

Aus der Medizinischen Klinik und Poliklinik IV
Klinikum der Ludwig-Maximilians-Universität München
Direktor: Prof. Dr. med. Martin Reincke

Klinische Entscheidungsfindung in der Endokrinologie am Beispiel des Cushing-Syndroms

Kumulative Habilitationsschrift
zur Erlangung der Venia Legendi für Experimentelle Innere Medizin
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

vorgelegt von
Dr. med. Leah Theresa Braun
aus Bonn

2023

Inhaltsverzeichnis

Hintergrund	3
Klinische Entscheidungsfindung – <i>Clinical Reasoning</i>	3
Endokrinologie und das Cushing-Syndrom	4
Synthesis – Klinische Entscheidungsfindung in der Endokrinologie	6
Zielsetzung der Habilitation	9
Resultate und Diskussion der eigenen Publikationen	10
Teil 1: Klinische Entscheidungsfindung in der Metaebene: Didaktische Modelle und angewandte Didaktik in der Endokrinologie	10
Evaluation des Diagnostikprozesses.....	10
<i>Bayesian Reasoning</i> in der Endokrinologie	17
Therapiefehler in der Endokrinologie.....	25
Teil 2: Klinische Entscheidungsfindung am konkreten Beispiel des Cushing-Syndroms	28
Die Diagnostik des Rezidivs beim Morbus Cushing.....	28
Speichelprofile zur Optimierung der Diagnosestellung	30
Zusammenfassung und Ausblick	32
Ausblick	35
Publikationsverzeichnis	36
Danksagung	42
Literaturverzeichnis	43

Hintergrund

A physician is obligated to consider more than a diseased organ, more even than the whole man - he must view the man in his world.

- *Harvey Cushing*

Klinische Entscheidungsfindung – *Clinical Reasoning*

Das Treffen klinischer Entscheidungen ist Hauptaufgabe des ärztlichen Handelns. Klinische Entscheidungen betreffen in der Inneren Medizin drei Bereiche: Diagnoseentscheidungen, Prozess- bzw. Prozedurenentscheidungen und Therapieentscheidungen, diese gehen fließend ineinander über in einer variablen zeitlichen Reihenfolge. Zwar besteht ein nicht zu vernachlässigender Zusammenhang zwischen der Richtigkeit des Diagnoseprozesses und der Richtigkeit der Diagnosestellung sowie der sich anschließenden Therapieentscheidung, doch werden im Einzelfall auch richtige Therapieentscheidungen trotz falscher Diagnosestellung getroffen oder richtige Diagnosen trotz eines fehlerhaften Diagnoseprozesses gefällt (1).

Clinical Reasoning hat in den letzten Jahren zunehmend an wissenschaftlicher Bedeutung gewonnen und auch in der öffentlichen Debatte treten Diagnosefehler immer mehr in den Fokus. Schätzungsweise ist jede 10. Diagnose eine Fehldiagnose (2), was es sehr wahrscheinlich macht, dass ein Großteil der Bevölkerung mindestens einmal im Leben mit einem Diagnosefehler konfrontiert wird – mit ganz unterschiedlichen Folgen, aber möglichem fatalen Ausgang (3). Therapiefehler sind in dieser Schätzung nicht eingerechnet. Reaktionen auf die Problematik sind eine zunehmende Organisation von Patientengruppen, beispielsweise in Selbsthilfegruppen. Ebenso führte dies auch zur Gründung der *Society to Improve Diagnosis in Medicine*, die sich eine wissenschaftliche Auseinandersetzung mit dieser Thematik zur Aufgabe gemacht hat und ein eigenes Journal – *Diagnosis* – herausgibt (4). Die bislang durchgeführten Untersuchungen zu den Ursachen von Diagnosefehlern bei Ärztinnen und Ärzten zeigen widersprüchliche Ergebnisse. Gemeinsam ist diesen Studien allerdings, dass neben systemischen Fehlern insbesondere kognitive Fehler im Diagnoseprozess eine Rolle spielen (5).

Auch in der medizinischen Ausbildung wird die Thematik *Clinical Reasoning* zunehmend berücksichtigt. Es besteht allerdings noch ein umfangreicher Forschungsbedarf bezüglich Lehrstrategien, Teststrategien und auch der Operationalisierung von Messmethoden für *Clinical Reasoning* (6). Die grundlegende Bedeutung des *Clinical Reasoning* wurde auch im nationalen kompetenzbasierten Lernzielkatalog der Medizin (NKML) in der Definition des Kompetenzbegriffs hervorgehoben: Im NKLM (7) wird in der Einführung zum Thema Kompetenzen, Rollen und Lernziele die folgende Definition ärztlicher Kompetenz angeführt:

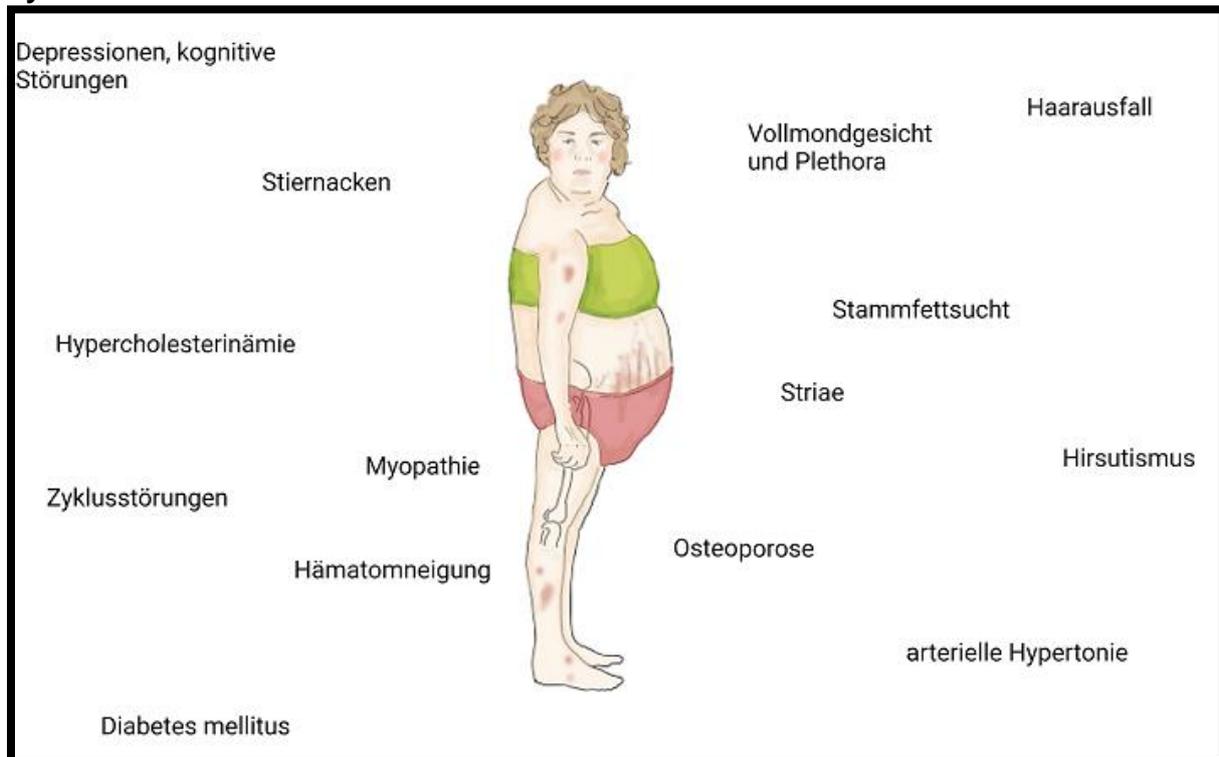
*„...professional competence is the habitual and judicious use of communication, knowledge, technical skills, **clinical reasoning**, emotions, values, and reflections in daily practice for the benefit of the individual and community being served...“ (Epstein & Hundert 2002, S. 226) (8).*

Obwohl die Bedeutung der Klinischen Entscheidungsfindung also offensichtlich ist, werden Forschungsarbeiten in diesem Bereich vorwiegend von nicht-ärztlichen Forscherinnen und Forschern (im Bereich Pädagogik, Psychologie und Medizindidaktik) und nicht von Ärztinnen und Ärzten geleistet. Es gibt ein hochkomplexes und fundiertes Forschungsfeld zum Thema Entscheidungsfindung in der Psychologie, deren Erkenntnisse aber selten auf das ärztliche Handeln übertragen werden und eine gewinnbringende Synthese der Kompetenzen und des Wissens bisher nicht stattfindet.

Endokrinologie und das Cushing-Syndrom

Die Endokrinologie ist ein Teilbereich der Inneren Medizin (9) mit sehr prävalenten ‚Volkskrankheiten‘ und einer in Summe großen Zahl von selteneren Erkrankungen. Die Diagnosestellung erfolgt im Wesentlichen anhand der Anamnese und der klinischen Untersuchung sowie der Biochemie und ist – im Gegensatz zu anderen Fachbereichen der Inneren Medizin – wenig von apparativen Untersuchungen abhängig. Es gibt verschiedene wichtige endokrinologische Erkrankungen, von denen viele – die Osteoporose, verschiedene Schilddrüsenerkrankungen oder die arterielle Hypertonie – sehr häufig sind, während andere Krankheitsbilder zu den selteneren Erkrankungen gehören. Hierzu zählt auch das endogene Cushing-Syndrom, welches eine sehr seltene, aber potenziell tödlich verlaufende Erkrankung ist (10). Patientinnen und Patienten mit einem Cushing-Syndrom können an einer Vielzahl verschiedener Symptome und Komorbiditäten leiden (Abbildung 1). Die diagnostische Herausforderung liegt allerdings darin, dass viele dieser Symptome auch mit anderen Erkrankungsbildern wie dem metabolischen Syndrom oder dem PCO-Syndrom (Polyzystisches Ovarsyndrom) überlappen (11).

Abbildung 1: Symptome und charakteristische Untersuchungsbefunde beim Cushing-Syndrom



Die Komplexität des Erscheinungsbildes sowie die Seltenheit der Erkrankung führen häufig zu verspäteten Diagnosen, wie eine multizentrische Studie und Metaanalyse mit über 5000 Patientinnen und Patienten bestätigte (12). Im Laufe des Diagnoseprozesses werden bis zu vier verschiedene Ärztinnen und Ärzte konsultiert (13), bis die korrekte Diagnose gestellt wird.

Wie kann nun eine frühzeitige und effiziente Diagnosestellung erfolgen? Es wurden bereits zahlreiche Studien durchgeführt, in denen Patientinnen und Patienten, die potenziell unter einem Cushing-Syndrom leiden könnten, – wie zum Beispiel Patientinnen und Patienten mit Diabetes mellitus (14, 15) oder Hypertonie (16, 17) – gescreent wurden. Die Studienergebnisse variieren hier stark zwischen dem gewählten Symptom/Komorbidität und der Prävalenz der Erkrankung. Letztlich wird von der *US Endocrine Society* bei erwachsenen Patienten ein Screening nur für folgende drei Patientengruppen empfohlen (18):

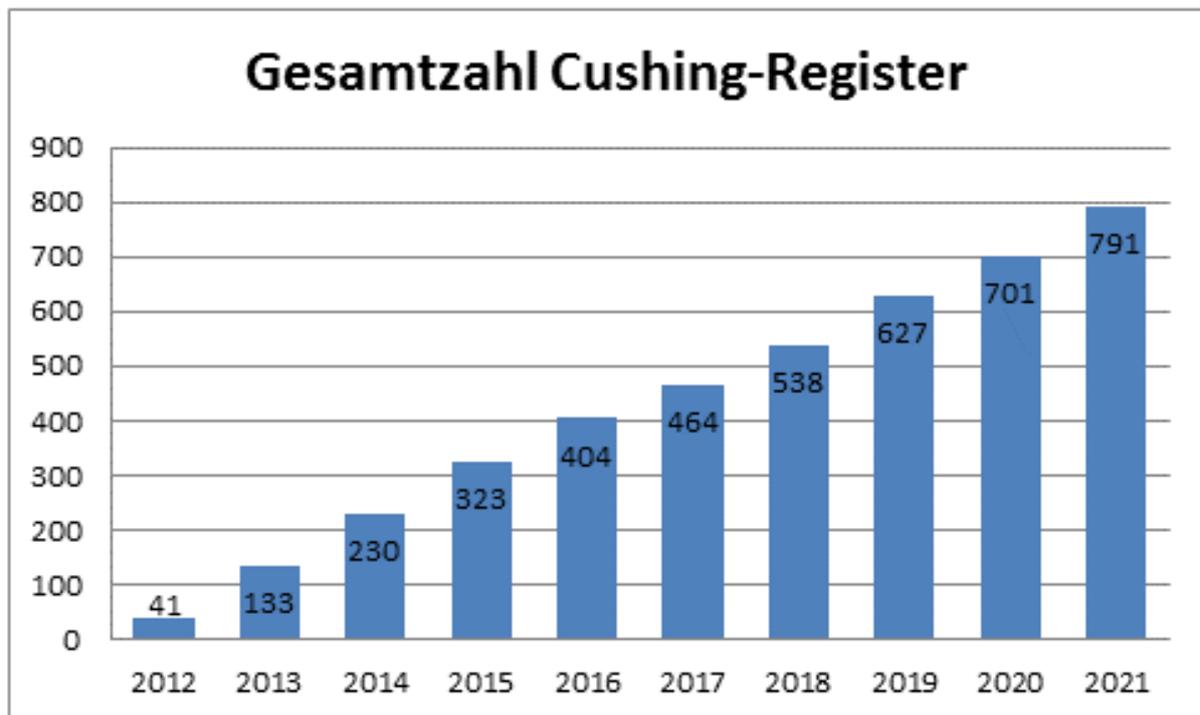
Patientinnen und Patienten.....

- mit multiplen und progressiven Symptomen, die für ein Cushing-Syndrom typisch sind
- mit adrenalem Adenom
- mit für ihr Alter ungewöhnlichen Symptomen (z.B. Osteoporose bei jungen Patienten)

Zur besseren Erforschung der Erkrankung wurde 2012 das *Deutsche Cushing-Register* gegründet. Patientinnen und Patienten mit Cushing-Syndrom werden prospektiv und

retrospektiv eingeschlossen und langfristig betreut. Es gibt verschiedene Studienzentren in Deutschland, wobei die meisten Patientinnen und Patienten in einer hierfür eigens geschaffenen Spezialambulanz am Standort München betreut werden. Zum jetzigen Zeitpunkt sind in München mehr als 700 Patienten in das Register eingeschlossen (Abbildung 2), wobei der Frauenanteil deutlich überwiegt. Alle an einem Cushing-Syndrom erkrankten werden einmal jährlich strukturiert anamnestiziert und untersucht.

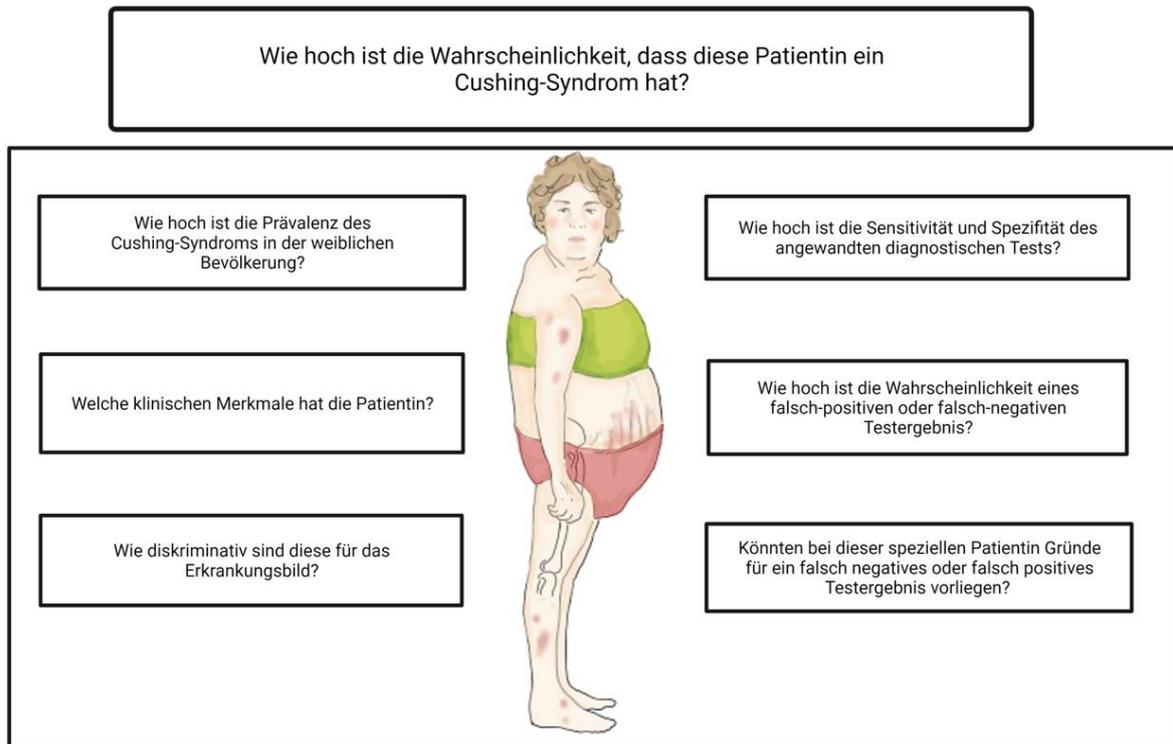
Abbildung 2: Patientenzahlen des *Deutschen Cushing-Registers* am Standort München



Synthesis – Klinische Entscheidungsfindung in der Endokrinologie

In der Literatur wird zwischen zwei möglichen Situationen unterschieden: *ambiguous risk* (19) und *objective risk conditions* (20). Während bei ersterem Szenario keine kalkulierbaren Wahrscheinlichkeiten vorliegen, die bei der Entscheidungsfindung helfen könnten, sind im zweiten Szenario – Beispiel Lotterie – die Wahrscheinlichkeiten des Outcomes prinzipiell berechenbar und damit objektiv zu erfassen. Die Klinische Entscheidungsfindung bewegt sich nun in einem Spannungsfeld zwischen diesen beiden Szenarien. Die Gesamtfragestellung – ob eine Erkrankung vorliegt – lässt sich häufig in Unterfragestellungen zerlegen, die leichter zu beantworten sind als die eigentliche Fragestellung. Auf das Cushing-Syndrom übertragen bedeutet dies beispielsweise folgendes:

Abbildung 3: Ambiguous und objective risk am Beispiel des Cushing-Syndroms



Diese Unterfragen können relativ präzise beantwortet werden, doch auch hierfür bedarf es eine weitere Kompetenz: das Lösen von Bayesian Aufgaben. Mit Hilfe der Sensitivität und Spezifität eines Tests sowie der Prävalenz einer Erkrankung lassen sich mit Hilfe der Formel nach Bayes (21) der positive und negative prädiktive Wert berechnen. Damit kann beantwortet werden, wie wahrscheinlich es ist, dass ein Patient mit einem positiven Testergebnis tatsächlich an einer bestimmten Erkrankung leidet. Aus einer Vielzahl von Studien ist allerdings bekannt, dass die Anwendung von *Bayesian Reasoning* sowohl für Ärztinnen und Ärzte als auch Medizinstudierende problematisch ist (22, 23). Ein klassisches Bayes-Problem ist das Mammographie-Screening (24):

„The probability of breast cancer is 1% for a woman of a particular age group who participates in a routine screening (a priori probability, also called prevalence $P(B)$). If a woman who participates in a routine screening has breast cancer, the probability is 80% that she will have a positive mammogram (sensitivity $P(M+|B)$). If a woman who participates in a routine screening does not have breast cancer, the probability is 9.6% that she will have a false-positive mammogram (false-alarm rate $P(M+|\neg B)$).

What is the probability that a woman who participates in a routine screening and has a positive mammogram actually has breast cancer?“

Medizinisches Personal kann diese Aufgabe in der Regel nicht lösen und neigt dazu, die Wahrscheinlichkeit für Brustkrebs bei einem positiven Mammogramm deutlich zu überschätzen (22). Die Aufgabe kann folgendermaßen gelöst werden:

$$\begin{aligned} P(B|M+) &= \frac{P(M+|B)P(B)}{P(M+|B)P(B) + P(M+|\neg B)P(\neg B)} \\ &= \frac{80\% \cdot 1\%}{80\% \cdot 1\% + 9.6\% \cdot 99\%} \\ &\approx 7.8\% \end{aligned}$$

Das Ergebnis von 7,8% mag überraschen, ist aber korrekt. Die Prävalenz einer Erkrankung hat maßgeblichen Einfluss auf den positiven und negativen prädiktiven Wert. So tendiert man in der klinischen Praxis dazu, bei einem positiven Testergebnis auch von einem sicheren Vorliegen der Erkrankung auszugehen. Gerade bei sehr seltenen Erkrankungen mit niedriger Prävalenz ist dies aber auch bei einer hohen Sensitivität oftmals nicht der Fall.

Was bedeutet dies nun konkret für die Endokrinologie?

In der Endokrinologie besteht eine starke Diskrepanz zwischen einerseits sehr häufigen „Volkskrankheiten“ wie Diabetes mellitus, Schilddrüsenerkrankungen und Osteoporose und andererseits sehr seltenen Erkrankungen, sogenannten *Rare diseases*, wie dem Cushing-Syndrom, dem Phäochromozytom oder seltenen Knochenstoffwechselerkrankungen und neuroendokrinen Tumoren. Dies ist von großer Bedeutung bei Entscheidungsszenarien in der täglichen klinischen Arbeit, weil verschiedene Prävalenzen Auswirkungen auf den positiven und negativen prädiktiven Wert haben. Dies macht es erforderlich, zwischen mathematisch sehr unterschiedlichen Situationen hin- und herzuwechseln. Die Diagnosefindung in der Endokrinologie ist zudem häufig dadurch gekennzeichnet, auf Such- und Bestätigungstests oder eine Kombination verschiedener biochemischer Tests zurückzugreifen. Im Gegensatz zu anderen Disziplinen der Inneren Medizin, in denen Entscheidungen oft von einem einzigen - hochsensitiven oder hochspezifischen - Laborparameter abhängen (Troponin, D-Dimer), ist dies in der Endokrinologie nur sehr selten der Fall. Die Kombination diverser Tests erhöht die Komplexität – und oft damit auch die Schwierigkeit – einer klinischen Situation und macht damit die Endokrinologie zu einem Fachbereich, welcher im besonderen Maße von Diagnosefehlern, Prozessfehlern und damit zum Teil auch Therapiefehlern betroffen sein kann.

Zielsetzung der Habilitation

Diese Habilitation soll zu einem besseren Verständnis für Diagnoseprozesse und klinische Entscheidungen in der Endokrinologie mit besonderem Fokus auf das Cushing-Syndrom beitragen.

Die Arbeit umfasst zwei Teilbereiche, die anhand meiner bisherigen Publikationen dargestellt werden:

- Klinische Entscheidungsfindung in der Metaebene: Didaktische Modelle und angewandte Didaktik in der Endokrinologie
- Klinische Entscheidungsfindung am konkreten Beispiel des Cushing-Syndroms

Die übergreifende Hypothese dieser Arbeit ist es, dass sich durch eine reflektierte Anwendung von medizindidaktischen Modellen, die klinische Entscheidungsfindung in verschiedenen Aspekten verbessern lässt.

Resultate und Diskussion der eigenen Publikationen

Teil 1: Klinische Entscheidungsfindung in der Metaebene: Didaktische Modelle und angewandte Didaktik in der Endokrinologie

In den folgenden Arbeiten stehen verschiedene Gruppen im Fokus: Medizinstudierende im vorklinischen und klinischen Studienabschnitt, Assistenzärzte und Ärztinnen und Fachärzte. Ob und in wie fern Erkenntnisse, die in einer Gruppe gewonnen wurden, auf die anderen übertragen werden können, ist studienindividuell zu diskutieren.

Evaluation des Diagnostikprozesses

Zugrunde liegende Publikationen:

Braun LT, Borrmann KF, Lottspeich C, Heinrich DA, Kiewewetter J, Fischer MR, Schmidmaier R. *Scaffolding clinical reasoning of medical students with virtual patients: effects on diagnostic accuracy, efficiency, and errors*. *Diagnosis (Berl)*. 2019 Jun 26;6(2):137-149. doi: 10.1515/dx-2018-0090. PMID: 30772868.

Braun LT, Borrmann KF, Lottspeich C, Heinrich DA, Kiewewetter J, Fischer MR, Schmidmaier R. *Guessing right - whether and how medical students give incorrect reasons for their correct diagnoses*. *GMS J Med Educ*. 2019 Nov 15;36(6):Doc85. doi: 10.3205/zma001293. PMID: 31844657; PMCID: PMC6905369.

Braun LT, Lenzer B, Kiewewetter J, Fischer MR, Schmidmaier R. *How case representations of medical students change during case processing - Results of a qualitative study*. *GMS J Med Educ*. 2018 Aug 15;35(3):Doc41. doi: 10.3205/zma001187. PMID: 30186951; PMCID: PMC6120161.

Braun LT, Lenzer B, Fischer MR, Schmidmaier R. *Complexity of clinical cases in simulated learning environments: proposal for a scoring system*. *GMS J Med Educ*. 2019 Nov 15;36(6):Doc80. doi: 10.3205/zma001288. PMID: 31844652; PMCID: PMC6905356.

Clinical Reasoning umfasst wie oben bereits dargestellt nicht nur das Diagnoseergebnis – die richtige oder falsche Diagnose –, sondern auch den gesamten Diagnoseprozess, der zur

Stellung der Diagnose führt. Ein Aspekt des Diagnoseprozesses ist die Diagnoseeffizienz. Als Diagnoseeffizienz kann man die Zeit, die bis zum Stellen einer richtigen Diagnose verwendet wird, definieren (25). Da im klinischen Alltag die Zeit häufig limitiert ist – sei es in der ambulanten Versorgung, sei es in der Notfallversorgung – ist Effizienz bei der Diagnosestellung essentiell. Die Diagnoseeffizienz kann leicht in online-basierten Fallformaten gemessen werden, in denen die Zeit bis zum Stellen der korrekten Diagnose aufgezeichnet wird. Möchte man allerdings einen tieferen Einblick in den Diagnoseprozess gewinnen, um eventuelle kognitive Fehler zu detektieren, bieten sich andere Verfahren an: Eine typische Methode ist die Anwendung sogenannter Laut-Denk-Protokolle, bei denen der Proband zeitgleich zur Bearbeitung eines Patientenfalls seine Gedanken laut ausspricht (26). Diese werden parallel protokolliert und können dann später qualitativ ausgewertet werden. Problematisch bei diesem Verfahren ist allerdings, dass es sehr aufwändig ist und daher nur eine kleine Zahl an Probanden untersucht werden kann.

In allen im Nachfolgenden beschriebenen Studien wurde daher eine andere Methode genutzt: In meiner ersten Studie mussten die Studierende (1) Patientenfälle auf der virtuellen Lernplattform CASUS (27) bearbeiten. Nach Bearbeitung eines Falls und Nennung einer Diagnose mussten die Studierenden zusätzlich eine Begründung schreiben, in welcher sie darlegen sollten, warum sie diese Diagnose getroffen haben. Insgesamt bearbeiteten 148 Medizinstudierende 15 verschiedene Patientenfälle. Es wurden nur die Diagnosebegründungen bei letztlich korrekt gestellter Diagnose ausgewertet, da Diagnosefehler bereits in einer vorherigen Studie eingehend analysiert wurden (28). In 1.135 Fällen wurde eine richtige Diagnose gestellt. Bei der überwiegenden Mehrheit (92%) lag der richtigen Diagnose auch ein korrekter Diagnoseprozess zugrunde, der sich in einer korrekten Diagnosebegründung widerspiegelte. 7% der Begründungen enthielten Fehler, während bei 1% der Diagnosen von den Studierenden angegeben wurde, dass sie die Diagnose erraten hätten. Beispiele für Diagnosebegründungen werden in Tabelle 1 gezeigt. Gründe für Fehler im Diagnoseprozess sind vorwiegend ein Mangel an pathophysiologischem Wissen (50%) sowie mangelnde diagnostische Skills (Befundung diagnostischer Tests wie Röntgenbilder) (1).

Tabelle 1: Beispiele für Diagnosebegründungen

Kategorie	Definition	Beispiel
Korrekte Begründung	Diagnose wird anhand der passenden Befunde in der Anamnese, der körperlichen Untersuchung oder der technischen Untersuchungen erklärt und enthält keine fehlerhaften Aspekte	„BGA zeigt ein erniedrigten CO ₂ Partialdruck und einen erhöhten PH. Die unauffälligen sonstigen Untersuchungen, sowie der Krampf aufgrund der Hyperventilation passen zu der Diagnose“ (Diagnose: Hyperventilation)
Fehlerhafte Begründung	Ein Teil oder die gesamte Diagnosebegründung ist fehlerhaft	„Lungenfunktion zeigt Restriktion, andere Diagnosen unwahrscheinlich“ (Diagnose: Akuter Asthmaanfall, in der Lungenfunktion ist eine Obstruktion zu sehen, keine Restriktion)
Richtig erratene Diagnosen	Die Probanden geben selbst an, die Diagnose geraten zu haben	„Ich hab keine Ahnung, um was für eine Diagnose es sich handeln könnte.“ (Diagnose: Pneumothorax)

aus: Braun LT al. GMS journal for medical education. 2019;36(6):Doc85.

Es ist wichtig, dass einer korrekt gestellten Diagnose auch ein korrekter Diagnoseprozess zugrunde liegt. Anderenfalls ist nicht davon auszugehen, dass ein anderer Patient mit gleicher Diagnose korrekt diagnostiziert werden kann. Erfreulicherweise zeigte sich in dieser Studie, dass richtig geratene Diagnosen nur sehr selten vorkommen. Fehlerhafte Diagnoseprozesse sind hingegen nicht selten. Diagnosebegründungen sind somit eine wichtige Methode, um die *Clinical Reasoning*-Kompetenz einzuschätzen.

Eine weitere Methode, um den Diagnoseprozess von Studierenden zu erfassen, wurde in meiner zweiten Studie genutzt: Es erfolgte die qualitative Auswertung von Fallrepräsentationen. Fallrepräsentationen sind innere Zusammenfassungen eines Patientenfalls, welche dem Studierenden oder der Ärztin oder dem Arzt helfen können, den Fall zu strukturieren und damit richtig zu lösen. Sie können bewusst eingesetzt werden, um den Diagnoseprozess zu verbessern und die Diagnoseeffizienz zu erhöhen (25). Gleichzeitig können aber auch die Fallrepräsentationen selbst ausgewertet werden, um einen Einblick in den Diagnoseprozess zu erhalten. Es kann codiert werden, welche Informationen – von allen im Fall in seiner Gesamtheit enthaltenen Informationen – in die Repräsentation aufgenommen werden. Ein Codierbeispiel zeigt Abbildung 4. In dieser Studie wurden insgesamt 516 Fallrepräsentationen von 43 Medizinstudierenden im klinischen Studienabschnitt ausgewertet. Die Studierende bearbeiten vier klinische Patientenfälle mit den Diagnosen Aortenklappenstenose, Urämie, Hyperventilation und AV-Knoten-Reentrytachykardie. Während jeder der vier Fallbearbeitungen wurden die Studierenden dreimal dazu aufgefordert, eine Fallrepräsentation zu schreiben. Dazu erhielten sie den folgenden Arbeitsauftrag: „Bitte fassen Sie den Fall so zusammen, wie Sie ihn Ihrem Oberarzt präsentieren würden.“ Diese

Aufforderung erhielten die Probanden nach der Anamnese, nach der körperlichen Untersuchung sowie ein letztes Mal nach den technischen Befunden, bevor sie die endgültige Diagnose stellten.

Abbildung 4: Codierung einer Fallrepräsentation

Repräsentation 3, Fall Urämie, falsche Diagnose: Herr Schuhmacher ist ein 82-jähriger Patient mit bekannter KHK, der sich aufgrund von zunehmender Dyspnoe und Erbrechen vorstellt. In der körperlichen Untersuchung fallen ein deutlich reduzierter AZ, sowie eine Hypertonie, Tachykardie und Tachypnoe auf. Das EKG zeigt keinen Hinweis auf einen Herzinfarkt, trotzdem Troponinerhöhung.

A	A	A	A	A	A	A	An	A
Geschlecht	Alter	Vorerkrankungen	Medikation	Alkohol-konsum	Nikotin-konsum	Aktueller Krankheitsbeginn	Verlauf	Dyspnoe
Herr	82-jähriger	bekannter KHK	x	x	x	x	zunehmender	Dyspnoe
A	A	A						
Müdigkeit	Übelkeit	Erbrechen						
x	x	Erbrechen						
KU	KU	KU	KU	KU	KU	KU	KU	
Vital-parameter	Allgemeinzustand	Ernährungszustand	Kardiologie	Pulmo	Abdomen	Neurostatus	Lymphknoten	
Hypertonie, Tachykardie und Tachypnoe	deutlich reduzierter AZ	x	x	x	x	x	x	
TU	TU	TU	TU	TU	TU	TU	TU	TU
Labor	EKG	UKG	Röntgen-Thorax	BGA	Urinstix	Lufu	Sonografie Abdomen	Bakteriologie
trotzdem Troponinerhöhung.	Das EKG zeigt keinen Hinweis auf einen Herzinfarkt	x	x	x	x	x	x	x

A=Anamnese, KU=Körperliche Untersuchung, TU=Technische Untersuchung, EKG=Elektrokardiogramm, UKG=Echokardiographie, BGA=Blutgasanalyse, Lufu=Lungenfunktionsprüfung

aus: Braun LT et al. GMS journal for medical education. 2018;35(3):Doc41.

Zu Beginn einer Fallbearbeitung schließen Studierende den Großteil der gegebenen Informationen in ihre Fallrepräsentation ein (66%). Je mehr Informationen hinzukommen, desto mehr wird selektiert: Nachdem alle Informationen geboten wurden, werden nur noch rund 38% der Informationen eingeschlossen. Die Selektion der relevanten Informationen gelingt also. Es zeigte sich allerdings kein Zusammenhang zwischen der Länge der Repräsentationen und der Lösungsrichtigkeit oder zwischen der Schwierigkeit des Falls und der Fallrepräsentation (29).

Um den Einfluss verschiedener Methoden auf die Diagnoserichtigkeit und den Diagnoseprozess zu untersuchen, wurde als drittes eine Interventionsstudie mit 148 Medizinstudierenden durchgeführt (30). Die Probandinnen und Probanden wurden in fünf Gruppen eingeteilt. Zwei Gruppen erhielten als Intervention die Aufforderung, während der Fallbearbeitung eine Fallrepräsentation zu schreiben; die anderen beiden Gruppen erhielten Scaffolds zur strukturierten Reflexion (31, 32). Die strukturierte Reflexion ist eine schon seit vielen Jahren angewandte Intervention, die sich in vielen Studien als hilfreich erwiesen hat, um die Diagnoserichtigkeit zu erhöhen. Es gibt verschiedene Ansätze zur strukturierten Reflexion; in meiner Studie wurde folgender gewählt: Die Studierenden sollten darüber nachdenken, welche drei Diagnosen in einem Fall am wahrscheinlichsten sind. Des Weiteren sollten sie alle Informationen angeben, die eine bestimmte Diagnose unterstützen, sowie alle Informationen, die gegen diese Diagnose sprechen. Zusätzlich sollten sie angeben, welche Informationen in dem Fall fehlen, aber die Diagnose noch weiter unterstützen würden.

Jeweils eine dieser beiden Gruppen erhielt zusätzlich nach der Fallbearbeitung Feedback zur Falllösung. Nach diesem Schema wurden fünf Fälle bearbeitet, danach schloss sich eine Testphase ohne Interventionen an, in denen weitere zehn Fälle bearbeitet wurden. Fünf der Testfälle wurden direkt im Anschluss bearbeitet, die anderen fünf Fälle eine Woche später. Die 5. Gruppe bearbeitete nur die Testfälle ohne vorherige Intervention. Die Fälle der Testphasen waren thematisch auf die Fälle der Lernphase abgestimmt (Abbildung 5).

Abbildung 5: Variation der Fallthemen

Lernphase	Testphase 1	Testphase 2
<ul style="list-style-type: none"> • Asthmaanfall • Pneumothorax • COPD • Myokarditis • Tuberkulose 	<ul style="list-style-type: none"> • Hyperventilation • Pneumothorax • COPD • Herzinsuffizienz • Choledocholithiasis 	<ul style="list-style-type: none"> • Lungenarterienembolie • COPD • Lungenkarzinom • Tuberkulose • Divertikulose

adaptiert aus: Braun LT et al. Diagnosis (Berlin, Germany). 2019;6(2):137-49.

Entgegen den Erwartungen verbesserte keine der Interventionen die Diagnoserichtigkeit oder Diagnoseeffizienz, sondern alle vier Gruppen bearbeiteten die Fälle ebenso gut wie die Kontrollgruppe. Dies könnte an der kurzen Interventionsphase liegen, ist aber widersprüchlich zu bisherigen Studienergebnissen (25, 33). Andere Erklärungen sind, dass beispielsweise das gegebene Feedback zu standardisiert oder zu kurz war. Ebenso könnte der hohe Komplexitätsgrad der Fälle eine Erklärung für den ausbleibenden Effekt der Interventionen sein.

Es wurden zusätzlich auch noch die Gründe für Diagnosefehler ausgewertet. In einer vorherigen Studie wurde bereits ein Codierschema für die Klassifikation kognitiver Fehler entwickelt (28). Nach diesem Schema können acht verschiedene Fehlerursachen unterschieden werden:

Fehlertyp	Definition
Unzureichendes Wissen	Mangelndes (pathophysiologisches) Wissen
Unzureichende diagnostische Fähigkeiten	Unzureichende Fähigkeit, technische Befunde (EKGs, Röntgenaufnahmen) zu interpretieren
Falsche Kontextgenerierung	Wichtige Fallaspekte werden nicht wahrgenommen
Überbewertung/Unterbewertung	Über- oder Unterbewertung der Relevanz von Fallinformationen
<i>Faulty Triggering</i>	Unangemessene Schlussfolgerungen
Verwechslung	Ähnlich wirkende Diagnosen werden verwechselt

Voreilige Schlussfolgerung	Andere mögliche Diagnosen werden nicht in Betracht gezogen; auf der Basis weniger Informationen wird rasch im Fallverlauf eine Diagnose gestellt
Keine Hypothese	Es wird keine finale Diagnose gestellt.

Insgesamt waren ein Mangel an diagnostischen Fähigkeiten (20%) sowie mangelndes Wissen (18%) die Hauptursachen für Diagnosefehler. Interessanterweise zeigte sich eine unterschiedliche Verteilung der Fehlerursachen zwischen den Gruppen: Studierende, die Feedback in der Interventionsphase erhielten, machten mehr Fehler aufgrund von Verwechslung von Diagnosen und voreiliger Schlussfolgerung. Dies unterstreicht, wie sehr das Diagnoseverhalten dadurch beeinflusst wird, mit welchen Diagnosen man vorab konfrontiert wird.

Ergänzend zu den obigen Studien publizierten wir in diesem Zusammenhang eine Methodenarbeit, in der ein Ratingverfahren zur Bewertung der Komplexität von klinischen Fällen vorgestellt wurde (34). In der Medizindidaktik werden Patientenfälle oder virtuelle Patientenfälle sowohl zu Lehrzwecken als auch zur Testung der Diagnosekompetenz eingesetzt. Hierbei werden oft die beiden Begrifflichkeiten *Schwierigkeit* und *Komplexität* vermischt, dabei lassen sich beide Begriffe sinnvoll differenzieren: Während die Schwierigkeit eines Falls von der diagnostizierenden Person abhängt – abhängig von Vorwissen, Vorerfahrung etc. - ist die Komplexität auf den Fall selbst bezogen (Abbildung 6). Je nach Informationsfülle und Widersprüchlichkeit der gebotenen Informationen kann ein Fall als mehr oder weniger komplex eingestuft werden. In dem vorgestellten Rating kann ein Fall mit maximal zehn Punkten (extrem komplex) und minimal mit einem Punkt (gar nicht komplex) bewertet werden. Es werden in fünf verschiedenen Kategorien Punkte vergeben: Anamnese, Körperliche Untersuchung, Technische Untersuchungen, Psychosoziale Aspekte, Diagnose & Nebendiagnosen. Ist eine Information im Fall enthalten, wird in der entsprechenden Kategorie ein Punkt gegeben. Es wird für alle zusätzlichen Informationen, die nicht im Zusammenhang mit dem eigentlichen Fall stehen und eine mögliche Nebendiagnose zulassen würden, ein weiterer Punkt gegeben. Das Scoringverfahren kann eingesetzt werden, um klinische Fälle in ihrer Komplexität für eine bestimmte Gruppe Medizinstudierender angemessen zu konzipieren. Des Weiteren ist es hilfreich, um Fälle, die zur Testung der Diagnosekompetenz genutzt werden, einheitlich und angemessen zu gestalten. Es könnte auch ein weiteres Problem in der medizindidaktischen Forschung lösen: Die Ergebnisse von Interventionsstudien lassen sich nicht immer an einem anderen Kollektiv Studierender replizieren. Dies könnte an unterschiedlichen Komplexitätsgraden der eingesetzten Fälle liegen. Um Studien methodisch besser durchführen zu können, wäre es sinnvoll, die Komplexitätsgrade der genutzten Fälle einheitlich anzugeben.

Abbildung 6: Differenzierung von Schwierigkeit und Komplexität

Statisch			Dynamisch	
niedrig ↑		Diagnose- prozess und Diagnose- ergebnis →Effizienz →Epistemische Aktivitäten →Fehlerarten →Korrekte Diagnose	hoch ↑	
Klinischer Fall →	Komplexität →		Schwierigkeit ←	Diagnostizierender mit bestimmter Diagnose- kompetenz ←
↓ hoch			↓ niedrig	

aus: Braun LT et al. GMS journal for medical education. 2019;36(6):Doc80.

Bayesian Reasoning in der Endokrinologie

Zugrunde liegende Publikationen:

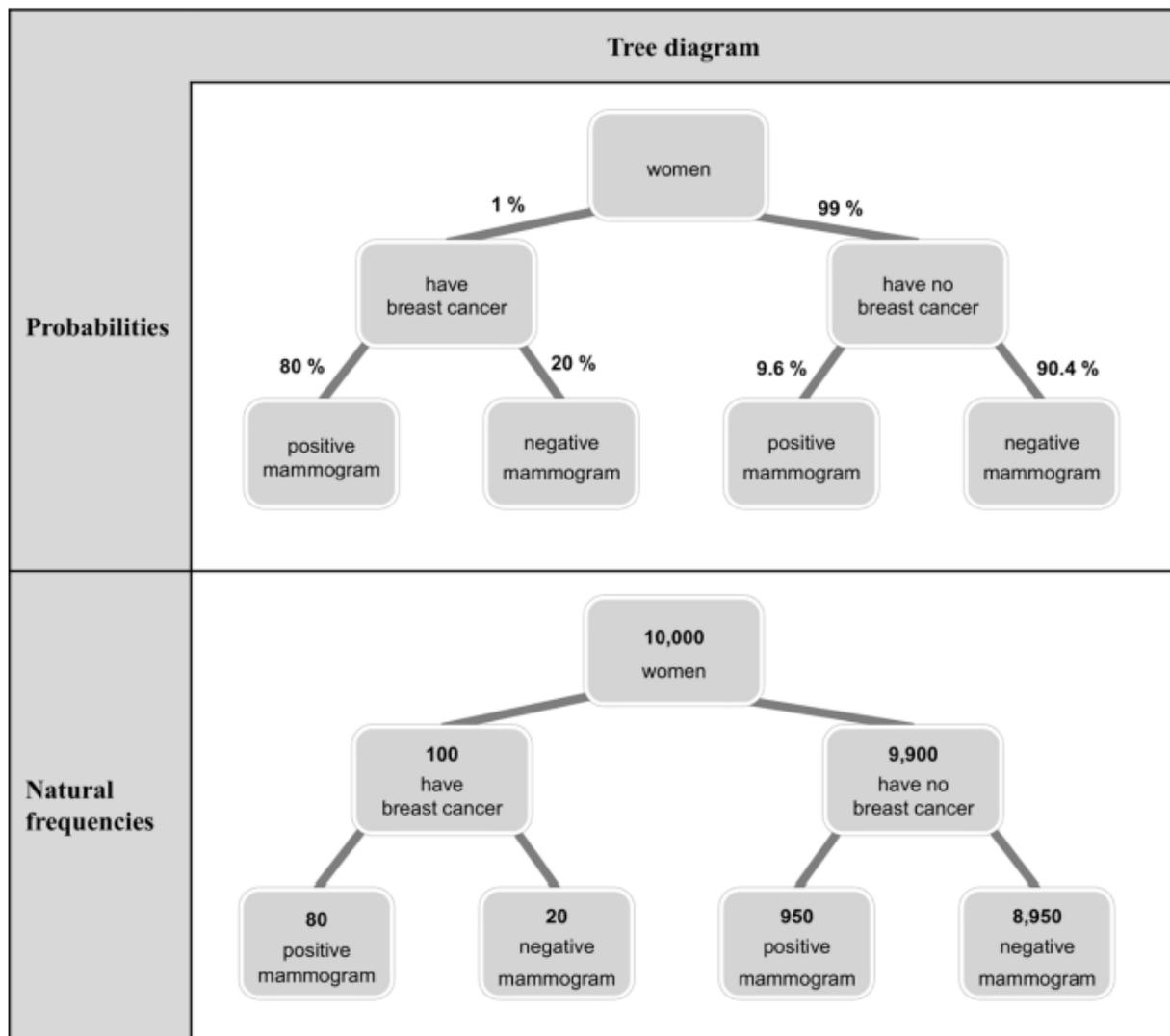
Binder K, Krauss S, Schmidmaier R, **Braun LT**. *Natural frequency trees improve diagnostic efficiency in Bayesian reasoning*. Adv Health Sci Educ Theory Pract. 2021 Feb 12. doi: 10.1007/s10459-020-10025-8. Epub ahead of print. PMID: 33599875.

Kunzelmann AK, Binder K, Fischer MR, Reincke M, **Braun LT***, Schmidmaier R*. Improving Diagnostic Efficiency with Frequency Double-Trees and Frequency Nets in Bayesian Reasoning. MDM Policy Pract. 2022 Mar 16;7(1):23814683221086623. doi: 10.1177/23814683221086623. PMID: 35321028; PMCID: PMC8935422.

Da es weithin bekannt ist, dass sowohl Ärztinnen und Ärzte als auch Medizinstudierende große Probleme bei der Lösung von Bayesian-Aufgaben haben, wurde in einer weiteren meiner Studien der Einfluss von verschiedenen Visualisierungsmethoden von Bayes-Aufgaben auf die Richtigkeit und Effizienz der Lösungen untersucht. Mit der Formel nach Bayes lässt sich anhand der Prävalenz einer Erkrankung sowie der Sensitivität eines Tests berechnen, wie wahrscheinlich es ist, dass ein Patient mit einem positiven Testergebnis tatsächlich unter einer bestimmten Erkrankung leidet. Je niedriger die Prävalenz einer Erkrankung ist, desto sensitiver muss ein Test sein, um die kranken Patientinnen und Patienten herauszufiltern. In Standardtextaufgaben, in denen die Wahrscheinlichkeiten für die genannten Parameter angegeben sind, sind die Lösungsraten extrem gering (35).

In meiner Studie bearbeiteten 111 Medizinstudierende jeweils vier verschiedene Bayes-Aufgaben. Sowohl die Lösungsrichtigkeit als auch die Zeit bis zur Diagnosestellung wurden ausgewertet. Die Studierenden wurden mit vier verschiedenen Aufgabentypen konfrontiert: Jeweils zwei Aufgaben mittels Baumdiagramm und zwei Aufgaben, in denen die Informationen als Text gegeben wurden. In beiden Aufgabentypen wurden die Informationen einmal als natürliche Häufigkeiten und einmal als Wahrscheinlichkeiten angegeben. Die Baumdiagramme, sowohl in der Version mit natürlichen Häufigkeiten als auch mit Wahrscheinlichkeiten, sind in Abbildung 7 gezeigt.

Abbildung 7: Das Baumdiagramm



aus: Binder K et al. *Advances in health sciences education: theory and practice*. 2021;26(3):847-63.

Die Abbildung zeigt das klassische Baumdiagramm mit Angabe der Zahlen als Wahrscheinlichkeiten und als natürliche Häufigkeiten. Dargestellt wird eine Gruppe von (10.000) Frauen, die am Mammographiescreening teilnehmen. In der zweiten Zeile ist die tatsächliche Prävalenz der Erkrankten aufgeführt; in der dritten Zeile die jeweiligen Testergebnisse der Mammographie bei den jeweiligen Frauen.

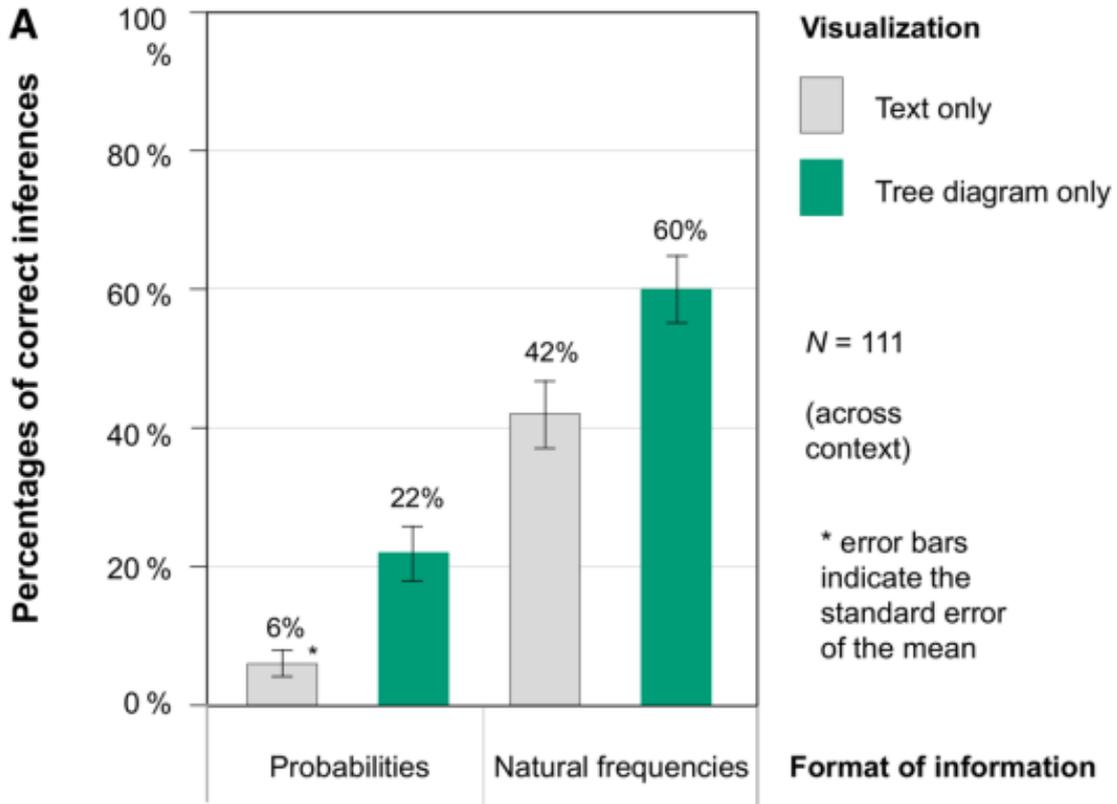
Somit bearbeitete jeder Studierende vier Aufgaben; die medizinischen Kontexte waren bei jeder Aufgabe unterschiedlich, sodass sich letztlich ein 2x2x4-Modell ergab. Die medizinischen Kontexte waren das Mammographiescreening, HIV, Trisomie 21 und Röteln. Eine Beispielaufgabe lautete folgendermaßen:

	Trisomy	
	Probability version	Natural frequency version
Medical situation	Imagine being a gynecologist in your own practice. For each pregnant woman, you perform a triple test for prenatal diagnosis between the 15th and 18th week of pregnancy in order to detect a possible trisomy 21 in the unborn child	
	You are currently counseling a pregnant woman who has received a positive test result in the triple test. This woman wants to know what this means for her unborn child	
	For your answer, only the following information is available, based on a sample of pregnant women who have also undergone a triple test	
Presentation of information	Text only	Text only
	Tree diagram only	Tree diagram only
Text	The probability that the unborn child has trisomy 21 is 0.18%	12 out of 6,760 unborn children have trisomy 21
	The likelihood that a pregnant woman will receive a positive triple test when the unborn child has trisomy 21 is 75%	In 9 of 12 unborn children with trisomy 21, the pregnant woman receives a positive triple test
	The likelihood that a pregnant woman will receive a positive triple test by mistake even though the unborn child does not have trisomy 21 is 5.9%	In 395 out of 6,748 unborn children without trisomy 21, the pregnant woman is mistakenly receiving a positive triple test
Tree diagram	Probability tree (in the version with a tree diagram)	Frequency tree (in the version with a tree diagram)
Question	What is the probability that her unborn child will actually have Trisomy 21?	How many unborn children with a positive triple test actually have trisomy 21?
	Answer: _____	Answer: ____ out of ____

aus: Binder K et al. *Advances in health sciences education: theory and practice*. 2021;26(3):847-63.

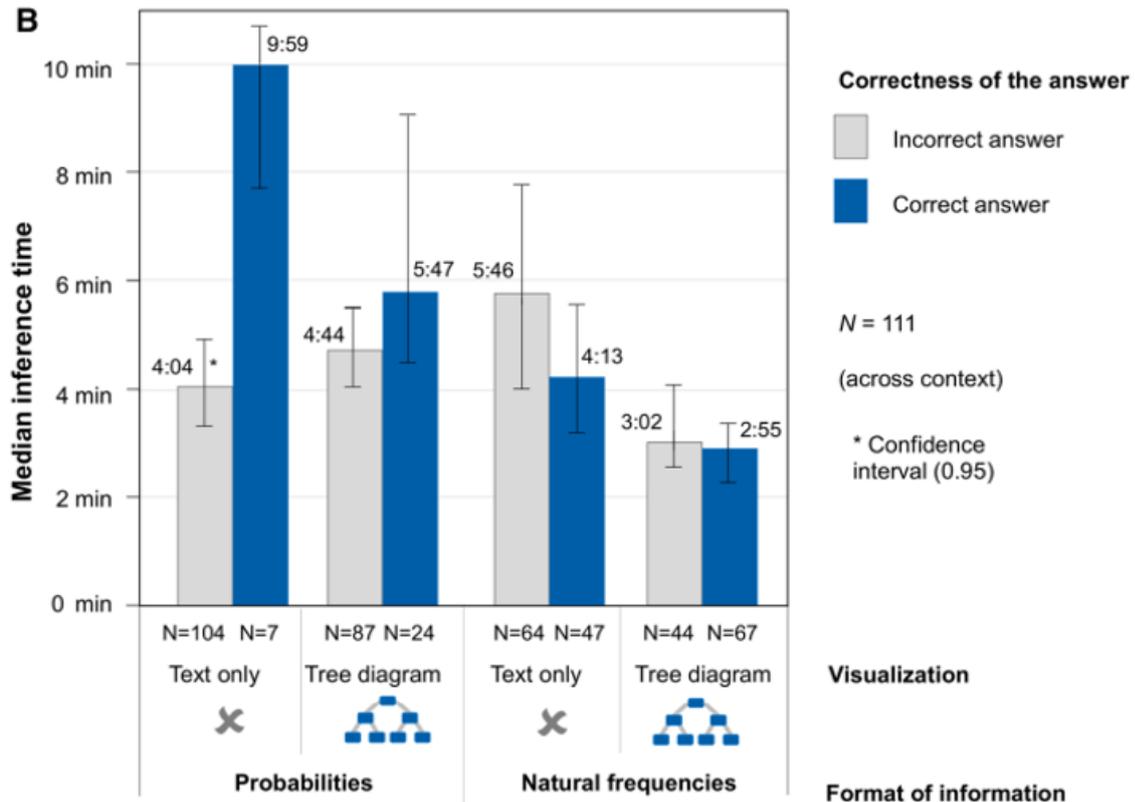
Die Angaben von Zahlen als natürliche Häufigkeiten und die Darstellung als Baumdiagramme tragen wesentlich zu einem besseren Verständnis von *Bayesian-Reasoning* bei: Die Studierenden konnten diesen Aufgabentyp am besten lösen, wobei sich deutlich zeigt, dass insbesondere der Faktor *Natürliche Häufigkeiten* die Lösungsrichtigkeit erhöht (Abbildung 8). Aufgaben mit Angabe der Zahlen als Wahrscheinlichkeiten können nur schlecht gelöst werden – unabhängig davon, ob die Zahlen als Text oder Baumdiagramm angeführt werden. Dennoch zeigte sich, dass selbst in der optimalen Kombination (Baumdiagramm + natürliche Häufigkeiten) nur 60% der Aufgaben korrekt gelöst wurden. Übertragen auf den klinischen Alltag ist dies natürlich ein deutlicher Zugewinn, aber lässt weiterhin Raum für Verbesserungen. Neben der Richtigkeit zeigte sich auch ein deutlicher Einfluss der Visualisierung auf die Geschwindigkeit der Bearbeitung. Für korrekt gelöste Aufgaben mit reinem Text und Angabe der Wahrscheinlichkeiten benötigten die Studierenden mehr als 3x so lange wie für Aufgaben in Baumformat mit Angabe der Wahrscheinlichkeiten als natürliche Häufigkeiten (Abbildung 9).

Abbildung 8: Lösungsrichtigkeiten bei den verschiedenen Bayes-Aufgaben



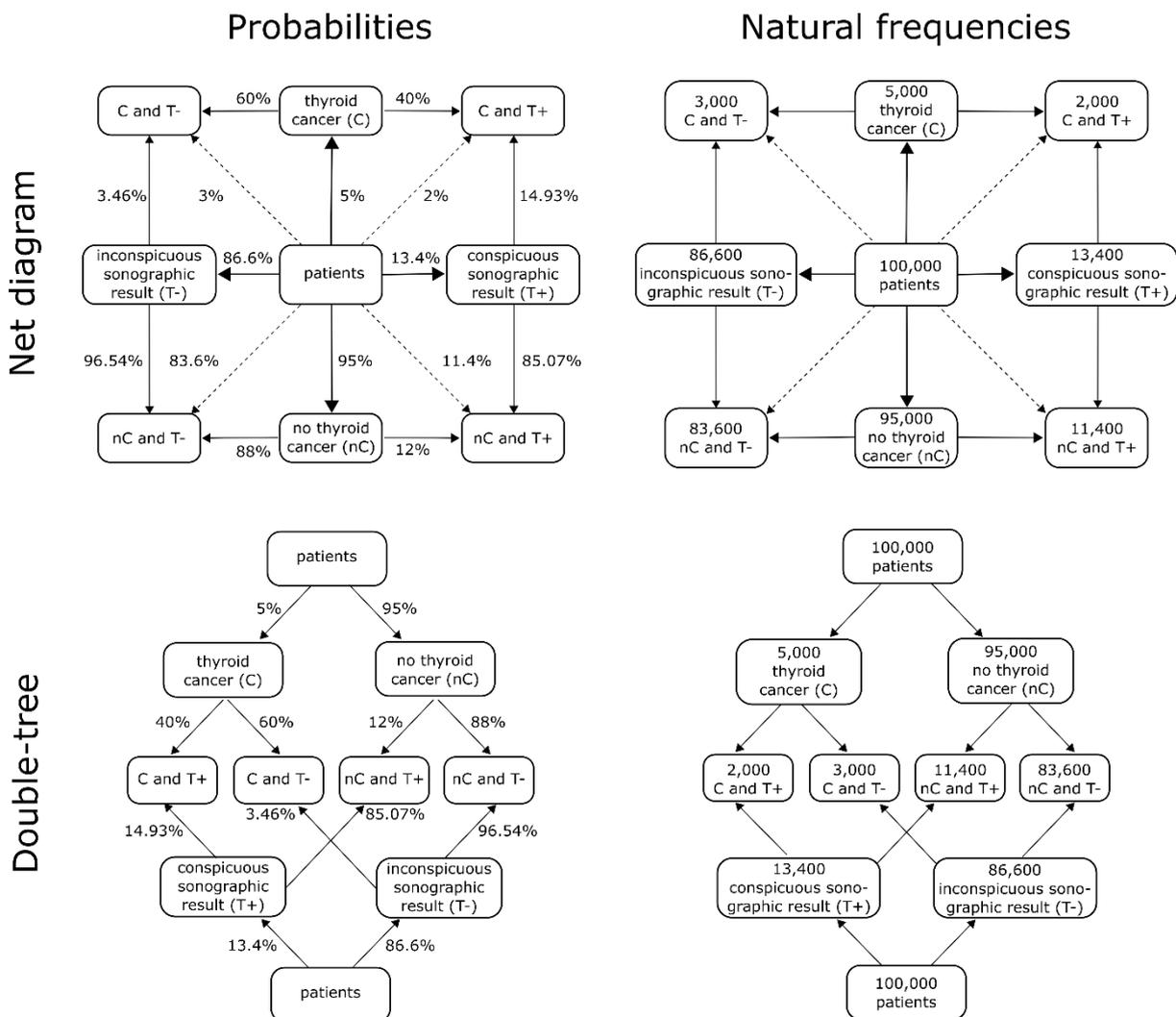
aus: Binder K et al. *Advances in health sciences education: theory and practice*. 2021;26(3):847-63.

Abbildung 9: Zeitlicher Aufwand je nach Falllösung



In einer weiteren Studie zum Thema *Bayesian Reasoning* bearbeiteten 142 Medizinstudierende von verschiedenen deutschen Universitäten ebenfalls vier verschiedene Aufgaben (36). Diesmal wurden allerdings zwei andere Visualisierungsmöglichkeiten gewählt: Der Doppelbaum sowie Netzdiagramme (Abbildung 10). Bei diesen Visualisierungsmöglichkeiten muss die Lösung nicht mehr errechnet werden, sondern kann direkt an den Diagrammen abgelesen werden. Auch in dieser Studie wurden Wahrscheinlichkeiten und natürliche Häufigkeiten verglichen. Als Fallkontexte wurden vier verschiedene Szenarien aus der Endokrinologie gewählt: Conn-Syndrom, Cushing-Syndrom, Schilddrüsenkarzinom, familiäre hypocalciurische Hyperkalzämie. Outcome-Variablen waren erneut die Richtigkeit der Lösung sowie die Zeit bis zum Stellen der richtigen Lösung.

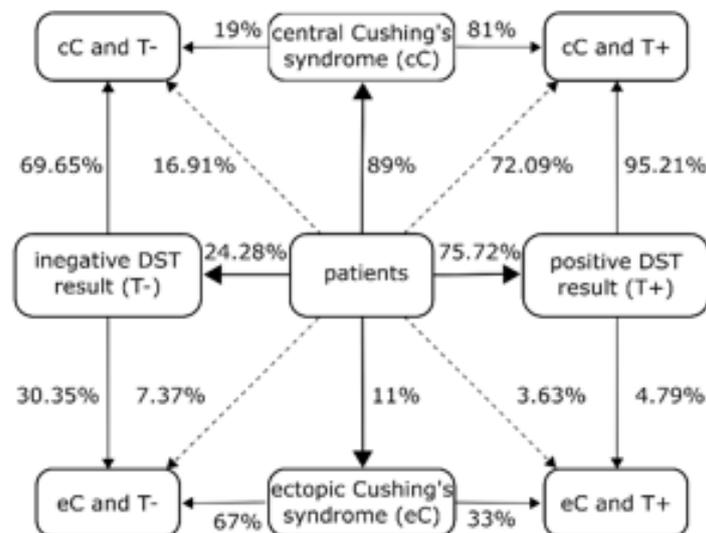
Abbildung 10: Doppelbaum und Netzdiagramme



aus: Kunzelmann AK et al. *MDM policy & practice*. 2022;7(1):23814683221086623.

Da hier der Kontext Cushing-Syndrom genutzt wurde, soll genauer auf dieses klinische Problem eingegangen werden. Das Cushing-Syndrom kann in ACTH-abhängige und ACTH-unabhängige Formen unterschieden werden. Das ACTH-abhängige Cushing-Syndrom kann weiter unterteilt werden in den wesentlich häufigeren Morbus Cushing (zentrales Cushing-Syndrom) und in das ektope Cushing-Syndrom. Dabei sind 89% der ACTH-abhängigen Cushing-Syndrome dem zentralen Cushing-Syndrom zuzuordnen. Die restlichen 11% sind ektope Cushing-Syndrome. Zur Subdifferenzierung zwischen den Formen können drei in Deutschland gängige Tests durchgeführt werden: der CRH-Test, der hochdosierte 8mg-Dexamethason-Suppressionstest sowie der Sinus-petrosus-Katheter, der allerdings ein invasiver Test ist. Der 8mg-Dexamethason-Hemmtest wird noch häufig durchgeführt, obwohl die Sensitivität gering ist. Dies wird aus dem Netzdiagramm ersichtlich: Während die Wahrscheinlichkeit für ein zentrales Cushing-Syndrom initial bei 89% liegt, wird sie durch ein positives Testergebnis lediglich auf 95% angehoben. Der klinische Nutzen ist somit von eher untergeordneter Bedeutung, dennoch wird der Test im Alltag noch häufig für Diagnose- und Managemententscheidungen herangezogen.

Abbildung 11: Der 8mg-Dexamethason-Hemmtest zur Subdifferenzierungsdiagnostik

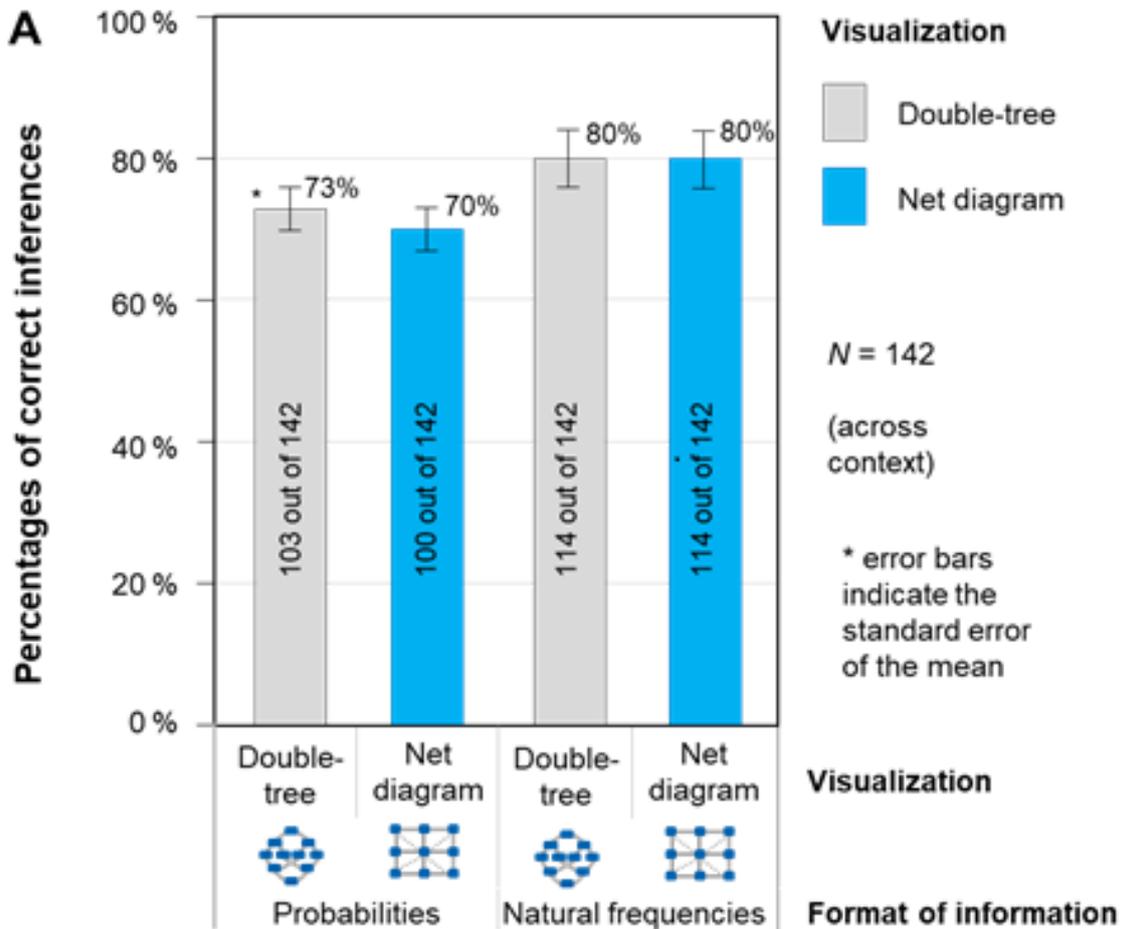


aus: Kunzelmann AK et al. *MDM policy & practice*. 2022;7(1):23814683221086623.

Bei diesen Visualisierungen zeigte sich eine deutlich höhere Lösungsrichtigkeit als bei Textaufgaben oder den einfachen Baumdiagrammen. Je nach Aufgabentyp wurde in 70-80% der Fälle die Aufgabe korrekt gelöst. Es zeigte sich auch hier eine geringe Überlegenheit der natürlichen Häufigkeiten im Vergleich zu Wahrscheinlichkeiten. Die Art der Visualisierung – sei es als Netz, sei es als Doppelbaum – führte nicht zu unterschiedlichen Ergebnissen (Abbildung 12). Auch der Aufgabenkontext hatte keinen Einfluss auf die Richtigkeit der

Lösung. Allerdings zeigte sich ein deutlicher Lerneffekt: Die erste Aufgabe, welche die Probandinnen und Probanden erhielten, wurde signifikant schlechter gelöst als die folgenden Aufgaben.

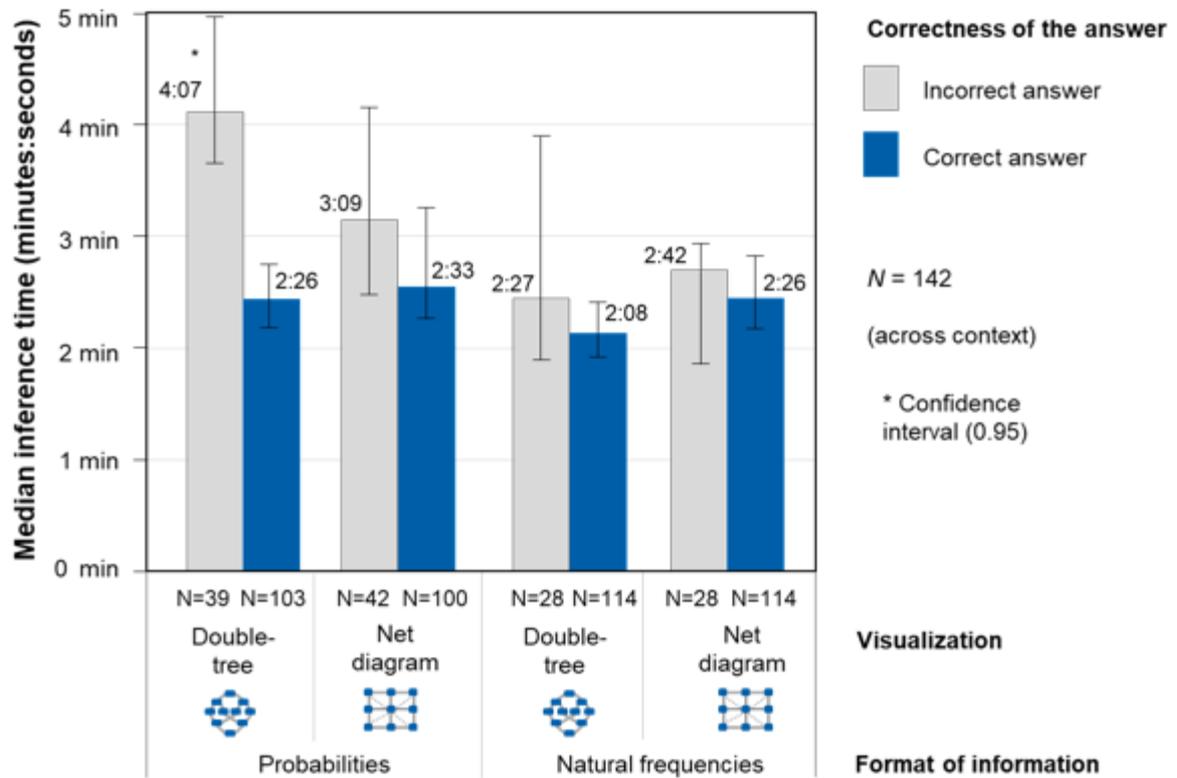
Abbildung 12: Lösungsrichtigkeit je nach Aufgabentyp



aus: Kunzelmann AK et al. *MDM policy & practice*. 2022;7(1):23814683221086623.

Die Bearbeitungszeiten schwankten gering zwischen korrekt und inkorrekt gelösten Aufgaben (Abbildung 13). Bei korrekt gelösten Aufgaben zeigten sich keine Unterschiede zwischen den Aufgabenformaten, sondern die Aufgaben konnten im Schnitt in 2:08 – 2:33 Minuten gelöst werden. Dies leuchtet ein, da in allen Fällen die korrekte Lösung einfach abgelesen werden konnte, was naturgemäß – wurde das Konzept einmal verstanden – sehr schnell gehen kann. Bei inkorrekt gelösten Aufgaben wandten die Studierenden insbesondere mehr Zeit auf die Aufgaben mit Wahrscheinlichkeiten auf. Zudem benötigten die Studierenden mehr Zeit für die Kontexte Cushing-Syndrom und Familiäre hypocalciurische Hyperkalzämie.

Abbildung 13: Zeitspanne zum Lösen der Aufgaben



aus: Kunzelmann AK et al. *MDM policy & practice*. 2022;7(1):23814683221086623.

Die beiden Studien zeigen, dass das Lösen von Bayes-Aufgaben problematisch sein kann, sich aber die Lösungsraten und auch die Zeiten zum Lösen der Aufgaben durch verschiedene Formen der Visualisierung und die Darstellung der Zahlen als natürliche Häufigkeiten verbessern lassen. Solche Bayes-Szenarien begegnen Ärztinnen und Ärzten immer wieder im klinischen Alltag. Der Einsatz von Häufigkeitsnetzen oder Doppelbäumen könnte hierbei das Verständnis für diese Aufgaben deutlich erhöhen und damit die Beratung der Patientinnen und Patienten verbessern.

Therapiefehler in der Endokrinologie

Zugrunde liegende Publikation:

Tausendfreund O*, **Braun LT***, Schmidmaier R. Types of therapeutic errors in the management of osteoporosis made by physicians and medical students. BMC Med Educ. 2022 Apr 27;22(1):323. doi: 10.1186/s12909-022-03384-w. PMID: 35473636.

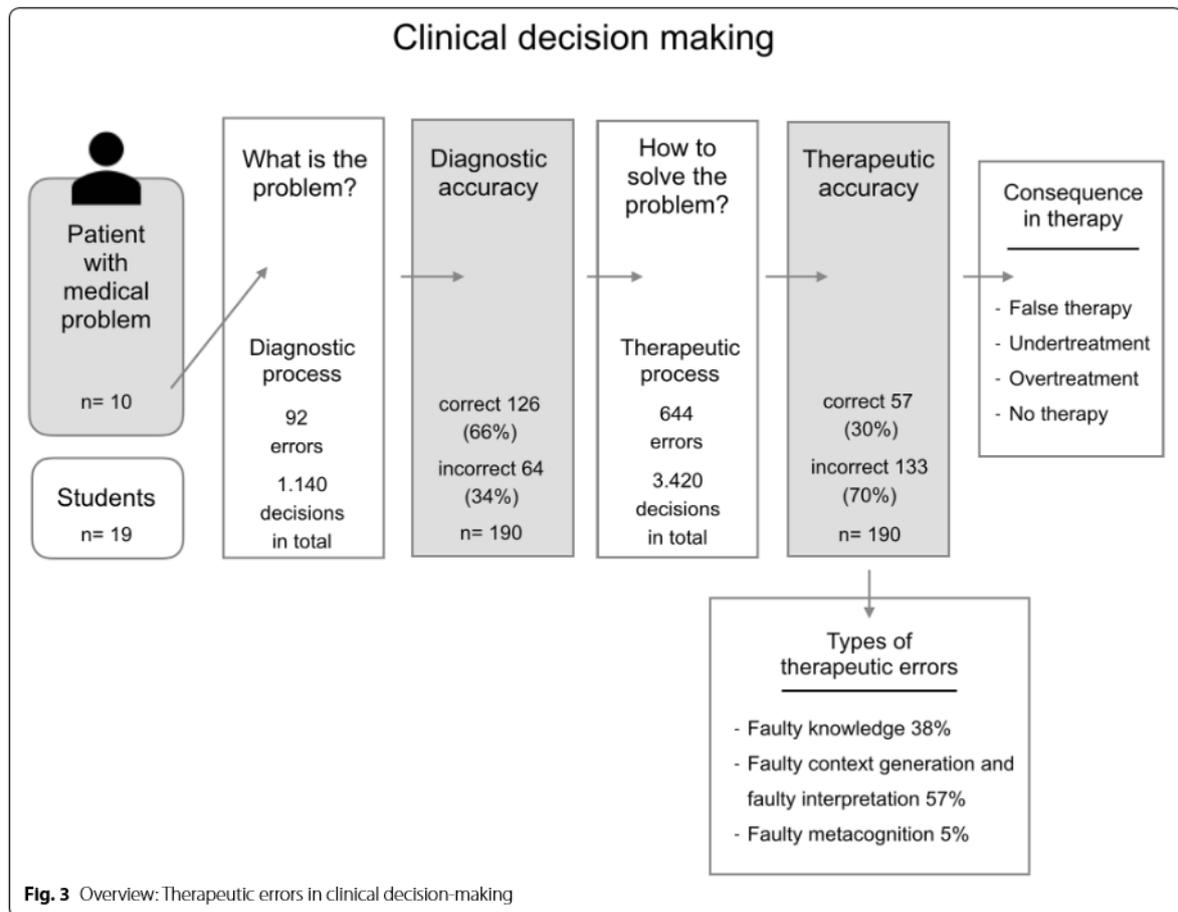
Der Diagnoseprozess endet nicht mit der korrekten Diagnosestellung, sondern nachdem eine Diagnose gestellt wurde, muss eine Therapieentscheidung getroffen werden. Therapieentscheidungen sind potenziell sogar noch komplexer als Diagnoseentscheidungen, da sie eine gemeinsame Entscheidungsfindung mit dem betroffenen Patienten/der Patientin (und in einigen Fällen auch mit den Angehörigen) verlangen (37). Zusätzlich stehen bei vielen Erkrankungen verschiedene Therapiekonzepte und Medikamente zur Verfügung, die einer sorgfältigen Abwägung bedürfen. Das Cushing-Syndrom kann beispielsweise neben der operativen Therapie auch medikamentös sowie durch die bilaterale Adrenalektomie oder eine Strahlentherapie behandelt werden (38).

Therapiefehler in der Inneren Medizin wurden bislang noch nicht umfassend analysiert, sodass wenig über die kognitiven und systemischen Ursachen bekannt ist. In meiner oben genannten Studie wurden daher Therapiefehler am Beispiel der Osteoporose unter kontrollierten Bedingungen analysiert (39). Osteoporose wurde als Kontext gewählt, da hier die Therapieempfehlungen ziemlich klar festgelegt sind, sodass Therapieentscheidungen leicht als „korrekt“ oder „inkorrekt“ klassifiziert werden können.

Studierende im klinischen Studienabschnitt sowie praktizierende Ärztinnen und Ärzte aus Fachbereichen, die potenziell im klinischen Alltag Patienten und Patientinnen mit Osteoporose behandeln müssen, bearbeiteten auf der elektronischen Lernplattform CASUS zehn verschiedene Patientenfälle. In jedem Fall mussten sie entscheiden, ob eine spezifische Osteoporosetherapie indiziert ist, wenn ja, welche Therapie für den jeweiligen Erkrankten am besten geeignet ist und zuletzt, ob Kontraindikationen gegen bestimmte medikamentöse Therapien vorliegen. Zudem sollten die Probandinnen und Probanden entscheiden, ob die Überweisung an einen Osteologen oder eine Osteologin zur weiteren Abklärung erforderlich ist.

Die Probanden und Probandinnen machten insgesamt bei 70% der Fälle einen oder mehrere Fehler. Die kognitiven Ursachen für Fehler ließen sich den folgenden Kategorien zuordnen: *faulty knowledge* (38%), *faulty context generation and interpretation* (57 %) und *faulty metacognition* (5%) (Abbildung 14).

Abbildung 14: Studienverlauf und Ergebnisse



aus: Tausendfreund O et al. BMC medical education. 2022;22(1):323.

Innerhalb dieser Kategorien wurde eine weitere Differenzierung vorgenommen: Die wichtigste Subkategorie, welche die meisten Fehler bedingt, ist „*Faulty interpretation of results resulting in undertreatment*“ (Abbildung 15).

Abbildung 15: Fehlerkategorien

Table 2 Cognitive contributions to therapeutic errors

Type	Students' therapeutic errors (N = 644)	Physicians' therapeutic errors (N = 223)
a. Faulty knowledge	247 (38%)	78 (35%)
<i>Lack of knowledge of a necessary therapeutic action</i>	119	29
<i>Lack of knowledge of a special indication</i>	96	30
<i>Lack of knowledge of contraindications</i>	32	19
b. Faulty context generation and interpretation	365 (57%)	128 (57%)
<i>Misidentification of information as a contraindication</i>	39	16
<i>Failure in recognizing contraindications</i>	7	2
<i>Underestimation of a finding in the process of considering patients' individual risk</i>	100	43
<i>Faulty interpretation of results resulting in "undertreatment"</i>	155	48
<i>Overestimation of a finding in the process of considering patients' individual risks</i>	30	7
<i>Faulty interpretation of results resulting in "overtreatment"</i>	34	9
<i>Failure to leave the common path of procedures</i>	0	3
c. Faulty metacognition	32 (5%)	17 (8%)
<i>Lack of confidence</i>	22	12
<i>Possible overconfidence</i>	10	5

aus: Tausendfreund O et al. *BMC medical education*. 2022;22(1):323.

Auch insgesamt betrachtet war die Therapiekonsequenz in den meisten Fällen eine falsche oder ungenügende Therapie, während eine Übertherapie nur in wenigen Fällen zu beobachten war. Dies scheint typisch für den Kontext Osteoporose zu sein, wie auch schon aus der Literatur bekannt (40). Für andere Therapieentscheidungen, wie zum Beispiel die Verschreibungsrate von Antibiotika, ist ein umgekehrter Effekt mit einer Tendenz zur Übertherapie zu sehen (41).

In dieser Studie konnten erstmals die kognitiven Ursachen für Therapiefehler systematisch untersucht werden. Damit bietet diese Studie einerseits die methodischen Grundlagen, um Therapiefehler auch in anderen Kontexten systematisch auszuwerten, und andererseits einen Ansatzpunkt, um Therapiefehler im Bereich der Osteoporose im klinischen Alltag zu vermeiden.

Teil 2: Klinische Entscheidungsfindung am konkreten Beispiel des Cushing-Syndroms

Die Diagnostik des Rezidivs beim Morbus Cushing

Zugrunde liegende Publikationen:

Braun LT, Zopp S, Vogel F, Honegger J, Rubinstein G, Schilbach K, Künzel H, Beuschlein F, Reincke M. *Signs, symptoms and biochemistry in recurrent Cushing disease: a prospective pilot study*. 2021 Sep;73(3):762-766. doi: 10.1007/s12020-021-02719-9. Epub 2021 Apr 19. PMID: 33871792; PMCID: PMC8325659.

Braun LT, Rubinstein G, Zopp S, Vogel F, Schmid-Tannwald C, Escudero MP, Honegger J, Ladurner R, Reincke M. *Recurrence after pituitary surgery in adult Cushing's disease: a systematic review on diagnosis and treatment*. *Endocrine*. 2020 Nov;70(2):218-231. doi: 10.1007/s12020-020-02432-z. Epub 2020 Aug 2. PMID: 32743767; PMCID: PMC7396205.

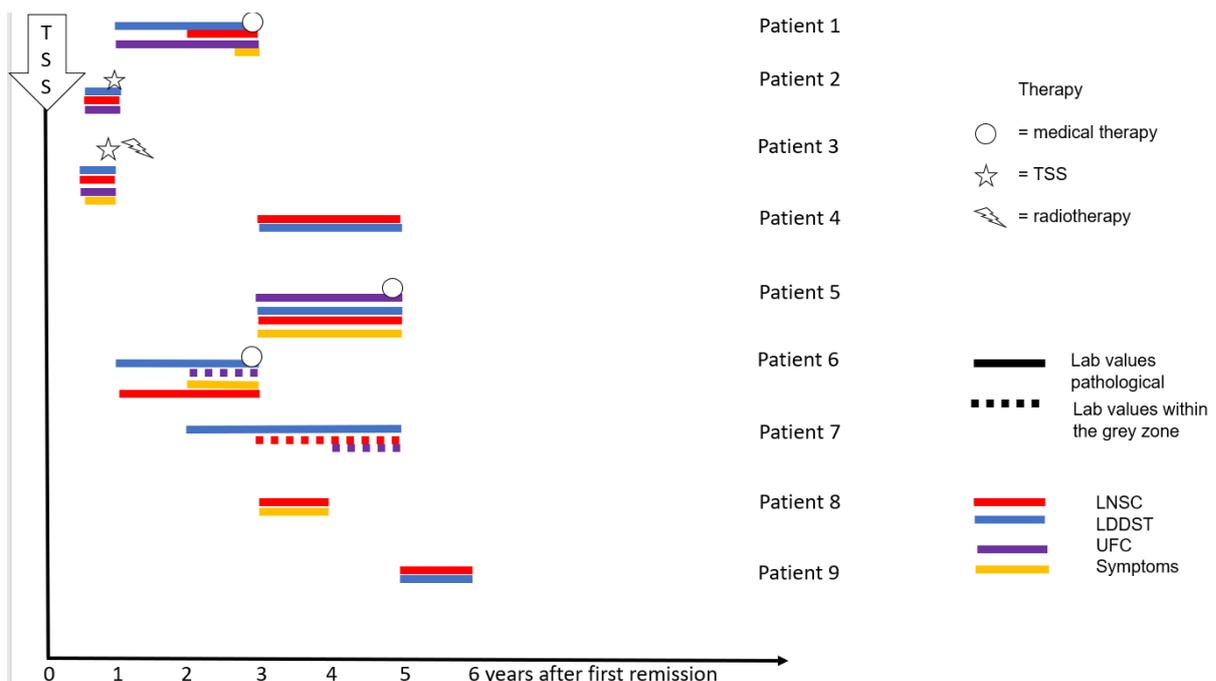
Rezidive beim Morbus Cushing stellen ein diagnostisches und therapeutisches Problem dar. Je nach Studie erleiden 5-20% der Patienten ein Rezidiv nach initial erfolgreicher Operation (42). Die meisten Rezidive treten innerhalb der ersten fünf Jahre nach Operation auf, es gibt allerdings auch spätere Rezidive, sodass bislang zumindest eine jährliche Nachsorge in den ersten zehn Jahren nach der Operation empfohlen wird (43). Sehr viele Studien untersuchten prognostische Marker, die als Anhaltspunkt für die Wahrscheinlichkeit eines Rezidivs herangezogen werden können. In einer Vielzahl dieser Studien sowie einer hierauf beruhenden Metaanalyse konnte gezeigt werden, dass insbesondere die postoperativen Hormonwerte – Cortisol und ACTH – ein wichtiger prognostischer Marker sind (44). Faktoren wie Alter, Geschlecht und Tumorgöße scheinen hingegen keinen Einfluss auf die Rezidivhäufigkeit zu haben (44). Die Erfahrung des Operateurs (45) sowie die präoperative Lokalisation des Tumors sind ebenfalls prognostisch relevante Faktoren (46).

Unklarheit besteht allerdings hinsichtlich der Frage, wie und in welchen Abständen Patienten mit Morbus Cushing Nachsorgeuntersuchungen erhalten sollten. Bislang gab es zu dieser Thematik keine prospektiven Studien, sondern lediglich diverse Expertenkonsensus. Ayala et. al. empfehlen Screeningintervalle, die vom initialen postoperativen Cortisol abhängig sind. Die Screeningintervalle liegen bei diesem Konsensus zwischen neun und zwölf Monate (47). Ein anderer Ansatz ist in dem Konsens von Fleseriu et. al. zu finden: Hier wird ein jährliches klinisches Screening empfohlen; ein biochemisches Screening sollte bei möglichen Cushing-

Symptomen oder einem neuerlichen Hypophysenadenom in der MRT durchgeführt werden (48). Es gibt allerdings noch komplexere Ansätze (Geer et al.): Je nach Zeit seit erster Operation, Hydrocortisonersatztherapie und Remissionsverlauf wird ein Screening in Abständen von zwei bis vier Monaten empfohlen (49).

Um nun zukünftig ein evidenzbasiertes Konzept durchführen zu können, wurde in meiner Pilotstudie der klinische und biochemische Verlauf von Patienten und Patientinnen mit Morbus Cushing untersucht. Die Studie ist Teil des 2012 gegründeten Deutschen Cushing Registers. 67 Patientinnen und Patienten mit Morbus Cushing wurden ab Diagnosezeitpunkt prospektiv verfolgt. Es wurde ein jährliches standardisiertes klinisches Screening durchgeführt; zudem erfolgte jährlich ein biochemisches Screening mittels Cortisolmessung im 24h-Sammelurin, Bestimmung des spätabendlichen Speichelcortisols und Durchführung des 1mg-Dexamethason-Hemmtests. Von diesen 67 Patienten erlitten 9 Patientinnen und Patienten während des Beobachtungszeitraums ein Rezidiv. Es zeigte sich, dass viele Patientinnen und Patienten biochemische Auffälligkeiten entwickeln, bevor erneut klinische Symptome eintreten (Abbildung 16). Bei einigen Erkrankten treten klinische Probleme erst Monate bis Jahre nach den biochemischen Auffälligkeiten auf. Diese Studie belegt, dass ein jährliches biochemisches Screening unbedingt erfolgen sollte, um Rezidive frühzeitig zu diagnostizieren (50). Der dadurch gewonnene zeitliche Vorsprung kann von immenser klinischer Bedeutung sein: Solange ein Patient oder eine Patientin keine bis wenige klinische Symptome hat, stehen viele therapeutische Optionen zur Verfügung.

Abbildung 16: Der klinische und biochemische Verlauf bei Rezidiven



aus: Braun LT et al. *Endocrine*. 2021.

Speichelprofile zur Optimierung der Diagnosestellung

Zugrunde liegende Publikation:

Braun LT, Vogel F, Zopp S, Rubinstein G, Schilbach K, Künzel H, Beuschlein F, Reincke M. Diurnal Salivary Cortisol Profiles in Patients with Cushing's Syndrome. *Exp Clin Endocrinol Diabetes*. 2022 Jan 17. doi: 10.1055/a-1719-5381. Epub ahead of print. PMID: 35038761.

Wie bereits erwähnt, beruht die Diagnostik des Cushing-Syndrom auf drei biochemischen Tests: der Cortisolmessung im 24h-Sammelurin, dem 1mg-Dexamethason-Hemmtest sowie dem spätabendlichen Speichelcortisol (18). Keiner dieser drei Tests bietet aber eine optimale Sensitivität oder Spezifität, sodass immer auf eine Kombination von mindestens zwei Tests zurückgegriffen werden muss. Gründe für falsch-positive und falsch-negative Ergebnisse sind in Tabelle 2 aufgeführt.

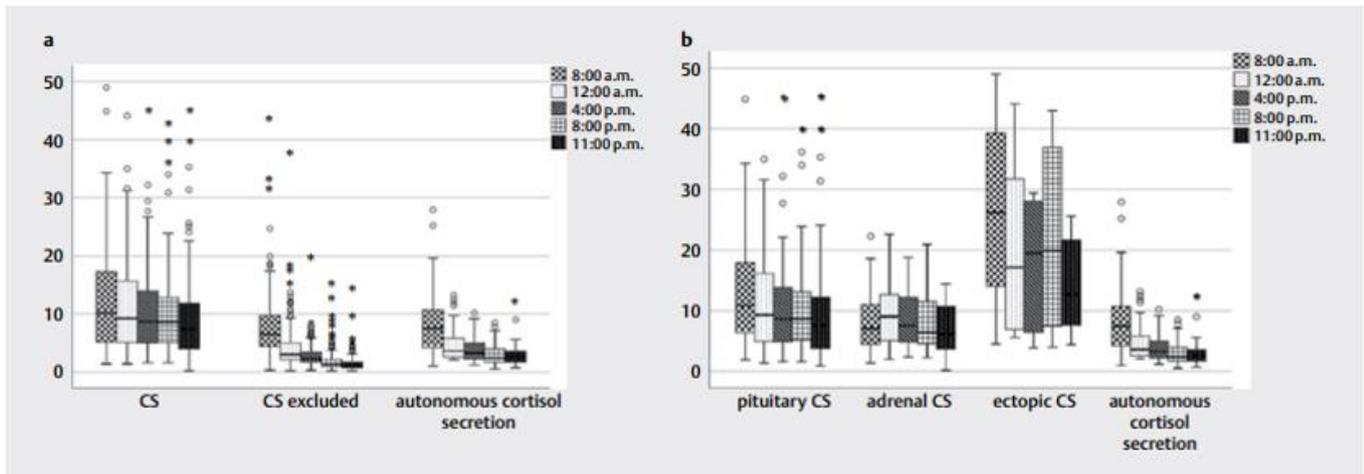
Tabelle 2: Störfaktoren für die biochemische Testung (adaptiert nach Braun et al. (11))

Screeningtest	Falsch positiv	Falsch negativ
1mg-Dexamethason-Hemmtest	Orale Kontrazeptiva, diverse Medikamente, rasche Verstoffwechslung des Dexamethasons, Resorptionsstörungen	Niereninsuffizienz, Leberversagen
Cortisol im 24h-Sammelurin	Assaystörungen, hohes Urinvolumen (> 5l), Medikamente, falsche Durchführung der Sammlung	Assaystörungen, Niereninsuffizienz, unvollständige Sammlung
Speichelcortisol	Depressionen, Schichtarbeit, Rauchen, Infektionen oder Blutungen im Mundbereich	Falsche Sammlung (z.B. unzureichende Speichelmenge)

Zudem ist bekannt, dass Patienten mit Cushing-Syndrom eine gestörte Cortisoltagessrhythmik haben. Diese lässt sich mit Hilfe eines Cortisoltagessprofils nachweisen. Bislang wurde dieses Wissen allerdings nicht zur Diagnosestellung oder Subdifferenzierungsdiagnostik verwendet. Hingegen werden Speichelprofile genutzt, um die medikamentöse Therapie mit Adrenostatika optimal einzustellen und dienen also in diesem Kontext der therapeutischen Überwachung (51). Auf Basis dieser Überlegungen stellt sich die Frage, ob ein Cortisoltagessprofil einen zusätzlichen diagnostischen Nutzen bringen könnte.

In meiner Studie wurden 428 Patienten und Patientinnen untersucht (52): Bei 278 Patienten wurde ein Cushing-Syndrom zwar initial klinisch vermutet, aber in der weiteren Diagnostik ausgeschlossen. Bei den verbleibenden 111 Patienten konnte ein Cushing-Syndrom bestätigt werden. Hiervon litten 75 Patienten unter einem Morbus Cushing, 27 an einem adrenalem Cushing-Syndrom und 9 an einem ektope Cushing-Syndrom. Es wurden weitere 39 Patienten mit autonomer Cortisolsekretion zum Vergleich eingeschlossen. Das Speichelcortisol wurde bei allen Patienten an fünf Zeitpunkten gemessen (8:00, 12:00, 16:00, 20:00 und 22:00). Es zeigten sich signifikant höhere Werte zu allen Tageszeitpunkten bei Patienten mit Cushing-Syndrom ($p \leq 0.001$). Die diagnostische Sicherheit bei der Unterscheidung zwischen Patienten mit und ohne Cushing-Syndrom ließ sich dadurch etwas verbessern (85 vs. 91% richtige Diagnosestellung). Zusätzlich zeigten sich innerhalb der Subtypen verschiedene Cortisolprofile: Patienten mit adrenalem Cushing-Syndrom haben eine starre Rhythmik, bei Patienten mit Morbus Cushing ist eine angedeutete Tagesrhythmik erkennbar bei allerdings insgesamt deutlich höheren Werten als bei Patienten mit adrenalem Cushing-Syndrom. Patienten mit ektope Cushing-Syndrom haben keine erkennbare Tagesrhythmik. Aufgrund großer interindividueller Unterschiede sind diese Resultate allerdings nur von eingeschränktem Nutzen in der täglichen klinischen Praxis.

Abbildung 17: Die Cortisoltagesrhythmen bei Patienten mit verschiedenen Subtypen des Cushing-Syndroms sowie autonomer Cortisolsekretion



aus: Braun LT et al. *Experimental and clinical endocrinology & diabetes : official journal, German Society of Endocrinology [and] German Diabetes Association*. 2022

Zusammenfassung und Ausblick

Die Klinische Entscheidungsfindung – *Clinical Reasoning* – ist eine hoch relevante ärztliche Aufgabe, die in einigen Fachbereichen sicherlich die größte Herausforderung darstellt. Der Prozess der Klinischen Entscheidungsfindung beginnt mit dem ersten Patientenkontakt, bei dem mögliche Diagnosen und Differentialdiagnosen überdacht werden, Untersuchungen initiiert und koordiniert werden müssen, eine finale Diagnose gestellt und letztlich ein Therapiekonzept – gemeinsam mit dem Patienten – festgelegt wird. Fehler während der Klinischen Entscheidungsfindung können fatale Folgen für den Patienten haben (53, 54).

In der Medizindidaktik nehmen daher Forschungsfragen zu dieser Thematik einen großen Raum ein, um die kognitiven Prozesse bei *Clinical Reasoning* besser zu verstehen und im Alltag berücksichtigen zu können. In den oben genannten Arbeiten wurden verschiedene Aspekte der Klinischen Entscheidungsfindung entweder aus didaktischer oder aus klinischer Perspektive beleuchtet.

Aus didaktischer Perspektive zeigten sich in den Studien zwei Hauptergebnisse:

Bayesian Reasoning – was elementar ist, um Testergebnisse korrekt interpretieren zu können – fällt Medizinstudierenden schwer. Wie bereits vorab bekannt, kann die Darstellung von Zahlenformaten als Natürliche Häufigkeiten im Gegensatz zu Wahrscheinlichkeiten die Lösungsrate verbessern. In Ergänzung zu diesen Ergebnissen konnte in diesen Studien gezeigt werden, dass sich auch die Diagnoseeffizienz durch den Einsatz von Natürlichen Häufigkeiten verbessern lässt. Zudem sind Doppelbäume und Häufigkeitsnetze Visualisierungsformen, welche die Richtigkeit bei der Bearbeitung deutlich erhöhen können. *Bayesian Reasoning* ist in seiner Wichtigkeit eine im klinischen Alltag bislang unterschätzte Kompetenz. Dabei ist das Verständnis grundlegend für das Treffen von Therapieentscheidungen und für die Beratung von Patienten. Die Ergebnisse zeigen eine Möglichkeit zur Unterstützung im klinischen Alltag auf. Natürlich kann nicht erwartet werden, dass für alle möglichen klinischen Szenarien beispielsweise Häufigkeitsnetze zur Verfügung stehen, aber allein das Bewusstsein für den Einfluss von Prävalenz einer Erkrankung und den Testsensitivitäten und -spezifitäten auf die Rate an falsch positiv und falsch negativen Ergebnissen ist essentiell. Nun muss der Transfer in die medizinische Ausbildung und klinische Praxis gelingen:

An allererster Stelle muss das Bewusstsein für die Relevanz von *Bayesian Reasoning* im Medizinstudium geschaffen werden. Bislang ist das Lehren und Lernen von *Bayesian Reasoning* nicht im Nationalem Kompetenzbasierten Lernzielkatalog aufgeführt, sodass es nicht per se in jedem Curriculum abgedeckt wird. Falls *Bayesian Reasoning* Bestandteil eines spezifischen Curriculums ist, sollten konsequent Natürliche Häufigkeiten und Visualisierungsmethoden genutzt werden, um *Bayesian Reasoning* zu lehren. Nur wenn diese

Kompetenz bereits im Medizinstudium vermittelt wird, kann dieses Wissen in die klinische Praxis übertragen werden. Durch ein korrektes Verständnis – beispielsweise für das Mammographieproblem – würden Patientinnen in der klinischen Praxis gegebenenfalls anders beraten werden. Auch aus der ökonomischen Perspektive ist dies wichtig, weil viele Screeningtests, so auch die Mammographie - relativ teuer sind, sodass das Kosten-Nutzen-Risiko hinterfragt werden muss (55, 56). Im klinischen Alltag wird oft der Fokus auf verpasste oder verspätet gestellte Diagnosen gesetzt, während die Konsequenzen einer Fehldiagnose, die zu einer möglicherweise unnötigen Überversorgung führen kann, häufig als weniger relevant wahrgenommen wird (56). An dieser Stelle sollten Ärztinnen und Ärzte sich in der klinischen Praxis immer wieder auch die Konsequenzen von Überdiagnostik und Übertherapie für die Patientinnen und Patienten vor Augen führen. Ein grundlegendes Verständnis für Statistik kann hierbei helfen. Allerdings ist aus einer Vielzahl an Studien bekannt, dass Wissenschaftlicher und Wissenschaftlerinnen häufig Probleme haben, Statistik korrekt anzuwenden oder zu interpretieren, sei es in Bezug auf den p-Wert (57), Überinterpretationen von Ergebnissen oder kleinen Fallzahlen (58), Datenmanagement oder schlichtweg auch bezüglich der Kommunikation von wissenschaftlichen Ergebnissen (59). Diese Mängel haben unmittelbare Konsequenzen im klinischen Alltag und sollte damit ein wesentlicher Bestandteil der medizinischen Ausbildung werden.

Neben diesen Aspekten zeigte sich in den qualitativen Studien zum **Clinical Reasoning**, dass bei der Bewertung der Diagnosekompetenz nicht nur das Ergebnis – die am Ende gestellte Diagnose – berücksichtigt werden sollte, sondern auch der Diagnoseprozess. Bei 10% der korrekt gestellten Diagnosen konnte durch Analyse der Diagnosebegründungen gezeigt werden, dass der Diagnoseprozess fehlerhaft ist. Natürlich bleibt unklar, inwiefern ein falscher Diagnoseprozess bei letztlich doch korrekter Diagnose, Konsequenzen für Patientinnen und Patienten haben kann. Darüber kann an dieser Stelle nur spekuliert werden, da weitere prospektive Studien nötig wären, um dieser Fragestellung nachzugehen. Es sollte allerdings vermehrt in medizindidaktischen Studien bedacht werden, dass nicht automatisch von einer korrekten Diagnose auf einen fehlerfreien Diagnoseprozess geschlossen werden kann. Gegebenenfalls wird hierdurch die Diagnosekompetenz von Studierenden überschätzt und Wissenslücken werden verkannt. Die verschiedenen qualitativen Analysen zeigten zudem, dass Fehler insbesondere durch mangelndes Wissen und nicht – wie man eventuell annehmen könnte – durch fehlgeleitete kognitive Prozesse, welche sich durch (strukturierte) Reflexionsformen verbessern ließen, zustande kommen. Diese Erkenntnis hat unmittelbaren Nutzen für die medizinische Weiterbildung: Fehler entstehen durch unzureichendes Wissen.

Aus der **klinischen Perspektive** zeigen sich ähnliche Schwierigkeiten. Dies wird besonders ersichtlich am Beispiel des Cushing-Syndroms, da es sich hierbei um eine Erkrankung handelt, die mehrere statistische und klinische Probleme aufweist (60):

- Die niedrige Prävalenz der Erkrankung
- Die variable klinische Präsentation
- Die suboptimale Spezifität und Sensitivität der diagnostischen Tests
- Die komplexe Subdifferenzierungsdiagnostik
- Die Komplexität der therapeutischen Entscheidungen

Mangelndes Wissen in Bezug auf die Epidemiologie, das klinische Erscheinungsbild, die komplexen diagnostischen Tests und ihre Interpretation sowie die verschiedenen Indikationen und Kontraindikationen der therapeutischen Optionen können zu Fehldiagnosen, Fehlentscheidungen und Fehltherapien führen.

Besonders eklatant wird dies in der Rezidivdiagnostik des Cushing-Syndroms. Bisherige Empfehlungen gehen weit auseinander hinsichtlich der Zeitintervalle und der Art der Nachsorge. Interessanterweise gibt es mehrere Experten/Konsensus-Empfehlungen zu dieser Thematik, ohne dass hierzu bislang prospektive Studien vorlagen. In unserer Studie wird deutlich, dass eine reine klinische Nachsorge nicht ausreichend ist, um Rezidive frühzeitig zu diagnostizieren. Naturgemäß sind allerdings multizentrische, prospektive Studien mit einem größeren Patientenkollektiv notwendig, um diese These weiter zu stützen.

Im übergeordneten Sinne ist diese Studie ein Beispiel dafür, dass aufgrund von mangelndem Wissen über den klinischen und biochemischen Verlauf der Erkrankung falsche Empfehlungen vertreten werden. Dieses Phänomen ist in der Medizin häufig: Als Beispiel sei hier eine aktuelle Thematik herausgegriffen: Die Vitamin D-Versorgung bei Patienten mit COVID-19. Frühzeitig im Laufe der Pandemie wurde eine Vitamin D zur Prophylaxe und Therapie einer COVID-Infektion aufgrund der limitierten Datenlage empfohlen (61, 62). In einer aktuellen Metaanalyse zeigte sich hingegen, dass es keine Evidenz für Empfehlungen bezüglich einer Vitamin D-Einnahme zur Therapie oder Prophylaxe dieser Erkrankung gibt (63).

Es ist selbstverständlich, dass klinische Empfehlungen ständig in Bezug auf die aktuelle Datenlage angepasst werden müssen. Die klinische Entscheidungsfindung ist ein dynamischer Prozess – sowohl auf einen individuellen Erkrankten bezogen als auch patientenübergreifend. Sich dies immer wieder bewusst zu machen, ist ausschlaggebend für eine gelungene Patientenbetreuung.

Ausblick

Zusammenfassend legt diese Habilitation den Grundstein für die Verbindung der didaktischen und klinischen Forschung, welche bis jetzt ein relatives Novum darstellt. Diese Arbeit ist aber nur ein erster Einblick in diese komplexe Thematik, sodass viele Fragen offen bleiben bzw. neu aufgetreten sind.

Die Arbeit gibt Anstoß für weitere – große – Forschungsprojekte und Forschungsfragen:

- Wie kann ein *Bayesian Reasoning*-Curriculum sinnvoll in die medizinische Ausbildung integriert werden?
- Kann *Bayesian Reasoning* die Patientenberatung und die Kommunikation von Testergebnissen im klinischen Alltag verändern?
- Wie kann *Clinical Reasoning*-Kompetenz im Medizinstudium sinnvoll geprüft werden?
- Welche Diagnostikfehler werden in spezifischen Fachbereichen – beispielsweise in der Endokrinologie – gemacht und wie können diese verhindert werden?

Diese Auswahl an offenen Forschungsfragen und deren Aufarbeitung könnten die bestehende Verbindung zwischen medizindidaktischer und klinischer Forschung in den kommenden Jahren vertiefen.

Publikationsverzeichnis

Bei den mit * versehenen Stellen handelt es sich um geteilte Erst- oder Letztautorenschaften.

Originalarbeiten als Erst- oder Letztautorin

2022

Tausendfreund O, **Braun LT***, Schmidmaier R. Types of therapeutic errors in the management of osteoporosis made by physicians and medical students. *BMC Med Educ.* 2022 Apr 27;22(1):323. doi: 10.1186/s12909-022-03384-w. PMID: 35473636.

Kunzelmann AK, Binder K, Fischer MR, Reincke M, **Braun LT***, Schmidmaier R. Improving Diagnostic Efficiency with Frequency Double-Trees and Frequency Nets in Bayesian Reasoning. *MDM Policy Pract.* 2022 Mar 16;7(1):23814683221086623. doi: 10.1177/23814683221086623. PMID: 35321028; PMCID: PMC8935422.

Braun LT, Vogel F, Zopp S, Rubinstein G, Schilbach K, Künzel H, Beuschlein F, Reincke M. Diurnal Salivary Cortisol Profiles in Patients with Cushing's Syndrome. *Exp Clin Endocrinol Diabetes.* 2022 Jan 17. doi: 10.1055/a-1719-5381. Epub ahead of print. PMID: 35038761.

2021

Braun LT, Zopp S, Vogel F, Honegger J, Rubinstein G, Schilbach K, Künzel H, Beuschlein F, Reincke M. *Signs, symptoms and biochemistry in recurrent Cushing disease: a prospective pilot study.* *Endocrine.* 2021 Apr 19. doi: 10.1007/s12020-021-02719-9. Epub ahead of print. PMID: 33871792.

Binder K, Krauss S, Schmidmaier R, **Braun LT.** *Natural frequency trees improve diagnostic efficiency in Bayesian reasoning.* *Adv Health Sci Educ Theory Pract.* 2021 Feb 12. doi: 10.1007/s10459-020-10025-8. Epub ahead of print. PMID: 33599875.

2020

Braun LT, Fazel J, Zopp S, Benedix S, Osswald-Kopp A, Riester A, Rubinstein G, Seidensticker M, Beuschlein F, Drey M, Bidlingmaier M, Schmidmaier R, Reincke M. *The Effect of Biochemical Remission on Bone Metabolism in Cushing's Syndrome: A 2-Year Follow-Up Study.* *J Bone Miner Res.* 2020 Sep;35(9):1711-1717. doi: 10.1002/jbmr.4033. Epub 2020 May 8. PMID: 32315096.

2019

Braun LT, Borrmann KF, Lottspeich C, Heinrich DA, Kieseewetter J, Fischer MR, Schmidmaier R. *Guessing right - whether and how medical students give incorrect reasons for their correct diagnoses*. GMS J Med Educ. 2019 Nov 15;36(6):Doc85. doi: 10.3205/zma001293. PMID: 31844657; PMCID: PMC6905369.

Braun LT, Lenzer B, Fischer MR, Schmidmaier R. *Complexity of clinical cases in simulated learning environments: proposal for a scoring system*. GMS J Med Educ. 2019 Nov 15;36(6):Doc80. doi: 10.3205/zma001288. PMID: 31844652; PMCID: PMC6905356.

Braun LT, Borrmann KF, Lottspeich C, Heinrich DA, Kieseewetter J, Fischer MR, Schmidmaier R. *Scaffolding clinical reasoning of medical students with virtual patients: effects on diagnostic accuracy, efficiency, and errors*. Diagnosis (Berl). 2019 Jun 26;6(2):137-149. doi: 10.1515/dx-2018-0090. PMID: 30772868.

2018

Braun LT, Lenzer B, Kieseewetter J, Fischer MR, Schmidmaier R. *How case representations of medical students change during case processing - Results of a qualitative study*. GMS J Med Educ. 2018 Aug 15;35(3):Doc41. doi: 10.3205/zma001187. PMID: 30186951; PMCID: PMC6120161.

2017

Braun LT, Zwaan L, Kieseewetter J, Fischer MR, Schmidmaier R. *Diagnostic errors by medical students: results of a prospective qualitative study*. BMC Med Educ. 2017 Nov 9;17(1):191. doi: 10.1186/s12909-017-1044-7. PMID: 29121903; PMCID: PMC5679151.

Braun LT, Zottmann JM, Adolf C, Lottspeich C, Then C, Wirth S, Fischer MR, Schmidmaier R. *Representation scaffolds improve diagnostic efficiency in medical students*. Med Educ. 2017 Nov;51(11):1118-1126. doi: 10.1111/medu.13355. Epub 2017 Jun 6. PMID: 28585351.

Originalarbeiten als Koautor/in

2022

Rubinstein G, Osswald A, **Braun LT**, Vogel F, Kroiss M, Pilz S, Deniz S, Aigner L, Knösel T, Bertherat J, Bouys L, Ladurner R, Riester A, Bidlingmaier M, Beuschlein F, Reincke M. The

role of adrenal venous sampling (AVS) in primary bilateral macronodular adrenocortical hyperplasia (PBMAH): a study of 16 patients. *Endocrine*. 2022 Mar 10. doi: 10.1007/s12020-022-03020-z. Epub ahead of print. PMID: 35274284.

2021

Vogel F, **Braun L**, Rubinstein G, Zopp S, Oßwald A, Schilbach K, Schmidmaier R, Bidlingmaier M, Reincke M. Metformin and Bone Metabolism in Endogenous Glucocorticoid Excess: An Exploratory Study. *Front Endocrinol (Lausanne)*. 2021 Oct 27;12:765067. doi: 10.3389/fendo.2021.765067. PMID: 34777259; PMCID: PMC8578886.

Armignacco R, Jouinot A, Bouys L, Septier A, Lartigue T, Neou M, Gaspar C, Perlemoine K, **Braun L**, Riester A, Bonnet-Serrano F, Blanchard A, Amar L, Scaroni C, Ceccato F, Rossi GP, Williams TA, Larsen CK, Allasonnière S, Zennaro MC, Beuschlein F, Reincke M, Bertherat J, Assié G. Identification of glucocorticoid-related molecular signature by whole blood methylome analysis. *Eur J Endocrinol*. 2022 Jan 13;186(2):297-308. doi: 10.1530/EJE-21-0907. PMID: 34914631; PMCID: PMC8789024.

Vogel F, **Braun L**, Rubinstein G, Zopp S, Benedix S, Schneider H, Ritzel K, Schilbach K, Schmidmaier R, Beuschlein F, Bidlingmaier M, Reincke M. *Patients with low IGF-I after curative surgery for Cushing's syndrome have an adverse long-term outcome of hypercortisolism-induced myopathy*. *Eur J Endocrinol*. 2021 May 4;184(6):813-821. doi: 10.1530/EJE-20-1285. PMID: 33830940.

Lottspeich C, **Braun LT**, Fischer MR, Schmidmaier R. *Simulated ward round training in the medical curriculum Munich*. *GMS J Med Educ*. 2021 Apr 15;38(4):Doc75. doi: 10.3205/zma001471. PMID: 34056064; PMCID: PMC8136345.

Heger E, Rubinstein G, **Braun LT**, Zopp S, Honegger J, Seidensticker M, Reincke M, Oßwald A. *Chemosensory dysfunction in Cushing's syndrome*. *Endocrine*. 2021 Apr 5. doi: 10.1007/s12020-021-02707-z. Epub ahead of print. PMID: 33821391.

Röcker N, Lottspeich C, **Braun LT**, Lenzer B, Frey J, Fischer MR, Schmidmaier R. *Implementation of self-directed learning within clinical clerkships*. *GMS J Med Educ*. 2021 Feb 15;38(2):Doc43. doi: 10.3205/zma001439. PMID: 33763528; PMCID: PMC7958912.

Vetrivel S, Zhang R, Engel M, Altieri B, **Braun L**, Osswald A, Bidlingmaier M, Fassnacht M, Beuschlein F, Reincke M, Chen A, Sbiera S, Riester A. *Circulating microRNA Expression in Cushing's Syndrome*. *Front Endocrinol (Lausanne)*. 2021 Feb 22;12:620012. doi: 10.3389/fendo.2021.620012. PMID: 33692756; PMCID: PMC7937959.

2020

Vogel F, **Braun LT**, Rubinstein G, Zopp S, Künzel H, Strasding F, Albani A, Riester A, Schmidmaier R, Bidlingmaier M, Quinkler M, Deutschbein T, Beuschlein F, Reincke M. *Persisting Muscle Dysfunction in Cushing's Syndrome Despite Biochemical Remission*. J Clin Endocrinol Metab. 2020 Dec 1;105(12):e4490–8. doi: 10.1210/clinem/dgaa625. PMID: 32882010; PMCID: PMC7538105.

Wurth R, Tirosh A, Kamilaris CDC, Camacho J, Faucz FR, Maria AG, Berthon A, Papadakis GZ, Nilubol N, Hamimi A, Gharib AM, Demidowich A, Zilbermint M, Eisenhofer G, **Braun L**, Reincke M, Stratakis CA, Hannah-Shmouni F. *Volumetric Modeling of Adrenal Gland Size in Primary Bilateral Macronodular Adrenocortical Hyperplasia*. J Endocr Soc. 2020 Oct 29;5(1):bvaa162. doi: 10.1210/jendso/bvaa162. PMID: 33305158; PMCID: PMC7716656.

Adolf C, **Braun LT**, Fuss CT, Hahner S, Künzel H, Handgriff L, Sturm L, Heinrich DA, Schneider H, Bidlingmaier M, Reincke M. *Spirolactone reduces biochemical markers of bone turnover in postmenopausal women with primary aldosteronism*. Endocrine. 2020 Sep;69(3):625-633. doi: 10.1007/s12020-020-02348-8. Epub 2020 Jun 27. PMID: 32594379. Handgriff L, Adolf C, Heinrich DA, **Braun L**, Nirschl N, Sturm L, Ladurner R, Ricke J, Seidensticker M, Bidlingmaier M, Reincke M. *The Impact of Glucocorticoid Co-Secretion in Primary Aldosteronism on Thyroid Autoantibody Titers During the Course of Disease*. Horm Metab Res. 2020 Jun;52(6):404-411. doi: 10.1055/a-1164-1944. Epub 2020 May 13. PMID: 32403151; PMCID: PMC7746512.

Rubinstein G, Osswald A, Hoster E, Losa M, Elenkova A, Zacharieva S, Machado MC, Hanzu FA, Zopp S, Ritzel K, Riester A, **Braun LT**, Kreitschmann-Andermahr I, Storr HL, Bansal P, Barahona MJ, Cosaro E, Dogansen SC, Johnston PC, Santos de Oliveira R, Raftopoulos C, Scaroni C, Valassi E, van der Werff SJA, Schopohl J, Beuschlein F, Reincke M. *Time to Diagnosis in Cushing's Syndrome: A Meta-Analysis Based on 5367 Patients*. J Clin Endocrinol Metab. 2020 Mar 1;105(3):dgz136. doi: 10.1210/clinem/dgz136. PMID: 31665382.

Müller LM, Kienitz T, Deutschbein T, Riester A, Hahner S, Burger-Stritt S, Berr CM, Oßwald A, **Braun L**, Rubinstein G, Reincke M, Quinkler M. *Glucocorticoid Receptor Polymorphisms Influence Muscle Strength in Cushing's Syndrome*. J Clin Endocrinol Metab. 2020 Jan 1;105(1):dgz052. doi: 10.1210/clinem/dgz052. PMID: 31613324.

Beckers MA, **Braun LT**, Epstein N, Fischer MR, Schmidmaier R. *Die Facharztprüfung Innere Medizin: Eine qualitative und quantitative Analyse von Gedächtnisprotokollen* [The specialist examination in internal medicine: a qualitative and quantitative analysis of minutes from memory]. Dtsch Med Wochenschr. 2020 Jan;145(2):e10-e17. German. doi: 10.1055/a-0970-6348. Epub 2019 Sep 23. PMID: 31546268.

Übersichtsartikel/Reviews

2022

Braun LT, Vogel F, Reincke M. Long-term morbidity and mortality in patients with Cushing's syndrome. *J Neuroendocrinol.* 2022 Mar 4:e131113. doi: 10.1111/jne.13113. Epub ahead of print. PMID: 35312199.

2021

Vogel F, **Braun L**, Reincke M. Morbidität und Mortalität beim Cushing-Syndrom [Morbidity and mortality in Cushing's syndrome]. *Internist (Berl).* 2022 Jan;63(1):34-42. German. doi: 10.1007/s00108-021-01222-7. Epub 2021 Dec 20. PMID: 34928397.

2020

Braun LT, Rubinstein G, Zopp S, Vogel F, Schmid-Tannwald C, Escudero MP, Honegger J, Ladurner R, Reincke M. *Recurrence after pituitary surgery in adult Cushing's disease: a systematic review on diagnosis and treatment.* *Endocrine.* 2020 Nov;70(2):218-231. doi: 10.1007/s12020-020-02432-z. Epub 2020 Aug 2. PMID: 32743767; PMCID: PMC7396205.

Braun LT, Reincke M. *What is the role of medical therapy in adrenal-dependent Cushing's syndrome?* *Best Pract Res Clin Endocrinol Metab.* 2020 May;34(3):101376. doi: 10.1016/j.beem.2020.101376. Epub 2020 Jan 31. PMID: 32063487.

2019

Braun LT, Riester A, Oßwald-Kopp A, Fazel J, Rubinstein G, Bidlingmaier M, Beuschlein F, Reincke M. *Toward a Diagnostic Score in Cushing's Syndrome.* *Front Endocrinol (Lausanne).* 2019 Nov 8;10:766. doi: 10.3389/fendo.2019.00766. PMID: 31787931; PMCID: PMC6856055.

Sonstige Veröffentlichungen

Braun LT, Reincke M. *The Effect of Biochemical Remission on Bone Metabolism in Cushing's Syndrome: A 2-Year Follow-Up Study.* *J Bone Miner Res.* 2021 Apr 26. doi: 10.1002/jbmr.4309. Epub ahead of print. PMID: 33900642.

Reincke M, **Braun LT**: Diagnostik und Differentialdiagnostik bei Morbus Cushing: Klassisches und Neues. *Endokrinologie Informationen, Sonderheft 2020. Mitteilungen der Deutschen Gesellschaft für Endokrinologie.*

Adolf C, **Braun LT**, Fuss CT, Hahner S, Künzel H, Handgriff L, Sturm L, Heinrich DA, Schneider H, Bidlingmaier M, Reincke M. Correction: Spironolactone reduces biochemical markers of bone turnover in postmenopausal women with primary aldosteronism. *Endocrine*. 2022 Jan;75(1):312. doi: 10.1007/s12020-021-02905-9. Erratum for: *Endocrine*. 2020 Sep;69(3):625-633. PMID: 34687447; PMCID: PMC8763797.

Braun LT, Schmidmaier R. *Dealing with cognitive dissonance: an approach*. *Med Educ*. 2019 Dec;53(12):1167-1168. doi: 10.1111/medu.13955. Epub 2019 Sep 18. PMID: 31532838.

Danksagung

Große Projekte entstehen selten allein, daher sei an dieser Stelle den vielen Personen gedankt, die mich in den letzten Jahren unterstützt haben:

Dank an:

- Herrn Professor *Martin Reincke* für die stete klinische und wissenschaftliche Förderung, für die Unterstützung im Rahmen meines Clinician Scientist Programms sowie für die konstante Ermutigung, über meine eigenen Grenzen hinauszuwachsen.
- Herrn Professor *Ralf Schmidmaier* für die nun schon seit meiner Doktorarbeit konstante Betreuung in allen Belangen. Die Freude an der wissenschaftlichen Arbeit verdanke ich ihm.
- Herrn Professor *Martin Fischer* für die langjährige Unterstützung – auch in den nötigen didaktischen Weiterbildungen.
- Frau *Stephanie Zopp*, der Study Nurse unserer Cushing-Ambulanz, ohne deren immerwährenden und besonnenen Einsatz keines der klinischen Forschungsprojekte möglich gewesen wäre.
- Meine *Kolleginnen und Kollegen* der endokrinologischen Ambulanz für die fortwährende Hilfe; hierbei ganz besonders Herrn Dr. Frederick Vogel für die gute wissenschaftliche Zusammenarbeit.
- Meine *Doktorandinnen und Doktoranden* für ihre Motivation und Ausdauer.
- Die *Else-Kröner-Fresenius-Stiftung* sowie die *Eva Luise und Horst Köhler Stiftung* für die finanzielle Unterstützung im Rahmen des Clinician Scientist Programms *RISE – Rare important Syndromes in Endocrinology*, ebenso gilt mein Dank den Koordinatorinnen des Programms.
- Unsere *Patientinnen, Patienten und Studierenden*, die bereitwillig die Forschungsprojekte unterstützt haben, und ohne die keines der Projekte möglich gewesen wäre.
- Wissenschaftliche *Kooperationen* sind eine Grundlage, um gute Forschungsprojekte durchzuführen. Herausgegriffen sei an dieser Stelle das Steroidlabor in Gießen unter der Leitung von *Herrn Professor Wudy* – vielen Dank für die freundliche und gewinnbringende Kooperation! Ebenso möchte ich hier die gute Zusammenarbeit mit Frau *Professor Karin Binder* von der Mathematikdidaktik der LMU hervorheben und mich sehr herzlich hierfür bedanken!

An dieser Stelle möchte ich auch von ganzem Herzen meinem Freund Benjamin danken, der mich immer unterstützt, jeden Erfolg mit mir feiert und mir Kraft für jede Aufgabe gibt. Mein persönlicher Dank gilt zudem meinen Eltern, die mich nun schon seit Jahren in jeglicher Hinsicht begleiten.

Literaturverzeichnis

1. Braun LT, Borrmann KF, Lottspeich C, Heinrich DA, Kiesewetter J, Fischer MR, et al. Guessing right - whether and how medical students give incorrect reasons for their correct diagnoses. *GMS journal for medical education*. 2019;36(6):Doc85.
2. Graber ML. The incidence of diagnostic error in medicine. *BMJ quality & safety*. 2013;22 Suppl 2(Suppl 2):ii21-ii7.
3. Olson APJ, Graber ML. Learning from tragedy - improving diagnosis through case reviews. *Diagnosis (Berlin, Germany)*. 2018;5(4):171-3.
4. Graber ML, Lippi G, Plebani M. Diagnosis is now indexed in PubMed. *Diagnosis (Berlin, Germany)*. 2018;5(1):1-2.
5. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Archives of internal medicine*. 2005;165(13):1493-9.
6. Braun L. Diagnosekompetenz bei Medizinstudierenden: Imu; 2019.
7. Fischer MR, Bauer D, Mohn K. Finally finished! National Competence Based Catalogues of Learning Objectives for Undergraduate Medical Education (NKLM) and Dental Education (NKLZ) ready for trial. *GMS Zeitschrift für medizinische Ausbildung*. 2015;32(3):Doc35.
8. Epstein RM, Hundert EM. Defining and assessing professional competence. *Jama*. 2002;287(2):226-35.
9. Wieviele Endokrinologen/Diabetologen arbeiten in Deutschland? | DGE Blog (endokrinologie.net)
10. Lacroix A, Felders RA, Stratakis CA, Nieman LK. Cushing's syndrome. *Lancet (London, England)*. 2015;386(9996):913-27.
11. Braun LT, Riester A, Oßwald-Kopp A, Fazel J, Rubinstein G, Bidlingmaier M, et al. Toward a Diagnostic Score in Cushing's Syndrome. *Frontiers in endocrinology*. 2019;10:766.
12. Rubinstein G, Osswald A, Hoster E, Losa M, Elenkova A, Zacharieva S, et al. Time to Diagnosis in Cushing's Syndrome: A Meta-Analysis Based on 5367 Patients. *The Journal of clinical endocrinology and metabolism*. 2020;105(3).
13. Kreitschmann-Andermahr I, Psaras T, Tsiogka M, Starz D, Kleist B, Siegel S, et al. From first symptoms to final diagnosis of Cushing's disease: experiences of 176 patients. *European journal of endocrinology*. 2015;172(6):X1.
14. Terzolo M, Reimondo G, Chiodini I, Castello R, Giordano R, Ciccarelli E, et al. Screening of Cushing's syndrome in outpatients with type 2 diabetes: results of a prospective multicentric study in Italy. *The Journal of clinical endocrinology and metabolism*. 2012;97(10):3467-75.

15. Reimondo G, Pia A, Allasino B, Tassone F, Bovio S, Borretta G, et al. Screening of Cushing's syndrome in adult patients with newly diagnosed diabetes mellitus. *Clinical endocrinology*. 2007;67(2):225-9.
16. Anderson GH, Jr., Blakeman N, Streeten DH. The effect of age on prevalence of secondary forms of hypertension in 4429 consecutively referred patients. *Journal of hypertension*. 1994;12(5):609-15.
17. Omura M, Saito J, Yamaguchi K, Kakuta Y, Nishikawa T. Prospective study on the prevalence of secondary hypertension among hypertensive patients visiting a general outpatient clinic in Japan. *Hypertension research : official journal of the Japanese Society of Hypertension*. 2004;27(3):193-202.
18. Nieman LK, Biller BM, Findling JW, Newell-Price J, Savage MO, Stewart PM, et al. The diagnosis of Cushing's syndrome: an Endocrine Society Clinical Practice Guideline. *The Journal of clinical endocrinology and metabolism*. 2008;93(5):1526-40.
19. Klibanoff P, Marinacci M, Mukerji S. A Smooth Model of Decision Making under Ambiguity. *Econometrica*. 2005;73(6):1849-92.
20. Schiebener J, Brand M. Decision Making Under Objective Risk Conditions—a Review of Cognitive and Emotional Correlates, Strategies, Feedback Processing, and External Influences. *Neuropsychology Review*. 2015;25(2):171-98.
21. Westbury CF. Bayes' rule for clinicians: an introduction. *Frontiers in psychology*. 2010;1:192.
22. Eddy D. M.(1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. *Judgment under uncertainty: Heuristics and biases*. 1982:249-67.
23. McDowell M, Jacobs P. Meta-analysis of the effect of natural frequencies on Bayesian reasoning. *Psychological bulletin*. 2017;143(12):1273.
24. Binder K, Krauss S, Schmidmaier R, Braun LT. Natural frequency trees improve diagnostic efficiency in Bayesian reasoning. *Advances in health sciences education : theory and practice*. 2021;26(3):847-63.
25. Braun LT, Zottmann JM, Adolf C, Lottspeich C, Then C, Wirth S, et al. Representation scaffolds improve diagnostic efficiency in medical students. *Medical education*. 2017;51(11):1118-26.
26. Kiesewetter J, Ebersbach R, Tsalas N, Holzer M, Schmidmaier R, Fischer MR. Knowledge is not enough to solve the problems - The role of diagnostic knowledge in clinical reasoning activities. *BMC medical education*. 2016;16(1):303.
27. Fischer MR, Schauer S, Gräsel C, Baehring T, Mandl H, Gärtner R, et al. [CASUS model trial. A computer-assisted author system for problem-oriented learning in medicine]. *Zeitschrift für ärztliche Fortbildung*. 1996;90(5):385-9.

28. Braun LT, Zwaan L, Kiesewetter J, Fischer MR, Schmidmaier R. Diagnostic errors by medical students: results of a prospective qualitative study. *BMC medical education*. 2017;17(1):191.
29. Braun LT, Lenzer B, Kiesewetter J, Fischer MR, Schmidmaier R. How case representations of medical students change during case processing - Results of a qualitative study. *GMS journal for medical education*. 2018;35(3):Doc41.
30. Braun LT, Borrmann KF, Lottspeich C, Heinrich DA, Kiesewetter J, Fischer MR, et al. Scaffolding clinical reasoning of medical students with virtual patients: effects on diagnostic accuracy, efficiency, and errors. *Diagnosis (Berlin, Germany)*. 2019;6(2):137-49.
31. Ibiapina C, Mamede S, Moura A, Eloi-Santos S, van Gog T. Effects of free, cued and modelled reflection on medical students' diagnostic competence. *Medical education*. 2014;48(8):796-805.
32. Mamede S, Schmidt HG. The structure of reflective practice in medicine. *Medical education*. 2004;38(12):1302-8.
33. Mamede S, Figueiredo-Soares T, Elói Santos SM, de Faria RMD, Schmidt HG, van Gog T. Fostering novice students' diagnostic ability: the value of guiding deliberate reflection. *Medical education*. 2019;53(6):628-37.
34. Braun LT, Lenzer B, Fischer MR, Schmidmaier R. Complexity of clinical cases in simulated learning environments: proposal for a scoring system. *GMS journal for medical education*. 2019;36(6):Doc80.
35. Hoffrage U, Gigerenzer G. Using natural frequencies to improve diagnostic inferences. *Academic medicine : journal of the Association of American Medical Colleges*. 1998;73(5):538-40.
36. Kunzelmann AK, Binder K, Fischer MR, Reincke M, Braun LT, Schmidmaier R. Improving Diagnostic Efficiency with Frequency Double-Trees and Frequency Nets in Bayesian Reasoning. *MDM policy & practice*. 2022;7(1):23814683221086623.
37. Cook DA, Durning SJ, Sherbino J, Gruppen LD. Management Reasoning: Implications for Health Professions Educators and a Research Agenda. *Academic medicine : journal of the Association of American Medical Colleges*. 2019;94(9):1310-6.
38. Rubinstein G, Osswald A, Zopp S, Ritzel K, Theodoropoulou M, Beuschlein F, et al. Therapeutic options after surgical failure in Cushing's disease: A critical review. *Best practice & research Clinical endocrinology & metabolism*. 2019;33(2):101270.
39. Tausendfreund O, Braun LT, Schmidmaier R. Types of therapeutic errors in the management of osteoporosis made by physicians and medical students. *BMC medical education*. 2022;22(1):323.
40. Ayub N, Faraj M, Ghatan S, Reijers JAA, Napoli N, Oei L. The Treatment Gap in Osteoporosis. *Journal of clinical medicine*. 2021;10(13).

41. Teixeira Rodrigues A, Roque F, Falcão A, Figueiras A, Herdeiro MT. Understanding physician antibiotic prescribing behaviour: a systematic review of qualitative studies. *International journal of antimicrobial agents*. 2013;41(3):203-12.
42. Petersenn S, Beckers A, Ferone D, van der Lely A, Bollerslev J, Boscaro M, et al. Therapy of endocrine disease: outcomes in patients with Cushing's disease undergoing transsphenoidal surgery: systematic review assessing criteria used to define remission and recurrence. *European journal of endocrinology*. 2015;172(6):R227-39.
43. Atkinson AB, Kennedy A, Wiggam MI, McCance DR, Sheridan B. Long-term remission rates after pituitary surgery for Cushing's disease: the need for long-term surveillance. *Clinical endocrinology*. 2005;63(5):549-59.
44. Roelfsema F, Biermasz NR, Pereira AM. Clinical factors involved in the recurrence of pituitary adenomas after surgical remission: a structured review and meta-analysis. *Pituitary*. 2012;15(1):71-83.
45. Barker FG, 2nd, Klibanski A, Swearingen B. Transsphenoidal surgery for pituitary tumors in the United States, 1996-2000: mortality, morbidity, and the effects of hospital and surgeon volume. *The Journal of clinical endocrinology and metabolism*. 2003;88(10):4709-19.
46. Chee GH, Mathias DB, James RA, Kendall-Taylor P. Transsphenoidal pituitary surgery in Cushing's disease: can we predict outcome? *Clinical endocrinology*. 2001;54(5):617-26.
47. Ayala A, Manzano AJ. Detection of recurrent Cushing's disease: proposal for standardized patient monitoring following transsphenoidal surgery. *Journal of neuro-oncology*. 2014;119(2):235-42.
48. Fleseriu M, Hamrahian AH, Hoffman AR, Kelly DF, Katznelson L. American association of clinical endocrinologists and American college of endocrinology disease state clinical review: diagnosis of recurrence in Cushing disease. *Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists*. 2016;22(12):1436-48.
49. Geer EB, Ayala A, Bonert V, Carmichael JD, Gordon MB, Katznelson L, et al. Follow-up intervals in patients with Cushing's disease: recommendations from a panel of experienced pituitary clinicians. *Pituitary*. 2017;20(4):422-9.
50. Braun LT, Zopp S, Vogel F, Honegger J, Rubinstein G, Schilbach K, et al. Signs, symptoms and biochemistry in recurrent Cushing disease: a prospective pilot study. *Endocrine*. 2021.
51. Yoshida K, Fukuoka H, Odake Y, Nakajima S, Tachibana M, Ito J, et al. Multiple Salivary Cortisol Measurements Are a Useful Tool to Optimize Metyrapone Treatment in Patients with Cushing's Syndromes Treatment: Case Presentations. *Frontiers in endocrinology*. 2017;8:375.

52. Braun LT, Vogel F, Zopp S, Rubinstein G, Schilbach K, Künzel H, et al. Diurnal Salivary Cortisol Profiles in Patients with Cushing's Syndrome. *Experimental and clinical endocrinology & diabetes : official journal, German Society of Endocrinology [and] German Diabetes Association*. 2022.
53. Graber ML, Berg D, Jerde W, Kibort P, Olson APJ, Parkash V. Learning from tragedy: the Julia Berg story. *Diagnosis (Berlin, Germany)*. 2018;5(4):257-66.
54. Graber ML, Olson APJ, Barnett T. Learning from tragedy - The Jessica Barnett story: challenges in the diagnosis of long QT syndrome. *Diagnosis (Berlin, Germany)*. 2021;8(3):392-7.
55. Gigerenzer G, Muir Gray JA. Launching the century of the patient. *Better doctors, better patients, better decisions: Envisioning health care 2020*: MIT Press; 2011. p. 3-28.
56. Wegwarth O, Gigerenzer G. Overdiagnosis and Overtreatment: Evaluation of What Physicians Tell Their Patients About Screening Harms. *JAMA Internal Medicine*. 2013;173(22):2086-8.
57. Nuzzo R. Scientific method: Statistical errors. *Nature*. 2014;506(7487):150-2.
58. Makin TR, de Xivry J-JO. Science Forum: Ten common statistical mistakes to watch out for when writing or reviewing a manuscript. *Elife*. 2019;8:e48175.
59. Brown AW, Kaiser KA, Allison DB. Issues with data and analyses: Errors, underlying themes, and potential solutions. *Proceedings of the National Academy of Sciences*. 2018;115(11):2563-70.
60. Braun LT, Riestler A, Osswald-Kopp A, Fazel J, Rubinstein G, Bidlingmaier M, et al. Toward a Diagnostic Score in Cushing's Syndrome. *Frontiers in endocrinology*. 2019;10:766.
61. Pereira M, Dantas Damascena A, Galvão Azevedo LM, de Almeida Oliveira T, da Mota Santana J. Vitamin D deficiency aggravates COVID-19: systematic review and meta-analysis. *Critical reviews in food science and nutrition*. 2022;62(5):1308-16.
62. Munshi R, Hussein MH, Toraih EA, Elshazli RM, Jardak C, Sultana N, et al. Vitamin D insufficiency as a potential culprit in critical COVID-19 patients. *Journal of medical virology*. 2021;93(2):733-40.
63. Bassatne A, Basbous M, Chakhtoura M, El Zein O, Rahme M, El-Hajj Fuleihan G. The link between COVID-19 and Vitamin D (VIVID): A systematic review and meta-analysis. *Metabolism: clinical and experimental*. 2021;119:154753.