung dieser Fähigkeit führen. Bei der gruppenspezifischen Betrachtung der fiktiven Fallbeispiele ist die Entwicklung einer wissenschaftlichen Fragestellung die stärkste Kategorie im Fall B. Das ist vermutlich auf die detaillierte Fallbeschreibung zurückzuführen. Inhaltlich erkundigt sich darin ein Gichtpatient nach diätischen Optionen und wünscht insbesondere Informationen über die Verträglichkeit konkreter Nahrungsmittel wie Schweinebraten oder Kaffee. Da der fiktive Patient bereits unterschiedliche Aspekte der Diät anspricht, berücksichtigen die Studenten eher alternative Interventionen und erreichen mehr Punkte in der Unterkategorie Comparison. Auch Patient und Intervention werden in der Beschreibung des Fallbeispiels so klar benannt, dass in diesen Unterkategorien konstant überdurchschnittlich hohe Punktzahlen erreicht werden. Vermutlich schneiden die Studenten im Fallbeispiel C deutlich schlechter ab, da der Patient darin nach einer Korrelation zwischen erhöhten Harnsäurewerten und kardiovaskulären Erkrankungen fragt. Das entspricht auf den ersten Blick nicht dem klassischen PICO-Schema. Dabei lassen sich auch aus dieser Aufgabenstellung die einzelnen Aspekte Patient, Intervention, Comparison und Outcome herausarbeiten. Sie werden lediglich nicht so offensichtlich benannt und entsprechen somit eher einem realen Fall. Indem die Studenten die Anwendung des PICO-Schemas auf die weniger offensichtlichen Fallschilderungen üben, sind sie besser auf die realen Fälle in den Hausarztpraxen vorbereitet.

4.3.3 Bonn Test Ergebnisse – Theorie: Acquire – Suchstrategie

Die Wortwahl ist die stärkste Leistung der Studenten im gesamten Bonn Test. Die intuitive Auswahl der für die Suchanfrage relevanten Inhalte ist möglicherweise auf den alltäglichen Umgang mit Suchmaschinen in Internet wie Google zurückzuführen. Allerdings könnten die Blockpraktikanten von einer Vertiefung der Lehre zu Medical Subject Headings und präzisierenden Optionen profitieren. In nur etwa der Hälfte der Beiträge werden ausschließliche MeSH verwendet. Im Fall von PubMed werden zwar auch verwandte Wörter automatisch zu den MeSH verlinkt, allerdings ist es generell wichtig bei der Suchanfrage Formulierungen zu verwenden, die das Ergebnis nicht unnötig einschränken. Durch den ausschließlichen Gebrauch von Medical Subject Headings kann diese Fähigkeit trainiert werden. Zusätzlich sollten die wichtigsten präzisierenden Optionen in Lernmodul deutlicher hervorgehoben werden, da bereits die Einschränkung der Suche nach Systematic Reviews bzw. die korrekte Verwendung von Operatoren die Qualität der Suchergebnisse verbessern und den zeitlichen Aufwand der Recherche reduzieren kann. Bei der Formulierung der Suchstrategie schneidet Fallbeispiel B in allen drei Unterkategorien, vor allem aber bei der Verwendung von präzisierenden Optionen, schlechter ab. Möglicherweise vernachlässigen die Studenten diese Funktion bei der Suchanfrage eher als bei den anderen Fallbeispielen, da im vergleichbar detailreichen Fall B die einzelnen Aspekte des PICO-Schemas scheinbar schon präzise genug benannt sind. Dadurch verzichten die Studenten aber unter anderem auf die Eingrenzung der vorgeschlagenen Abstracts auf besonders hochwertige Studien mit Hilfe des Review-Filters. Die Blockpraktikanten sollten in den Seminaren also nochmals darauf aufmerksam gemacht werden, dass die Qualität der vorgeschlagenen Rechercheergebnisse mit Hilfe der präzisierenden Optionen beeinflusst werden kann.

4.3.4 Bonn Test Ergebnisse – Theorie: Acquire – Abstract

Möglicherweise ist das Kapitel des Lernmoduls zu den unterschiedlichen Studiendesigns und Recherchequellen zu ausführlich, sodass die wichtigen Kernaussagen untergehen. Es gelingt zwar nahezu allen Studenten ein inhaltlich geeignetes Abstract zu wählen, doch nur in etwa zwei Drittel der Beiträge wird dabei auf die höchste verfügbare Evidenzstufe

zugegriffen. Sowohl im Lernmodul, als auch in den Seminaren sollte also nochmals betont werden, dass für die Frage nach der idealen Therapie ein Systematic Review am besten geeignet ist. Für andere Fragestellungen können auch andere Studiendesigns von Vorteil sein. Wird beispielsweise die Auswahl und Interpretation des besten Diagnostiktests erfragt, bietet sich am ehesten eine prospektive Studie an, bei der ein neues Testverfahren im Blindversuch gegen den bestehenden Goldstandard erprobt wird. Richtet sich die Frage hingegen nach der zu erwartenden Prognose für den Krankheitsverlauf des Patienten, liefert eine Kohorten Studie die aussagekräftigsten Daten (Duke University Medical Center Library und Health Sciences Library at the University of North Carolina at Chapel Hill, 2014). Diese Unterschiede sollten den Studenten klar sein, um die bestmögliche Auswahl an Studienergebnissen treffen zu können, die letztendlich in der Therapiefindung berücksichtigt werden. Die im Vergleich zu den anderen Fällen höhere Punktzahl der Kategorie Acquire - Abstract im Fall A lässt sich dadurch begründen, dass es viele hochwertige Studien zur Wirkung von Allopurinol bei Niereninsuffizienz gibt und diese Thematik sogar in der Leitlinie berücksichtigt wird. Somit ist die Wahrscheinlichkeit auf eine hochwertige Studie zu stoßen höher als bei einem weniger erforschten Thema. Zusammenfassend erbringen die Studenten bei der Formulierung der Suchstrategie und der Auswahl des Abstracts also bereits sehr starke Leistungen. Eine Vertiefung der geeigneten Studiendesigns sowie der Strategien zur Präzisierung der Suchanfrage kann diese Fähigkeiten jedoch verbessern und den Umfang der Rechercheergebnisse auf ein realistisches Maß reduzieren.

4.3.5 Bonn Test Ergebnisse – Theorie: Appraise

Die Evaluation der Rechercheergebnisse ist bei der Anwendung von EbM entscheidend darüber, ob die Erkenntnisse der ausgewählten Studien valide, in sich stimmig und vor allem auf den Patienten übertragbar sind. Auch wenn dieser Schritt bei den häufig verwendeten, vorab durch Experten evaluierten Informationsquellen wie der Cochrane Library, UptoDate oder auch Systematic Reviews auf PubMed in den Hintergrund rückt, stellt er die zentrale Fähigkeit der Evidenzbasierten Medizin dar. Nur wenn die Studenten lernen, die ihnen angebotenen Informationen unabhängig von der Quelle kritisch zu hinterfragen, kann eine Generation von Ärzten ausgebildet werden, die sich an den

individuellen Eigenschaften und Bedürfnissen des Patienten orientiert und weniger anfällig für äußere Einflüsse zum Beispiel durch die Pharmaindustrie ist. Das Lernmodul schult die Blockpraktikanten bezüglich der Beurteilung von Validität, Größe und Präzision beziehungsweise Übertragbarkeit und Anwendbarkeit einer Studie. Diese Fähigkeit ist entsprechend der Ergebnisse des Bonn Tests noch ausbaufähig. So treten gehäuft Probleme bei der Zuordnung der Evidenzstufe auf. Konkrete Beispiele für jede Evidenzstufe könnten dabei helfen, diese Hierarchie besser zu verstehen. Da dies den Rahmen des Lernmoduls überschreiten würde, bietet sich eine Lerneinheit in den Seminaren an, bei der die Studenten die Evidenzstufen konkreten Abstracts zuordnen sollen. Weiterhin zeigen sich größere Probleme bei der Untersuchung von Übertragbarkeit und Anwendbarkeit der Rechercheergebnisse. Auch diese Fähigkeit sollte idealerweise anhand von weiteren Beispielen geübt werden, bei denen die Blockpraktikanten die Unterschiede zwischen den Eigenschaften der Studienteilnehmer und dem konkreten Patienten erkennen müssen. Die Fähigkeit zur Evaluation von Rechercheergebnissen sollte also vor allem unter dem Aspekt der Übertragbarkeit, sowie der Hierarchie der Evidenzstufen vertieft werden. Dass die Evaluation der Rechercheergebnisse am besten im Fallbeispiel C gelingt ist vermutlich darauf zurückzuführen, dass die Studenten sich hierbei eher mit der Materie auseinandersetzen. Die Thematik der Korrelation von erhöhten Harnsäurewerten und kardiovaskulären Erkrankungen ist weniger erforscht als die Anwendung von Allopurinol bei Niereninsuffizienz oder diätischen Empfehlungen bei Hyperurikämie und erzielt somit weniger hochwertige Rechercheergebnisse. Im Ausgleich setzen sich die Studenten ausführlicher mit möglichen Mängeln der Studien und den Parallelen oder Differenzen zu ihrem eigenen Patienten auseinander. Passend zu dieser Erklärung schneidet Fallbeispiel A am schlechtesten ab. Obwohl im Fall A die inhaltlich und qualitativ besten Abstracts ausgewählt werden, gehen die Blockpraktikanten im weiteren Verlauf am seltensten auf diese Eigenschaften ein. Da es unwahrscheinlich ist, dass die ausgewählten Abstracts zufällig nahezu ausschließlich die Höchstpunktzahl erzielen, sind sich die Studenten deren Inhalt und Qualität vermutlich so sicher, dass sie die Bewertung von Validität, Größe und Präzision bzw. Übertragbarkeit und Anwendbarkeit häufig im Bonn Test vernachlässigen. Die bewusste Überprüfung dieser Kriterien ist jedoch unerlässlich, da sie nicht nur beweisend für das Verständnis dieser Merkmale ist, sondern auch Mängel in Design, Durchführung oder auch Auswertung in

der zugrundeliegenden Studie aufdecken kann. Daher sollten diese Eigenschaften auch bei augenscheinlich hochwertigen Studien überprüft werden.

4.3.6 Bonn Test Ergebnisse – Theorie: Apply

Obwohl die Studenten überwiegend in der Lage sind einen in sich validen Therapieansatz aus den Rechercheergebnissen abzuleiten, haben insbesondere diejenigen Studenten dabei Probleme, die auch in den vorherigen Kategorien Unsicherheiten zeigten. Da die Kategorie Apply auf den Ergebnissen der Schritte Ask, Acquire und Appraise beruht, kann zum Beispiel eine ungünstig formulierte Fragestellung oder Suchstrategie, sowie ein inhaltlich unpassendes Abstract, das gar nicht auf den Patienten übertragbar ist, dazu führen, dass die daraus gezogenen Schlüsse nicht mehr auf das ursprüngliche Problem eingehen. Somit sollte ein intensiveres Training der ersten drei Schritte auch zu einem besseren Ergebnis bei der Anwendung führen. Dass der Großteil der Studenten 0 Punkte in der Unterkategorie externe Validität erzielt ist weniger auf mangelnde Kompetenzen, sondern viel mehr auf den Erwartungshorizont selbst zurückzuführen. Eine hervorragende externe Validität wird dabei definiert als Berücksichtigung möglicher Restriktionen der Anwendung, zum Beispiel die Übertragbarkeit auf den Patienten oder die praktische Umsetzbarkeit. Dabei können gegebenenfalls Verbesserungsvorschläge für die Umsetzung der Studie von einem sehr guten Verständnis für Probleme und Schwachstellen zeugen. Da die Studenten jedoch in der vorausgehenden Kategorie Appraise die Übertragbarkeit der Studie auf den Patienten prüfen, gehen die meisten von ihnen nicht mehr zusammenfassend auf die externe Validität ein. Dieses Missverständnis lässt sich möglicherweise bereits durch eine angepasste Fragestellung des Bonn Tests beheben. Die Fragestellung lautet bisher „Sind die Ergebnisse aus der Studie auf Ihren Patienten übertragbar?“. Der Zusatz „Wie können mögliche Restriktionen behoben werden?“ könnte die Blockpraktikanten an dieser Stelle für die Überprüfung der externen Validität sensibilisieren. Zudem sollte die Aufteilung in interne und externe Validität in das Lernmodul übernommen werden, damit die Blockpraktikanten eher einschätzen können, auf welche Punkte sie bei der abschließenden Beurteilung der Recherche und bei der sinnvollen Therapiewahl achten müssen. Insgesamt sollte sich die Fähigkeit Apply also durch eine gezielte Schulung der vorherigen Schritte Ask, Acquire und Appraise sowie

eine Ergänzung des Lernmoduls um die Definitionen zu interner und externer Validität und gegebenenfalls präziseren Formulierungen des Bonn Tests verbessern lassen. Die weniger klar formulierten PICO-Elemente und die seltener in Studien thematisierte Fragestellung im Fall C führen scheinbar trotz der missverständlichen Aufgabenstellung dazu, dass die Blockpraktikanten ihre Rechercheergebnisse kritischer hinterfragen und explizit bei der Auswahl der Therapie berücksichtigen. Sie sollten zukünftig dafür sensibilisiert werden auch scheinbar offensichtliche Lösungsansätze kritisch zu überprüfen und so auf die individuellen Bedürfnisse des Patienten abzustimmen.

4.3.7 Bonn Test Ergebnisse – Theorie: Assess

Die Blockpraktikanten sind zwar größtenteils in der Lage eine sinnvolle Evaluationsmaßnahme zu wählen, mangels klinischer Erfahrung treten allerdings häufig Probleme bei der Festlegung eines Zeitintervalls auf. Auch in diesem Fall bietet sich eine Anpassung der Fragestellung im Bonn Test an. Das korrekt gewählte Zeitintervall dient im Erwartungshorizont der Differenzierung zwischen hervorragenden und starken Leistungen, wird im Lernmodul aber lediglich in den vorgestellten Fallbeispielen besprochen. Da die Studenten selten ausreichend klinische Erfahrung haben, um beispielsweise die zeitlichen Abstände zwischen Laborkontrollen einschätzen zu können, lassen sie diese häufig ganz weg. Im Bonn Test könnte man also nicht nur die Maßnahme, sondern auch den zeitlichen Rahmen für die Verlaufsevaluation der gewählten Behandlung direkt erfragen. Das Seminar könnte durch einen Crashkurs zu Verlaufskontrollen typischer Krankheitsbilder erweitert werden, um den Studenten ein Gefühl dafür zu vermitteln, wie solche Intervalle typischerweise aussehen könnten. Da die korrekte Wahl eines Zeitintervalls jedoch eher ein Zeichen klinischer Erfahrung ist, könnte der Schwerpunkt wieder mehr auf evidenzbasierte Medizin gelegt werden, indem der Erwartungshorizont durch eine zweite Unterkategorie erweitert wird. Inhaltlich bietet sich dafür eine Reevaluation der anfangs festgelegten Ziele an. Dies entspricht zwar dem Outcome des PICO-Schemas vom Anfang des Bonn Tests, aber es hilft einen direkten Vergleich herzustellen, ob die anfänglich festgelegten Ziele beibehalten wurden und das gewählte Therapiekonzept überhaupt in der Lage ist diese zu erreichen. Zur Verbesserung der Kategorie Assess könnte also ein Intensivkurs zu Verlaufskontrollen

typischer Krankheitsbilder beitragen. Auch hier könnte die Aufgabenstellung des Bonn Tests präziser auf den Erwartungshorizont abgestimmt werden, indem auch das Zeitintervall für die Evaluationsmaßnahmen erfragt wird. Eine Ergänzung des Erwartungshorizonts um die Reevaluation der Ziele könnte der Kontrolle des festgelegten Outcomes dienen. Dass die Kategorie Assess den Studenten mit dem Fallbeispiel C schwerer fällt als der Gesamtheit lässt sich auf die anfänglichen Probleme in der Kategorie Ask zurückführen. Gelingt es den Studenten an dieser Stelle nicht ein relevantes Outcome festzulegen, so ist auch die gewählte Evaluationsmaßnahme häufig nicht zielführend in der Erhebung der notwendigen Daten. Durch eine Verbesserung des Schrittes Ask ist also ebenfalls eine gesteigerte Leistung in der Kategorie Assess zu erwarten.

4.3.8 Bonn Test: Verbesserungsmöglichkeiten

Die gesammelten Ideen zur Weiterentwicklung des Lehrkonzepts in den unterschiedlichen Kategorien des 5-A-Schritteprozesses werden im Folgenden noch einmal tabellarisch zusammengefasst (**Tab. 3**).

Tab. 3: Zusammenfassung von Verbesserungsvorschlägen für zukünftige Anwendungen des Lehrkonzepts.

Kategorie	Verbesserungsvorschläge
Ask	<ul style="list-style-type: none"> • Unterkategorien Comparison und Outcome vertiefen • Formulierung einer Fragestellung bei nicht-klassischer Fallkonstellation kann besser auf reale Fälle vorbereiten
Acquire – Suchstrategie	<ul style="list-style-type: none"> • Unterkategorien MeSH und Präzisierende Optionen vertiefen
Acquire – Abstract	<ul style="list-style-type: none"> • Recherchequellen in Bezug auf Systematic Reviews vertiefen • Geeignete Studiendesigns für andere Fragestellungen, z. B. nach der besten Diagnostik, vorstellen
Appraise	<ul style="list-style-type: none"> • Unterkategorie Übertragbarkeit vertiefen • Hierarchie der Evidenzstufen verdeutlichen
Apply	<ul style="list-style-type: none"> • Anpassung der Fragestellung im Bonn Test (Betonung der externen Validität) • Aufteilung in interne und externe Validität im Lernmodul
Assess	<ul style="list-style-type: none"> • Anpassung der Fragestellung im Bonn Test (zeitlichen Rahmen für die Verlaufskontrolle der gewählten Behandlung direkt erfragen) • Intensivkurs zu Verlaufskontrollen typischer Krankheitsbilder entwickeln • Ergänzung des Erwartungshorizonts um die Reevaluation der anfangs festgelegten Ziele

4.4 Universelle Gültigkeit (Bonn Test – Praxis)

Auch wenn die durchschnittliche Punktzahl des Bonn Tests in der Praxis unter der Punktzahl für die Anwendung in der Theorie liegt, so übertrifft sie dennoch mit Abstand die nötige Mindestpunktzahl zum Bestehen. Das zeigt, dass sowohl die Inhalte des Lernmoduls, als auch Aufgabenstellung und Erwartungshorizont von der Theorie in die Praxis übertragen werden können. Die durchschnittlich geringere Punktzahl in einigen

Kategorien ist vermutlich auf die Komplexität der realen Patientenfälle zurückzuführen, die noch mehr Übung erfordert, um den 5-A-Schritteprozess systematisch anzuwenden. Dass die stärkste und schwächste Kategorie jeweils für Theorie und Praxis identisch sind spricht ebenfalls dafür, dass der Bonn Test die EbM-Fähigkeiten angemessen widerspiegelt. In der Kategorie Ask verbessern sich die Studenten leicht. Obwohl sie im ersten Durchlauf des Bonn Tests noch eher Probleme mit komplexeren Fallbeschreibungen aufweisen, scheint sich dies im zweiten Durchgang dadurch zu relativieren, dass sie die Fragestellung selbstständig auswählen. Dadurch setzen sie sich intensiver mit Behandlungsoptionen und Therapiezielen auseinander und sind insbesondere für die Unterkategorien Intervention, Comparison und Outcome sensibilisiert. Bei der Formulierung einer Suchstrategie schneiden die komplexen Fälle aus der Hausarztpraxis wie zuvor schon das detailreichere Fallbeispiel B schlechter als die anderen fiktiven Fälle ab. Ebenso entspricht die Kategorie der Abstractauswahl den bereits starken Leistungen der Studenten im ersten Durchlauf. Lediglich die Punkte für die Unterkategorie Qualität fallen etwas geringer aus, da die Studienlage einiger gewählter Themen weniger evident ist. Obwohl die Studenten besonders hochwertige Studien zur Beantwortung ihrer Fragestellungen auswählen, vernachlässigen sie die separate Evaluation dieser Rechercheergebnisse. Da die Auswahl des Abstracts aber in beiden Durchläufen die größte Stärke der Blockpraktikanten darstellt, ist davon auszugehen, dass sie durchaus in der Lage sind hochwertige Studien zu erkennen und lediglich die Evaluation für selbstverständlich halten oder weniger ausführlich notieren. In der Kategorie Apply schneiden die Studenten in der Praxis besser ab als in der Theorie. Das spricht ebenfalls dafür, dass sie sich in der Realität stärker mit der Therapie auseinandersetzen. Bei der Kategorie Assess stoßen sie dabei hingegen vermehrt an ihre Grenzen bei der Festlegung von Evaluationsmaßnahmen mangels klinischer Erfahrung. Das weite Themenspektrum, das insgesamt durch die selbstgewählten Fragen abgedeckt wird, verdeutlicht das Potential der praktischen Anwendung von Evidenzbasierter Medizin.

4.5 Akzeptanz

Die Akzeptanz der Studenten für die Erweiterung des Blockpraktikums ist soweit gegeben, dass die zukünftige Implementierung in das Curriculum möglich ist. Einzelne gruppenspezifische Abweichungen bei der Evaluation, insbesondere durch die dritte Gruppe, lassen sich auf eine gewisse Gruppendynamik und den späteren Zeitpunkt der Durchführung im Semester zurückführen. Insgesamt sprechen sich die Studenten einvernehmlich für die Etablierung des Lehrkonzepts während eines früheren Abschnitts des Studiums aus.

4.5.1 Studentische Zufriedenheit – Lernmodul

Obwohl Systematic Reviews zu EbM-Schulungen gezeigt haben, dass eine Lehre im Blended Learning Stil den klassischen Konzepten vorzuziehen ist und allgemein auch von den Studenten favorisiert wurde (Ilic et al., 2015), zeigen die Blockpraktikanten der Bonner Machbarkeitsstudie hierzu ein äußerst gemischtes Meinungsbild. Im Gespräch mit den Studenten wird diesbezüglich häufig der Zeitaufwand des Lernmoduls als Grund angegeben. Eben dieser Faktor scheint auch maßgeblich Einfluss auf die negativeren Evaluationen der dritten Gruppe zu haben. Das Blockpraktikum im Fachgebiet Allgemeinmedizin wird im 6. klinischen Semester und somit unmittelbar vor dem 2. Staatsexamen durchgeführt. Der Projektablauf ist bei allen drei Gruppen identisch, aber weil die dritte Gruppe den spätesten Termin hat, sind besonders in dieser Gruppe einige Teilnehmer aufgrund der zusätzlichen Arbeit neben der Examensvorbereitung unzufrieden. Grundsätzlich erfordert das gewählte Lehrkonzept ein höheres Maß an eigenständiger Arbeit als bei der passiven Teilnahme an einem Vortrag zum Thema EbM. Da das Ziel der Machbarkeitsstudie jedoch die Vermittlung praktischer Fähigkeiten ist, die ebenfalls zukünftig aktiv angewendet werden sollen, ist die zeitlich aufwendigere Auseinandersetzung mit den Inhalten des Lernmoduls durchaus gerechtfertigt. Tatsächlich verhindert auch im Praxisalltag der Zeitaufwand einer ausführlichen Recherche häufig die Etablierung sinnvoller Neuerungen aus wissenschaftlichen Erkenntnissen. In den letzten Jahrzehnten hat sich die Publikationsrate von wissenschaftlich relevanten Artikeln auf dem Gebiet der Medizin vervielfacht. Während in

den späten 70er Jahren jeden Tag etwa 14 RCTs veröffentlicht wurden, waren es 2010 schon 75 RCTS und 11 Systematic Reviews (Bastian et al., 2010). Dazu kommen nochmals jeden Tag zahlreiche Veröffentlichungen niederer Evidenz, denen Ärzte neben ihrer praktischen Tätigkeit kaum gerecht werden können. Eine regelmäßige Bearbeitung wissenschaftlicher Fragestellungen mit Hilfe des Lernmoduls ist daher empfehlenswert, um die Abläufe besser zu verinnerlichen und evidenzbasierte Medizin effizienter anwenden zu können. Im weiteren Verlauf ist auch die Nutzung vorab evaluierter Recherchequellen wie der Cochrane Library sinnvoll und kann zeitlich in den Praxisalltag einfacher integriert werden. Unabhängig von ihrer Bewertung des Lehrprojekts sprechen sich die Studenten aller drei Gruppen für eine Implementierung in das Curriculum zu einem früheren Zeitpunkt des Studiums aus. Dadurch könnten die Kommilitonen der unteren Semester bereits frühzeitig von der Evidenzbasierten Medizin beispielsweise bei der Recherche für eine eigene Dissertation profitieren. Aus organisatorischen Gründen war die Durchführung des Projekts im Rahmen der Machbarkeitsstudie zunächst nur im 6. Klinischen Semester möglich. Zukünftig wäre die Etablierung zu einem früheren Zeitpunkt z. B. als Zusatz des Seminars Allgemeinmedizin im 2. Klinischen Semester wünschenswert.

4.5.2 Studentische Zufriedenheit – Seminare

Insgesamt sind die Blockpraktikanten zufrieden mit der Umsetzung des Seminars, geben aber an von den Inhalten weniger profitiert zu haben als von dem Lernmodul. Um dieser Rückmeldung der Studenten gerecht zu werden, sollten vermehrt neue inhaltliche Impulse im Vergleich zum Lernmodul vermittelt werden. So könnte im Rahmen zukünftiger Durchläufe in den Seminaren ein größerer Schwerpunkt auf die praktische Anwendung der Fähigkeiten gelegt werden. Dazu ermöglicht die Auswertung des Bonn Tests der drei Gruppen eine gezielte Analyse von häufigen Schwächen bei der Anwendung von EbM. Diese können für die Entwicklung kompakter Lerneinheiten genutzt werden, die es weiteren Blockpraktikumsgruppen ermöglichen aus den Fehlern ihrer Vorgänger zu lernen. Durch eine stetige Anpassung der Seminarinhalte nach Auswertung des Bonn Tests ist das Lehrkonzept so kurzfristig an neu auftretende Probleme adaptierbar. Jede Weiterentwicklung der Inhalte ermöglicht dadurch eine praxisnähere Lehre, die den

Studenten die praktische Anwendung von EbM erleichtern kann. Die gruppendifferente Evaluation bei der Abwägung des Blended Learning Stils gegen konventionelle Unterrichtsmodelle spiegelt deutlich die Gruppendynamik, insbesondere der dritten Gruppe, wider, die sich bereits in Bezug auf das Lernmodul manifestierte und in diesem Kontext diskutiert wird.

4.5.3 Studentische Zufriedenheit – Blockpraktikum

Überwiegend fühlen sich die Studenten bei der praktischen Anwendung von evidenzbasierter Medizin im Rahmen des Blockpraktikums sicher, was gemeinsam mit den positiven Ergebnissen des Bonn Test für den Erfolg der Machbarkeitsstudie spricht. Die meisten Probleme haben die Studenten bei der Erhebung der Patientendaten. Dies deckt sich mit den Ergebnissen einer Studie aus Tübingen zur Erhebung von Patientendaten in Hausarztpraxen durch Blockpraktikanten (Moßhammer et al., 2011). Grund für die problematische Datenerhebung war in dieser Studie häufig die fehlende Erfahrung im Umgang mit unterschiedlicher Praxissoftware. Der Erfolg ist dabei abhängig von der Unterstützung und Anleitung des Blockpraktikanten durch das Praxisteam, worum im Anschreiben an die Lehrärzte nochmals ausdrücklich gebeten werden sollte. Der Großteil der Blockpraktikanten bewertete die Zusammenarbeit mit dem Lehrarzt positiv und hilfreich bei der Entwicklung eines geeigneten Therapiekonzepts. Obwohl die Unterstützung durch den Lehrarzt kein eigentlicher Teil des Lehransatzes ist, auf den das Institut für Hausarztmedizin einen Einfluss hat, wird er dennoch in der Evaluation durch die Studenten berücksichtigt. Erst dadurch werden die drei Säulen, die die evidenzbasierte Medizin stützen, Attitude, Skills und Knowledge, vervollständigt (Beyer M et al., 2005). Selbst wenn die Blockpraktikanten die Bereitschaft sich ein Leben lang fortzubilden (Attitude) mitbringen und sich die nötigen Fähigkeiten (Skills) zur Anwendung von EbM mit Hilfe des Lernmoduls aneignen, so kann ihr Fachwissen (Knowledge) noch nicht mit dem eines ausgebildeten Hausarztes mit jahrelanger Berufserfahrung konkurrieren. Hier können die Blockpraktikanten von der internen Evidenz, also dem Fachwissen und der Erfahrung, des Lehrarztes profitieren. Gleichzeitig kann der Blockpraktikant, sofern er die Anwendung von EbM beherrscht, zusätzliche externe Evidenz, also Erkenntnisse aus der Literaturrecherche für den Lehrarzt erarbeiten und

evaluieren. Im Idealfall entsteht ein Arbeitsverhältnis von dem nicht nur Student und Lehrarzt profitieren, sondern auch der Patient, der immer im Mittelpunkt der evidenzbasierten Medizin steht. Neue wissenschaftliche Erkenntnisse könnten so schneller Einfluss auf die Therapie und somit auf die Qualität der Behandlung nehmen. Nach Westfall et al. dauert es noch durchschnittlich 17 Jahre, damit nur 14 % der neusten wissenschaftlichen Erkenntnisse zur Alltagsroutine im klinischen Alltag werden (Westfall et al., 2007). Wichtigste Voraussetzung für solch eine positive Zusammenarbeit ist dabei die Mitarbeit der Lehrärzte. In der Tübingen Studie wurde dies als Problem herausgearbeitet, da viele Lehrärzte das kritische Hinterfragen ihrer Therapieansätze negativ empfanden. Im Rahmen der Machbarkeitsstudie bewertet der Großteil der Blockpraktikanten die Zusammenarbeit mit dem Lehrarzt jedoch als positiv und hilfreich bei der Entwicklung eines geeigneten Therapiekonzepts. Bei den schlechteren Evaluationen gilt es im Einzelfall den Grund für die gescheiterte Zusammenarbeit zu klären. Die Sicht der Lehrärzte, die dabei eine wichtige Rolle spielt, wird durch eine andere Doktorandin am Institut für Hausarztmedizin im Rahmen ihrer Dissertation beleuchtet.

4.6 Stärken und Schwächen

Im Rahmen dieser Machbarkeitsstudie wird die Anwendbarkeit einer Lehreinheit zum praktischen Einsatz von Evidenzbasierter Medizin lediglich geprüft. Der Erwartungshorizont erlaubt bereits eine Leistungsdifferenzierung in unterschiedlichen Fähigkeiten der EbM, die in sich schlüssig erscheint. Dennoch sind weitere Studien wie RCTs notwendig um die Validität von Lernmodul, Bonn Test und Erwartungshorizont zu bestätigen. Auch wenn nicht alle Studenten das Konzept im Blended Learning Modell einem klassischen Frontalunterricht vorziehen, so ist dies doch die ideale Möglichkeit die Basisausbildung in Evidenzbasierter Medizin trotz des eingeschränkten zeitlichen Rahmens effizient zu vermitteln. Insbesondere das online Lernmodul führt anschaulich und interaktiv durch das notwendige Grundlagenwissen. Der Zugang über E-Campus ermöglicht eine hohe Flexibilität in Vorbereitung auf das Blockpraktikum. Wie jedes andere Selbststudium ist das online Lernmodul aber auch mit einem nötigen Maß an Eigeninitiative verbunden. Es liegt dabei in der Verantwortung der Studenten die Rahmenbedingungen für eine ruhige Lernatmosphäre zu schaffen. Das inhaltliche

Kernstück des Lehrkonzepts, der 5-A-Schritteprozess, hat sich bereits im Rahmen zahlreichen EbM-Schulungen bewährt, steht aber teilweise in der Kritik dafür den komplexen Vorgang der Evidenzbasierten Medizin auf ein zu simples Schema herunterbrechen zu wollen (Booth, 2009). Tatsächlich ist der 5-A-Schritteprozess aber bereits ein eher ausführliches Werkzeug, das im täglichen Gebrauch häufig nur teilweise Anwendung findet, und zum Beispiel durch vorabevaluierte Recherchequellen verkürzt wird. Gerade für eine erste Einführung in die EbM ist der 5-A-Schritteprozess aber ein ideales Modell, um die Anwendung der wichtigsten Fähigkeiten systematisch zu üben. Durch einen Pretest des Konzepts im Vorfeld, konnte die benutzerfreundliche Anwendung weiterhin verbessert werden. Zusätzlich ermöglichen die Schlussfolgerungen aus den Ergebnissen des Bonn Tests eine kontinuierliche Weiterentwicklung des Lehrkonzepts. Ebenso wie für das Lernmodul ist auch der Bonn Test auf die Eigenverantwortlichkeit der Studenten angewiesen. Damit die Ergebnisse wirklich vergleichbar sind, sollte der Test eigentlich unter identischen Ausgangsbedingungen durchgeführt werden, was allerdings den organisatorischen Rahmen des Blockpraktikums sprengen würde, da am Institut für Hausarztmedizin beispielsweise nicht ausreichend online Zugänge für die Literaturrecherche zur Verfügung stehen. Die fiktiven Fallbeispiele A, B und C sind aber randomisiert und werden den Studenten willkürlich vor Antritt des Blockpraktikums zugeteilt. Die offenen Formulierungen des Bonn Tests und seines entsprechenden Erwartungshorizonts sind mit einer aufwendigeren Korrektur verbunden, die zunächst eine Einarbeitung in die Thematik EbM erfordert. Gleichzeitig stellen die offenen Formulierungen die wohl größte Stärke des Bonn Tests dar, da sie die Evaluation der praktischen Anwendung von EbM erst ermöglichen. Dadurch kann der Bonn Test im Gegensatz zu den bisher gängigen Evaluationsinstrumenten auf jede beliebige Fragestellung angewendet werden.

5. Zusammenfassung

Insgesamt bestätigt die vorliegende Studie, dass die Erweiterung des humanmedizinischen Studiums im Fachbereich für Allgemeinmedizin um eine Lehreinheit zur Evidenzbasierten Medizin machbar ist und die Studenten die erlernten Inhalte erfolgreich in der Praxis anwenden können. Dadurch liefert diese Arbeit einen wichtigen Beitrag zur Ausbildung zukünftiger Ärzte in der Anwendung von EbM, da andere Schulungsmodelle sich bisher auf die Vermittlung theoretischer Grundlagen beschränken. Die Patientenversorgung kann durch eine EbM-basierte Therapie verbessert werden, da wissenschaftliche Neuerungen bisher nur selten oder zeitverzögert Anwendung im klinischen Alltag finden. Das online Lernmodul, zwei tutorengestützte Seminare in Kleingruppen, der Bonn Test und sein Erwartungshorizont stellen ein in sich geschlossenes System dar, mit dem die wichtigsten EbM-Fähigkeiten systematisch im Selbststudium erlernt und anschließend angewendet und evaluiert werden können. Der Bonn Test ist angelehnt an den Fresno Test, ein validiertes Evaluationsinstrument in EbM, aber prüft anstelle des theoretischen Verständnisses die praktische Anwendung. Die Zufriedenheit der Studenten mit der Erweiterung des Blockpraktikums wird mit zusätzlichen Evaluationsbögen erhoben. In drei Gruppendurchläufen gelingt den Studenten jeweils die Bearbeitung eines fiktiven Fallbeispiels und anschließend die Entwicklung einer klinischen Fragestellung auf der Grundlage eines echten Patientenfalls in den Hausarztpraxen. Die erreichten Punktzahlen im Bonn Test für die theoretischen und die praktischen Fallbeispiele erlauben dabei eine gezielte Analyse von Stärken und Schwächen der notwendigen Fähigkeiten. Den ersten Daten zu Folge scheint der Bonn Test dabei ein adäquates Evaluationsinstrument für die praktische Anwendung von EbM darzustellen. Mit Hilfe der ausgewerteten Bonn Tests und studentischen Rückmeldungen kann das Lehrkonzept für zukünftige Einsätze gezielt verbessert werden. Im Rahmen weiterer Studien, vorzugsweise RCTs, sollte die Validität des Tests weiterhin untersucht werden. Der verwendete Lehransatz ist nach Auswertung aller Ergebnisse ein vielversprechendes Modell zur Implementierung von Evidenzbasierter Medizin im Curriculum, das Medizinstudenten gezielt auf die Anwendung von EbM im Praxisalltag vorbereiten kann. Weitere Untersuchungen des Lehrkonzepts, insbesondere zu einem früheren Zeitpunkt während des Medizinstudiums, sind empfehlenswert.

6. Abbildungsverzeichnis

Abbildung 1: Der 5-A-Schritteprozess	10
Abbildung 2: Die „3 Pfeiler“ der EbM	21
Abbildung 3: Rekrutierung von auswertbaren Bonn Tests zu fiktiven Fallbeispielen	36
Abbildung 4: Gesamtpunktzahl der fiktiven Fallbeispiele	37
Abbildung 5: Ergebnisse der unvollständig eingereichten Bonn Tests	38
Abbildung 6: Durchschnittlich erreichte Punktzahl der fiktiven Fallbeispiele A, B und C in den Kategorien des 5-A-Schritteprozesses	39
Abbildung 7 A-D: Hierarchische Darstellung der Leistungen in den Kategorien des 5-A-Schritteprozesses für die fiktiven Fallbeispiele	40
Abbildung 8 A: Ergebnisse aller fiktiven Fallbeiträge in der Kategorie Ask	41
Abbildung 8 B: Punkteverteilung der Unterkategorien Patient, Intervention, Comparison und Outcome	42
Abbildung 8 C: Punkteverteilung der Kategorie Ask in Fall A, B und C	42
Abbildung 9 A: Ergebnisse aller fiktiven Fallbeiträge in der Kategorie Acquire – Suchstrategie	43
Abbildung 9 B: Punkteverteilung der Unterkategorien Wortwahl, MeSH und Präzisierende Optionen	43
Abbildung 9 C: Punkteverteilung der Unterkategorie Präzisierende Optionen in Fall A, B und C	43
Abbildung 10 A: Ergebnisse aller fiktiven Fallbeiträge in der Kategorie Acquire – Auswahl des Abstracts	44

Abbildung 10 B: Punkteverteilung der Unterkategorien Inhalt und Qualität	44
Abbildung 10 C: Punkteverteilung der Kategorie Acquire – Auswahl des Abstracts für Fall A im Vergleich zu allen Beiträgen	45
Abbildung 11 A: Ergebnisse aller fiktiven Fallbeiträge in der Kategorie Appraise	46
Abbildung 11 B: Punkteverteilung der Unterkategorien Validität, Größe und Präzision und Übertragbarkeit	46
Abbildung 11 C: Punkteverteilung der Kategorie Appraise in Fall A, B und C	46
Abbildung 12 A: Ergebnisse aller fiktiven Fallbeiträge in der Kategorie Apply	47
Abbildung 12 B: Punkteverteilung der Unterkategorien interne Validität und externe Validität	48
Abbildung 12 C: Punkteverteilung der externen Validität für Fall A, B und C	48
Abbildung 13 A: Ergebnisse aller fiktiven Fallbeiträge in der Kategorie Assess	49
Abbildung 13 B: Punkteverteilung der Kategorie Assess für Fall C im Vergleich zu allen Beiträgen	49
Abbildung 14: Vergleich der Ergebnisse in fiktiven und realen Fallbeispielen	50
Abbildung 15 A-B: Hierarchische Darstellung der mittleren erreichten Punktzahl in allen Kategorien für die fiktiven und die realen Fälle	50
Abbildung 16: Ergebnisse in allen Kategorien und Unterkategorien für die fiktiven und realen Fälle im Vergleich	51
Abbildung 17 A: Mittelwert und Punkteverteilung für die studentische Zufriedenheit mit dem online Lernmodul statt Frontalunterricht	53
Abbildung 17 B-D: Mittelwert und Punkteverteilung für die studentische Zufriedenheit mit der inhaltlichen (B), didaktischen (C) und technischen (D) Umsetzung	54

Abbildung 17 E: Mittelwert und Punkteverteilung für die studentische Einschätzung der Verbesserung des Grundlagenwissens	55
Abbildung 17 F-G: Mittelwert und Punkteverteilung für die studentische Einschätzung der Anwendung von EbM-Fähigkeiten (F) und des Nutzens des Bonn Tests (G)	56
Abbildung 18 A: Mittelwert und Punkteverteilung für die studentische Einschätzung des Mehrgewinns durch die Seminare	57
Abbildung 18 B und C: Mittelwert und Punkteverteilung für die studentische Zufriedenheit mit Tutoren (B) und Kleingruppen (C)	58
Abbildung 19 A-B: Mittelwert und Punkteverteilung für die studentische Einschätzung der Entwicklung eines Fallbeispiels (A) und eines Patientenprofils (B)	59
Abbildung 19 C-D: Mittelwert und Punkteverteilung für die studentische Einschätzung der Erhebung von Patientendaten (C) und der Literaturrecherche (D)	60
Abbildung 19 E-F: Mittelwert und Punkteverteilung für die studentische Einschätzung der Gelegenheit (E) für und des Mehrgewinnes (F) durch das Gespräch mit dem Lehrarzt	61
Abbildung 20: Mittelwert und Punkteverteilung für die Gesamtnote	62

7. Tabellenverzeichnis

Tabelle 1: Ablauf des EbM-Lehrkonzepts	19
Tabelle 2: Auflistung der Themen für selbstgewählte Fragestellungen	52
Tabelle 3: Verbesserungsvorschläge für das Lehrkonzept	75

8. Literaturverzeichnis

Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day. *PLoS medicine* 2010; 7: e1000326

Bero L. The Cochrane Collaboration. *JAMA* 1995; 274: 1935

Bethesda M D. PubMed Help. 2005; Verfügbar unter <https://www.ncbi.nlm.nih.gov/books/NBK3827/>

Beyer M, Bergold M, Donner-Banzhoff N, Falck-Ytter Y, Gensichen J, Gerlach F M, Koncczny N, Lelgemann M, Lümann D, Ochsendorf F, Schulze J, Weberschock T. Curriculum Evidenzbasierte Medizin im Studium. 2005; Verfügbar unter <http://www.ebm-netzwerk.de/pdf/curricula-zertifizierung/curriculum-ebm-im-studium.pdf> (Zugriffsdatum: 02.11.2017)

Blümle A, Meerpohl JJ, Wolff R, Antes G. Evidenzbasierte Medizin und systematische Übersichtsarbeiten. *MKG-Chirurg* 2009; 2: 86–92

Booth A. EBLIP five-point-zero. *Health information and libraries journal* 2009; 26: 341–344

Cardarelli R, Seater MM. Evidence-based medicine, part 4. An introduction to critical appraisal of articles on harm. *The Journal of the American Osteopathic Association* 2007; 107: 310–314

Cardarelli R, Virgilio RF, Taylor L. Evidence-based medicine, part 2. An introduction to critical appraisal of articles on therapy. *The Journal of the American Osteopathic Association* 2007; 107: 299–303

Cochrane Deutschland, Hrsg. Systematische Übersichtsarbeiten der Cochrane Library. 2016; Verfügbar unter <http://www.cochrane.de/de/systematische-uebersichtsarbeiten> (Zugriffsdatum: 01.06.2016)

Cochrane Deutschland, Hrsg. Von der Evidenz zur Empfehlung (Klassifikationssysteme). Verfügbar unter www.cochrane.de/de/evidenz-empfehlung#contact (Zugriffsdatum: 08.04.2016)

Cook DJ, Jaeschke R, Guyatt GH. Critical appraisal of therapeutic interventions in the intensive care unit. *Journal of intensive care medicine* 1992; 7: 275–282

Deutsches Netzwerk Evidenzbasierte Medizin e.V. (DNEbM e.V.), Hrsg. Definitionen. 2011; Verfügbar unter: www.ebm-netzwerk.de/was-ist-ebm/grundbegriffe/definitionen (Zugriff am 01.01.2019)

Duke University Medical Center Library, Health Sciences Library at the University of North Carolina at Chapel Hill, Hrsg. Introduction to Evidence-Based Practice tutorial. 2014; Verfügbar unter <http://guides.mclibrary.duke.edu/ebmtutorial> (Zugriffsdatum: 02.11.2017)

Du Prel J-B, Röhrig B, Blettner M. Critical appraisal of scientific articles. Deutsches Arzteblatt international 2009; 106: 100–105

Du Prel J-B, Röhrig B, Hommel G, Blettner M. Choosing statistical tests. Deutsches Arzteblatt international 2010; 107: 343–348

Du Prel JB, Röhrig B, Blettner M. Statistical Methods in Medical Research. Deutsches Arzteblatt international 2009; 106: 99

Engel B, Prautzsch H. Häufige Gichtanfalle und Chronische Gicht in der hausärztlichen Versorgung. Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin, Hrsg. 2013

Explanation of the 2011 Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence (Background Document). Verfügbar unter <http://www.cebm.net/index.aspx?o=5653> (Zugriffsdatum: 07.11.2017)

Francke AL, Smit MC, Veer AJE de, Mistiaen P. Factors influencing the implementation of clinical guidelines for health care professionals. BMC medical informatics and decision making 2008; 8: 38

Fritsche L. Do short courses in evidence based medicine improve knowledge and skills? BMJ (Clinical research ed.) 2002; 325: 1338–1341

General Medical Council, Hrsg. Outcomes for graduates. Tomorrow's Doctors 2015

Greenhalgh T, Howick J, Maskrey N. Evidence based medicine. BMJ (Clinical research ed.) 2014; 348: g3725

Guyatt G, Cairns J, Churchill D. Evidence-based medicine. A new approach to teaching the practice of medicine. JAMA 1992; 268: 2420–2425

Heneghan C. Reflections on David Sackett's time at the Centre for Evidence-Based Medicine. 2015; Verfügbar unter <http://www.cebm.net/david-sackett-and-the-centre-for-evidence-based-medicine/> (Zugriffsdatum: 02.11.2017)

Hill J, Bullock I, Alderson P. A summary of the methods that the National Clinical Guideline Centre uses to produce clinical guidelines for the National Institute for Health and Clinical Excellence. Annals of internal medicine 2011; 154: 752–757

Howick J, Chalmers I, Glasziou P, Greenhalgh T, Heneghan C, Liberati A, Moschetti I, Phillips B, Thornton H. The 2011 Oxford CEBM Levels of Evidence (Introductory Document). Verfügbar unter <http://www.cebm.net/index.aspx?o=5653> (Zugriffsdatum: 07.11.2017)

Ilic D, Nordin RB, Glasziou P, Tilson JK, Villanueva E. A randomised controlled trial of a blended learning education intervention for teaching evidence-based medicine. *BMC medical education* 2015; 15: 39

Kabisch M, Ruckes C, Seibert-Grafe M, Blettner M. Randomized controlled trials. *Deutsches Arzteblatt international* 2011; 108: 663–668

Lai NM, Teng CL, Nalliah S. Assessing Undergraduate Competence in Evidencebased Medicine: A Preliminary Study on the Correlation Between Two Objective Instruments. *Educ Health* 2012; 25: 33–39

Lecher S, Satzinger W, Trojan A, Koch U. Patientenorientierung durch Patientenbefragungen als ein Qualitätsmerkmal der Krankenversorgung. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 2002; 45: 3–12

Madea B, Preuß J, Mußhoff F, Dettmeyer R. Behandlungsfehlervorwürfe bei Arzneimitteltherapie - Gutachterliche Aspekte, Kausalität. 2006; 77–99

Moßhammer D, Roos MJ, Kronenthaler A, Lorenz G, Eissler M, Joos S. Students' performing of practical research tasks for their scientific qualification - an approach within the family practice internship in undergraduate education. *GMS Zeitschrift für medizinische Ausbildung* 2011; 28: Doc24

Patel MR, Schardt CM, Sanders LL, Keitz SA. Randomized trial for answers to clinical questions. *Journal of the Medical Library Association: JMLA* 2006; 94: 382–387

Pope C. Resisting evidence: the study of evidence-based medicine as a contemporary social movement. *Health* 2003; 7: 267–282

PUBMED.DE, Hrsg. Literatursuche, Wissenswertes für die Literatursuche im WWW. Verfügbar unter <https://www.pubmed.de/literatursuche/literatursuche/> (Zugriffsdatum: 02.11.2017)

Ramos KD, Schafer S, Tracz SM. Validation of the Fresno test of competence in evidence based medicine. *British Medical Journal* 2003; 326: 319–321

Ressing M, Blettner M, Klug SJ. Systematic literature reviews and meta-analyses. *Deutsches Arzteblatt international* 2009; 106: 456–463

Röhrig B, Du Prel J-B, Blettner M. Study design in medical research. *Deutsches Arzteblatt international* 2009; 106: 184–189

Röhrig B, Du Prel J-B, Wachtlin D, Kwiecien R, Blettner M. Sample size calculation in clinical trials. *Deutsches Arzteblatt international* 2010; 107: 552–556

Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine. *BMJ (Clinical research ed.)* 1996; 312: 71–72

Schilling K, Wiecha J, Polineni D, Khalil S. An interactive web-based curriculum on evidence-based medicine. *Family medicine* 2006; 38: 126–132

Schranz DA, Dunn MA. Evidence-based medicine, part 3. An introduction to critical appraisal of articles on diagnosis. *The Journal of the American Osteopathic Association* 2007; 107: 304–309

Thomas RE, Kreptul D. Systematic review of evidence-based medicine tests for family physician residents. *Family medicine* 2015; 47: 101–117

Universitätsklinikum Freiburg / Deutsches Cochrane Zentrum, Hrsg. *Manual Systematische Literaturrecherche für die Erstellung von Leitlinien*. 2013

U.S. National Library of Medicine, Hrsg. *PubMed Tutorial*. 2001; Verfügbar unter https://www.nlm.nih.gov/bsd/disted/pubmedtutorial/015_010.html (Zugriffsdatum: 02.11.2017)

Virgilio RF, Chiapa AL, Palmarozzi EA. Evidence-based medicine, part 1. An introduction to creating an answerable question and searching the evidence. *The Journal of the American Osteopathic Association* 2007; 107: 295–297

Universität Bonn, Hrsg. *Vorlesungsverzeichnis*. 2018; Verfügbar unter https://basis.uni-bonn.de/qisserver/rds?state=wsearchv&search=1&subdir=veranstaltung&veranstaltung.dtxt=Evidenzbasiert&veranstaltung.semester=20172&P_start=0&P_anzahl=10&P.sort=&_form=display (Zugriffsdatum: 27.02.2018)

Westfall JM, Mold J, Fagnan L. Practice-based research--"Blue Highways" on the NIH roadmap. *JAMA* 2007; 297: 403–406

9. Anlagen

9.1 Anlage I: Druckversion des Online Lernmoduls

Evidenzbasierte Medizin (EbM)

Was ist EbM?

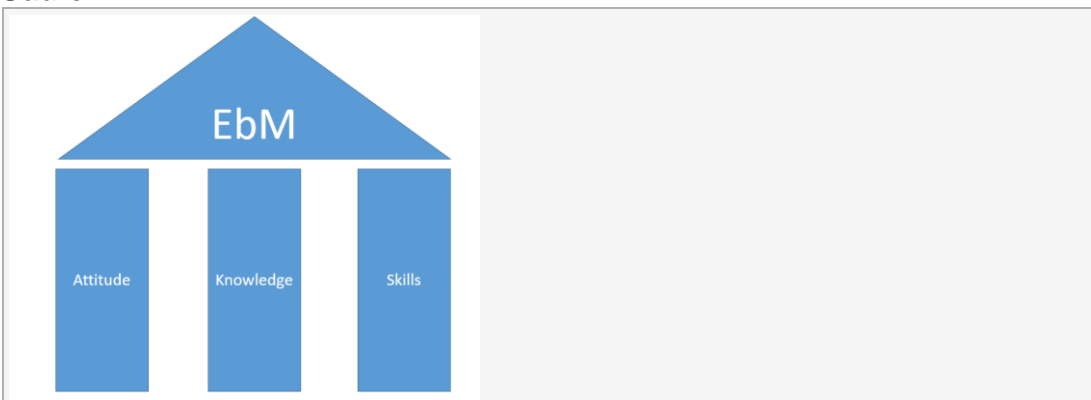
Definition:

„Evidenzbasierte Medizin ist die gewissenhafte, explizite und umsichtige Anwendung der besten verfügbaren Evidenz um Behandlungsentscheidungen zum Wohle des individuellen Patienten zu treffen.“



David L. Sackett (Gründungsdirektor des Centre for Evidence-Based Medicine des National Health Service an der Universität Oxford)

EbM Säulen



EbM wird von drei Säulen getragen. Das Patientenwohl hat in allen Kategorien höchste Priorität!

Attitude

Die Bereitschaft des Arztes sich ein Leben lang fortzubilden

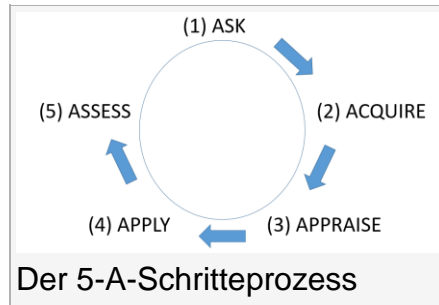
Knowledge

- „Interne Evidenz“ = durch das Medizinstudium erworbenes Fachwissen
- „Externe Evidenz“ = Erkenntnisse aus der Literaturrecherche (Muss vor der Anwendung evaluiert werden!)

Skills

Der 5-A-Schritte-Prozess fasst alle wichtigen Fähigkeiten für die Anwendung von EbM zusammen.

Der 5-A-Schritteprozess



Ask = Formulierung einer wissenschaftlichen Frage

Acquire = Literaturrecherche

Appraise = Kritische Evaluation der Rechercheergebnisse

Apply = Auswahl der Behandlung

Assess = Evaluation der Behandlung

Wählen Sie unter den folgenden Möglichkeiten die „Pfeiler“ der EbM aus.

- Examination
- Attitude
- Research
- Knowledge
- Skills
- Auswerten

Ordnen Sie den einzelnen Schritten des 5-A-Schritteprozesses jeweils die richtige Definition zu.

Positionen der Terme zurücksetzen

Appraise	Auswahl der Behandlung
Ask	Formulierung einer wissenschaftlichen Frage
Assess	Evaluation der Rechercheergebnisse
Apply	Evaluation der Behandlung
Acquire	Literaturrecherche

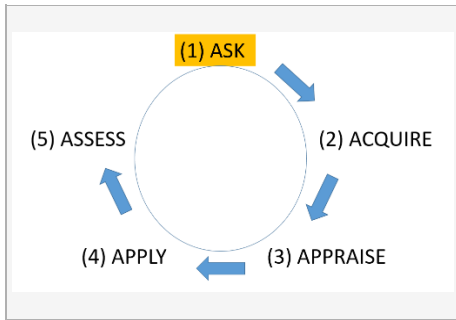
Auswerten

Bringen Sie die einzelnen Punkte des 5-A-Schritteprozesses jeweils in die richtige Reihenfolge.

Assess Apply Acquire Appraise Ask

Auswerten

Ask



Eine wissenschaftliche Fragestellung sollte mit Hilfe des PICO-Schemas gestellt werden:

P = „Patient Problem“: Relevante Merkmale des Patienten (z. B. Haupterkrankung und Begleiterkrankungen, Alter, Geschlecht oder ethnische Herkunft)

I = „Intervention“: Angestrebte Therapie (z. B. Medikamente, OP)

C = „Comparison“: Alternative Behandlung

O = „Outcome“: Erwünschte Wirkung

Fallbeispiel

Die 42-jährige Sekretärin Frau Schiller klagt über immer wiederkehrende Schmerzen im unteren Rücken. Vor lauter Überstunden komme sie zurzeit nicht mehr zum Sport, aber in der „Bunten“ habe sie gelesen, dass Akupunktur ja auch ein echter Geheimtipp sei und viel weniger Zeit in Anspruch nehme.

P = „Patient Problem“: Weibliche 42-jährige Patientin mit häufig sitzender Tätigkeit und chronischen Schmerzen im Lendenwirbelbereich

I = „Intervention“: Akupunktur

C = „Comparison“: Sport

O = „Outcome“:

- Primärer Endpunkt: Schmerzreduktion
- Sekundärer Endpunkt: Zeitersparnis

Wofür steht das „P“ im „PICO-Schema“?

Process
 Personality
 Patient Problem
 Privatpatient
 Progress

?

Wofür steht das „I“ im „PICO-Schema“?

Intervention
 Illness
 Income
 Integration
 Intention

Auswerten

Wofür steht das „C“ im „PICO-Schema“ zur Formulierung einer wissenschaftlichen Fragestellung?

Calculation
 Comparison
 Compromise
 Compliance
 Cofactor

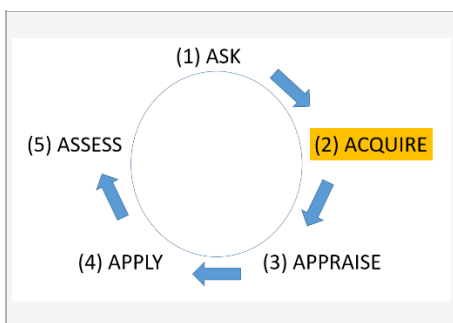
Auswerten

Wofür steht das „O“ im „PICO-Schema“?

Other treatment
 Outcome
 Off-Label
 Order
 Origin

Auswerten

Acquire

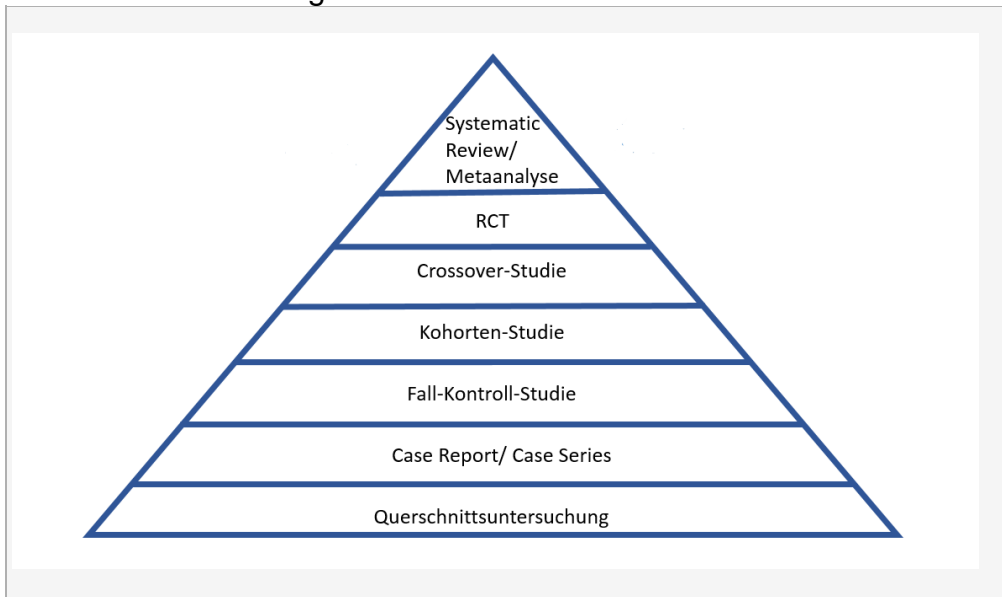


Eine systematische Literaturrecherche sollte entsprechend der folgenden Schritte ablaufen:

1. Auswahl von Studientyp und Recherchequellen
2. Festlegung von Suchvokabular und –strategie, Durchführung der ersten Suche
3. Sichtung der Ergebnisse (und ggf. angepasste zweite Suche)
4. Dokumentation der Recherche

Studientypen

Es ist wichtig die unterschiedlichen Studientypen und ihre Anfälligkeit für systematische Fehler und Verzerrungen zu kennen.



Die Literaturrecherche konzentriert sich zunächst auf die Spitze der Pyramide und wird jeweils auf die darunter liegende Stufe ausgeweitet falls sich keine Evidenz findet.

Zur Erinnerung:

Systematic Review = Inhaltliche Zusammenfassung und Evaluation methodisch hochwertiger Studien (z. B. RCTs) durch Experten

Metaanalyse = Statistische Zusammenführung der Ergebnisse verschiedener hochwertiger Studien

RCT (Randomized Controlled Trial) = Vergleich einer Interventionsgruppe mit einer Kontrollgruppe und Minimierung systematischer Fehler durch Randomisierung und Verblindung

CAVE: Goldstandard zur Überprüfung der Wirksamkeit und Sicherheit einer Therapie, dennoch muss die externe Evidenz für jeden Fall kritisch hinterfragt werden

Crossover-Studie = Jeder Patient erhält zu unterschiedlichen Zeitpunkten jeweils ein Medikament oder ein Placebo und dient so als eigene Kontrolle

CAVE: Zu kurz gewählte Auswaschphase verursacht Überhang-Effekt („Carryover“)

Kohorten-Studie = Prospektiver Vergleich einer Gruppe, die einem bestimmten Reiz ausgesetzt ist, mit einer Vergleichsgruppe, die diesem Faktor nicht ausgesetzt ist

CAVE: Andere Faktoren als die gesuchte Variable können unterschiedliche Ausgänge bei Interventions- und Kontrollgruppe provozieren

Fall-Kontroll-Studie = Retrospektiver Vergleich von Betroffenen und Nicht-Betroffenen, um mögliche Auslöser für eine Krankheit zu identifizieren

CAVE: - Statistische Zusammenhänge allein beweisen keine Kausalität

- *Fehlende oder fehlerhafte Erinnerungen nehmen starken Einfluss*

Case Report/ Case Series = Deskriptive Darstellung einzelner Fallberichte bzw. Sammlungen ähnlicher Fälle

CAVE: Geringe statistische Validität

Querschnittsuntersuchung = Beziehung zwischen einer Krankheit und anderen Faktoren zu einem bestimmten Zeitpunkt

CAVE: Geeignet für z. B. Diagnostik Tests, aber ungeeignet für vergleichende Studien z. B. zu Therapieformen

Recherchequellen

Im Folgenden werden einige der wichtigsten Recherchequellen vorgestellt.

Häufig werden aus Zeitgründen vorab evaluierte Quellen wie die Cochrane Library (<http://www.cochranelibrary.com/>) genutzt. Zum besseren Verständnis des 5-A-Schritte-Prozesses wird im weiteren Verlauf dieses Lernmoduls die Literaturrecherche mit PubMed (<https://www.pubmed.de/gateway/nlm-pubmed/>) durchgeführt.

Recherchequellen	Relevanz	Validität
Literaturdatenbanken (z.B. MEDLINE, Embase)	Die beiden größten Datenbanken für biomedizinische Literatur (umfassen alle Studientypen)	Texte müssen durch den Leser evaluiert werden
Cochrane Library	Spezialisierung auf Systematic Reviews (Inhaltliche Zusammenfassung der Studienergebnisse aus randomisierten, kontrollierten Studien (RCT))	Vorabevaluation durch Experten
EBP-Quellen („Pre-appraised Resources“ z.B. ACP Journal Club, Clinical Evidence, UpToDate)	Zusammenfassung klinisch hochwertiger Studien	Vorabevaluation durch Experten
Leitliniendatenbanken (z.B. AWMF, NGC) oder Leitlinien-Erstellern (z.B. NICE, SIGN)	Besonders relevant für Fragen zu Therapiemaßnahmen	Vorabevaluation durch Experten

Ordnen Sie die folgenden Studientypen entsprechend ihrer Empfindlichkeit für systematische Fehler und Verzerrungen. Nennen Sie den empfindlichsten Studientyp zuerst.

SystematicReview Fall-Kontroll-Studie Crossover-Studie KohortenStudie RCT

Auswerten

Wählen Sie für die beiden folgenden Definitionen jeweils den richtigen Studientypen aus.

a) Zusammenfassung und Evaluation methodisch hochwertiger Studien

b) Jede Testperson erhält zu unterschiedlichen Zeitpunkten jeweils ein Medikament oder ein Placebo und dient so als eigene Kontrolle

Crossover-Studie
 Kohorten Studie
 Querschnittsstudie
 RCT
 Systematic Review

Auswerten

Bei welchem Studientyp verursacht eine zu kurz gewählte Auswaschphase einen Überhang-Effekt?

Systematic Review
 Querschnittsstudie
 Crossover-Studie
 Kohorten Studie
 Case Report

Auswerten

Suchstrategien für MEDLINE & PubMed

Die Recherche sollte in drei Schritten erfolgen:

- 1) Umwandlung der PICO-Elemente in eine Suchanfrage
- 2) Berücksichtigung von „Medical Subject Headings“ (Schlagwortvokabular)
- 3) Ggf. Verwendung von Filtern und Operatoren

Vom PICO-Schema zur Suchanfrage

Bei der Umwandlung der PICO-Elemente in eine Suchanfrage bei PubMed ist sowohl der Inhalt als auch die Wortwahl entscheidend.

Zur Erinnerung:

P = „Patient Problem“: Weibliche 42-jährige Patientin mit häufig sitzender Tätigkeit und chronischen Schmerzen im Lendenwirbelbereich

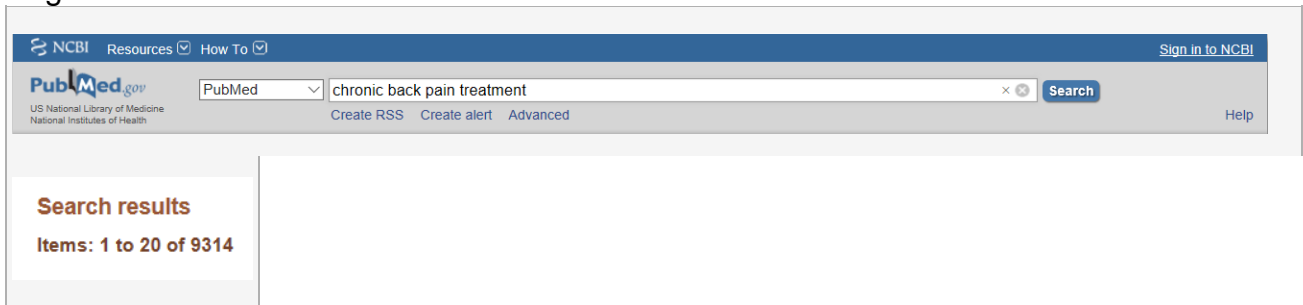
I = „Intervention“: Akupunktur

C = „Comparison“: Sport

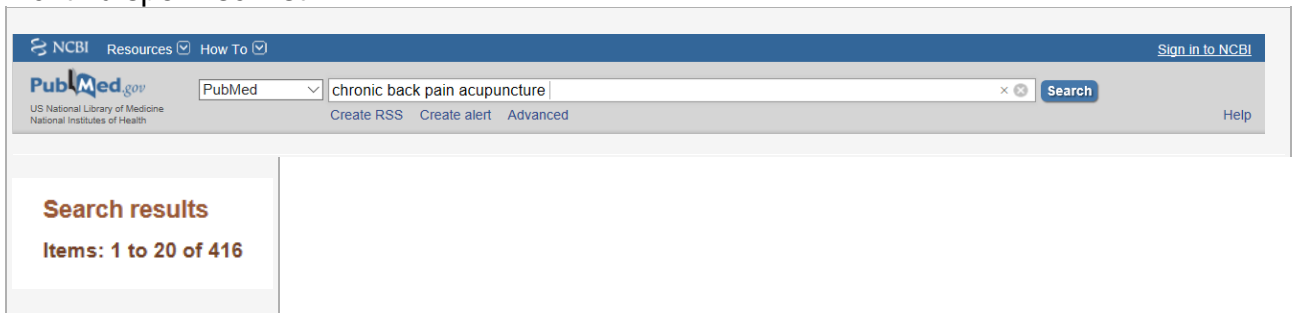
O = „Outcome“:

- Primärer Endpunkt: Schmerzreduktion
- Sekundärer Endpunkt: Zeitersparnis

Die Suchsprache ist bei PubMed wie in den meisten anderen online Recherchequellen Englisch.



Bei der Umwandlung einer PICO-Frage in eine PubMed-Suchanfrage muss abgewogen werden, welche Informationen übernommen werden sollen, damit die Suche präzise aber nicht zu spezifisch ist.



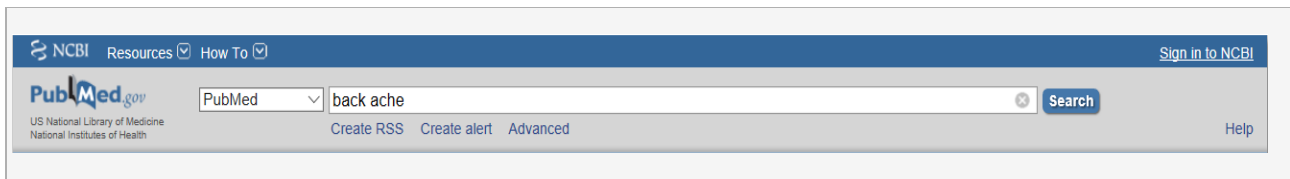
Berücksichtigung von „Medical Subject Headings“

In PubMed kann nach Stichwörtern (engl.: keyword) oder Schlagwörtern (engl.: descriptor) gesucht werden.

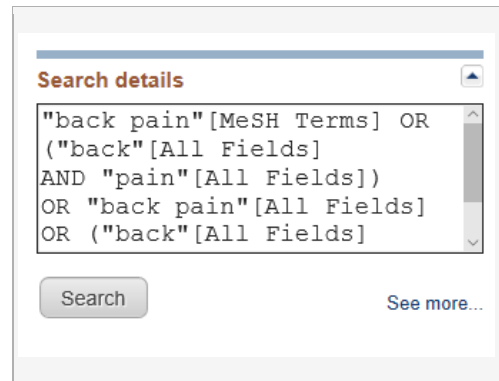
Stichwörter kommen in Titel oder Abstract vor und sind besonders relevant für die Suche nach aktuellen Artikeln

Schlagwörter fassen den Inhalt eines Artikels zusammen und werden automatisch mit Synonymen verknüpft, sodass unterschiedliche Wörter mit derselben Bedeutung auch zum selben Suchergebnis führen

„**Medical Subject Headings**“ (**MeSH**) ist ein von der U.S.-amerikanischen National Library of Medicine entwickeltes Schlagwortvokabular. Es ist hierarchisch unterteilt in 16 Hauptkategorien und untergeordnete Begriffe. Zusätzlich können „Subheadings“ vergeben werden, um Aspekte näher zu beschreiben.

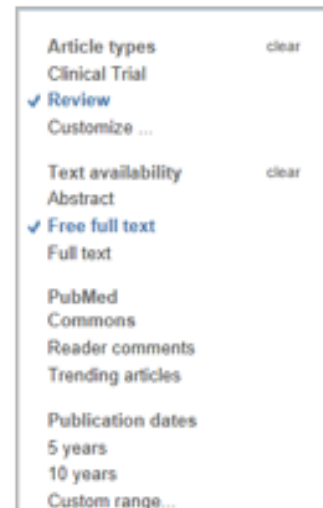
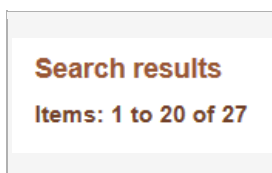
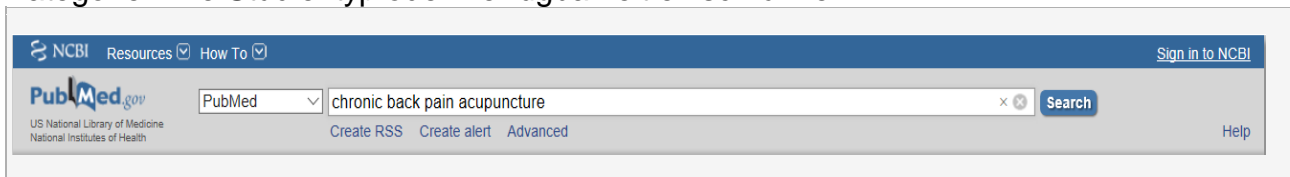


Die Suche nach Stichwörtern wie „back ache“ wird auf PubMed häufig automatisch mit den entsprechenden MeSH Terms wie „back pain“ verknüpft. Details können in der rechten Spalte unter „Search details“ eingesehen werden.



Filter

Filter können über die Filterleiste ausgewählt werden und Ergebnisse in verschiedenen Kategorien wie Studientyp oder Verfügbarkeit einschränken.



Operatoren

Erzielt die Suchanfrage zu viele Treffer, kann das Ergebnis eingeschränkt werden, indem Suchbegriffe durch spezifischere Alternativen ersetzt und zusätzliche Begriffe oder Operatoren verwendet werden. Operatoren verknüpfen die Suchbegriffe von links nach rechts, es sei denn diese automatische Reihenfolge wird durch das Setzen von Klammern verändert.

Die wichtigsten Operatoren sind:

- AND (Suche nach beiden Begriffen)
- OR (Suche nach mindestens einem der Begriffe)
- NOT (Schließt den zweiten Begriff aus der Suche aus)

The screenshot shows the PubMed search interface. The search bar contains the query "chronic AND back pain NOT prolapse". The interface includes the NCBI logo, "Resources" and "How To" dropdowns, a "Sign in to NCBI" link, and a "Search" button. Below the search bar, there are links for "Create RSS", "Create alert", and "Advanced".

Anschließend werden die Überschriften überflogen und das zu der Fragestellung am besten passende Abstract ausgewählt.

The screenshot shows the PubMed search results page for the query "chronic back pain acupuncture". The search bar contains the query "chronic back pain acupuncture". The interface includes the NCBI logo, "Resources" and "How To" dropdowns, a "Sign in to NCBI" link, and a "Search" button. Below the search bar, there are links for "Create RSS", "Create alert", and "Advanced".

The search results are displayed in a list format. The first result is "Noninvasive Treatments for Low Back Pain [Internet]". The second result is "Acupuncture for low back pain: an overview of systematic reviews.". The third result is "Traditional Chinese medicine for neck pain and low back pain: a systematic review and meta-analysis.". The search results are filtered by "Review" and "Free full text". The search details show the query: "(chronic[All Fields] AND ("back pain"[MeSH Terms] OR ("back"[All Fields] AND "pain"[All Fields])) OR "back pain"[All Fields])".

Dokumentation

Das Suchergebnis wird in Zitatform dokumentiert:

Autor/in (jeweils mit ausgeschriebenem Nachnamen und Anfangsbuchstabe des Vornamens). Titel. Verlag. Jahr: Seitenzahl bzw. Internetpfad (ggf mit Zugriffsdatum: xx.xx.xxxx)


Liu L, Skinner M, McDonough S, Mabire L, Baxter GD. Acupuncture for low back pain: an overview of systematic reviews. Evid Based Complement Alternat Med. 2015; 2015:328196. doi: 10.1155/2015/328196

Der folgende Link bietet optional weitere Informationen zum Umgang mit PubMed:

http://www.ncbi.nlm.nih.gov/books/NBK3827/#pubmedhelp.How_do_I_search_PubMed


Welche Schritte sollten bei der Recherche (Acquire) berücksichtigt werden?

- Umwandlung der PICO-Elemente in eine Suchanfrage
- Ggf. Verwendung von Filtern und Operatoren
- Berücksichtigung von „Medical Subject Headings“ (Schlagwortvokabular)
- Länge des Artikels
- Bekanntheitsgrad der Autoren




Welche der folgenden Aussagen treffen zu?

- "MeSH" steht für "Medical Subject Headings"
- Schlagwörter kommen in Titel oder Abstract vor
- "MeSH" ist ein von der U.S.-amerikanischen National Library of Medicine entwickeltes Schlagwortvokabular
- Stichwörter sind besonders relevant für der Suche nach aktuellen Artikeln, die noch nicht verschlagwortet worden sind
- Stichwörter werden automatisch mit Synonymen verknüpft

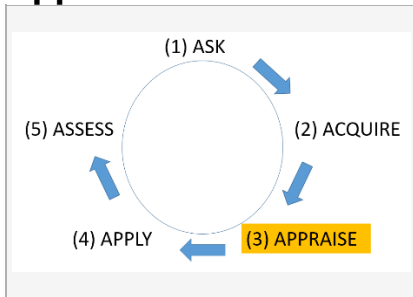


Wählen Sie die wichtigsten Operatoren aus.

- OR
- WHY
- AND
- NOT
- NEITHER



Appraise



Die Evaluation einer Studie sollte systematisch nach 3 Aspekten erfolgen:

1. Validität
2. Größe und Präzision
3. Übertragbarkeit & Anwendbarkeit

Validität

Die Validität der Rechercheergebnisse ist abhängig vom Studientypen. Je hochwertiger die Studie, desto höher ist die entsprechende Evidenzstufe.

Evidenzstufe	Validitätskriterien
Ia)	Systematic Review, das auf mindestens einer randomisierten kontrollierten Studie (RCT) basiert
Ib)	Randomisierte kontrollierte Studie (RCT)
IIa)	Hochwertige Studie ohne Randomisierung (z.B. Crossover-Studie)
IIb)	Hochwertige Studie eines anderen Typs quasi experimenteller Studien (z.B. Kohorten-Studie)
III)	Mehr als eine methodisch hochwertige nichtexperimentelle Studie (z.B. Fall-Kontroll-Studie)
IV)	Expertenmeinung oder beschreibende Studien (z.B. Case Report/ Case Series)

CAVE: Ein hochwertiges Studienmodell kann inhaltliche Mängel aufweisen oder schlecht durchgeführt worden sein. Dies muss in der Evaluation berücksichtigt werden.

Größe und Präzision

Die Beurteilung einer wissenschaftlichen Veröffentlichung erfolgt in drei Kategorien.

Kriterien
Design (z.B. Studientyp, Fragestellung, Messverfahren, Beobachtungseinheit, Fallzahlplanung, Studienpopulation)
Durchführung (z.B. festgelegter Studienablauf, ggf. Abweichungen vom ursprünglich geplanten Ablauf, adäquate Verblindung oder Randomisierung)
Auswertung (z.B. nachvollziehbare Schlussfolgerungen aus gewonnenen Studienergebnissen, Objektivität, Berücksichtigung systematischer Fehler, Übereinstimmen mit anderen Studien)

Übertragbarkeit und Anwendbarkeit

Kriterien

Neben der internen Validität sollte auch die externe Validität der Studie, also die **Übertragbarkeit der Ergebnisse von der Studienpopulation auf die Zielpopulation** kritisch hinterfragt werden.

Die **Kosten-Nutzen-Relation** für die Behandlung ist besonders in Hinblick auf die individuellen Endpunkte des Patienten abzuwägen.

Korrekte Verwendung von **Literaturangaben** oder etwaige **Interessenkonflikte** der Autoren

Beispiel für die Untersuchung eines Abstracts auf die Studienqualität (Appraise)

The screenshot shows the PubMed interface for the abstract 'Acupuncture for low back pain: an overview of systematic reviews'. The page includes the following elements:

- Header:** NCBI logo, 'Resources' and 'How To' dropdowns, 'Sign in to NCBI' link, 'PubMed' search bar, and 'Advanced' search options.
- Notification:** A blue banner stating: 'NCBI will be testing https on public web servers from 10:00-11:00 AM EDT (14:00-15:00 UTC) on Thursday, October 13. You may experience problems with NCBI services during that time. Please plan accordingly. [Read more.](#)'
- Format:** 'Abstract' dropdown menu.
- Comment:** 'See comment in PubMed Commons below'.
- Citation:** 'Evid Based Complement Alternat Med. 2015;2015:328196. doi: 10.1155/2015/328196. Epub 2015 Mar 4.'
- Title:** 'Acupuncture for low back pain: an overview of systematic reviews.'
- Authors:** 'Liu L¹, Skinner M¹, McDonough S², Mabire L¹, Baxter GD¹.'
- Author information:** Expandable section.
- Abstract:**

Objective. As evidence of the effectiveness of acupuncture for low back pain (LBP) is inconsistent, we aimed to critically appraise the evidence from relevant systematic reviews. **Methods.** Systematic reviews of randomized controlled trials (RCTs) concerning acupuncture and LBP were searched in seven databases. Internal validity and external validity of systematic reviews were assessed. Systematic reviews were categorized and high quality reviews assigned greater weightings. Conclusions were generated from a narrative synthesis of the outcomes of subgroup comparisons. **Results.** Sixteen systematic reviews were appraised. Overall, the methodological quality was low and external validity weak. For acute LBP, evidence that acupuncture has a more favorable effect than sham acupuncture in relieving pain was inconsistent; it had a similar effect on improving function. For chronic LBP, evidence consistently demonstrated that acupuncture provides short-term clinically relevant benefits for pain relief and functional improvement compared with no treatment or acupuncture plus another conventional intervention. **Conclusion.** Systematic reviews of variable quality showed that acupuncture, either used in isolation or as an adjunct to conventional therapy, provides short-term improvements in pain and function for chronic LBP. More efforts are needed to improve both internal and external validity of systematic reviews and RCTs in this area.
- PMID:** [25821485](#) **PMCID:** [PMC4364128](#) **DOI:** [10.1155/2015/328196](#)
- Links:** '[PubMed] Free PMC Article' and social media icons (Facebook, Twitter, LinkedIn).
- Right sidebar:**
 - Send to:** Dropdown menu.
 - Full text links:** 'FREE Full text article at Hindawi' and 'PMC Full text'.
 - Save items:** 'Add to Favorites' button.
 - Similar articles:**
 - Review:** Effectiveness of acupuncture for low back pain: a system [Spine (Phila Pa 1976). 2008]
 - Review:** Advice for the management of low back pain: a systematic review of ran [Man Ther. 2007]
 - Review:** Is immediate imaging important in managing low back pain? [J Athl Train. 2011]
 - Review:** Efficacy of spinal manipulation and mobilization for low back pain and [Spine J. 2004]
 - Review:** Clinical critical qualitative evaluation of the selected rand [Zhong Xi Yi Jie He Xue Bao. 2010]
 - Actions:** 'See reviews...' and 'See all...' links.

NCBI Resources How To Sign in to NCBI

PubMed US National Library of Medicine National Institutes of Health

Advanced Search Help

NCBI will be testing https on public web servers from 10:00-11:00 AM EDT (14:00-15:00 UTC) on Thursday, October 13. You may experience problems with NCBI services during that time. Please plan accordingly. [Read more.](#)

Format Abstract Send to

See comment in PubMed Commons below

Evid Based Complement Alternat Med. 2015;2015:328196. doi: 10.1155/2015/328196. Epub 2015 Mar 4.

Acupuncture for low back pain: an overview of systematic reviews.

Liu L¹, Skinner M¹, McDonough S², Mabire L¹, Baxter GD¹.

Author information

Abstract

Validität: Evidenzstufe Ia

Objective. As evidence of the effectiveness of acupuncture for low back pain (LBP) is inconsistent, we aimed to critically appraise the evidence from relevant systematic reviews. Methods. Systematic reviews of randomized controlled trials (RCTs) concerning acupuncture and LBP were searched in seven databases. Internal validity and external validity of systematic reviews were assessed. Systematic reviews were categorized and high quality reviews assigned greater weightings. Conclusions were generated from a narrative synthesis of the outcomes of subgroup comparisons. Results. Sixteen systematic reviews were appraised. Overall, the methodological quality was low and external validity weak. For acute LBP, evidence that acupuncture has a more favorable effect than sham acupuncture in relieving pain was inconsistent; it had a similar effect on improving function. For chronic LBP, evidence consistently demonstrated that acupuncture provides short-term clinically relevant benefits for pain relief and functional improvement compared with no treatment or acupuncture plus another conventional intervention. Conclusion. Systematic reviews of variable quality showed that acupuncture, either used in isolation or as an adjunct to conventional therapy, provides short-term improvements in pain and function for chronic LBP. More efforts are needed to improve both internal and external validity of systematic reviews and RCTs in this area.

PMID: [25821485](#) PMCID: [PMC4364128](#) DOI: [10.1155/2015/328196](#)

[PubMed] [Free PMC Article](#)

Full text links

FREE full text article at Hindawi PMC Full text

Save items

Add to Favorites

Similar articles

Review Effectiveness of acupuncture for low back pain: a system [Spine (Phila Pa 1976). 2008]

Review Advice for the management of low back pain: a systematic review of ran [Man Ther. 2007]

Is immediate imaging important in managing low back pain? [J Athl Train. 2011]

Review Efficacy of spinal manipulation and mobilization for low back pain and [Spine J. 2004]

Clinical critical qualitative evaluation of the selected rand [Zhong Xi Yi Jie He Xue Bao. 2010]

See reviews... See all...

NCBI Resources How To Sign in to NCBI

PubMed US National Library of Medicine National Institutes of Health

Advanced Search Help

NCBI will be testing https on public web servers from 10:00-11:00 AM EDT (14:00-15:00 UTC) on Thursday, October 13. You may experience problems with NCBI services during that time. Please plan accordingly. [Read more.](#)

Format Abstract Send to

See comment in PubMed Commons below

Evid Based Complement Alternat Med. 2015;2015:328196. doi: 10.1155/2015/328196. Epub 2015 Mar 4.

Acupuncture for low back pain: an overview of systematic reviews.

Liu L¹, Skinner M¹, McDonough S², Mabire L¹, Baxter GD¹.

Author information

Abstract

Größe und Präzision

Objective. As evidence of the effectiveness of acupuncture for low back pain (LBP) is inconsistent, we aimed to critically appraise the evidence from relevant systematic reviews. Methods. Systematic reviews of randomized controlled trials (RCTs) concerning acupuncture and LBP were searched in seven databases. Internal validity and external validity of systematic reviews were assessed. Systematic reviews were categorized and high quality reviews assigned greater weightings. Conclusions were generated from a narrative synthesis of the outcomes of subgroup comparisons. Results. Sixteen systematic reviews were appraised. Overall, the methodological quality was low and external validity weak. For acute LBP, evidence that acupuncture has a more favorable effect than sham acupuncture in relieving pain was inconsistent; it had a similar effect on improving function. For chronic LBP, evidence consistently demonstrated that acupuncture provides short-term clinically relevant benefits for pain relief and functional improvement compared with no treatment or acupuncture plus another conventional intervention. Conclusion. Systematic reviews of variable quality showed that acupuncture, either used in isolation or as an adjunct to conventional therapy, provides short-term improvements in pain and function for chronic LBP. More efforts are needed to improve both internal and external validity of systematic reviews and RCTs in this area.

PMID: [25821485](#) PMCID: [PMC4364128](#) DOI: [10.1155/2015/328196](#)

[PubMed] [Free PMC Article](#)

Full text links

FREE full text article at Hindawi PMC Full text

Save items

Add to Favorites

Similar articles

Review Effectiveness of acupuncture for low back pain: a system [Spine (Phila Pa 1976). 2008]

Review Advice for the management of low back pain: a systematic review of ran [Man Ther. 2007]

Is immediate imaging important in managing low back pain? [J Athl Train. 2011]

Review Efficacy of spinal manipulation and mobilization for low back pain and [Spine J. 2004]

Clinical critical qualitative evaluation of the selected rand [Zhong Xi Yi Jie He Xue Bao. 2010]

See reviews... See all...

NCBI Resources How To Sign in to NCBI

PubMed.gov PubMed Search

US National Library of Medicine National Institutes of Health Advanced Help

NCBI will be testing https on public web servers from 10:00-11:00 AM EDT (14:00-15:00 UTC) on Thursday, October 13. You may experience problems with NCBI services during that time. Please plan accordingly. [Read more.](#)

Format Abstract - Send to -

See comment in PubMed Commons below

Evid Based Complement Alternat Med. 2015;2015:328196. doi: 10.1155/2015/328196. Epub 2015 Mar 4.

Acupuncture for low back pain: an overview of systematic reviews.

Liu L¹, Skinner M¹, McDonough S², Mabire L¹, Baxter GD¹.

Author information

Übertragbarkeit und Anwendbarkeit

Abstract

Objective. As evidence of the effectiveness of acupuncture for low back pain (LBP) is inconsistent, we aimed to critically appraise the evidence from relevant systematic reviews. Methods. Systematic reviews of randomized controlled trials (RCTs) concerning acupuncture and LBP were searched in seven databases. Internal validity and external validity of systematic reviews were assessed. Systematic reviews were categorized and high quality reviews assigned greater weightings. Conclusions were generated from a narrative synthesis of the outcomes of subgroup comparisons. Results. Sixteen systematic reviews were appraised. Overall, the methodological quality was low and external validity weak. For acute LBP, evidence that acupuncture has a more favorable effect than sham acupuncture in relieving pain was inconsistent; it had a similar effect on improving function. For chronic LBP, evidence consistently demonstrated that acupuncture provides short-term clinically relevant benefits for pain relief and functional improvement compared with no treatment or acupuncture plus another conventional intervention. Conclusion. Systematic reviews of variable quality showed that acupuncture, either used in isolation or as an adjunct to conventional therapy, provides short-term improvements in pain and function for chronic LBP. More efforts are needed to improve both internal and external validity of systematic reviews and RCTs in this area.

PMID: 25821485 PMCID: PMC4364128 DOI: 10.1155/2015/328196

[PubMed] Free PMC Article

Full text links: FREE full text article at Hindawi, PMC Full text

Save items: Add to Favorites

Similar articles: Effectiveness of acupuncture for low back pain: a system [Spine (Phila Pa 1976). 2008], Advice for the management of low back pain: a systematic review of ran [Man Ther. 2007], Is immediate imaging important in managing low back pain? [J Athl Train. 2011], Efficacy of spinal manipulation and mobilization for low back pain and [Spine J. 2004], Clinical critical qualitative evaluation of the selected rand [Zhong Xi Yi Jie He Xue Bao. 2010]


See reviews... See all...

Ordnen Sie den Evidenzstufen jeweils das richtige Validitätskriterium zu.

Positionen der Terme zurücksetzen

Ia	mehr als eine methodisch hochwertige nichtexperimentelle Studie
Ib	RCT
IIa	Expertenmeinung oder beschreibende Studien
IIb	hochwertige quasi-experimentelle Studie
III	hochwertige Studie ohne Randomisierung
IV	Systematic Review auf Basis von RCTs

Auswerten



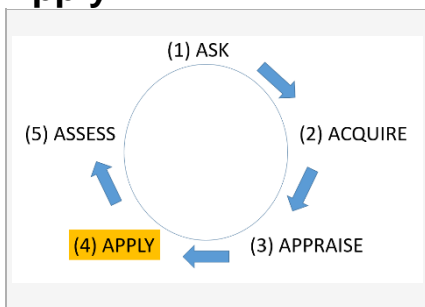
Nach welchen Kriterien beurteilen Sie Größe und Präzision einer Studie?

- Impact
- Durchführung
- Auswertung
- Autoren
- Design

Welche Kriterien berücksichtigen Sie bezüglich der Übertragbarkeit und Anwendbarkeit einer Studie?

- Vergleichbare Studien- und Zielpopulation
- Korrekte Verwendung von Literaturangaben
- Kosten-Nutzen-Relation
- Bekanntheitsgrad des Autors
- Interessenkonflikte des Autors

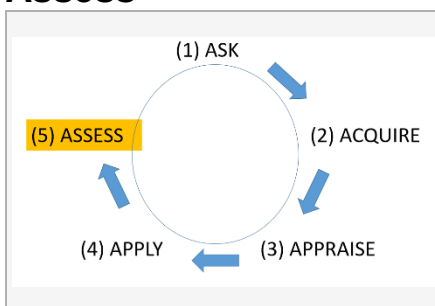
Apply



Die Anwendung der Rechercheergebnisse ist ein Prozess, der sich auf die vorherigen Punkte „Ask“, „Acquire“ und „Appraise“ stützt.

Die Therapieempfehlung sollte sorgfältig evaluiert werden und zu den individuellen Bedürfnissen des Patienten passen.

Assess



Die Evaluation der Behandlung dient der Beurteilung des gesamten 5-A-Schritteprozesses. Die Patientenproblematik ist dabei jeweils Start- und Endpunkt des Prozesses. Bleibt der Erfolg der Therapie aus, wird der 5-A-Schritteprozess erneut mit einer angepassten Fragestellung durchlaufen.

Im Folgenden wird der 5-A-Schritteprozess anhand eines Fallbeispiels durchlaufen...

Fallbeispiel

Herr Müller (73 Jahre, 1,70m, 87 kg), chronischer Gichtpatient, stellt sich in der Hausarztpraxis vor, in der Sie zurzeit famulieren.

Nachdem drei Gichtanfällen im vergangenen Jahr hatte seine Frau ihn vor einem halben Jahr doch noch überreden können zum Arzt zu gehen, wo im Rahmen der körperlichen Untersuchung Tophi im rechten Großzehengrundgelenk aufgefallen waren.

Die Blutanalyse hatte einen Harnsäurewert von 11,5 mg/dl ergeben.

Damals war Herr Müller über wichtige Diätempfehlungen zur Vermeidung von Gichtanfällen aufgeklärt und zunächst mit 100 mg Allopurinol therapiert worden.

Drei Monate später zur Kontrolle des HS-Spiegels lag der Wert bei 10.7 mg/dl, sodass die Dosis auf 300 mg angepasst wurde. Der aktuelle HS-Spiegel liegt bei 6mg/dl.

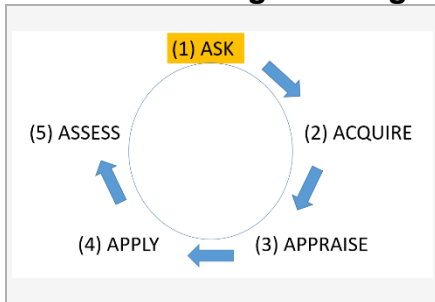
"Das Allopurinol bekommt mir gar nicht gut!", klagt Herr Müller. "Ständig habe ich Durchfälle und meinen Appetit habe ich auch verloren. Gibt es da keine bessere Alternative?"

1. Fassen Sie den Fall in kurzen Stichpunkten zusammen.

Herr Müller, 73 J., männl., 1,70 m, 87 kg, chron. Gichtpatient,

- vor einem Jahr: drei Gichtanfälle
- vor 6 Monaten: Tophi im rechten Großzehengrundgelenk rechts und erhöhter Harnsäurespiegel (11,5 mg/dl) -> Diätempfehlungen + 100mg Allopurinol
- vor 3 Monaten: 10,7 mg/dl Harnsäure -> Dosisanpassung auf 300 mg Allopurinol
- aktuell: 6mg/dl Harnsäure, Allopurinolunverträglichkeit (Appetitverlust + Durchfall)

2. Welche Fragestellung leiten Sie aus dem vorliegenden Fall ab?



P = „Patient Problem“: Männlicher adipöser Patient (73 Jahre) mit Appetitverlust und Durchfällen in Folge von Allopurinol-Behandlung zur harnsäuresenkenden Therapie von chronischer Gicht

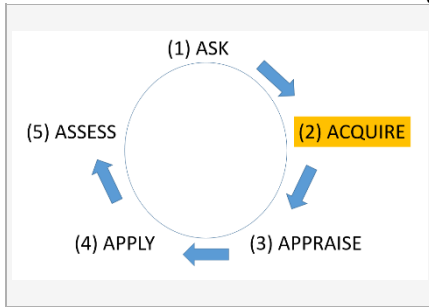
I = „Intervention“: Umstellung der harnsäuresenkenden Therapie

C = „Comparison“: Weiterbehandlung mit Allopurinol

O = „Outcome“:

- Primärer Endpunkt: Senkung des Harnsäurespiegels (auf unter 6,5mg/dl)
- Sekundärer Endpunkt: Reduktion der Nebenwirkungen (Appetitverlust und Durchfälle)

3. Wie lautet ihre Suchanfrage bei Pubmed?



Welche Wortwahl ist für die Suchanfrage am besten geeignet?

- "chronic gout treatment"
- "73 years old gout patient diarrhea loss of appetite allopurinol reduction of hyperuricemia alternative medication"
- "hyperuricemia treatment NOT allopurinol"
- "allopurinol adverse drug effects hyperuricemia treatment"

Auswerten

Machen Sie sich bewusst, dass es nicht nur eine richtige Suchstrategie gibt, sondern viele unterschiedliche Formulierungen zielführend sein können. Dennoch kann eine präzise, aber nicht zu spezifische Suchanfrage, möglichst unter Verwendung von PubMeds Schlagwortvokabular "MeSH", deutlich bessere Ergebnisse erzielen.

4. Welche zusätzlichen Optionen bei PubMed nutzen Sie, um die Suchanfrage möglichst präzise zu gestalten?

Welche der folgenden Optionen nutzen Sie um ihre Suchanfrage weiter einzuschränken?

- Filter: "Controlled Clinical Trial"
- Filter: "Comperative study"
- Operator: "NOT"
- Filter: "Review"

Auswerten

Filter bieten sich an, um die Anzahl der Suchergebnisse weiter einzuschränken und den Fokus zunächst auf besonders hochwertige Studientypen zu legen (Spitze der Pyramide). Finden Sie unter diesen Studientypen keine zufriedenstellende Antwort auf Ihre Frage, kann die Suche auf den nächstbesten Studientypen ausgeweitet werden.

Bei der Verwendung von Operatoren gilt manchmal "weniger ist mehr", da ein überflüssiger Gebrauch relevante Suchergebnisse ausschließen kann.

5. Welches Abstract deckt Ihre zuvor formulierte Fragestellung am besten ab?

z. B.: Li S, Yang H, Guo Y, Wei F, Yang X, Li D, Li M, Xu W, Li W, Sun L, Gao Y, Wang Y. Comperative efficacy and safety of urate-lowering therapy for the treatment of hyperuricemia: a sytematic review and network meta- analysis. Sci Rep. 2016 Sep 8;6:33082. doi: 10.1038/srep33082.

Sci Rep. 2016 Sep 8;6:33082. doi: 10.1038/srep33082.

Comparative efficacy and safety of urate-lowering therapy for the treatment of hyperuricemia: a systematic review and network meta-analysis.

Li S^{1,2}, Yang H^{1,2}, Guo Y¹, Wei F³, Yang X⁴, Li D², Li M², Xu W^{4,5}, Li W³, Sun L¹, Gao Y¹, Wang Y¹.

Author information

Abstract

The prevalence of hyperuricemia and gout has been increasing, but the comparative effectiveness and safety of different treatments remain uncertain. We aimed to compare the effectiveness and safety of different treatments for hyperuricemia using network meta-analysis methodology. We systematically reviewed fifteen randomized controlled trials (involving 7,246 patients through January 2016) that compared the effects of different urate-lowering drugs (allopurinol, benzbromarone, febuxostat, pegloticase and probenecid) on hyperuricemia. Drug efficacy and safety, as outcomes, were measured by whether the target level of serum urate acid was achieved and whether any adverse events occurred, respectively. We derived pooled effect sizes expressed as odds ratios (ORs) and 95% confidence intervals (CIs). The efficacy and safety of the drugs were ranked by cumulative ranking probabilities. Our findings show that febuxostat, benzbromarone, probenecid, pegloticase, and allopurinol were all highly effective at reducing the risk of hyperuricemia compared to placebo. Febuxostat had the best efficacy and safety compared to the other drugs. Furthermore, febuxostat 120 mg QD was more effective at achieving urate-lowering targets (OR: 0.17, 95% CI: 0.12-0.24) and safer (OR: 0.72, 95% CI: 0.56-0.91) than allopurinol.

PMID: 27605442 PMCID: PMC5015109 DOI: 10.1038/srep33082

[PubMed - in process] **Free PMC Article**

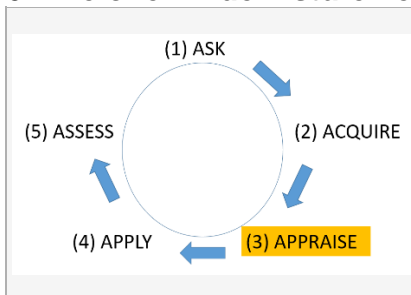


Auch hier gilt: Es gibt nicht nur ein richtiges Abstract und der Rest ist falsch.

Überfliegen Sie die Überschriften nach Übereinstimmungen mit Ihrer Fragestellung und lesen Sie sich einzelne Abstracts durch. Achten Sie auf die Hochwertigkeit der zugrundeliegenden Studie, einer Übereinstimmung der Inhalte mit ihrer Fragestellung und der Anwendbarkeit von Therapieempfehlungen.

Wenn Sie nicht auf Anhieb fündig werden, können Sie die Wortwahl anpassen und/oder nach einem weniger hochwertigen Studientypen suchen.

6. Welche Evidenzstufe liegt bei Ihrem Abstract vor?




Zur Erinnerung:

Evidenzstufe	Validitätskriterien
Ia)	Systematic Review, das auf mindestens einer randomisierten kontrollierten Studie (RCT) basiert
Ib)	Randomisierte kontrollierte Studie (RCT)
IIa)	Hochwertige Studie ohne Randomisierung (z.B. Crossover-Studie)
IIb)	Hochwertige Studie eines anderen Typs quasi experimenteller Studien (z.B. Kohorten-Studie)
III)	Mehr als eine methodisch hochwertige nichtexperimentelle Studie (z.B. Fall-Kontroll-Studie)
IV)	Expertenmeinung oder beschreibende Studien (z.B. Case Report/ Case Series)

Welche Evidenzstufe liegt bei Ihrem Abstract vor?

Ia
 IIa
 Ib
 III
 IIb
 IV



7. Welche Information zu Design, Durchführung & Auswertung gewinnen Sie aus dem Abstract?

Sci Rep. 2016 Sep 8;6:33082. doi: 10.1038/srep33082.

Comparative efficacy and safety of urate-lowering therapy for the treatment of hyperuricemia: a systematic review and network meta-analysis.

Li S^{1,2}, Yang H^{1,2}, Guo Y¹, Wei F³, Yang X⁴, Li D², Li M², Xu W^{4,5}, Li W³, Sun L¹, Gao Y¹, Wang Y¹.




Author information

Abstract

The prevalence of hyperuricemia and gout has been increasing, but the comparative effectiveness and safety of different treatments remain uncertain. We aimed to compare the effectiveness and safety of different treatments for hyperuricemia using network meta-analysis methodology. We systematically reviewed fifteen randomized controlled trials (involving 7,246 patients through January 2016) that compared the effects of different urate-lowering drugs (allopurinol, benzbromarone, febuxostat, pegloticase and probenecid) on hyperuricemia. Drug efficacy and safety, as outcomes, were measured by whether the target level of serum urate acid was achieved and whether any adverse events occurred, respectively. We derived pooled effect sizes expressed as odds ratios (ORs) and 95% confidence intervals (CIs). The efficacy and safety of the drugs were ranked by cumulative ranking probabilities. Our findings show that febuxostat, benzbromarone, probenecid, pegloticase, and allopurinol were all highly effective at reducing the risk of hyperuricemia compared to placebo. Febuxostat had the best efficacy and safety compared to the other drugs. Furthermore, febuxostat 120 mg QD was more effective at achieving urate-lowering targets (OR: 0.17, 95% CI: 0.12-0.24) and safer (OR: 0.72, 95% CI: 0.56-0.91) than allopurinol.


PMID: 27605442 PMCID: PMC5015109 DOI: 10.1038/srep33082

[PubMed - in process] **Free PMC Article**

Welche der folgenden Informationen können Sie mit Hilfe des Abstracts beantworten?

Die Studienpopulation und Herr Müller stimmen in ihren Eigenschaften überein: selbes Alter, gleiche Grunderkrankung, Allopurinolunverträglichkeit
 Nachvollziehbare Schlussfolgerung aus Studienergebnissen: Alle Medikament waren wirksamer als das Placebo und Febuxostat hat in Bezug auf Wirksamkeit und Sicherheit am besten abgeschnitten
 Fragestellung: Vergleich von verschiedenen Urat-reduzierenden Medikamenten bei Hyperurikämie in Bezug auf Wirkung und Sicherheit
 Studientyp: Meta-Analyse mit Bezug auf 15 RCTs



Nicht jedes Abstract wird alle nötigen Informationen zur Beurteilung der durchgeführten Studie liefern. Berücksichtigen Sie diese Lücken bei Ihrer Evaluation der Therapieempfehlung.

Zum Teil können Links zu den entsprechenden Volltexten weitere Informationen liefern.

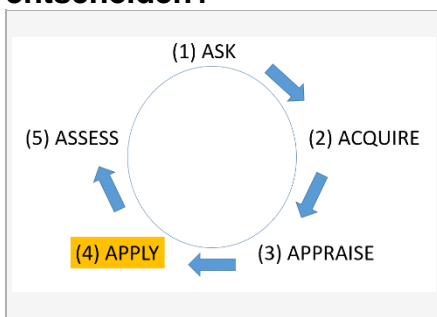
8. Sind die Ergebnisse aus der Studie auf Ihren Patienten übertragbar?

Febuxostat scheint eine alternative Hyperurikämiebehandlung zu sein, die das Allopurinol bei Herr Müller ersetzen könnte. Hier ist jedoch eine weitere Recherche notwendig, um zu überprüfen, ob der Patient für diese Therapie in Frage kommt und worauf sich die bessere "Sicherheit" von Febuxostat gegenüber Allopurinol bezieht.

Der Vorteil von Systematic Reviews ist, dass sie die Evidenz aus mehreren hochwertigen Studien bereits zusammenfassen.

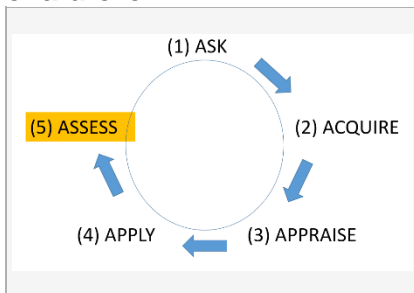
Trotzdem kann es nötig sein mehrere Systematic Reviews zu einem Thema zu vergleichen oder einzelne Aspekte in einer weiteren Suchanfrage zu recherchieren.

9. Für welche Behandlung würden Sie sich aufgrund der evaluierten Evidenz entscheiden?



Bestätigt sich im Zuge der weiteren Recherche, dass Febuxostat eine besser verträgliche alternative zu Allopurinol ist, kann die Therapie auf dieses Medikament umgestellt werden.

10. Mit welcher Maßnahme würden Sie Ihre gewählte Behandlung im Verlauf evaluieren?



Auch nach Therapieumstellung sollte Herr Müller regelmäßig (wie bisher alle 3 Monate) zur Laborkontrolle erscheinen und bei Auftreten von Nebenwirkungen jederzeit den Kontakt zur Praxis aufnehmen.

9.2 Anlage II: Aufgabenstellung der Fallbeispiele A, B und C

Es folgen die fiktiven Fallbeispiele A, B und C, die die Studenten am Ende des Lernmoduls bearbeiten sollen.

Fallbeispiel A

Herr Peters, 74 Jahre, leidet an einer seit Jahren bekannten Niereninsuffizienz.

Kürzlich traten zusätzlich erste Symptome einer Gicht bei ihm auf. Im Rahmen der Routinekontrolle seiner Nierenwerte zeigt sich ein deutlicher Anstieg der Harnsäure, sodass eine Harnsäuresenkende Therapie mit Allopurinol erwogen wird.

"Geht das denn überhaupt mit meinen Nieren?", will er wissen.

Fallbeispiel B

Herr Breuer, 75 Jahre alt und pensionierter Koch, klagt über zunehmende Schmerzen in den Gelenken.

Bei der körperlichen Untersuchung fallen Tophi an der linken Ohrmuschel auf. Das Labor bestätigt den Verdacht auf Gicht mit einem erhöhten Harnsäurespiegel.

"Gicht?", fragt Herr Breuer Sie. "Das heißt nie wieder Schweinebraten?! Und was ist mit meinem Kaffee am Morgen? Was darf ich denn überhaupt noch essen und trinken?"

Fallbeispiel C

Im Rahmen Ihrer Hausarztfamulatur sollen Sie mit einem Patienten seine Laborergebnisse besprechen.

Herr Schmitz (68 Jahre, 87kg, 1,70m) klagt gelegentlich über Rückenschmerzen und nächtlichen Harndrang, ist aber ansonsten klinisch unauffällig. Bei einer Routineuntersuchung fielen jedoch erhöhte Harnsäurewerte auf.

"Bei meinem Bruder war das genauso.", berichtet Herr Schmitz. "Bei dem hat der Wert auch nicht gestimmt und ein paar Monate später ist er an einem Herzinfarkt gestorben. Macht mein Herz jetzt auch schlapp?"

9.3 Anlage III: Erwartungshorizont des Bonn Tests

Tab. 1: Erwartungshorizont des Schrittes Ask.

	Patient	Intervention	Comparison	Outcome
Hervorragend (6 Pkt.)	Mehrere relevante Eigenschaften des Patienten werden berücksichtigt z. B. „Gichtpatient mit Appetitverlust und Durchfällen unter Allopurinol bei harnsäuresenkender Therapie“	Spezifische Intervention wird genannt z. B. „Umstellung der harnsäuresenkenden Therapie“	Alternative zur Intervention der Wahl wird genannt z. B. „Weiterbehandlung mit Allopurinol“	Objektiver/ für den Patienten bedeutender Ausgang der Intervention wird genannt z. B. „nebenwirkungsarme Senkung des Harnsäurespiegels“
Stark (4 Pkt.)	Eine relevante Eigenschaft des Patienten wird genannt z. B. „Gichtpatient“/ „Allopurinol-unverträglichkeit“	-Typ der Intervention wird genannt z. B. „Medikamentenumstellung“ -Intervention und Comparison werden nicht differenziert	-Spezielle Kontrollgruppe wird genannt z. B. „keine Therapieumstellung“ -Comparison und Intervention werden nicht differenziert	- Unspezifischer Ausgang z. B. „Blutwerte“/ „keine Nebenwirkung“ -Krankheitsorientierter Ausgang z. B. „Harnsäurespiegel“ (ohne Maß für klinische Relevanz)
Limitiert (2 Pkt.)	Einzelne Eigenschaft wird genannt, trägt aber nicht zur spezifischen Recherche bei z. B. „Patient“	Intervention wird genannt, trägt aber nicht zur spezifischen Recherche bei z. B.	Alternative Intervention wird genannt, trägt aber nicht zur spezifischen Recherche bei	Ausgang wird genannt, trägt aber nicht zur spezifischen Recherche bei z. B. „Effekte“/ „Veränderter Ausgang“

		„Methode“/ “Optionen“/ „Behandlung“	z.B. „vergleichen mit anderen Methoden“	
Keine Evidenz (0 Pkt.)	Nichts davon trifft zu	Nichts davon trifft zu	Nichts davon trifft zu	Nichts davon trifft zu

Tab. 2: Erwartungshorizont des Schrittes Acquire – Suchstrategie.

	Wortwahl	MeSH	Präzisierende Optionen
Hervor- ragend (8 Pkt.)	Ausschließlich die mar- kantesten PICO-Elemente werden übernommen z. B. „allopurinol adverse drug effects hyperuricemia treatment“/ „allopurinol intolerance alternative therapy“	Ausschließlich MeSH werden verwendet z. B. „allopuri- nol adverse drug effects hyperuricemia treatment“	Filter oder Operatoren werden sinnvoll verwendet z. B. „Filter: Review“, „allopurinol AND hyperuricemia“
Stark (6 Pkt.)	Zu viele oder zu wenige Informationen werden übernommen z. B. „73 years old gout patient diarrhea loss of appetite allopurinol reduction of hyperuricemia alternative medication“/ „chronic gout treatment“	Überwiegend werden MeSH verwendet z. B. „allopurinol intolerance alternative therapy“	-Weniger passender Filter wird verwendet oder verwendeter Operator ist überflüssig z. B. „Filter:Comment“/ „therapy AND mediation“ -Grundlegende Maßnahmen zur Präzisierung von Suchanfragen werden genannt, aber nicht auf die eigene Suchanfrage bezogen

Limitiert (3 Pkt.)	Unspezifische Fragestellung, die nicht zielführend für Problemstellung ist z. B. „Medikamentenumstellung“	Überwiegend werden keine MeSH verwendet z. B. „medication intolerance alternative“	nicht sinnvoller Operator schließt mögliche relevante Suchergebnisse aus z. B. „hyperuricemia therapy NOT allopurinol“
Keine Evidenz (0 Pkt.)	Nichts davon trifft zu	Nichts davon trifft zu	Nichts davon trifft zu

Tab. 3: Erwartungshorizont des Schrittes Acquire – Auswahl des Abstracts.

	Inhalt	Qualität
Hervorragend (12 Pkt.)	Das gewählte Abstract beantwortet die ursprüngliche Fragestellung und die Schlussfolgerung für eine Therapieempfehlung ist auf den Patienten übertragbar	Die höchste verfügbare Evidenzstufe wurde gewählt und die Informationen zu Studiendesign und -durchführung bestätigen die Hochwertigkeit
Stark (8 Pkt.)	Das gewählte Abstract beantwortet die ursprüngliche Fragestellung, die Übertragbarkeit der Studienpopulation auf den Patienten ist allerdings fragwürdig	Die höchste verfügbare Evidenzstufe wurde gewählt, einzelne Aspekte zu Studiendesign und -durchführung werten die Qualität jedoch ab
Limitiert (4 Pkt.)	Das gewählte Abstract befasst sich mit der Thematik des „Patient Problem“, liefert aber keine direkte Antwort auf die Fragestellung	Eine weniger hochwertige Evidenzstufe wurde gewählt bzw. Studiendesign und -durchführung weisen starke Mängel auf
Keine Evidenz (0 Pkt.)	Nichts davon trifft zu	Nichts davon trifft zu

Tab. 4: Erwartungshorizont des Schrittes Appraise.

	Validität	Größe und Präzision	Übertragbarkeit und Anwendbarkeit
Hervorragend (8 Pkt.)	Korrekte Zuordnung der Evidenzstufe z. B. „Ia“	Alle im Abstract genannten Qualitätsmerkmale bezüglich Design, Durchführung und Auswertung werden aufgelistet	Abweichungen bzw. Parallelen zwischen Studienkollektiv und Patient oder andere mögliche Einschränkungen für die Übertragbarkeit auf den Patienten (falls vorhanden) werden mit Beispiel genannt z. B. „Keine Angaben zu Kriterien des Studienkollektivs -> Anwendbarkeit auf Patienten fraglich“
Stark (6 Pkt.)	-Geringfügige Verwechslung der Evidenzstufen z. B. „Ib“ anstelle von „Ia“ -Studientyp ist aus Abstract nicht deutlich ableitbar	Die meisten im Abstract genannten Qualitätsmerkmale bezüglich Design, Durchführung und Auswertung werden aufgelistet	Unspezifische Nennung von Problemen bei der Übertragung, die nicht weiter ausgeführt werden z. B. „Kosten-Nutzen-Abschätzung“
Limitiert (3 Pkt.)	-Validitätskriterium wird beschrieben aber keine Evidenzstufe wird benannt -Die falsche Evidenzstufe wird benannt	Nur ein geringer Anteil der im Abstract genannten Qualitätsmerkmale bezüglich Design, Durchführung und Auswertung werden aufgelistet	Übertragbarkeit wird erwähnt, trägt aber nicht zur Evaluation des Abstracts bei z. B. „anwendbar“/ „übertragbar“

Keine Evidenz (0 Pkt.)	Nichts davon trifft zu	Nichts davon trifft zu	Nichts davon trifft zu
------------------------	------------------------	------------------------	------------------------

Tab. 5: Erwartungshorizont des Schrittes Apply.

	Interne Validität	Externe Validität
Hervorragend (12 Pkt.)	Die ausgewählte Therapieempfehlung ist schlüssig aus den evaluierten Rechercheergebnissen abgeleitet	Mögliche Restriktionen der Anwendung wie die Übertragbarkeit auf den Patienten oder die praktische Umsetzbarkeit werden berücksichtigt und ggf. Verbesserungsvorschläge gemacht z. B. „geringe Fallzahlen -> ggf. Wiederholung durch größere RCTs“
Stark (8 Pkt.)	Die ausgewählte Therapieempfehlung ist auf die Thematik des Abstracts zurück zu führen, wird darin aber nicht eindeutig benannt	Die meisten Aspekte zu Übertragbarkeit und Anwendbarkeit werden bei der Therapieempfehlung berücksichtigt
Limitiert (4 Pkt.)	-Die ausgewählte Therapieempfehlung ist zu unspezifisch, um Fragestellung beantworten zu können z. B. „Medikamentenumstellung“ -Die Therapieempfehlung wird zwar richtig abgeleitet aus dem Abstract, beantwortet aber nicht die ursprüngliche Fragestellung	Externe Validitätskriterien werden nur unspezifisch benannt z. B. „nicht übertragbar“/ „schlecht umsetzbar“
Keine Evidenz (0 Pkt.)	Nichts davon trifft zu	Nichts davon trifft zu

Tab. 6: Erwartungshorizont des Schrittes Assess.

	Evaluationsmaßnahmen
Hervorragend (24 Pkt.)	Die Evaluation erfolgt nach einem festgelegten Zeitintervall und bietet die Möglichkeiten Schäden für den Patienten frühzeitig zu erkennen und abzuwenden z. B. „Routine-kontrolle mit Labor in 3 Monaten und Kontaktaufnahme mit Praxis bei Auftreten von Nebenwirkungen“
Stark (18 Pkt.)	-Die Evaluation erfolgt in einem festgelegten Zeitintervall, das jedoch zu groß gewählt wird oder nicht die Möglichkeit bietet kurzfristig auftretende Probleme zu erkennen und zu beheben z. B. „Routinekontrolle mit Labor einmal im Jahr“ -Es wird kein Zeitintervall genannt, aber die gewählten Evaluationsmaßnahmen sind sinnvoll
Limitiert (10 Pkt.)	Die Evaluationsmaßnahme ist nicht zielführend für die Erhebung der nötigen Daten z. B. „Sprechstunde“ statt „Labor“
Keine Evidenz (0 Pkt.)	Nichts davon trifft zu