Aus dem Institut für Radiologie und Kinderradiologie (CCM)
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Sicherheit und Selbstwirksamkeit junger Ärzte – Beitrag des Medizinstudiums

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Fabian Stroben

aus Hamm

Datum der Promotion: 01.03.2019

1. Inhaltsverzeichnis

2. Zusammenfassung/Abstract

Einführung: Ärztinnen und Ärzte fühlen sich zu Beginn ihrer Weiterbildung unsicher, was zu Verzögerungen und Fehlern in Diagnose und Behandlung führen und so das Patientenwohl gefährden kann. Die Stärkung einer zu niedrigen Selbstsicherheit könnte dies verbessern. Die Inzidenz medizinischer Fehler, u.a. durch mangelhafte Teamarbeit und Kommunikation verursacht, wird auf ca. 9% eingeschätzt. Das Training der Teamarbeit mittels Simulation kann Fehler verringern und wird bereits in der Weiterbildung eingesetzt. Um Medizinstudierende bestmöglich auf ihre spätere Arbeit vorzubereiten, werden Rahmenlehrpläne konzipiert, die die zu erlernenden Kompetenzen während des Studiums festlegen. Ein zukunftssicheres Curriculum gäbe dabei die Möglichkeit, Studierende besser vorzubereiten und so ebenfalls die Selbstsicherheit zu stärken. Die Dissertation untersucht, ob sich mittels Notfallsimulationen Selbstsicherheit steigern lässt, entwickelt Hilfsmittel zum Training der Teamarbeit und untersucht die Delphi-Methode als eine Möglichkeit einen Rahmenlehrplan auf Zukunftssicherheit zu untersuchen.

Methodik: Es wurde ein simulierter Nachtdienst für Studierende konzipiert, in dem diese Notfallsituationen trainieren und Feedback erhalten. Die Änderung der Selbstsicherheit wurde mittels Fragebögen erfasst. Zudem wurde ein Feedbackinstrument zur Teamarbeit entwickelt und ein Assessmentinstrument zum Erfassen der Teamarbeit adaptiert. Darüber hinaus wurde das „Konsensusstatement praktische Fertigkeiten" der Gesellschaft für medizinische Ausbildung mit der Delphi-Methode auf Zukunftssicherheit überprüft und versucht, neue Trends in der Ausbildung zu antizipieren.

Ergebnisse: Vor Beginn der Notfallsimulation waren Studierende eher unsicher, danach fühlten sich Studierende deutlich sicherer. Der Effekt ist unabhängig von der Rolle, die Studierende während der Simulation inne haben. Zur Verbesserung des Feedbacks der Teamarbeit wurde der „TeamTAG" entwickelt, der die Beobachtung der Teamarbeit vereinfacht und gut anwendbar ist. Das Assessmentinstrument „TEAM" zur Beobachtung der Teamarbeit wurde übersetzt und als valide und reliabel bewertet. Mittels Delphi-Methode konnten 288 Lernziele des Konsensusstatement untersucht werden, 147 davon werden als relevant angesehen. Es wurden außerdem elf Thesen zur zukünftigen Entwicklung der medizinischen Praxis untersucht, von denen 7 als eher wahrscheinlich eintreffend bewertet worden sind.

Schlussfolgerung: Simulation ist als Methode der Stärkung der Selbstsicherheit geeignet. TeamTAG und TEAM funktionieren als Instrumente zur Untersuchung der Teamarbeit in Simulationen, mögliche Effekte des TeamTAG auf die Teamarbeit der Studierenden werden aktuell untersucht. Die Delphi-Methode kann genutzt werden, um Rahmenlehrpläne auf ihre

Zukunftssicherheit zu untersuchen, die Bewertung von Thesen kann zudem helfen, zukünftige Trends zu antizipieren um Medizinstudierende besser auf Ihre Arbeit vorzubereiten.

English Version:

Introduction: Junior doctors feel unsafe when starting into postgraduate training. This might lead to delayed diagnosis, treatment, errors and endanger patient safety. Strengthening self-efficacy might improve these problems. Incidence of medical errors e.g. insufficient teamwork and communication is estimated with about 9%. Training of teamwork might reduce these errors and is already applied in postgraduate training. To prepare medical students as well as possible, outcome frameworks are conceptualized to determine required competences during studies. A future-proofed outcome framework would give the possibility to improve students' postgraduate preparedness and – as consequence – their self-efficacy. This thesis examines the use of emergency simulations to strengthen self-efficacy, develops tools for training teamwork and investigates the Delphi method as a possibility to check an outcome framework for future reliability.

Methods: A simulated nightshift for medical students was conceived in which students trained emergency scenarios and underwent feedback. Changes in self-efficacy were measured with questionnaires. Furthermore, a feedback tool for teamwork was developed and a tool for assessing teamwork adapted. At last, the "consensus statement on practical skills" by German Association for Medical Education was examined for future reliability and to anticipate new trends in education.

Results: Before emergency simulations, students were rather unsafe, afterwards students felt way more secure. This effect is independent of the role students have during simulation. For fostering teamwork feedback, the TeamTAG was developed, which is feasible and simplifies observing teamwork. The Assessment tool "TEAM" for observing teamwork was translated and proofed as valid and reliable. With Delphi methods, 288 learning goals of the consensus statement were checked with judging 147 as relevant. Eleven assumptions for future developments in medical practice were examined – seven were judged as possibly occurring.

Take-home points: Simulation is able to improve self-efficacy. TeamTAG and TEAM work as tools for examining teamwork in simulation. Possible effects of the use of TeamTAG on students' teamwork are currently examined. The Delphi method can be used to check outcome frameworks for future reliability. Judging of assumptions might further help to anticipate trends and prepare students better for postgraduate training.

3. Einführung

Selbstsicherheit junger Ärztinnen und Ärzte

Verschiedene Studien haben zeigen können, dass junge Ärztinnen und Ärzte sich bei Beginn ihrer Weiterbildung unsicher fühlen.[1,2] Dieser Befund zeigt sich insbesondere in der Notfallmedizin[3] und differiert zwischen verschiedenen Ländern,[2,4,5] was auf einen Einfluss der Ausbildungssysteme hindeutet. In Deutschland fühlen sich ca. 60% der jungen Ärzte ungenügend auf die klinische Praxis vorbereitet.[4]

Als mögliche Folge kann das Patientenwohl durch medizinische Fehler gefährdet werden, da eine geringe Selbstsicherheit im diagnostischen Prozess zu einer falschen oder verzögerten Behandlung führen kann. Die persönliche und situative Sicherheit hat Auswirkungen darauf, mit welcher Wahrscheinlichkeit eine Person eine Handlung tatsächlich ausführt,[6] eine gestärkte Sicherheit kann dabei die Wahrscheinlichkeit erhöhen, dass die Handlung tatsächlich erfolgt[7,8] und so konkrete (positive) Auswirkungen auf das Patientenwohl haben. Darüber hinaus wird eine geringe Selbstsicherheit als Grund für psychische Gesundheitsprobleme der jungen BehandlerInnen diskutiert.[9]

Medizinische Fehler und ihre Auswirkungen

Das Auftreten medizinischer Fehler wird mit einer Inzidenz um ca. 9% beschrieben, mögliche Folgen können das Patientenwohl gefährden.[10] In Notaufnahmen wird sogar von doppelt so hohen Fallzahlen berichtet.[11] In den USA wird, abgeleitet aus post-mortem Analysen, der Anteil von Fehlern im diagnostischen Prozess mit Auswirkungen auf den Tod des Patienten mit 10% geschätzt.[12] Daneben sind vor allem Fehler im der Bereich der *human factors*[13] und bei der Medikation[14] verbreitet. Als *human factors* werden unterschiedliche Teilbereiche summiert, beispielsweise individuelle Faktoren wie Situationsbewusstsein, Beharrlichkeit und Aufgabenmanagement, aber auch Faktoren der Teamarbeit und Kommunikation und arbeitsplatzbezogene Faktoren.[15] Ca. 45% der Fehler werden als vermeidbar angesehen[10] und erste Studien zeigen, dass die Verbesserung der Teamarbeit Fehler verringern kann.[16] Im Bereich der ärztlichen Weiterbildung hat sich das Training der *human factors* (meist im Rahmen des *Crisis Ressource Managements – CRM*) mittels Simulation bereits etabliert und konnte erste positive Effekte zeigen.[17]

Simulation als Unterrichtsformat

Simulation, also die interaktive Darstellung einer komplexen Kasuistik durch Simulatoren, die Vitalfunktionen vortäuschen und mit den Lernenden interagieren können, wird als Unterrichtsformat bereits weitreichend eingesetzt[18] und zeigt gute Ergebnisse sowohl in der

Verbesserung technischer und kommunikativer Fertigkeiten[19] als auch zum Training der *human factors*.[20,21] Debriefing bzw. Feedback als gemeinsame Diskussion mit bzw. durch die/den Betreuende/n wird als der entscheidende Erfolgsfaktor angesehen, unklar sind aber bis heute die genauen Mechanismen.[22] Daneben wurden unter anderem eine an die Lernenden angepasste Schwierigkeit, wiederholender und longitudinaler Einsatz der Simulationen, Individualisierung und Variierung der Aufgaben identifiziert.[23] Da ähnliche Faktoren, nämlich Feedback, praktisches Training im Wechsel mit theoretischer Ausbildung und Training in Diagnostik die Selbstsicherheit der Lernenden verbessern können,[4,24] erscheint die Simulation neben dem Training der *human factors* auch ideal zur Verbesserung der Selbstsicherheit von Medizinstudierenden und jungen ÄrztInnen.

Da insbesondere in Deutschland die Selbstsicherheit junger ÄrztInnen sehr gering ist, soll die Simulation als Möglichkeit evaluiert werden, um eine Verbesserung der Selbstsicherheit von Medizinstudierenden zu erreichen.

Darüber hinaus soll auf diesen Erfahrungen aufgebaut werden und notfallmedizinische Simulationen im Medizinstudium auf Ihren Nutzen zur Verbesserung der Teamarbeit (als ein Teil der *human factors*) evaluiert werden. Insbesondere wird dabei der Einfluss verschiedener Debriefing-Methoden evaluiert.

Die Rolle von *outcome frameworks* auf die Sicherheit

Neben der Adressierung der oben genannten Unsicherheit junger ÄrztInnen durch konkrete Interventionen scheint ein weiterer Grund für die Unsicherheit die fehlende Integration mutmaßlich relevanter Fertigkeiten in die Curricula und Prüfungen zu sein. Dies ergaben eine Absolventenbefragung der Universität Köln (Stosch C et al., unveröffentlich) und eine nationale Befragung.[25] Aus diesem Zweck wurden in den letzten Jahren vermehrt Rahmenlehrpläne, sog. *outcome frameworks*, veröffentlicht, die kompetenzorientiert zu erwerbende Fähigkeiten und Kenntnisse beschreiben.[26,27] Ein typisches Beispiel sind die CanMeds,[28] in Deutschland wurde vor kurzem der „Nationale Kompetenzbasierte Lernzielkatalog Medizin" (NKLM) veröffentlicht.[29] Kritikpunkte der *outcome frameworks* sind unter anderem die eingeschränkte Vergleichbarkeit untereinander[30] und die fehlende Spezifität für die tatsächliche Tätigkeit – Ärztinnen und Ärzte fühlen sich so zu Beginn der klinischen Tätigkeit nicht sicher.[31] Da die Curriculumsentwicklung zudem ein langwieriger Prozess ist,[32] kann begründet angenommen werden, dass eine Orientierung am aktuellen Stand der Medizin unter Umständen nicht mit den rasanten Entwicklungen der Medizin Schritt hält und jungen ÄrztInnen am Beginn ihrer Weiterbildung nicht die notwendigen

Fertigkeiten vermittelt wurden und dies einen Beitrag zur bestehenden geringen Selbstsicherheit leistet.
Zur Analyse von zukünftigen, noch unsicheren Entwicklungen und zur Erfassung von Handlungsoptionen hat sich die sogenannte Delphi-Methode als ein wertvolles Instrument in unterschiedlichen Kontexten[33] – auch der Humanmedizin[34] - erwiesen und aggregiert in einem mehrstufigen Verfahren Gruppenmeinungen zur Bewertung von Sachverhalten.[35]
In der hier vorliegenden Dissertation soll untersucht werden, ob durch Nutzung der Delphi-Methode die Untersuchung eines *outcome frameworks* auf ihre Zukunftsfähigkeit möglich ist. Als *outcome framework* dient konkret das „Konsensusstatement Praktische Fertigkeiten im Medizinstudium“,[36] der einen Grundstein für die Entwicklung des NKLM legte.

Insgesamt werden drei übergeordnete Fragestellungen im Rahmen dieser Dissertation untersucht:

1. Ist Simulation prinzipiell geeignet um die Selbstsicherheit Medizinstudierender zu verbessern?
2. Kann ein erweitertes Notfalltraining für Medizinstudierende eine Stärkung der Sicherheit im Bereich der Teamarbeit bzw. *human factors* erreichen?
3. Kann mithilfe der Delphi-Methodik die Zukunftsausrichtung eines *outcome frameworks* untersucht werden?

4. Methodik

Simulation als Stärkung der Selbstsicherheit Studierender

Zur Adressierung der primären Fragestellung wurde die Intervention „Nachtdienst: Bist du bereit“ konzipiert und erstmals im September 2013 durchgeführt. Die detaillierte Darstellung lässt sich im Artikel[37] nachlesen, im Folgenden findet sich eine kurze Zusammenfassung der Methodik und Umsetzung:

An dieser ca. sechsstündigen Simulation können Studierende im letzten Jahr ihres Studiums teilnehmen. Die Teilnehmenden durchlaufen in Gruppen zu fünf Personen verschiedene Simulationen mit häufigen Szenarien der klinischen Notfallmedizin. Umgesetzt werden die Fälle mittels Simulatoren und durch den Einsatz von SimulationspatientInnen (SPs). Betreut werden die Studierenden von TutorInnen des Lernzentrums, die die Gruppen organisatorisch und die Fälle inhaltlich betreuen. Feedback wird von TutorInnen, KommilitonInnen und SPs als sog. Multi-Source-Feedback, ein Feedback aus möglichst vielen Perspektiven zur Schaffung eines Lernerfolgs (Teamarbeit, Kommunikation, Verhalten) für die

Teilnehmenden, gegeben.[38]

Teilnehmende

Teilnahmeberechtigt an der Simulation sind maximal 30, sich freiwillig meldende Studierende der Charité-Universtitätsmedizin im letzten Abschnitt ihres Studiums, dem praktischen Jahr (PJ). Die Verteilung der Plätze erfolgte nach dem „first come-first served“ Prinzip, die Gruppen wurden geschlechterstratifiziert und randomisiert zusammengestellt. Alle Teilnehmenden wurden vorher mündlich und schriftlich über die Studie aufgeklärt und gaben ihr schriftliches Einverständnis zur Datenerhebung.

Bei jedem Fall definieren die Gruppen eine/n Teilnehmende/n als Teamleitung, die/der für Entscheidungen verantwortlich ist und das Team koordiniert, und je nach Fall zwei bis vier Mitglieder, die der Teamleitung assistieren. Passive Teilnehmende beobachten Kommunikation und Inhalt und werden dabei mittels Checklisten unterstützt. Die Aufteilung der Rollen wechselt nach jedem Fall.

Fälle

Bei den zu bearbeitenden Fällen wurde insbesondere auf eine breite Auswahl, eine gute Umsetzung mittels Simulation und ein adäquater Schwierigkeitsgrad für die Studierenden geachtet. Jeder Fall adressiert eine spezifische Fachdisziplin und hat einen individuellen Fokus des Feedbacks. Eine detaillierte Beschreibung der Fälle mit Hinterlegung der genutzten Leitlinien finden sich als „table 1“ und „Additional file 1 & 2“ im entsprechenden Artikel,[37] die Tabelle 1 gibt einen kurzen Überblick über die Fälle und den Feedbackfokus.

Tabelle 1. Fälle des „Nachtdienst: Bist du bereit? 2013“

Disziplin	Kasuistik	Fokus des Feedbacks
Kardiologie	Patient mit **Myokardinfarkt und ST-Streckenhebung (STEMI)** + Herzrhythmusstörungen	Stringentes Arbeiten und schnelle Verlegung
Neurologie	Patientin mit ischämischem **Schlaganfall** (Aphasie + Hemiparese) mit Kontraindikation für Lyse	Umgang mit Kommunikationsbarrieren in der Notfallsituation
Chirurgie 1	**Alkohol**-intoxikierter Patient mit **Kopfplatzwunde** nach Fahrradunfall	Einschätzung intoxikierter Patienten
Pulmologie	Patient mit **exazerbierter chronisch obstruktiver Lungenerkrankung** (viral), nicht kooperativ, starker Nikotinabusus	differentialdiagnostisches Denken, Umgang mit nicht-kooperativem Patienten
Chirurgie 2	Mitarbeiter mit **Milzruptur** nach Sturz in Klinik, Erstversorgung am Unfallort und Versorgung im Schockraum	Systematische Arbeit im Schockraum nach Traumaleitlinie und unter Anleitung
Anästhesie	**Reanimation** bei Kammerflimmern bei Diagnose STEMI	Arbeit unter beengten Verhältnissen, klare Kommunikation bei Reanimation
Urologie	Patientin mit unkompliziertem **Harnwegsinfekt** bei Erstdiagnose Gravidität	Umgang mit schwangeren Patientinnen, Behandlung einer ambulanten Patientin

Tabelle adaptiert aus [37]

Datenerhebung September 2013

Die Datenerhebung erfolgte als prospektive Längsschnittstudie ohne Kontrollgruppe. Die Teilnehmenden füllten vor Beginn der Nacht einen Fragebogen aus, der neben demographischen Daten und potentiellen Confoundern die Selbsteinschätzung der Studierenden bzgl. Sicherheit in den verschiedenen klinischen Disziplinen erfragte.

Nach Durchlaufen jedes Szenarios und vor Beginn des Feedbacks schätzten alle Teilnehmenden ihre Sicherheit bezogen auf den bearbeiteten Fall ein, Teammitglieder und Beobachtende schätzten außerdem die Sicherheit ihrer Teamleitung ein.

Fünf Tage nach dem Nachtdienst erfolgte online (LimeSurvey; Hamburg, Deutschland) eine Nachbefragung zur Selbsteinschätzung der Studierenden bzgl. Sicherheit in den verschiedenen klinischen Disziplinen analog zur Befragung zum Beginn der Nacht.

Außerdem wurden nach jeder Simulation und als Abschlussevaluation am Ende der Nacht allgemeine Qualitätskriterien zur Feedbackgüte, Organisation, Versorgung und zu den einzelnen Fällen und ihrer Umsetzung erfragt.

Alle Evaluationen nutzten 7-stufige Likert-Skalen (+3: stimme voll zu // -3: stimme gar nicht zu) und Freitext-Kommentarmöglichkeiten. Das Sicherheitsgefühl wurde ebenfalls als 7-stufige Likert-Skala (+3: sehr sicher // -3: sehr unsicher) erfragt.

Stärkung der Teamarbeit durch Simulation

Das bestehende Konzept wurde 2016 um einen Fokus auf *human factors* erweitert, um Studierenden so einen Einblick in die strukturierte und sichere Teamarbeit zu geben.

Aus diesem Grund wurden die TutorInnen zur Durchführung des Teamarbeit-fokussierten Debriefings fortgebildet: Es wurde eine strukturierte Schulung entwickelt und den TutorInnen eine Checkliste zur Verfügung gestellt. Genaue Informationen zu diesen Punkten findet sich im Artikel,[39] es folgt hier eine kurze Zusammenfassung der Studie:

Entwicklung der Checkliste „TeamTAG“

Vorlage für die Checkliste bildeten die 15 CRM-Prinzipien, die in Deutschland weit verbreitet sind.[15] Die Forschungsgruppe wählte aus diesen Prinzipien sechs aus, die einer Beobachtung zugänglich und für den Nachtdienst passend sind. Ergänzt wurden die Prinzipien um behaviorale Anker, also konkrete Verhaltensweisen, die bei Ausführung des Prinzips gezeigt werden und die Beobachtung erleichtern. Die Checkliste trägt den Namen

„Teamwork Techniques Analysis Grid (TeamTAG)“ und ist als „online supplementary material“ zum Artikel[39] verfügbar.

Training des Debriefings für TutorInnen

Alle TutorInnen, die während des Nachtdienstes Teamarbeit-fokussiertes Debriefing anleiten, werden vorher in einem neu konzipierten Training geschult: Neben einer allgemeinen Einführung zu CRM und *human factors* werden alle 15 CRM-Prinzipien an Videobeispielen erläutert (adaptiert aus dem Konzept des *frame of reference training*).[40] Die Videobeispiele wurden eigens angefertigt und zeigen standardisierte Reanimationsszenarien mit unterschiedlich guter Teamarbeit. Außerdem wurde den TutorInnen eine teilstandardisierte Struktur zur Anleitung des Debriefings vermittelt.[41]

Studienaufbau & Datenerhebung Januar 2016

Im Januar 2016 wurde der TeamTAG erstmalig zur Unterstützung des Debriefings der GruppentutorInnen während des „Nachtdienst: Bist du bereit?“ eingesetzt und in einer Anwendungsstudie ohne Kontrollgruppe auf Anwendbarkeit und Akzeptanz für die TutorInnen untersucht. 2016 wurden 35 Studierende in sieben Gruppen für den Nachtdienst rekrutiert, die sonstige Durchführung erfolgte analog zu oben beschriebener in 2013.

Die Evaluation des TeamTAG erfolgte mittels Fragebogen nach Ende der Simulationen und erfragt die Anwendbarkeit der Checkliste, die wahrgenommene Schwierigkeit der Aufgabe und lässt die TutorInnen zudem bewerten ob ausreichend Zeit zum Debriefing zur Verfügung stand. Die Bewertung des Feedbacks durch die Teilnehmenden erfolgte analog zur 2013 durchgeführten Methodik direkt nach jedem Fall und am Ende des Nachtdienstes. Die Evaluationen nutzen 7-stufige Likert-Skalen (+3: stimme voll zu // -3: stimme gar nicht zu) und Freitext-Kommentarmöglichkeiten.

Studienaufbau 2017

Basierend auf der guten Bewertung des TeamTAG (Vgl. auch Ergebnisteil und andere Studien)[39,42] wurde für den folgenden Nachtdienst 2017 eine kontrolliert-randomisierte Experimentalstudie konzipiert, die die Wirksamkeit des TeamTAG auf den Lernerfolg der Studierenden bezüglich CRM und die konkrete Anwendung von Teamarbeit evaluiert. Der Rahmen des Nachtdienstes bleibt wie oben beschrieben, die Teilnehmenden erhalten Feedback zu ihrer Teamarbeit durch die TutorInnen: In der Interventionsgruppe wird ein teilstandardisiertes Debriefing unter Nutzung des TeamTAG durchgeführt, hierbei sollen alle im TeamTAG enthaltenen CRM-Prinzipien in den ersten fünf Szenarien thematisiert werden und bei unzureichender Besserung der Teamarbeit wieder aufgegriffen werden. Die TutorInnen der Kontrollgruppen leiten lediglich ein teilstandardisiertes Debriefing an und können Schwerpunkte, inwiefern CRM-Prinzipien besprochen werden, nach individueller Abwägung setzen. Alle TutorInnen durchlaufen das gleiche Training (analog zu 2016), danach erfolgt lediglich eine kurze, separate Information zu der Kondition, in die die TutorInnen randomisiert worden sind. Die Randomisierung erfolgt stratifiziert nach Vorausbildung und Fortschritt im Studium.

Es werden insgesamt 20 Teilnehmende (vier Gruppen) in die Interventionsgruppe und 15 Teilnehmende (drei Gruppen) in die Kontrollgruppe randomisiert. Teilnehmende werden schriftlich und mündlich über die Durchführung der Studie aufgeklärt und geben vor Beginn der Studie ihre schriftliche Zustimmung zur Datenerhebung.

Datenerhebung Januar 2017

Zur Erfassung der Vorerfahrungen werden zu Beginn der Simulationsnacht von allen Teilnehmenden Confounder und Vorerfahrungen in der Notfallmedizin erfragt. Innerhalb der Gruppen werden Vorerfahrung zu Prinzipen der Teamarbeit und gemeinsam notfallmedizinische Problemstellungen diskutiert (adaptiert als *team readiness assurance test* aus der Methodik des *team based learning* [43,44]).

Die Teamarbeit wird mittels des standardisierten „Team Emergency Assessment Measure (TEAM)“ Fragebogens durch zwei unabhängige RaterInnen erfasst.[45] Da der Fragebogen noch nicht auf Deutsch verfügbar ist, wurde dieser übersetzt und in einer Vorstudie validiert (siehe Ergebnisteil und Studie[39]). In der Auswertung wird die Leistung der Studierenden vom ersten zum letzten Szenario und zwischen Interventions- und Kontrollgruppe verglichen. Es ist davon auszugehen, dass sich zwar alle Gruppen im Laufe der Nacht verbessern, die Steigerung aber bei den Interventionsgruppen höher zu erwarten ist. Für die zugehörige Fallzahlberechnung sei abermals auf den Artikel[39] verwiesen.

Nach jeder Station geben die Studierenden außerdem die Zufriedenheit mit dem Feedback der verschiedenen TutorInnen und SP an. Nach Ende des Nachtdienstes werden die Studierenden gefragt, welche Prinzipien im Laufe der Nacht diskutiert worden sind und für wie relevant sie diese halten. Annahme der Studie ist, dass Studierende der Interventionsgruppe mehr Prinzipien diskutieren und diese höher werten als in der Kontrollgruppe, da durch Nutzen des TeamTAG ein spezifischeres und höherwertiges Feedback möglich ist.

Die Evaluationen nutzen 7-stufige Likert-Skalen (+3: stimme voll zu // -3: stimme gar nicht zu) und Freitext-Kommentarmöglichkeiten. Der TEAM-Fragebogen enthält insgesamt 11 Skalen zur Teamarbeit mit jeweils 5-stufigen Likert-Skalen und eine allgemeine Teamarbeitsbewertung mit einer Skala von 1-10.

Untersuchung eines Outcomeframeworks unter dem Aspekt der Zukunftsausrichtung

Zur beispielhaften Untersuchung des “Konsensusstatements praktische Fertigkeiten“[36] auf seine Zukunftsfähigkeit wurde dieses mit der Delphi-Methodik untersucht. Die Methodik der durchgeführten Delphi-Studie ist ausführlich im Artikel[46] beschrieben, hier erfolgt lediglich eine kurze Zusammenfassung der relevanten Teile:

Zur Vorbereitung der Delphi-Studie wurden teilstrukturierte Interviews mit Stakeholdern durchgeführt, um Experteneindrücke zu eventuell eintretenden Trends in der Medizin zu identifizieren, welche vorher in einer Literaturrecherche ermittelt wurden. Die befragten Experten sind in Tabelle 1 des Artikels[46] dargestellt, die Ergebnisse wurden qualitativ mittels induktiver Kategorisierung nach Mayring[47] ausgewertet, von der Arbeitsgruppe konsentiert und zu elf Thesen aggregiert.

In der sich anschließenden Arbeitsphase wurden zum einen die 289 Lernziele des Konsensusstatements auf ihre Zukunftsfähigkeit durch KlinikerInnen bewertet und zum anderen die elf Zukunftsthesen auf ihre Eintrittswahrscheinlichkeit eingeschätzt. Es wurden ca. 8000 ÄrztInnen der Universitätskliniken und niedergelassene ÄrztInnen per Mail eingeladen teilzunehmen. Teilnehmenden ÄrztInnen bewerteten dabei nicht alle 289 Lernziele, sie wurden in eine von zehn möglichen Gruppen randomisiert und bewerteten so ca. 30 Lernziele pro Gruppe. Der Delphi-Methodik folgend wurde ein Teil der Lernziele des Konsensusstatements (Auswahl nach Mittelwert/Standardabweichung und Meinung der Forschungsgruppe) in einer zweiten Delphi-Runde unter Vorlage der Erstrundenergebnisse wieder vorgelegt. Auch hier folgte eine Randomisierung in Gruppen, sodass jeder Experte ca. 50 Lernziele erneut bewertete.

Die Bewertung aller Lernziele und Thesen erfolgte auf 4-stufigen Likert-Skalen (1: hohe Relevanz // 4: keine Relevanz). Die Befragung wurde als Onlinebefragung (LimeSurvey) realisiert. Demographische Daten wurden auf freiwilliger Basis erhoben um die ExpertInnen mit der Gesamtheit aller zu diesem Zeitpunkt in Deutschland tätigen ÄrztInnen vergleichen zu können, hierbei wurden die Daten der offiziellen Statistik der Bundesärztekammer 2014 genutzt.[48]

5. Ergebnisse

Simulation als Stärkung der Selbstsicherheit Studierender

Am Nachtdienst 2013 nahmen insgesamt 30 Studierende (davon 20 weiblich) teil. Vor Beginn der Nacht waren die Studierenden eher unsicher (Mean = -0.34), dies änderte sich direkt nach den Fällen (Mean = 0.95) und im Follow-Up nach fünf Tagen (Mean = 0.66), die Effektgröße ist groß (Cohen's *d*= 1.86). Die genaue Darstellung der Änderung des Sicherheitsgefühls zwischen Beginn und Follow up ist in Tabelle 2 dargestellt.

Tabelle 2. Sicherheitsgefühl und Änderung von vor Beginn des Nachtdienstes zum Follow Up

Fachdisziplin	Sicherheitsgefühl vor Beginn des Nachtdienstes (Mittelwert & SD)	Sicherheitsgefühl im Follow-Up nach 5 Tagen (Mittelwert & SD)	*p*-Wert
Gesamt	-0.34 (0.49)	0.66 (0.59)	0.001 *
Anamnese	1.27 (1.02)	1.72 (0.9)	0.035 *
Anästhesie	0.14 (1.06)	1.17 (0.62)	<0.001 *
Urologie	-0.77 (1.25)	0.28 (1.53)	0.013 *
Kardiologie	-0.1 (1.06)	0.28 (1.13)	0.145
Pulmologie	-0.4 (0.97)	0.11 (1.13)	0.07
Chirurgie	0.13 (1.33)	0.83 (1.3)	0.101
Neurologie	-0.47 (1.07)	0.22 (1.11)	0.1

7-stufige Likert-Skala (+3: sehr sicher // -3: sehr unsicher) für jedes Item. * p<0.05, SD = Standardabweichung. Tabelle adaptiert aus [37].

Mittels Varianzanalyse (ANOVA mit Messwiederholung) wurde außerdem der Einfluss der Rolle während der Simulation auf die Veränderung des Sicherheitsgefühls untersucht. Hier zeigt sich, dass die angegebene Sicherheit nach Ende des Szenarios unabhängig von der Rolle während des Szenarios ist ($F(2,52) = 0.123$; $p = 0.884$). Die Einschätzung der Sicherheit der Teamleitung durch beobachtende und aktive Teilnehmende ist nicht unterschiedlich zu der Selbsteinschätzung der Teamleitung ($F(2,52) = 2.055$; $p = 0.138$) und damit unabhängig von der Rolle – in der Beobachtung der Sicherheit der Teamleitung zeigt sich jedoch ein Zusammenhang zwischen der eigenen Sicherheit der aktiven Teammitglieder und der eingeschätzten Sicherheit ($r = 0.61$; $p < 0.001$), dies besteht für passive Beobachter nicht.

Das Feedback, der Einsatz von Simulation/SPs und die Betreuung während der Nacht wurde insgesamt sehr positiv durch die Teilnehmenden bewertet, für Details sei auf die Originalpublikation[37] verwiesen.

Stärkung der Teamarbeit durch Simulation

Zur Verbesserung des Teamarbeit-fokussierten Debriefings wurde der „TeamTAG“ erarbeitet. Dieser umfasst insgesamt sechs der 15 bekannten CRM-Prinzipien und wurde in der Arbeitsgruppe konsentiert. Die enthaltenen CRM-Leitlinien umfassen a) Antizipiere & plane voraus, b) setze Prioritäten dynamisch, c) Reevaluiere die Situation immer wieder, d) Hilfe anfordern, lieber früh als spät, e) Übernimm die Führungsrolle/ sei ein gutes Teammitglied und f) Kommuniziere sicher und effektiv. Der gesamte TeamTAG ist in [39] veröffentlicht.

Zum Nachtdienst 2016 wurde der konzipierte Nachtdienst erstmals auf seine Anwendbarkeit untersucht und die TutorInnen (n = 7), die den TeamTAG nutzen, nachbefragt. Hier zeigte sich, dass die Anwendung als gut bewertet wurde (Mean = 1.9; SD = 0.9) und sowohl bei der Beobachtung der Zusammenarbeit der Teilnehmenden (Mean = 2.3; SD = 0.8) und als auch beim Geben des Feedbacks hilfreich ist (Mean = 2.3; SD = 0.5). Hingegen wurde die zur Verfügung stehende Zeit zum Geben des Feedbacks als heterogen erachtet (Mean = -0.3; SD = 1.1). Teilnehmende (n = 35) fanden das Feedback der GruppentutorInnen nützlich (Mean = 1.7; SD = 1.0).

In der Vorplanung der Experimentalstudie 2017 wurde zur Messung der Teamarbeit der TEAM verwendet. Da dieser nicht auf Deutsch vorlag, wurde dieser strukturiert durch zwei unabhängige Personen übersetzt und durch einen Muttersprachler wieder rückübersetzt. Die Ergebnisse wurden gemeinsam konsentiert. Das genaue Vorgehen ist in der veröffentlichen Studie[39] dargestellt, hier findet sich ebenfalls die übersetzte Fassung des TEAM. Vor Einsatz im Nachtdienst 2017 wurde eine Voruntersuchung zur Testung der Fragebogengüte durchgeführt. Hier wurde der TEAM durch vier unabhängige Rater (Untersuchung der Reliabilität) anhand von zwei unterschiedlich schweren Reanimationsszenarien (Untersuchung der Validität), die auf Video vorlagen, genutzt um die Teamarbeit zu bewerten. Tabelle 3 zeigt die Bewertung der Rater.

Tabelle 3: Voruntersuchung des Fragebogens, Ergebnisse der Rater-Übereinstimmung

Szenario	Mittelwert TEAM-Summenscore	Standardabweichung TEAM-Summenscore	Intra-Klassen Korrelation
Reanimation 1: gute Teamarbeit	42.3	1.3	0.99
Reanimation 2: schlechte Teamarbeit	22.5	3.1	0.85

TEAM = Team Emergency Assessment Measure; TEAM-Summenscore gebildet aus den 11 Skalen des Fragebogens, mögliche Werte: 0-44. Tabelle erstellt aus [39].

Untersuchung eines *outcome frameworks* unter dem Aspekt der Zukunftsausrichtung

Zur Teilnahme an der Delphi-Umfrage zur Zukunftssicherheit eines *outcome frameworks* meldeten sich insgesamt 738 ExpertInnen an, von diesen konnten 594 Datensätze zur Bewertung der Lernziele (19.5% Dropout) und 651 Datensätze zur Bewertung der Zukunftsthesen (11.8% Dropout) verwendet werden. Die Auswertung der demographischen

Daten zeigt, dass ein Großteil (87.9%) der ExpertInnen an einem Maximalversorger beschäftigt sind und über mindestens ein Jahr Berufserfahrung verfügen (96.0% der Teilnehmenden). Insgesamt waren 26 Disziplinen vertreten. An der zweiten Delphi-Runde nahmen noch 314 ExpertInnen teil, hier konnten 188 vollständige Datensätze (Dropout 40.1%) beachtet werden. Die genaue demographische Darstellung der ExpertInnen und ein Vergleich zur Statistik der Bundesärztekammer findet sich im Artikel,[46] Tabelle 2.

Eine Übersicht der Bewertung der elf Zukunftsthesen durch o.g. ExpertInnengruppe findet sich in Tabelle 4, insgesamt sind sieben der elf Thesen eher angenommen worden.

Tabelle 4. Antworten der Experten zu den elf Zukunftsthesen.

These	**N**	**Mittelwert**	**Zustimmung**
Aufgrund des demographischen Wandels besitzen spezielle Kenntnisse und Fertigkeiten im kommunikativen Umgang mit an Demenz erkrankten Patienten einen erhöhten Stellenwert für alle Fachbereiche der Erwachsenenmedizin im Jahr 2025.	626	1.7	551
Im Jahr 2025 werden bisher ausschließlich ärztliche Tätigkeiten auch von nichtärztlichen Berufsgruppen durchgeführt und abgerechnet.	651	1.8	509
Fernüberwachung von Patienten, Konsultationen über Video-Telefonie und Übertragung von Laborwerten durch den Einsatz von Internet und Smartphone sind im Jahr 2025 akzeptiert und werden bei der Mehrheit der Patienten eingesetzt.	626	1.9	476
Der Arzt im Jahr 2025 ist ein Gesundheitsmanager, dessen Ausbildung um grundlegende Kenntnisse der Organisation und Betriebswirtschaftslehre erweitert werden müssen.	651	1.9	522
Neue Möglichkeiten der Diagnose und Therapie durch innovative Anwendungen im IT-Bereich führen zukünftig zu weniger physischen Kontakt zwischen behandelndem Arzt und Patient.	651	2.1	449
Das primäre Kriterium der Auswahl von Versorgungs- und Behandlungsmöglichkeiten von Patienten im Jahr 2025 sind finanzielle Aspekte.	651	2.2	285
Der Informationsgradient zwischen Arzt und Patient nimmt weiter ab. Deshalb entscheidet nicht die ärztliche Autorität, sondern seine Fähigkeit zur kommunikativen Vermittlung und Argumentation in Bezug auf Diagnose und Therapie zukünftig über die Behandlung.	626	2.2	402
Das wichtigste Werkzeug des Arztes im Jahr 2025 sind seine Hände.	651	2.8	196
Im Jahr 2025 ist die medizinische Grundversorgung durch Haus- und Fachärzte überwiegend mittels mobile Versorgungskonzepte, wie beispielsweise Tagespraxen, Hausbesuche oder Busse sichergestellt, anstatt durch lokal ansässige Praxen.	626	2.9	164
Der Arzt im Jahr 2025 ist austauschbar in seiner Person und wird von den Patienten vor allem in seiner Funktion aufgesucht: der Zugang zu Therapie und Diagnostik und nicht mehr der persönliche Kontakt sind entscheidend.	651	3.0	171
Anamnese und Diagnose werden von zertifizierten IT-Systemen automatisiert durchgeführt. Bei Bedarf werden speziell geschulte Ärzte hinzugezogen.	626	3.0	157

Bewertung auf 4er Likert Skala (1 = sehr wahrscheinlich […] 4 = sehr unwahrscheinlich), Zustimmung aggregiert aus Punktwerten 1-2 der 4-stufigen Likert-Skala. Tabelle adaptiert aus [46].

Die 289 Lernziele des Konsensusstatements wurden in zwei Runden bewertet. Akzidentiell wurde ein Lernziel nicht in die Befragung übernommen. In der ersten Befragungsrunde wurden 240 Lernziele als relevant, 47 als irrelevant und eines indifferent bewertet. Von diesen Lernzielen wurden 103 (71 relevante, 31 irrelevante und ein irrelevantes Lernziel/e) in die zweite Runde übernommen. Insgesamt sind in beiden Runden von 288 Lernzielen 231 Lernziele als relevant und 57 Lernziele als irrelevant bewertet worden. Die Lernziele wurden ebenfalls getrennt nach Wahl- und Kernlernzielen untersucht – hier fällt auf, dass insgesamt 54.5% der nicht-relevanten Lernziele Wahllernziele sind. Eine genauere Darstellung der Bewertung aller Lernziele und weitere Untersuchungen von Untergruppen findet sich im Artikel[46] und im Online Appendix des Artikels.

6. Diskussion

Wie in der bereits publizierten Literatur[1–3] schätzen sich PJ-Studierende an der Charité in der Anwendung ihres Wissen als eher unsicher ein.[37] Die zu evaluierenden Instrumente der Simulation und des Teamwork-Trainings konnten in mehreren aufeinander aufbauenden Durchführungen erfolgreich untersucht werden.[37,39] Hier zeigt sich, dass die Methodik der Simulation hohe positive Effekte auf teilnehmende Studierende hat. So geben Teilnehmende des Nachtdienstes im September 2013 nach Absolvieren des Simulationstrainings direkt nach den Szenarien und nach fünf Tagen deutliche Steigerungen der Selbstsicherheit an.[37] Darüber hinaus lassen sich noch weitere wichtige Effekte ableiten: Die Zunahme der Selbstsicherheit ist unabhängig von der Rolle der Studierenden während der Simulation – Beobachtende geben einen genauso hohen Zuwachs an wie Agierende.[37] Dieser Effekt wurde bereits beim Erlernen praktischer Fertigkeiten beschrieben,[49] erfolgreich angewendet[50] und kann nun erweitert werden: Lehrende können Beobachtungsaufgaben während Simulationstrainings verteilen, ohne Einbußen an den subjektiven Trainingseffekten ihrer Teilnehmenden befürchten zu müssen. Es zeigt sich außerdem, dass Beobachtende die Selbstsicherheit der Teamleitung einschätzen können, was a.e. durch die Wahrnehmung des situativen Verhaltens und behavioraler Indikatoren möglich ist.[37] Da Teamarbeit in der Lage ist, die Diagnosekompetenz der BehandlerInnen zu verbessern,[51] könnte eine Wahrnehmung von Unsicherheit innerhalb eines Team eine Möglichkeit bieten, um Behandlungsfehler vorherzusehen. Diese These wird aktuell in einer anderen Studie untersucht.[11]

Das Training der Teamarbeit i.S. von *human factors*, welches bisher vor allem in der ärztlichen Weiterbildung Anwendung gefunden hat, wurde ebenfalls innerhalb von Simulationen untersucht.[17] Während die Auswertung der randomisierten Untersuchung zum Zeitpunkt des Abschlusses dieser Promotion noch aussteht, konnten bereits einige Schritte zu einer konkreten und effektiven Anwendung während Notfallsimulationen PJ-Studierender

geleistet werden: Es wurde der TeamTAG als ein Hilfsmittel zur Verbesserung des Debriefings entwickelt, der in einer ersten Erhebung als hilfreich bewertet wurde.[39] Der allgemeine Nutzen von Checklisten (oder sog. *cognitive aids*) wurde bereits in anderen Kontexten hinreichend gezeigt[52,53] und scheint insbesondere bei der Anwendung von noch unerfahrenen Betreuenden hilfreich.[54] Deswegen kann in dem hier geplanten Kontext von einem konkreten Nutzen für die Teilnehmenden ausgegangen werden kann.
Da die vorherigen Erhebungen auf Eigenangaben der Studierenden basierten und bisher keine objektiven Bewertungen der Leistung erhoben wurden, wird auch dies im Rahmen der nächsten Erhebung untersucht. Mit dem TEAM wurde ein exzellent evaluiertes Instrument zur Messung der Teamarbeit ausgewählt, im Rahmen dieser Dissertation übersetzt und in einer ersten Erhebung auf die Testgütekriterien untersucht: Hier zeigt sich eine ausreichende Inhaltsvalidität durch die unterschiedliche Bewertung der beiden Szenarien und eine sehr gute Reliabilität.[39] Die erzielten Ergebnisse lassen sich gut mit den Ergebnissen der bisherigen Validierungen des französischen[55] und englischen TEAM[45,56] vergleichen und legen nahe, dass auch die deutsche Version des TEAM bestens geeignet ist, um Teamarbeit zu messen und diese in Zukunft als objektivere Maßzahl sowohl für Studien und Qualitätssicherung als auch zu Lehrzwecken einsatzbereit ist, was von anderen AutorInnen ausdrücklich gefordert wurde.[16,57]

Neben der konkreten Integration neuer Methoden zur Stärkung der Selbstsicherheit kommt aber auch der Anpassung des Curriculums und der sinnvollen Weiterentwicklung eine wichtige Rolle zu. Um Sicherheit zu vermitteln, sollten diese spezifisch auf die spätere Tätigkeit zutreffen.[31] Möglichkeiten zur Antizipation von Trends in der Medizin wurden aber bisher nicht untersucht. Aus diesem Grund wurde die Delphi-Methodik eingesetzt und konnte sowohl zukünftige Trends wie die Zunahme von telemedizinischen Anwendungen vorhersagen als auch die bestehenden Lernziele eines *outcome frameworks* untersuchen.[46] Die getätigten Aussagen der Umfrage sind dabei keinesfalls spekulativ – telemedizinische Anwendungen finden bereits Anwendung in der Praxis[58] und haben zur Formulierung eines eigenen Lernzielkatalogs beigetragen.[59] Weitere konkrete Anwendungsmöglichkeiten lassen sich aus der Diskussion der entsprechenden Studie und ihrem Anhang ableiten.[46]

Es müssen einige Limitationen angemerkt werden, diese setzen sich aus den Einzelstudien zusammen:[37,39,46]
Allen Studien ist die Modellhaftigkeit ihrer Ergebnisse gemein. Die Generalisierbarkeit steht aufgrund des monozentrischen Studiendesigns aus, dementsprechend sollten weitere Studien versuchen die Ergebnisse zu bestätigen, obwohl sich diese bereits gut in die Literatur

einfügen. Darüber hinaus sollten die Studienergebnisse an weiteren und größeren Populationen erprobt werden, wie es z.B. durch eine curriculare Implementierung in bestehende Notfallkurse geschehen könnte.
An den Nachtdienst-Studien konnten jeweils nur wenige Probanden teilnehmen, was vor allem finanzielle und logistische Gründe hat. In der Datenerhebung 2013 wurde deswegen eine post-hoc Poweranalyse durchgeführt,[37] für die Datenerhebung der randomisierten Studie wurde die nötige Fallzahl vorher berechnet.[39]
Die Verbindung zwischen Selbstsicherheit und konkreten Änderungen der Leistung wird kontrovers diskutiert.[60,61] Demgegenüber steht die Aussage, dass eine hohe Selbstsicherheit die Chance erhöht, eine Aufgabe erfolgreich zu lösen.[62] Diese postulierte Verbindung konnte in Studien bereits erfolgreich gezeigt werden.[50,63] Um diese Kritik antizipieren zu können, wurde mit dem TEAM ein sinnvolles Messinstrument übersetzt, welches objektivere Messwerte der Teamarbeit generiert. Die durchgeführte Validierung ist aufgrund der Fallzahl ebenfalls als vorläufig zu betrachten, die Ergebnisse korrespondieren aber bereits gut mit der Literatur.
Durch die Nutzung einer Checkliste, wie sie der TeamTAG darstellt, kann die Feedbackqualität zwar erhöht werden, eine allgemeine Aussage über die Feedbackgüte bezogen bspw. auf Struktur, Inhalt und Zentrierung auf die Lernnenden wird dabei nicht erfasst. Zukünftige Studien könnten dies beispielsweise mit dem *„Observational Structured Assessment of Debriefing“* kontrollieren,[64] in dem hier vorliegenden Design wurde aus organisatorischen Gründen auf eine sorgsame Randomisierung und Schulung der TutorInnen gesetzt.
Die durchgeführte Delphi-Studie weist ebenfalls einige Einschränkungen auf: So sind kognitive Bias bei dieser Art von Umfrage typisch, die Befragten sind nicht immer in der Lage zwischen rationalen Einschätzungen und subjektiven Meinungen zu unterscheiden.[65] Des Weiteren kann die teilnehmende Studienpopulation kritisiert werden – hier zeigt sich ein großer Anteil an Ärztinnen und Ärzten, die an einem Maximalversorger tätig sind, außerdem ist die Allgemeinmedizin unter- und die Anästhesie/Intensivmedizin überrepräsentiert, was die stark positive Bewertung technischer Neuerungen begründen könnte.[46] Es sollte in zukünftigen Studien auf eine repräsentativere Auswahl geachtet werden oder gezielte Kollektive angesprochen werden. Diese Kritikpunkte adressieren v.a. die konkreten Ergebnisse der Studie, die Wertigkeit der Methodik ist dadurch aber nicht beeinträchtigt.

Zusammenfassend sind die Ergebnisse dieser Dissertation in der Lage, neue Instrumente zur Antizipation der bestehenden Unsicherheit junger Ärztinnen und Ärzte in Deutschland liefern. So lässt sich der Wert von medizinischen Notfallsimulation für die Lehre auch unter dem

Gesichtspunkt einer Steigerung der Selbstsicherheit betonen. Zur Implementierung von *human factors* in diese Simulationen konnten erste wichtige Instrumente geliefert werden: So wurde eine hilfreiche Checkliste zur Feedbackunterstützung erarbeitet und ein Instrument zur Qualitätsüberprüfung der Teamarbeit übersetzt und validiert. Beide Instrumente sind für weitere Studien nützlich und sollten hier ebenfalls auf ihre Wirksamkeit untersucht werden. Zeitgleich darf die curriculare Perspektive nicht vernachlässigt werden: Der Nutzen der Delphi-Methode wurde gezeigt, die ersten Ergebnisse und die Methode als Gesamtes können nun auf konkrete Fragestellungen der zukunftsorientierten Curriculumsentwicklung angewandt werden.

7. Literaturverzeichnis

1. Morgan PJ, Cleave-Hogg D. Comparison between medical students' experience, confidence and competence. Med Educ. 2002;36:534–9.
2. Goldacre MJ, Taylor K, Lambert TW. Views of junior doctors about whether their medical school prepared them well for work: questionnaire surveys. BMC Med Educ. 2010 Jan;10(1):78.
3. Tallentire VR, Smith SE, Skinner J, Cameron HS. The preparedness of UK graduates in acute care: a systematic literature review. Postgrad Med J. 2012 Jul;88(1041):365–71.
4. Ochsmann EB, Zier U, Drexler H, Schmid K. Well prepared for work? Junior doctors' self-assessment after medical education. BMC Med Educ. 2011 Jan;11(1):99.
5. Cave J, Woolf K, Jones A, Dacre J. Easing the transition from student to doctor: how can medical schools help prepare their graduates for starting work? Med Teach. 2009 May;31(5):403–8.
6. Woods JL, Pasold TL, Boateng B a, Hensel DJ. Adolescent health care and the trainee: roles of self-efficacy, standardized patients, and an adolescent medicine rotation. Simul Healthc. 2013;8(6):359–67.
7. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol. 2001 Jan;52:1–26.
8. Conner M, Norman P. Predicting health behaviour. Conner M, Norman P, editors. Vol. 24, Safety Science. Open University Press; 2005. 408 p.
9. Tyssen R, Vaglum P. Mental health problems among young doctors: an updated review of prospective studies. Harv Rev Psychiatry. 2002;10(3):154–65.
10. de Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review. Qual Saf Health Care. 2008;17:216–23.
11. Hautz SC, Schuler L, Kämmer JE, Schauber SK, Ricklin ME, Sauter TC, Maier V, Birrenbach T, Exadaktylos A, Hautz WE. Factors predicting a change in diagnosis in patients hospitalised through the emergency room: a prospective observational study. BMJ Open. 2016;6(5):e011585.
12. Balogh EP, Miller BT, Ball JR, editors. Improving Diagnosis in Health Care. Washington, DC: The National Academies Press; 2015.
13. Risser DT, Rice MM, Salisbury ML, Simon R, Jay GD, Berns SD. The Potential for Improved Teamwork to Reduce Medical Errors in the Emergency Department. Ann Emerg Med. 1999;34(3):373–83.
14. Leape LL. Error in medicine. JAMA. 1994;272(23):1851–7.
15. Rall M, Oberfrank S. "Human factors" und "crisis resource management"–Erhöhung der

Patientensicherheit. Unfallchirurg. 2013;116:892–9.
16. Manser T. Teamwork and patient safety in dynamic domains of healthcare: a review of the literature. Acta Anaesthesiol Scand. 2009;53:143–51.
17. Brydges R, Hatala R, Zendejas B, Erwin PJ, Cook DA. Linking Simulation-Based Educational Assessments and Patient-Related Outcomes: A Systematic Review and Meta-Analysis. Acad Med. 2015;90(2):246–56.
18. Fandler M, Habersack M, Dimai HP. Have "new" methods in medical education reached German-speaking Central Europe: a survey. BMC Med Educ. 2014;14(1):172.
19. Orledge J, Phillips WJ, Murray WB, Lerant A. The use of simulation in healthcare: from systems issues, to team building, to task training, to education and high stakes examinations. Curr Opin Crit Care. 2012 Aug;18(4):326–32.
20. Morey JC, Simon R, Jay GD, Wears RL, Salisbury M, Dukes KA, Berns SD. Error reduction and performance improvement in the emergency department through formal teamwork training: Evaluation results of the medteams project. Health Serv Res. 2002;37(6):1553–81.
21. Theilen U, Fraser L, Jones P, Leonard P, Simpson D. Regular in-situ simulation training of paediatric Medical Emergency Team leads to sustained improvements in hospital response to deteriorating patients, improved outcomes in intensive care and financial savings. Resuscitation. 2017;115:61–7.
22. Cheng A, Eppich W, Grant V, Sherbino J, Zendejas B, Cook DA. Debriefing for technology-enhanced simulation: a systematic review and meta-analysis. Med Educ. 2014;48:657–66.
23. Cook DA, Hamstra SJ, Brydges R, Zendejas B, Szostek JH, Wang AT, Erwin PJ, Hatala R. Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. Med Teach. 2013;35(1):e867-98.
24. Hattie J, Timperley H. The Power of Feedback. Rev Educ Res. 2007 Mar 1;77(1):81–112.
25. Damanakis A. Ein kleiner Statusreport zu Skills Labs in der D-A-CH-Region. In: Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA). München, Germany: German Medical Science GMS Publishing House; 2011.
26. Harris P, Snell L, Talbot M, Harden RM. Competency-based medical education: implications for undergraduate programs. Med Teach. 2010;32:646–50.
27. Busing N, Slade S, Rosenfield J, Gold I, Maskill S. In the Spirit of Flexner: Working Toward a Collective Vision for the Future of Medical Education in Canada. Acad Med. 2010;85(2):340–8.
28. Frank J, Snell L, Sherbino J, editors. CanMEDS 2015 Physician Competency Framework. Ottawa: Royal College of Physicians and Surgeons of Canada; 2015.
29. 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 Z Med Ausbild. 2015;32(3):1–5.
30. Hautz SC, Hautz WE, Feufel MA, Spies CD. Comparability of outcome frameworks in medical education: Implications for framework development. Med Teach. 2015;37(11):1051–9.
31. Matheson C, Matheson D. How well prepared are medical students for their first year as doctors? The views of consultants and specialist registrars in two teaching hospitals. Postgrad Med J. 2009 Nov;85(1009):582–9.
32. Kern DE, Thomas PA, Hughes MT, editors. Curriculum development for medical education: a six- step approach. Baltimore: Johns Hopkins University Press; 1998.
33. Rowe G, Wright G. The Delphi technique as a forecasting tool: issues and analysis. Int J Forecast. 1999;15:353–75.
34. Sinha IP, Smyth RL, Williamson PR. Using the Delphi Technique to Determine Which Outcomes to Measure in Clinical Trials : Recommendations for the Future Based on a Systematic Review of Existing Studies. PLoS One. 2011;8(1):e1000393.

35. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs. 2000;32(4):1008–15.
36. Schnabel KP, Boldt PD, Breuer G, Fichtner A, Karsten G, Kujumdshiev S, Schmidts M, Stosch C. A consensus statement on practical skills in medical school - a position paper by the GMA Committee on Practical Skills. GMS Z Med Ausbild. 2011;28(4):Doc58.
37. Stroben F, Schröder T, Dannenberg KA, Thomas A, Exadaktylos A, Hautz WE. A simulated night shift in the emergency room increases students ' self-efficacy independent of role taking over during simulation. BMC Med Educ. 2016;16:177.
38. Wood L, Hassell A, Whitehouse A, Bullock A, Wall D. A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design. Med Teach. 2006;28(7):e185-91.
39. Freytag J, Stroben F, Hautz WE, Eisenmann D, Kämmer JE. Improving patient safety through better teamwork: How effective are different methods of simulation debriefing? Protocol for a pragmatic, prospective and randomized study. BMJ Open. 2017;7:e015977.
40. Feldman M, Lazzara EH, Vanderbilt AA, DiazGranados D. Rater Trainings to Support High-Stakes Simulation-Based Assessments. J Contin Educ Health Prof. 2012;33(4):279–88.
41. Sawyer T, Eppich W, Brett-Fleegler M, Grant V, Cheng A. More Than One Way to Debrief–A Critical Review of Healthcare Simulation Debriefing Methods. Simul Healthc. 2016;11(3):209–17.
42. Freytag J, Stroben F. Teaching CRM principles to undergraduates using a peer-teaching approach. In: Conferece of the Association for Medical Education in Europe (AMEE) Barcelona, Spain. 2016.
43. Park H, Kim C, Park J, Park E. Effects of team-based learning on perceived teamwork and academic performance in a health assessment subject. Collegian. 2015;22(3):299–305.
44. Parmelee D, Michaelsen LK, Cook S, Hudes PD, Parmelee D. Team-based learning: A practical guide: AMEE Guide No. 65. Med Teach. 2016;34(5):e275–87.
45. Cooper S, Cant R, Porter J, Sellick K, Somers G, Kinsman L, Nestel D. Rating medical emergency teamwork performance: Development of the Team Emergency Assessment Measure (TEAM). Resuscitation. 2010;81(4):446–52.
46. Dannenberg KA, Stroben F, Schroeder T, Thomas A, Hautz WE. The future of practical skills in undergraduate medical education – an explorative Delphi-Study. GMS J Med Educ. 2016;33(4):Doc62.
47. Mayring P. Qualitative Inhaltsanalyse: Grundlagen und Techniken. 12 Auflage. Weinheim Basel: Beltz; 2015.
48. Bundesärztekammer. Ärztestatistik 2014 [Internet]. 2014 [cited 2017 Nov 13]. Available from: http://www.bundesaerztekammer.de/ueber-uns/aerztestatistik/aerztestatistik-2014/
49. Stegmann K, Pilz F, Siebeck M, Fischer F. Vicarious learning during simulations: is it more effective than hands-on training? Med Educ. 2012 Oct;46(10):1001–8.
50. Bloch SA, Bloch AJ. Simulation training based on observation with minimal participation improves paediatric emergency medicine knowledge, skills and confidence. Emerg Med J. 2015;32:195–202.
51. Hautz WE, Kämmer JE, Schauber S, Spieß CD, Gassmeier W. Diagnostic Performance by Medical Students Working Individually or in Teams. JAMA. 2015;313(3):303–4.
52. Marshall S. The Use of Cognitive Aids During Emergencies in Anesthesia: A Review of the Literature. Anesth Analg. 2013;117(5):1162–71.
53. Haugen AS, Softeland E, Almeland SK, Sevdalis N, Vonen B, Eide GE, Nortvedt MW, Harthug S. Effect of the World Health Organization Checklist on Patient Outcomes: A Stepped Wedge Cluster Randomized Controlled Trial. Ann Surg. 2015;261(5):821–8.
54. Cheng A, Hunt EA, Donoghue A, Nelson-McMillan K, Nishisaki A, LeFlore J, Eppich W, Moyer M, Brett-Fleegler M, Kleinman M, Anderson J, Braga M, Kost S, Stryjewski

G, Min S, Podraza J, Lopreiato J, Hamilton MF, Stone K, Reid J, Hopkins J, Manos J, Duff J, Richard M, Investigators for the E. Examining Pediatric Resuscitation Education Using Simulation and Scripted Debriefing: A Multicenter Randomized Trial. JAMA Pediatr. 2013;167(6):528–36.

55. Maignan M, Koch F-X, Chaix J, Phellouzat P, Binauld G, Collomb Muret R, Cooper SJ, Labarere J, Danel V, Viglino D, Debaty G. Team Emergency Assessment Measure (TEAM) for the assessment of non-technical skills during resuscitation: Validation of the French version. Resuscitation. 2016;101:115–20.
56. Cooper S, Cant R, Connell C, Sims L, Porter JE, Symmons M, Nestel D, Ying S. Measuring teamwork performance: Validity testing of the Team Emergency Assessment Measure (TEAM) with clinical resuscitation teams. Resuscitation. 2016;101:97–101.
57. Valentine MA, Nembhard IM, Edmondson AC. Measuring Teamwork in Health Care Settings: A Review of Survey Instruments. Med Care. 2015;53(4):e16-30.
58. Lundell S, Holmner Å, Rehn B, Nyberg A, Wadell K. Telehealthcare in COPD: A systematic review and meta-analysis on physical outcomes and dyspnea. Respir Med. 2015 Jan;109(1):11–26.
59. Röhrig R, Stausberg J, Dugas M. Development of national competency-based learning objectives “Medical Informatics” for undergraduate medical education. Methods Inf Med. 2013 Jan;52(3):184–8.
60. Barnsley L, Lyon PM, Ralston SJ, Hibbert EJ, Cunningham I, Gordon FC, Field MJ. Clinical skills in junior medical officers: a comparison of self-reported confidence and observed competence. Med Educ. 2004 Apr;38(4):358–67.
61. Duns G, Weiland T, Crotty B, Jolly B, Cuddihy H, Dent A. Self-rated preparedness of Australian prevocational hospital doctors for emergencies. Emerg Med Australas. 2008 Apr;20(2):144–8.
62. Eva K, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. Acad Med. 2005;80(10 Suppl):S46–54.
63. Schubert A, Tetzlaff JE, Tan M, Ryckmann V, Mascha E. Consistency, Inter-rater Reliability, and Validity of 441 Consecutive Mock Oral Examinations in Anesthesiology. Anesthesiology. 1999 Mar;91:288–98.
64. Arora S, Ahmed M, Paige J, Nestel D, Runnacles J, Hull L, Darzi A, Sevdalis N. Objective structured assessment of debriefing: bringing science to the art of debriefing in surgery. Ann Surg. 2012;256(6):982–8.
65. Gerhold L. Standards und Gütekriterien der Zukunftsforschung: ein Handbuch für Wissenschaft und Praxis. Wiesbaden: Springer VS; 2015.

8. Anteilserklärung

Fabian Stroben hatte folgenden Anteil an den folgenden Publikationen:

Publikation 1: Stroben F*, Schroeder T, Dannenberg KA, Thomas A, Exadaktylos A, Hautz WE: „A Simulated Night Shift in the Emergency Room Increases Students' Self-efficacy Independent of Role Taking over during Simulation“. BMC Med Educ 2016, 16:177 (*Erstautorenschaft, IF = 1.312)
Beitrag im Einzelnen: Korrespondierender Autor, Konzeption und Design der Studie, Durchführung der Datensammlung und Interpretation der Daten, Anfertigen des Manuskriptes.

Publikation 2: Freytag J*, Stroben F*, Hautz WE, Eisenmann D, Kämmer JE: "Improving patient safety through better teamwork: How effective are different methods of simulation debriefing? Protocol for a pragmatic, prospective and randomized study". BMJ Open 2017, 7:e015977 [IF = 2.562 (Top Journal in "Medicine, General & Internal"); *geteilte Erstautorenschaft]
Beitrag im Einzelnen: Korrespondierender Autor, Konzeption und Design der Studie, Übersetzen des TEAM-Instruments, Durchführung der Vorstudie mit JF, Durchführung der Datenanalyse und Interpretation. Anfertigen des Manuskriptes.

Publikation 3: Dannenberg KA, Stroben F, Schroeder T, Thomas A, Hautz WE: „The future of practical skills in undergraduate medical education – an explorative Delphi-Study“. GMS J Med Educ 2016, 33(4):Doc62 (IF = nicht berechnet)
Beitrag im Einzelnen: Analyse und Interpretation der Daten, Anfertigen des Manuskriptes.

Unterschrift, Datum und Stempel der betreuenden Hochschullehrerin

____

Unterschrift des Doktoranden

____

9. Eidesstattliche Versicherung

„Ich, Fabian Stroben, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: *Sicherheit und Selbstwirksamkeit junger Ärzte – Beitrag des Medizinstudiums* selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren beruhen, sind als solche in korrekter Zitierung (siehe „Uniform Requirements for Manuscripts (URM)" des ICMJE -*www.**icmje**.org*) kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) entsprechen den URM (s.o) und werden von mir verantwortet.

Meine Anteile an den ausgewählten Publikationen entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Betreuer/in, angegeben sind. Sämtliche Publikationen, die aus dieser Dissertation hervorgegangen sind und bei denen ich Autor bin, entsprechen den URM (s.o) und werden von mir verantwortet.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§156,161 des Strafgesetzbuches) sind mir bekannt und bewusst."

Datum Unterschrift

10. Druckexemplare der ausgewählten Publikationen

Publikation 1 S. 26 – 32

„A Simulated Night Shift in the Emergency Room Increases Students' Self-efficacy Independent of Role Taking over during Simulation“

https://doi.org/10.1186/s12909-016-0699-9 (Open Access Artikel)

Publikation 2 S. 33 – 42

“Improving patient safety through better teamwork: How effective are different methods of simulation debriefing? Protocol for a pragmatic, prospective and randomized study”

https://doi.org/10.1136/bmjopen-2017-015977 (Open Access Artikel)

Publikation 3 S. 43 – 54

„The future of practical skills in undergraduate medical education – an explorative Delphi-Study“

https://doi.org/10.3205/zma001061 (Open Access Artikel)

Stroben *et al. BMC Medical Education* (2016) 16:177
DOI 10.1186/s12909-016-0699-9

BMC Medical Education

RESEARCH ARTICLE Open Access

A simulated night shift in the emergency room increases students' self-efficacy independent of role taking over during simulation

Fabian Stroben[1,2*], Therese Schröder[1,2], Katja A. Dannenberg[1,3], Anke Thomas[2], Aristomenis Exadaktylos[4] and Wolf E. Hautz[4]

Abstract

Background: Junior doctors do not feel well prepared when they start into postgraduate training. High self-efficacy however is linked to better clinical performance and may thus improve patient care. What factors affect self-efficacy is currently unknown. We conducted a simulated night shift in an emergency room (ER) with final-year medical students to identify factors contributing to their self-efficacy and thus inform simulation training in the ER.

Methods: We simulated a night in the ER using best educational practice including multi-source feedback, simulated patients and vicarious learning with 30 participants. Students underwent 7 prototypic cases in groups of 5 in different roles (leader, member and observer). Feeling of preparedness was measured at baseline and 5 days after the event. After every case students recorded their confidence dependent of their role during simulation and evaluated the case.

Results: Thirty students participated, 18 (60 %) completed all surveys. At baseline students feel unconfident (Mean −0.34). Feeling of preparedness increases significantly at follow up (Mean 0.66, $p = 0.001$, $d = 1.86$). Confidence after simulation is independent of the role during simulation ($F(2,52) = 0.123$, $p = 0.884$). Observers in a simulation can estimate leader's confidence independent of their own ($r = 0.188$, $p = 0.32$) while team members cannot ($r = 0.61$, $p < 0.001$).

Conclusions: Simulation improves self-efficacy. The improvement of self-efficacy is independent of the role taken during simulation. As a consequence, groups can include observers as participants without impairing their increase in self-efficacy, providing a convenient way for educators to increase simulation efficiency. Different roles can furthermore be included into multi-source peer-feedback.

Keywords: Medical education, Undergraduate education, Simulation-based education, Emergency medicine, High-fidelity simulation, Self-assessment, Self-efficacy

Background

Junior doctors do not feel well prepared when they start into postgraduate training [1–3] independent of their objective performance [4]. Next to the accuracy of a diagnosis, adequate confidence in this diagnosis however is a necessity for safe and effective patient care. Too little confidence in an accurate diagnosis may harm patients through the delay of necessary treatment and unnecessary and potentially harmful additional investigations.

While the relationship between confidence and tendency to act applies to all of medicine, it is especially relevant to emergency medicine, where delayed action may have severe consequences.

Situational confidence (or self-efficacy) is a key factor to determine what actions one may take [5]. As an individual's reliance on personal abilities to succeed in a given challenge, self-efficacy increases the likelihood of that individual's actions actually occurring [6, 7]. By contrast, low

* Correspondence: fabian.stroben@charite.de
[1]Lernzentrum (Skills Lab), Charité-Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany
[2]Department of Gynecology and Obstretics CCM & CVK, Charité-Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany
Full list of author information is available at the end of the article

self-efficacy and resulting distress is argued to contribute to mental health problems [8, 9].

Several factors have been identified to influence self-reported feelings of preparedness. The percentage of graduates not feeling well prepared for clinical work differs strongly between countries [3, 10, 11] implying a great impact of educational systems and practices. Factors known to contribute to higher feelings of preparedness include frequent and immediate feedback [12], theoretical education counterbalanced with practice training, good skills education and training in diagnostic decision-making [10].

Simulation is a teaching format that may (and should) contain all four of those elements [13, 14] and thus should affect individual feelings of preparedness besides the well-known effects on objective performance [15]. Another teaching format known to increase self-efficacy includes observational or vicarious learning which is as effective as hands-on training in the acquisition of practical skills [16].

The aim of our study was to develop a best practice simulation session and evaluate the effect of simulation on the development of students' feelings of preparedness. We further aimed to identify factors within the simulation that affect confidence and feelings of preparedness in order to design a well-balanced simulation, budgeting both costs and educational effectiveness. To identify such factors we focused on the role students take over during the simulation and differences between self-reported confidence and confidence judged by peers.

Methods

Study design

A six-hour simulation session took place in 2013 at Charité – Universitätsmedizin Berlin as a night shift in a simulated emergency room (ER). The ER consisted of several rooms and an ambulance vehicle. Each room hosted a different simulated case of a total of seven. We invited students in their final year of medical school to participate.

Participants were randomized into teams of five. Each team rotated through each of the scenarios, thus seeing seven different patients, each presenting a typical ER case. Each group was staffed with a peer tutor who counseled on teamwork in between scenarios, helped with logistics and ensured participants filled in evaluations. Each room was staffed with a case tutor who ran the simulation scenario. Before starting each scenario, the group decided on a team leader, team members and observers. Feedback was given after each scenario. Figure 1 illustrates the study design.

At the beginning of the night, participants completed a questionnaire on possible confounder and self-reported feeling of preparedness in different medical specialties together with an informed consent form. Directly after each scenario, all active members recorded their confidence individually before feedback was given. Furthermore team members and observers estimated the team leader's confidence. After feedback, participants evaluated the quality of the given feedback and of the simulation overall. At the end of the night, participants filled in a final evaluation focusing on overall quality of the simulations. Participants were further asked to complete a second questionnaire of self-reported feelings of preparedness five days after the event in an online survey similar to the first questionnaire.

All evaluations (forms available upon request) were conducted using Likert scales for each item ranging from "totally agree" to "totally disagree". We coded the responses on a numerical scale ranging from +3 to −3 with 0 equating "neutral". All but the last evaluation forms were filled in on paper during the simulation, the final questionnaire "feeling of preparedness 2" was conducted online using lime survey software.

Participants

Medical students who had completed their fifth year of medical school (around 600 total) were invited to take part in the night shift. In Germany, five years of medical studies are conducted at university, the sixth and final year is spent in internships at different hospitals. Participants were chosen on a first come-first served basis through an online registration. 30 participants were randomized into six teams stratified by gender using a computer-generated randomization list. Participants were greeted in a general assembly and informed about the course of events of the night. After completing the written informed consent including participant information, information about opportunities to withdraw and possible consequences of withdrawl (none), they split into teams. The study was approved by the institutional office for data protection at Charité Berlin and deemed exempt from ethical review under local legislation, because it does not involve patients.

Cases

Cases were drafted following national and international guidelines and chosen based on learning objectives from a German national consensus statement [17, 18]. Each case represented a common ER patient. There were more diagnostic investigations available than necessary per case in order to ensure an uninterrupted simulation. A checklist was developed for each scenario to guide feedback by peer observers and case tutors.

All cases started with a presentation by the case tutor who enacted ER staff reporting a patient to the on-call physicians. Simulated Patients (SP) or simulators were placed as required by the scenario. Guideline-oriented therapy including airway-management was possible in all cases. I.v. medication and/or oxygen could be administered

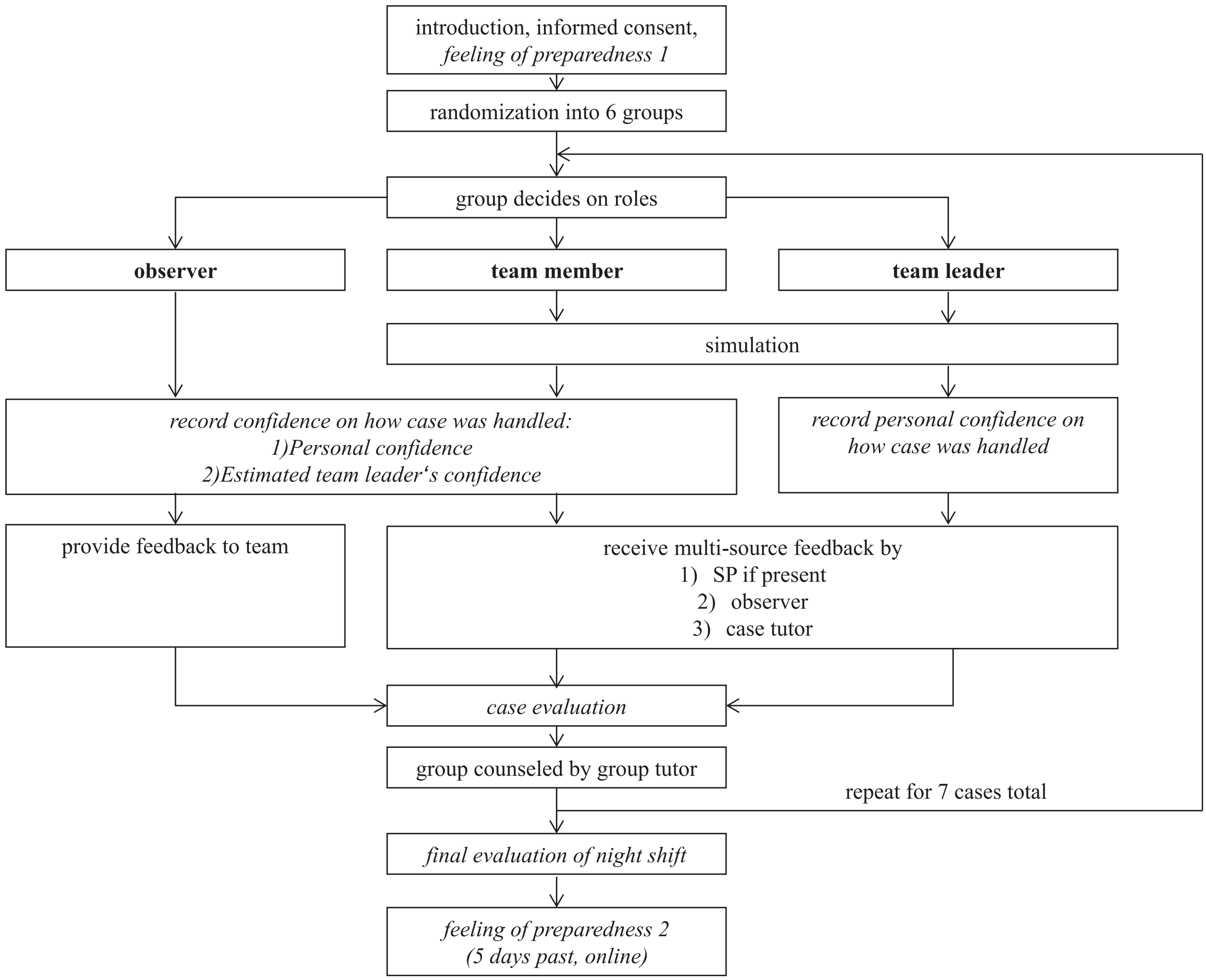

Fig. 1 Study design. Measurements in italic, roles in bold

if required. An overview of the developed cases is provided in Table 1.

Implementation

SPs were trained for five case scenarios (pulmonology, cardiology, neurology, urology and surgery 2, see Table 1). To guarantee an appropriate level of fidelity both hybrid simulations and mechanical simulators were used [14, 19]. To represent a real time course of events, all laboratory orders and radiology inquiries had to be requested by phone and/or in written form. The operational headquarters delayed their answer depending on the requested examination. Participants finished a scenario by arranging for the patient to be transferred to a ward or to be discharged. Each scenario lasted approximately 30–45 min including feedback. Additional technical details of the simulation, a detailed description of every case and used guidelines are provided as supporting information (see Additional files 1 and 2).

Roles of participants

For each scenario students took one of three roles:

- the *team leader* was responsible for the entire process – coordinating the team, choosing the right diagnosis and treating the patient accordingly.
- the *team member* was an active part of the group and supported the team leader throughout the process of finding the right diagnosis and treating the patient.
- the *team observer* observed the team using a checklist and provided feedback afterwards.

Roles within the group changed with each scenario so that at the end of the night each student had at least

Table 1 Cases and simulation settings

Discipline	Diagnosis (guidelines as sources)	Mode of simulation	Anticipated course of management
Pulmology	Exacerbated COPD	SP with examination possible	Chest X-Ray, blood-gas analysis, continous monitoring
Neurology	Ischemic media-stroke	SP with examination possible	CCT, continous monitoring
Cardiology	STEMI & non-sustained ventricular tachykardia	SP with examination possible	12-channel ECG, enzymes, continous monitoring
Anaesthesia	Ventricular fibrillation following STEMI	simulator-based approach	continous monitoring, ACLS
Surgery 1	Hemodynamic instable ruptured spleen	simulator-based approach with advanced monitoring	ATLS with FAST, continous monitoring
Urology	Urinary tract infection & pregnancy	SP with examination and sonography possible	urin test, ultrasound and gynaecological referal
Surgery 2	Head laceration	SP with examination and preparation of wound possible	Stitching of the wound

once taken on each role. Each group could freely develop their teamwork throughout the night shift. A peer tutor supervised and counseled the group.

Multi-source feedback

We used multi-sourced feedback [20] given by observers with specific assignments:

- the *SP* focused on communication using the Calgary-Cambridge Observation Guide (CCOG) to guide his or her feedback [21, 22].
- the *team observer* gave checklist-based feedback in order to provide the team with external observations but also to increase active monitoring of the simulation for his or her personal learning effect.
- the *case tutor* focused on the decision-making process with regards to medical content using case-specific checklists.
- the *peer tutor* focused on general teamwork and the development of team dynamics and gave feedback in a distinct setting to separate it from the case scenarios.

All tutors are trained in giving feedback and have extensive experience in peer teaching. Participants had experience giving and receiving feedback through curricular events. All SPs are trained regularly.

Statistical analysis

Collected data were analyzed with IBM SPSS Statistics 21.0 (SPSS Inc, Chicago, IL, USA). All data were first analyzed descriptively (mean, standard deviation). Confounder for feeling of preparedness were analyzed with Mann-Whitney-U-Tests, for differences between feeling of preparedness 1 and feeling of preparedness 2 we used a related sample Wilcoxon signed-rank test. For analysis of role and confidence we conducted a repeated measures analysis of variance (ANOVA), results of which we report as F- and p-values. Correlations between roles were analyzed with Pearson-correlations. Significance was defined as $p < 0.05$, Cohen's d was calculated as effect size. We further used G*Power, version 3.1.9.2 [23] to calculate the power achieved. We determined a gain of 0.51 on the Likert scale used from before to after the simulation as the smallest meaningful difference, because such a change would imply that participants chose one point better on the scale slightly more often than expected by chance. The primary dataset is provided as supporting information (see Additional file 3).

Results

A total of 30 students (20 female) participated in the simulation. Three participants had previous medical experience as a paramedic (2) or nurse (1). All 30 available places were booked up after 30 min in the online registration.

Feeling of preparedness

Participants feel rather ill prepared to care for patients before the simulation regardless of specialty (Mean −0.34) with no significant differences between gender ($p = 0.075$) or age ($p = 0.9$).

Right after each case students feel confident in their actions and with how they handled the case (Mean 0.95). All participants completed all surveys during the event (100 % response rate), 18 of the 30 participants (60 %) completed the online survey five days after the simulation and showed a significant increase in their general feeling of preparedness compared to before the simulation ($p = 0.001$). Participants now report to generally feel prepared (Mean 0.66); the effect is large ($d = 1.86$).

We analyzed these overall effects for every implemented discipline during simulation and found significant increases in the feeling of preparedness in anesthesiology, urology and taking history (see Table 2). The power of this study to detect a change in feeling of preparedness of 0.51 or greater was 99.79 %.

Table 2 Feeling of preparedness and change from before to five days after simulation

Discipline	Feeling of prepardness Baseline (Mean & SD)	Feeling of prepardness Follow Up (Mean & SD)	*p*-value
Overall	−0.34 (0.49)	0.66 (0.59)	0.001**
Taking History	1.27 (1.02)	1.72 (0.9)	0.035*
Anaesthesiology	0.14 (1.06)	1.17 (0.62)	<0.001***
Urology	−0.77 (1.25)	0.28 (1.53)	0.013*
Cardiology	−0.1 (1.06)	0.28 (1.13)	0.145
Pulmonology	−0.4 (0.97)	0.11 (1.13)	0.07
Surgery	0.13 (1.33)	0.83 (1.3)	0.101
Neurology	−0.47 (1.07)	0.22 (1.11)	0.1

Likert scales from +3 (totally agree) to −3 (totally disagree) we used for each item. *$p < 0{,}05$, **$p = 0{,}001$, ***$p < 0{,}001$

Role and confidence

In a repeated measures ANOVA with case as the within subject and role as the between subject factor, the self-reported confidence of participants is independent of their role during the simulation ($F(2,52) = 0.123$ $p = 0.884$). Both, team members and observers, are equally capable of judging the team leader's confidence independent of their own role ($F(2,52) = 2.055$ $p = 0.138$). How an active team member judges the team leader's confidence is in part predicted by his or her own confidence ($r = 0.61$; $p < 0.001$) while the confidence of the team leader judged by the passive observers does not correlate to the observer's personal confidence ($r = 0.188$; $p = 0.32$).

General evaluation

The simulation was evaluated very positively. Students were especially satisfied with how their peer tutors cared for them (Mean 2.93), how the SPs portrayed the patients, the difficulty of the scenarios and their opportunity to apply knowledge learned in medical school (all Mean > 2.7). The quality of the simulation was judged as very good (Mean 2.58). The ratings of each scenario right after the case correspond to the overall evaluation of the night shift.

Students reported to take most out of the feedback given by the case tutors (Mean 2.5) and slightly less out of the feedback by SPs and observing team members (both Mean 2.0).

Discussion

In line with previous findings [11, 24], especially in acute care [2], this study identifies a low feeling of preparedness among medical school graduates with results comparable to previous German [10] and British [3] studies. Our results provide evidence that even a relatively short simulation lasting just one night is effective in increasing students self-efficacy significantly as we observed an overall effect size of $d = 1.86$. Cohen himself suggested to classify effects as small when *Cohen's* $d > 0.2$, as medium when $d > 0.5$ and as large when $d > 0.8$ [25].

Intentionally including phases with observational tasks instead of active participation into the simulation may very well explain the simulations large increase simulation efficiency. Stegmann et al. previously demonstrated that hands-on-learning is as efficient as vicarious learning in the acquisition of complex manual skills [16] and Bloch and Bloch successfully used this method in ER-training sessions [26]. Active observation however is a requirement for vicarious learning [27] and giving feedback further enhances it [28]. Our results show that the effect of vicarious learning extends beyond knowledge and skills acquisition and affects situational confidence and ultimately the feeling of preparedness which we found to be unrelated to a learner's role during simulation. This provides a convenient opportunity for educators to increase group size in simulation with distributed, changing roles among participants and can influence the ratio of staff vs. participants to a more economic one. Furthermore, a recent study has demonstrated a large increase in diagnostic accuracy if patients are diagnosed by teams instead of individuals [29], further increasing the necessity to train medical staff in collaboration and to improve familiarity between ER-teammates, which was found to be surprisingly low in a recent observational study [30].

Training in the night may also be beneficial – nighttime hours are a neglected part of physicians training and may help to better prepare medical graduates for clinical settings [31] and reduce subjective stress of residents working on nighttime [24].

The observation that students significantly gained confidence in history taking may be explained by the facts that a) history taking was required in all cases presented during the night shift and students thus had ample opportunity to practice and b) history taking is directly observable to fellow students and tutors and participants may thus have received plenty of feedback regarding their interviewing skills. We can however only speculate as to why students feeling of preparedness improved for some (i.e. anesthesiology and urology) but not other (i.e. cardiology, pulmonology, surgery and neurology) disciplines and reasons might be discipline-specific. The change in urology may well be attributed to the fact that students hear little to nothing about this discipline during their course of studies [17, 18], while the increased feeling of preparedness in anesthesiology may be due to the high prevalence of algorithms in this discipline. However, the factors that determine changes in the feeling of preparedness warrant further study.

Beyond their implication for simulation practice, our results may also effect future studies of physician confidence. The observation that a team leaders self-reported confidence is not significantly different from his or her confidence judged by observers indicates an equivalence of self-reported and behavioral indicators of situational confidence. This finding further justifies the use of both measures in research on situational confidence, elsewhere termed self monitoring [32]. How the previous experimental finding, that discrepancies in confidence between team members is predictive of team failure [30], translates to real-world medical practice is currently explored in different studies [33]. Although we also did not find differences between team leader's confidence and their confidence judged by team members, team members account of the leader's confidence correlates to their own and should thus not be regarded as a valid measure.

Limitations

Because of the high personal effort and costs per participant, only a small number of students were included into the night shift simulation and this pilot study. This might be one reason for non-significant changes in feeling of preparedness in some disciplines. Achieved power however was adequate, thus implying that increasing sample size would likely only lead to the identification of irrelevant findings.

Further, one could argue that the feeling of preparedness is not necessarily linked to objective performance [34], an aspect discussed controversially [35] since self-efficacy is known to become a self-fulfilling prophecy by actually raising the chances of success on a given task [36]. In line with this model of self-efficacy, Bloch [26] and Schubert [37] both found good performance to be associated with high levels of self-reported feelings of preparedness.

Conclusion

Best-practice simulation increases the feeling of preparedness in medical students but remains expensive in the conceptual process. Assigning participants to different roles during simulation is a convenient way to increase group size. These roles have no negative influence on the increase in self-efficacy and provide an opportunity for implementing multi-source peer-feedback. The feeling of preparedness of the active team members and leader also is apparent to observers and can be used as part of a debriefing after a simulation.

Additional files

Additional file 1: Case descriptions (used Guidelines indicated as references). (PDF 87 kb)

Additional file 2: Technical details. Diagnosis and technical details for every case. (PDF 70 kb)

Additional file 3: Dataset of the study. (XLS 65 kb)

Abbreviations

ANOVA, analysis of variance; CCOG, Calgary-Cambridge observation guide; ER, emergency room; SP, simulated patient

Acknowledgments

We want to acknowledge all involved tutors of the Lernzentrum, especially Katharina Ziegeler, Leopold Vorwerk and Malte Stuhlert for participation in conception and organisation of the study. We also want to thank the Simulated Patients Program of Charité especially Bea Kampel for her support in the use of standardized patients. Furthermore we would like to acknowledge Kerry Goulston, Sydney, and Sebastian Höfer and Miriam Rüsseler, both from Frankfurt, for their constructive critique on an earlier version of the manuscript.

Funding

The study was partly funded by an unrestricted grant from Deutsche Ärztefinanz to the Skills Lab at Charite. The sponsor did not interfere with the conception and conduction of the study, analysis of data or production of the manuscript.

Availability of data and materials

The dataset supporting the conclusions of this article is included within the article (and its additional files).

Authors contributions

All authors have read and approved the manuscript. FS, KAD and WEH were responsible for the study conception and design. KAD, FS and TS performed the data collection. WEH performed the data analysis. All authors contributed to data interpretation. FS, TS and KAD were responsible for the drafting of the manuscript. All authors made critical revisions to the manuscript for important intellectual content. WEH obtained funding. WEH and AT provided administrative, technical or material support. WEH, AT and AE supervised the study.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The study was approved by the institutional office for data protection at Charité Berlin and deemed exempt from ethical review under local legislation, because it does not involve patients.

Author details

[1]Lernzentrum (Skills Lab), Charité-Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany. [2]Department of Gynecology and Obstretics CCM & CVK, Charité-Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany. [3]Department of Emergency Medicine at Campus Benjamin Franklin, Charité-Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany. [4]Universitäres Notfallzentrum, Inselspital Bern, 3010 Bern, Switzerland.

Received: 15 April 2016 Accepted: 16 June 2016
Published online: 15 July 2016

References

1. Morgan PJ, Cleave-Hogg D. Comparison between medical students' experience, confidence and competence. Med Educ. 2002;36:534–9.
2. Tallentire VR, Smith SE, Skinner J, Cameron HS. The preparedness of UK graduates in acute care: a systematic literature review. Postgrad Med J. 2012;88:365–71.

3. Goldacre MJ, Taylor K, Lambert TW. Views of junior doctors about whether their medical school prepared them well for work: questionnaire surveys. BMC Med Educ. 2010;10:78.
4. Tokuda Y, Goto E, Otaki J, Jacobs J, Omata F, Obara H, Shapiro M, Soejima K, Ishida Y, Ohde S, Takahashi O, Fukui T. Undergraduate educational environment, perceived preparedness for postgraduate clinical training, and pass rate on the National Medical Licensure Examination in Japan. BMC Med Educ. 2010;10:35.
5. Woods JL, Pasold TL, Boateng BA, Hensel DJ. Adolescent health care and the trainee: roles of self-efficacy, standardized patients, and an adolescent medicine rotation. Simul Healthc. 2013;8:359–67.
6. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol. 2001;52:1–26.
7. Conner M, Norman P: Predicting Health Behaviour. Volume 24. Berkshire, England: Open University Press; 2005.
8. Tyssen R, Vaglum P. Mental health problems among young doctors: an updated review of prospective studies. Harv Rev Psychiatry. 2002;10:154–65.
9. Hsu K, Marshall V. Prevalence of Depression and Distress in a Large Sample of Canadian Residents, Interns, and Fellow. Am J Psychiatry. 1987;144:1561–6.
10. Ochsmann EB, Zier U, Drexler H, Schmid K. Well prepared for work? Junior doctors' self-assessment after medical education. BMC Med Educ. 2011;11:99.
11. Cave J, Woolf K, Jones A, Dacre J. Easing the transition from student to doctor: how can medical schools help prepare their graduates for starting work? Med Teach. 2009;31:403–8.
12. Hattie J, Timperley H. The Power of Feedback. Rev Educ Res. 2007;77:81–112.
13. Cook DA, Hamstra SJ, Brydges R, Zendejas B, Szostek JH, Wang AT, Erwin PJ, Hatala R. Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. Med Teach. 2013;35:e867–98.
14. Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach. 2005;27:10–28.
15. Cortegiani A, Russotto V, Montalto F, Iozzo P, Palmeri C, Raineri SM, Giarratano A. Effect of High-Fidelity Simulation on Medical Students' Knowledge about Advanced Life Support: A Randomized Study. PLoS One. 2015;10:e0125685.
16. Stegmann K, Pilz F, Siebeck M, Fischer F. Vicarious learning during simulations: is it more effective than hands-on training? Med Educ. 2012;46:1001–8.
17. Schnabel KP, Boldt PD, Breuer G, Fichtner A, Karsten G, Kujumdshiev S, Schmidts M, Stosch C. A consensus statement on practical skills in medical school - a position paper by the GMA Committee on Practical Skills. GMS Z Med Ausbild. 2011;28:Doc58.
18. Blaum WE, Dannenberg KA, Friedrich T, Jarczewski A, Reinsch AK, Ahlers O. The practial use of the consensus statement on practical skills in medical school–a validation study. GMS Z Med Ausbild. 2012;29:Doc58.
19. May W, Park J, Lee J. A ten-year review of the literature on the use of standardized patients in teaching and learning: 1996–2005. Med Teach. 2009;31:487–92.
20. Wood L, Hassell A, Whitehouse A, Bullock A, Wall D. A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design. Med Teach. 2006;28:e185–91.
21. Kurtz SM, Silverman JD. The Calgary-Cambridge Referenced Observation Guides: an aid to defining the curriculum and organizing the teaching in communication training programmes. Med Educ. 1996;30:83–9.
22. Kurtz S, Silverman J, Benson J, Draper J. Marrying Content and Process in Clinical Method Teaching: Enhancing the Calgary–Cambridge Guides. Acad Med. 2003;78:802–9.
23. Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007;39:175–91.
24. Brennan N, Corrigan O, Allard J, Archer J, Barnes R, Bleakley A, Collett T, de Bere SR. The transition from medical student to junior doctor: today's experiences of Tomorrow's Doctors. Med Educ. 2010;44:449–58.
25. Cohen J. A power primer. Psychol Bull. 1992;112:155–9.
26. Bloch SA, Bloch AJ. Simulation training based on observation with minimal participation improves paediatric emergency medicine knowledge, skills and confidence. Emerg Med J. 2015;32:195–202.
27. Chi MTH, Roy M, Hausmann RGM. Observing tutorial dialogues collaboratively: insights about human tutoring effectiveness from vicarious learning. Cogn Sci. 2008;32:301–41.
28. Grierson LEM, Barry M, Kapralos B, Carnahan H, Dubrowski A. The role of collaborative interactivity in the observational practice of clinical skills. Med Educ. 2012;46:409–16.
29. Hautz WE, Kämmer JE, Schauber S, Spieß CD, Gassmeier W. Diagnostic Performance by Medical Students Working Individually or in Teams. JAMA. 2015;313:303–4.
30. Patterson PD, Pfeiffer AJ, Lave JR, Weaver MD, Abebe K, Krackhardt D, Arnold RM, Yealy DM. How familiar are clinician teammates in the emergency department? Emerg Med J. 2015;32:258–62.
31. Hanson JT, Pierce RG, Dhaliwal G. The New Education Frontier: Clinical Teaching at Night. Acad Med. 2013;89:1–4.
32. Eva K, Regehr G. "I'll never play professional football" and other fallacies of self-assessment. J Contin Educ Health Prof. 2008;28:14–9.
33. Hautz SC, Schuler L, Kämmer JE, Schauber SK, Ricklin ME, Sauter TC, Maier V, Birrenbach T Exadaktylos A, Hautz WE. Factors predicting a change in diagnosis in patients hospitalised through the emergency room: a prospective observational study. BMJ Open. 2016;6:e011585.
34. Barnsley L, Lyon PM, Ralston SJ, Hibbert EJ, Cunningham I, Gordon FC, Field MJ. Clinical skills in junior medical officers: a comparison of self-reported confidence and observed competence. Med Educ. 2004;38:358–67.
35. Duns G, Weiland T, Crotty B, Jolly B, Cuddihy H, Dent A. Self-rated preparedness of Australian prevocational hospital doctors for emergencies. Emerg Med Australas. 2008;20:144–8.
36. Eva K, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. Acad Med. 2005;80(10 Suppl):S46–54.
37. Schubert A, Tetzlaff JE, Tan M, Ryckmann V, Mascha E. Consistency, Inter-rater Reliability, and Validity of 441 Consecutive Mock Oral Examinations in Anesthesiology. Anesthesiology. 1999;91:288–98.

BMJ Open

Improving patient safety through better teamwork: how effective are different methods of simulation debriefing? Protocol for a pragmatic, prospective and randomised study

Julia Freytag,[1] Fabian Stroben,[2,3] Wolf E Hautz,[3] Dorothea Eisenmann,[2,4] Juliane E Kämmer[5,6]

To cite: Freytag J, Stroben F, Hautz WE, *et al.* Improving patient safety through better teamwork: how effective are different methods of simulation debriefing? Protocol for a pragmatic, prospective and randomised study. *BMJ Open* 2017;**7**:e015977. doi:10.1136/bmjopen-2017-015977

► Prepublication history and additional material for this paper are available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2017-015977).

JF and FS contributed equally.

Received 16 January 2017
Revised 2 May 2017
Accepted 18 May 2017

For numbered affiliations see end of article.

Correspondence to
Fabian Stroben; fabian.stroben@charite.de

ABSTRACT

Introduction Medical errors have an incidence of 9% and may lead to worse patient outcome. Teamwork training has the capacity to significantly reduce medical errors and therefore improve patient outcome. One common framework for teamwork training is crisis resource management, adapted from aviation and usually trained in simulation settings. Debriefing after simulation is thought to be crucial to learning teamwork-related concepts and behaviours but it remains unclear how best to debrief these aspects. Furthermore, teamwork-training sessions and studies examining education effects on undergraduates are rare. The study aims to evaluate the effects of two teamwork-focused debriefings on team performance after an extensive medical student teamwork training.

Methods and analyses A prospective experimental study has been designed to compare a well-established three-phase debriefing method (gather–analyse–summarise; the *GAS method*) to a newly developed and more structured debriefing approach that extends the GAS method with *TeamTAG* (teamwork techniques analysis grid). TeamTAG is a cognitive aid listing preselected teamwork principles and descriptions of behavioural anchors that serve as observable patterns of teamwork and is supposed to help structure teamwork-focused debriefing. Both debriefing methods will be tested during an emergency room teamwork-training simulation comprising six emergency medicine cases faced by 35 final-year medical students in teams of five. Teams will be randomised into the two debriefing conditions. Team performance during simulation and the number of principles discussed during debriefing will be evaluated. Learning opportunities, helpfulness and feasibility will be rated by participants and instructors. Analyses will include descriptive, inferential and explorative statistics.

Ethics and dissemination The study protocol was approved by the institutional office for data protection and the ethics committee of Charité Medical School Berlin and registered under EA2/172/16. All students will participate voluntarily and will sign an informed consent after receiving written and oral information about the study. Results will be published.

Strengths and limitations of this study

- The study design builds on established principles of teaching and assessing teamwork.
- The study will be one of the first to explore the effects of teamwork-focused debriefing on team performance with undergraduate medical students.
- The study will be embedded in a well-established simulation setting with proven efficacy.
- The study will be a pragmatic, randomised comparison of two debriefing methods.
- Only a single centre will be studied.
- Feedback quality will not be externally evaluated.

INTRODUCTION

Medical errors and adverse events occur with an incidence of about 9% and can seriously harm patients.[1 2] Error rates in emergency settings are even reported to be twice as high.[3–5] Most medical errors originate from human factors and teamwork[6] or medication errors[7] and about half of all medical errors are considered preventable.[1 7]

Empirical evidence[6 8–11] suggests that improving teamwork may be key to reducing medical error. Yet, although teamwork and patient safety are prominent objectives in many national outcome frameworks,[12–14] these topics are insufficiently represented in undergraduate education and are rarely assessed, even though validated teamwork assessment tools exist.[15 16] Consequently, about 60% of junior doctors in Germany reported feeling inadequately prepared for clinical practice[17] and almost half of the residents in a Canadian survey reported feeling overwhelmed when leading a resuscitation team.[18]

In addition, common interventions targeting the quality of teamwork and human factors, such as simulation training and crisis resource management (CRM) training, have produced a variety of effects.[19 20] In both simulation and CRM training, debriefing is considered crucial to enhancing learning[21] but little is known about how best to debrief. In fact, the widely differing effects of simulation may very well result from differences in debriefing. A feasible and beneficial debriefing method, particularly for undergraduates, could lead to more effective simulation sessions and thus ease the transition into clinical practice for junior doctors. This could ultimately lead to a reduction of medical errors and thus improved patient outcome. In this study we will compare the effects of two different debriefing methods on team performance and the acquisition of teamwork skills during teamwork simulations for medical students.

Training and debriefing

The concept of CRM was originally derived from safety training in aviation and has been adapted to the healthcare sector, another high-stakes environment.[22] The idea of CRM is to guide individuals and teams in emergency situations (crises), encouraging them to use all available resources to manage the situation effectively and prevent critical incidents from occurring in the first place. CRM training has been shown to be a potent tool to improve teamwork and—as a consequence—patient safety.[23–25] In our study, elements of CRM set the framework for teamwork training and debriefing during an emergency room simulation.

Simulation debriefing is defined as a bidirectional and interactive discussion after a simulation in which participants reflect on their actions and analyse their performance.[21] Feedback is a central process element of debriefing that is often used as a conversational technique especially in participants with little experience in debriefing.[26] Feedback is defined as the delivery of information to improve reasoning or behaviour compared with defined performance standards,[26 27] and it is critical in improving learning.[21] How best to integrate feedback into debriefing, what specific aspects to address and how to structure debriefing to foster learning are, however, still unknown.[21 28] The goal of this study is thus to evaluate the potential benefit of preselecting certain aspects to be discussed during debriefing and of structuring debriefing with the help of a cognitive aid. To this end, we will compare a well-established debriefing method to a more structured and feedback-focused method to evaluate their effects on teamwork, learning opportunities, feasibility and helpfulness for participants (and instructors). We will focus on two debriefing methods, the *gather–analyse–summarise* (GAS) method and the GAS method plus a cognitive aid:

1. The GAS method: This debriefing method consists of three parts: gathering, analysing and summarising.[29 30] The GAS method is one of many similar three-step debriefing structures[26] and has been used, for example, in simulation courses run by the American Heart Association.[30] During the first phase (gather), participants are given the opportunity to report their thoughts on the simulated situation. They are encouraged to exchange their views on what actually happened to establish a shared mental model of the situation. This model can afterwards be used to discuss the simulation in a learner-centred way (analyse). During this process, questions tailored towards specific learning objectives are used to facilitate participants' reflection on and analysis of their actions and induce learning. Finally, the debriefing is summed up and critically reviewed by the team and its instructor (summarise).[26 29] Topics discussed during the debriefing using this method are mostly self-selected by the team and instructor, which makes this method highly flexible. A possible drawback with regard to teamwork (or any other specific learning objective) is that its potential to enhance the quality of teamwork is influenced by the instructor's level of experience.[26] A typical question to start the debriefing with the gather step might be 'How do you feel now?' followed in the analysis step by 'What worked well?' or 'Do you see any opportunities for improvement?' The summarise step might be initiated by 'What we learned from this session….'
2. The GAS method plus a cognitive aid: This newly developed debriefing method uses the GAS structure detailed above and additionally provides the instructors with a cognitive aid to structure the debriefing in more detail. It further provides a selection of important aspects to address during debriefing. Cognitive aids are 'structured pieces of information designed to enhance cognition and adherence to…best practices.'[31] Cognitive aids have been shown to be beneficial in different areas of medicine.[32–34] Moreover, cognitive aids are useful for debriefing: Instructors' use of a cognitive aid may improve participants' acquisition of behavioural and cognitive outcomes after simulation—especially so with novice instructors.[35] In practice, such aids are often a pocket card, script or poster.

We will use a specific cognitive aid called 'TeamTAG' (teamwork techniques analysis grid) to foster observation and feedback relevant to teamwork. TeamTAG is a guideline for structuring the feedback process during debriefing and remembering what to address during the analysis step of the GAS method. The TeamTAG lists teamwork-relevant CRM principles together with descriptions of behavioural anchors that serve as directly observable patterns of teamwork and provides space for notes (see online supplementary information). The TeamTAG can be printed on a single sheet of paper (A4) and filled in during observation of the simulation. After the simulation, instructors have the flexibility to set priorities for debriefing based on their observations and structured notes. The debriefing itself will follow the same structure as under the GAS method. However, the TeamTAG might, for example, remind instructors that team leaders

 Freytag J, *et al*. *BMJ Open* 2017;**7**:e015977. doi:10.1136/bmjopen-2017-015977

'allocate roles & tasks' or are responsible for 'monitoring progress' (according to the CRM principle 'exercise leadership and followership'). These aspects might be specifically addressed by group instructors to improve group reflection during the analysis step.

Hypotheses

First, we assume that the GAS method plus TeamTAG will be a more effective debriefing tool than the common GAS method alone and will lead to the discussion of more teamwork-relevant principles. Debriefing using the GAS method plus TeamTAG should thus result in more learning opportunities for teams and ultimately in improved team performance. This hypothesis is based on the fact that the TeamTAG is concise and guides observation and feedback with practical examples. Using these examples during observation may help focus the observers' attention[36] and result in the team discussing more teamwork-relevant CRM principles. In undergraduate education, instructors are often novices and vary considerably regarding how experienced they are in debriefing. Because novices were shown to benefit more from structured debriefing scripts than more experienced instructors,[35] we consider our environment (see the Methods and analysis section) ideal for detecting differences between the two debriefing methods if they exist.

Hypothesis 1a: Participants who receive debriefing based on the GAS method plus TeamTAG will show a greater improvement in team performance than those who discuss the simulation according to the common GAS method alone.

Hypothesis 1b: Participants who receive debriefing based on the GAS method plus TeamTAG will report discussing a higher number of CRM principles than participants who are debriefed with the GAS method alone.

Second, we expect that teams receiving debriefing based on the GAS method plus TeamTAG will perceive teamwork skills as more important after the simulation event, which should increase their sensitivity to a culture of safety and the likelihood of changing their behaviour.[37 38] Moreover, perceiving the content of the debriefing as more important should lead to higher overall satisfaction with and perception of helpfulness of the debriefing.

Hypothesis 2a: Participants who receive debriefing based on the GAS method plus TeamTAG will report a higher level of perceived importance of teamwork principles than those who are debriefed according to the common GAS method.

Hypothesis 2b: Participants who receive debriefing based on the GAS method plus TeamTAG will report higher satisfaction with and helpfulness of the debriefing they received than those who are debriefed according to the GAS method alone.

Third, we will focus on the satisfaction of the instructors as a measure of feasibility and efficiency. We expect higher satisfaction when they use the GAS method plus TeamTAG as it might facilitate more structured feedback and it provides a better opportunity for instructors to address the learning objectives of their participants.

Hypothesis 3: Instructors who use the GAS method plus TeamTAG will report higher levels of feasibility and efficiency of their debriefing than instructors who use the GAS method alone.

METHODS AND ANALYSIS

This investigation is designed as a prospective experimental superiority study with intervention and control groups receiving debriefing during a simulation training based on either the GAS method plus TeamTAG or the GAS method alone, respectively. The study will be executed during an emergency department (ED) simulation at Charité Medical School, Berlin, Germany, on 14 January 2017. The ED simulation has been implemented at the local skills laboratory since 2013 on a peer-led basis. The main goal of this extensive, 8-hour night-shift simulation training is to give students the opportunity to experience being the person in charge of a patient's healthcare. This event takes place once a year, with about 35 students in their final year of medical studies participating voluntarily. Participants are recruited via newsletter and advertising posters. The students act in randomly assigned teams of five and self-select into different roles (team leader, team member, observer), which they switch during the night. Simulated patients and high-fidelity simulators are used to create realistic case simulations; simulated radiological and laboratory services are provided. One of the main goals of the event is to improve students' confidence in working with medical emergencies in an ED over the course of the night.[39] The simulation was awarded a project prize by the German Association for Medical Education in 2016.

Each student team has to work on six simulated cases. Each case is staffed with a case instructor who is responsible for the simulation and provides technical help. Each student team is accompanied by a group instructor who guides the participants during the night. After every case, multisource feedback is provided by simulated patients, observing participants and case instructors. As part of our study, in 2017 participants will additionally receive a teamwork-based debriefing by the group instructors after every case in one of two conditions (GAS method vs GAS method plus TeamTAG). Additionally, the quality of teamwork will be rated by trained raters throughout the night.

As group instructors we will choose experienced peer teachers who are advanced in their healthcare studies (medicine, nursing) and have completed emergency room courses/electives during their studies. Peer teachers at Charité Medical School Berlin frequently give courses in clinical skills training and simulator-based emergency medicine trainings for other medical students. All group instructors undergo extensive feedback training during their studies and are furthermore trained in working with and debriefing groups.

Development of the TeamTAG as cognitive aid

As a basis for this study, the TeamTAG guideline was developed with the goal of having a feasible and time-efficient feedback instrument that supports teaching basic teamwork skills to participants. Two investigators (JF and FS) developed the TeamTAG guidelines that present six common CRM principles,[22 40] each accompanied by the description of behavioural anchors. The six principles are (1) anticipate and plan ahead, (2) set priorities dynamically, (3) call for help early, (4) exercise leadership and followership, (5) communicate effectively and (6) re-evaluate repeatedly. The TeamTAG can be found in the online supplementary material. The CRM principles and their behavioural anchors were chosen to fit the following criteria: (A) simulation setting, (B) presumed skills of participants, (C) experience of instructors and (D) observability. The tool was reviewed and adjusted by an experienced group of anaesthesiologists, emergency medicine physicians, simulation instructors and peer tutors, all experienced in medical education and simulation-based learning. In a prestudy, feasibility for instructors was examined (see the Preliminary results section) but not compared with an approach without the TeamTAG.

Team performance measurement

To measure team performance, we will use the Team Emergency Assessment Measure (TEAM).[15] TEAM is an assessment tool that has been applied to both clinical and simulation environments.[15 16 41] It consists of 11 items belonging to the three subscales leadership, teamwork and task management. Example items are 'the team leader maintained a global perspective' and 'the team prioritized tasks', measured on a 5-point Likert scale of 0 (*never*) to 4 (*always*). Additionally, it includes an overall rating of team performance (range: 1 (*very poor performance*) to 10 (*very good performance*)).

As there was no German version of the TEAM, the English version was translated into German using elements of the TRAPD (translation, review, adjudication, pretest, documentation) methodology.[42] Two investigators (JF and FS) independently translated the TEAM into German in parallel, reviewed the results and consented to one version, which was translated back by a native English speaker. This new version was compared with the original TEAM and agreed to by both investigators and the native speaker. All steps of the translation were documented.

After the TEAM was translated, we developed a rater training. The training involves three aspects that are important in preparation for accurately assessing a certain behaviour or skill[1 43]: a *rater error training* in which information is provided on typical rating errors to raise awareness and prevent them,[2] a *performance dimension training* to teach raters about the targeted dimensions, including definitions and videotaped examples, and[3] a *frame-of-reference training*, in which videotaped examples showing teamwork of different levels of quality are assessed and discussed. All raters who will be responsible for TEAM ratings in this study (case instructors and additional raters) will receive this rater training and additional written material on teamwork and how to use the TEAM.

Group instructors debriefing training

Before data collection, all group instructors will receive a teamwork-related training and additional written material with information about how to provide feedback and conduct debriefings and about human factors in general and CRM in particular, which is intended to serve as a framework for discussing all teamwork aspects during debriefing. The training will include videos showing good and bad examples of teamwork and will be followed by discussions about opportunities for debriefing in these specific situations (adapted from *frame of reference training*[43]). After this training, which will be the same for all group instructors, the instructors will be randomly assigned, stratified by level of academic education and additional professional training (eg, nurse or paramedic), to the two conditions. The two groups will receive separate instruction from the investigators: The intervention group instructors will be told to discuss their groups' performance with the help of the TeamTAG and to focus on each CRM principle of the TeamTAG at least once during the first five cases (ie, one or two principles per case) so that by case 6 all CRM principles will have been debriefed and team performance during case 6 can be compared between conditions. Furthermore, they will be instructed to re-evaluate their previous focus of debriefing after each case if behaviour does not change sufficiently from their perspective. The order of chosen topics can be varied by the instructors and should be adjusted to observed difficulties in teamwork during the simulation. The control group instructors will be advised to give feedback regarding whatever teamwork-related aspect they deem important during the first five cases and also to re-evaluate the teamwork if needed. Instructors will stay with their groups during the whole simulation event to guarantee coordinated, consistent and longitudinal feedback.

Data collection

Upon arrival, every student participant will create an individual anonymised study code, which will be entered on every form and questionnaire and will allow us to link all measurements during the course of the night. Students will also track their role (leader, member, observer) after every case to allow subgroup analyses in relation to these roles. Figure 1 depicts the data collection procedure during the night-shift simulation.

Before starting the simulation, all 35 participants will be asked to fill in a first questionnaire that assesses possible confounders such as demographic data, professional training as a nurse or paramedic, or any training in teamwork/human factors. Next, students will be randomly assigned to seven groups via a computer-generated algorithm by the principal investigator. Four groups will serve as intervention groups and the remaining three

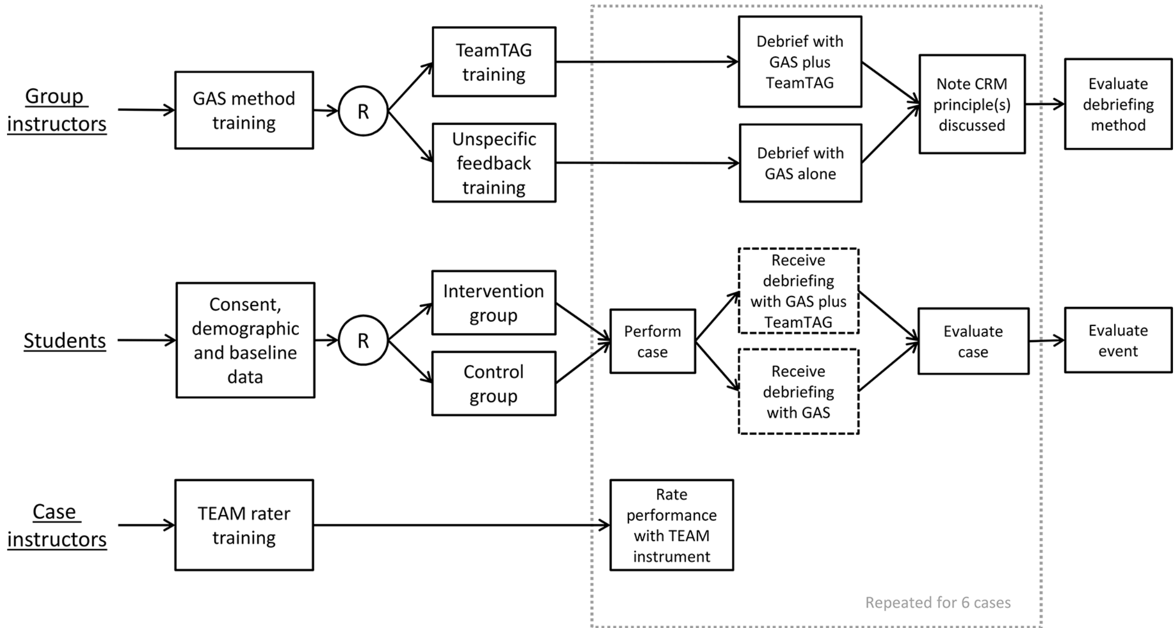

Figure 1 Study flow chart. CRM, crisis resource management; GAS, gather–analyse–summarise; R, randomisation; TEAM, Team Emergency Assessment Measure; TeamTAG, teamwork techniques analysis grid.

as controls; participants will not know to which condition they are assigned. After randomisation, all groups will gather separately and will be asked to discuss already known principles of teamwork and 15 multiple-choice questions concerning emergency medicine. A recent study showed that the results of such discussions are linked to team performance.[44]

During the simulation, all groups will face six simulations where teamwork will be measured and teamwork-related feedback provided. All cases depict common emergency situations where the participation of an emergency team in the emergency room is needed. Table 1 gives a brief overview of the diagnoses of the six cases and challenges for teamwork.

During every case, team performance will be measured using the TEAM,[16] which will be filled in by the case instructors and an additional rater. The two TEAM raters will be blind to the debriefing condition the group is assigned to.

After every case (duration about 30 min), debriefing will start (duration about 20 min) with checklist-based feedback from the simulated patients (focus: communication skills, empathy) and the case instructors and peer observers (focus: factual knowledge, diagnostic skills). As the last part of the debriefing process, the teamwork-related debriefing will be conducted by the group instructor using the GAS method with or without the support of the TeamTAG depending on the experimental condition. The strict timing, which will be centrally coordinated, will be necessary for a smooth transition of groups between cases and to ensure that the total length of the simulation does not exceed 8 hours.

After the debriefing process, all group members will be asked to evaluate the case and rate how helpful the debriefing was. Group instructors in both conditions will track the main topics of their teamwork debriefing in a debriefing protocol as free text. After the simulation, the content of these debriefing protocols will be clustered

Table 1 Teamwork-relevant cases presented in the emergency department simulation

Case	Diagnosis	Challenges for teamwork
1	Exacerbated COPD	Conflict management, control of emotions due to challenging patient
2	Ischaemic stroke of middle cerebral artery	Task management, communication with colleagues Manage aphasic patient
3	STEMI and non-sustained ventricular tachycardia	Patient deterioration (cardiac arrhythmia) during care
4	Ventricular fibrillation following STEMI	Team leadership, structured ACLS
5	Haemodynamically unstable ruptured spleen	Set priorities in evaluation and management, structured ATLS
6	Head laceration with ethanol intoxication	Manage agitated patient

ACLS, advanced cardiac life support; ATLS, advanced trauma life support; COPD, chronic obstructive pulmonary disease; STEMI, ST-elevation myocardial infarction.

independently (JF and FS) and matched with CRM principles.

Right after the last case of the night, all participants will fill in a final evaluation, which will ask them to list all the CRM principles on which they received feedback during the night. Participants will also evaluate the importance of each principle for their future work as physicians and provide a general evaluation of the night. Every group tutor will rate the feasibility, efficiency and difficulty of providing feedback.

Collected data

1. Baseline characteristics: The data collected on the first questionnaire and the results of group and teamwork discussions will be used to compare the baseline between the two conditions. Discussion results will be analysed qualitatively to identify differences in knowledge and in the personal definition of good teamwork at the beginning of the night. Furthermore, the TEAM scores during the first simulation case will serve as the baseline team performance.
2. Hypothesis 1 measurement (team performance, number of CRM principles discussed): Team performance will be evaluated using the 11 items of the translated TEAM. Similar to previous studies,[15 16 41 45 46] we will analyse ratings on the item level (range: 0–4), the sum score (range: 0–44) and the overall rating per case (range: 1–10). The number of CRM principles discussed will be derived from two sources, namely, the debriefing protocols of the group instructors and participants.
3. Hypothesis 2 measurement (importance, satisfaction, helpfulness): Estimated relevance of the CRM principles learnt and overall satisfaction with the simulation will be evaluated on 7-point Likert scales at the end of the night. Helpfulness of the debriefing from the different providers (simulated patient, peer, case tutor and group tutor) will be rated by participants after every case on a 7-point Likert scale.
4. Hypothesis 3 measurement (instructor ratings): Debriefing evaluation of the group instructors (feasibility, efficiency and difficulty of providing feedback) will be measured with 7-point Likert scales and as free-text answers at the end of the night.
5. Other measures: The general evaluation form will ask participants to rate pleasure, quality of instruction during the night, difficulty of cases and possibility of applying knowledge on 7-point Likert scales.

All 7-point Likert scales will be coded from +3 (*strongly agree*) to −3 (*strongly disagree*). All data collection forms will be available upon request.

Analyses

Data will be analysed in SPSS 24 and R using descriptive, inferential and explorative statistics. We conducted a calculation of power for our primary research question (team performance). Recent studies, reporting mainly data for well-trained and experienced teams, showed TEAM sum scores up to 40.[45 46] Only one study provided data for less experienced teams with a TEAM sum score of 21.[45] On the basis of these results and data from a prestudy (see the TeamTAG section in the Preliminary results section), we expect a TEAM sum score of about 20 for an untrained team and a score of around 40 for teams that receive a training related to teamwork skills and/or have a lot of experience in this area. These scores indicate a potential increase due to training of up to 20 points on the TEAM sum score. As a relevant training effect for a single training event such as ours, we estimate a gain in the TEAM sum score of 11 points (ie, one point per item). Using the SD from the last published study on the TEAM[46] (SD=4.4) and $\alpha<0.05$, we have determined that about six teams are needed to detect a significant difference between the conditions with a power of 80%. Missing data will be handled using pairwise deletion.

1. Baseline characteristics: Discussion results of the intervention and control groups will be compared using qualitative methods and confounder analysis (demographics, prior training) with parametric and non-parametric tests for testing equivalence. The TEAM scores (single items, sum score, overall score) from the first simulation case will be compared between conditions using multilevel analyses to take the hierarchical structure of data into account.
2. Analyses for hypothesis 1: The TEAM scores (single items, sum score, overall score) of the intervention and control groups during the sixth simulation case will be compared using multilevel analyses. The development of team performance over the six cases will be analysed using descriptive statistics and plotting ‘training curves’ for each team. The total number of CRM principles discussed in the control and intervention groups will be compared using a multilevel model.
3. Analyses for hypothesis 2: The participants’ ratings of the feedback’s helpfulness, the importance of CRM principles and satisfaction with the debriefing will be compared between the control and intervention groups using multilevel models.
4. Analyses for hypothesis 3: Group instructors’ evaluations of the instrument will be examined descriptively.
5. Other measures: The general evaluation will be examined in a descriptive way.

Methodological limitations

Group instructors will not be observed while debriefing due to our limited labour force. Therefore, we cannot be sure the quality of the debriefing will be comparable among the seven participating groups. Further studies could use debriefing assessment tools such as the Observational Structured Assessment of Debriefing tool,[47] which might help distinguish between effects of overall debriefing quality and our approach. In our study, we will try to address this limitation with extensive group instructor training to ensure an equal qualification level regarding debriefing and with a randomisation of instructors to conditions. Furthermore, participants will be asked to state the debriefing topic and to rate the quality of debriefing after every simulation case, which will be reported in later publications.

The time for debriefing after every case will be relatively short due to the design of our 8-hour simulation, where all groups will rotate through six cases to give participants a broad overview of emergency medicine and application areas of CRM. To use this limited time most productively, we have added additional specifications for debriefing (eg, focus on one or two principles per debriefing session, as described in the Methods and analysis section) because some instructors stated in a prestudy that the time allowed for debriefing was not sufficient. Future studies could investigate whether results of this study hold if all CRM principles are being discussed and thus repeated after every case/more often during the night and if time for debriefing is longer. Until now, there has been no strong evidence for the superiority of a longer debriefing.[21]

The study will focus only on short-term effects of two different debriefing approaches. Further research should investigate long-term effects on performance or changes in behaviour during clinical practice. A last limitation of this study is that it is a single-centre study and so results might be limited to local circumstances.

Data sharing statement

Data analysis will be conducted by the investigator's team (data management team). As the study is not a clinical trial, a data-monitoring team is not needed. The anonymised full data set will be published together with the journal publication or using the Dryad Data Repository (Durham, NC, USA) as required by the journal's guidelines. Data will furthermore be stored in the local data repository at Charité Medical School Berlin according to the local guidelines for good scientific practice.

PRELIMINARY RESULTS

Validation of the German TEAM

The German TEAM can be found in the online supplementary information. As a preliminary validation, inter-rater correlation was checked between three investigators (JF, FS and DE) and an external expert on two videotaped resuscitations. Both resuscitations were simulation based and had similar factual content; however, the first simulation showed good teamwork and the second intermediate teamwork performance. The videotaped simulations were used for group instructors' debriefing training and for validity testing of the German TEAM.

Intraclass correlation coefficients were .99 for the first resuscitation (mean TEAM score=42.3, *SD*=1.3) and .85 for the second (mean TEAM score=22.5, *SD*=3.1), which indicates excellent inter-rater agreement. For this reason, we consider the German TEAM a valid instrument for assessing team performance in our study.

TeamTAG

A first version of TeamTAG was used in a prestudy, conducted during the previous simulated night shift in 2016. In this prestudy, all instructors (n=7) used TeamTAG as part of their debriefing (similar to the GAS method plus TeamTAG). They were asked to rate the feasibility and helpfulness of the TeamTAG (7-point Likert scale; −3 to +3), as well as whether time for debriefing was sufficient (7-point Likert scale; −3 (*strongly insufficient*) to +3 (*strongly sufficient*)). Furthermore, they could comment on specific aspect of the guideline they liked or disliked (free-text answers). All participants were asked how useful the instructors' feedback was (7-point Likert scale; −3 to +3).

Instructors rated the guideline as a feasible tool (*M*=1.9, *SD*=0.9) and stated that it helped them in both observing and giving feedback to the participants of the simulation ($M_{observe}$=2.3, *SD*=0.8; $M_{feedback}$=2.3, *SD*=0.5). They had a heterogeneous view of the adequacy of time available for debriefing (*M*=−0.3, *SD*=1.1) The participants declared having found the feedback to be useful (*M*=1.7, *SD*=1.0).

ETHICS AND DISSEMINATION

The study protocol was designed according to the Declaration of Helsinki, the local guidelines for good scientific practice at Charité Medical School Berlin and the ICMJE (International Committee of Medical Journal Editors) recommendations. The study protocol was approved by the institutional office for data protection (AZ 737/16) and the ethics committee at Charité Medical School Berlin (EA2/172/16).

All participants and instructors will provide informed consent. Because the simulation is already a well-known event at Charité Medical School Berlin and receives official teaching funds, participants who refuse to take part in our study must have a chance to participate nevertheless. In this case, students will not provide the informed consent prior to randomisation; instead, an independent 'no-study' group will then be created, which will be identical to the control group but without any teamwork debriefing. We do not expect any harm for students who undergo the intervention.

Publication

Results of the study will be presented during national and international scientific meetings. The authors aim to publish all results in a peer-reviewed journal. Part of the protocol has been previously presented at the Research in Medical Education (RIME) conference in Duesseldorf, Germany, in March 2017 and was awarded the RIME Award: Best Research Protocol 2017.[48]

Author affiliations
[1]Simulated Patients Program, Charité Medical School Berlin, Berlin, Germany
[2]Lernzentrum (Skills Lab), Charité Medical School Berlin, Berlin, Germany
[3]Department of Emergency Medicine, Inselspital, University of Bern, Bern, Switzerland
[4]Department of Anesthesiology and Operative Intensive Care Medicine CCM & CVK, Charité Medical School Berlin, Berlin, Germany

[5]Progress Test Medizin, Charité Medical School Berlin, Berlin, Germany
[6]Max Planck Institute for Human Development, Center for Adaptive Rationality, Berlin, Germany

Acknowledgements The authors would like to acknowledge Hanno Heuzeroth and David Steinbart for support in conducting the prestudy and Simon Cooper for support in using and translating the TEAM. Furthermore, the authors thank all instructors and physicians, especially Tobias Deselaers, for organising the simulation and their willingness to participate as raters and instructors, and such. In addition, we would like to acknowledge Jane Runnacles (London), Suzanne Bentley (New York) and Christopher Timmis (Wolverhampton) for their constructive critique on an earlier version of the manuscript. The authors thank Anita Todd for editing the manuscript.

Contributors JF and FS translated the TEAM, designed the study and will be responsible for conduction. DE, WEH and JEK contributed to the study design. JEK supervised the study design and will supervise conduction. JF and FS are responsible for data analyses. JF, FS and JEK wrote the manuscript. JF and FS conducted the prestudy. DE is responsible for funding and local administration at CharitÃ© Medical School Berlin and heads the steering committee. All authors carefully read the manuscript, made critical and substantial revisions, and gave their approval for publication.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. The investigators JF and FS are funded by the German Federal Ministry of Research and Education (BMBF). The sponsor did not interfere with the conception and conduction of the study, data analysis or production of the manuscript.

Competing interests WEH received financial compensation for educational consultancy from the AO Foundation, Zurich, Switzerland. All other authors report no competing interests.

Ethics approval Ethics committee of Charité Medical School Berlin.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. de Vries EN, Ramrattan MA, Smorenburg SM, *et al*. The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care* 2008;17:216–23.
2. Kohn LT, Corrigan JM, Donaldson MS, eds. *To err is human: building a Safer Health System*. Washington: National Academies Press, 2000.
3. Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. *Am J Med* 2008;121(5 Suppl):S2–S23.
4. Schwartz A, Elstein AS. *et al*Clinical reasoning in medicine. In: Higgs J, Jones M, Loftus S, Christensen N, . eds. *Clinical reasoning in the health professions*. 3rd ed. Oxford (UK, 2008:223–34.
5. Hautz SC, Schuler L, Kämmer JE, *et al*. Factors predicting a change in diagnosis in patients hospitalised through the emergency room: a prospective observational study. *BMJ Open* 2016;6:e011585.
6. Risser DT, Rice MM, Salisbury ML, *et al*. The Potential for Improved Teamwork to Reduce Medical Errors in the Emergency Department. *Ann Emerg Med* 1999;34:373–83.
7. Leape LL. Error in medicine. *JAMA* 1994;272:1851–7.
8. Hautz WE, Kämmer JE, Exadaktylos A, *et al*. How thinking about groups is different from groupthink. *Med Educ* 2017;51:229.
9. Hautz WE, Kämmer JE, Schauber SK, *et al*. Diagnostic performance by medical students working individually or in teams. *JAMA* 2015;313:303–4.
10. Knaus WA, Draper EA, Wagner DP, *et al*. An evaluation of outcome from intensive care in major medical centers. *Ann Intern Med* 1986;104:410–8.
11. Manser T. Teamwork and patient safety in dynamic domains of healthcare: a review of the literature. *Acta Anaesthesiol Scand* 2009;53:143–51.
12. Frank J, Snell L, Sherbino J, eds. *CanMEDS 2015 Physician Competency Framework*. Ottawa: Royal College of Physicians and Surgeons of Canada, 2015.
13. Kaba A, Wishart I, Fraser K, *et al*. Are we at risk of groupthink in our approach to teamwork interventions in health care? *Med Educ* 2016;50:400–8.
14. Kiesewetter J, Gutmann J, Drossard S, *et al*. The Learning Objective Catalogue for Patient Safety in Undergraduate Medical Education–A Position Statement of the Committee for Patient Safety and Error Management of the German Association for Medical Education. *GMS J Med Educ* 2016;33:1–14.
15. Cooper S, Cant R, Connell C, *et al*. Measuring teamwork performance: Validity testing of the Team Emergency Assessment Measure (TEAM) with clinical resuscitation teams. *Resuscitation* 2016;101:97–101.
16. Cooper S, Cant R, Porter J, *et al*. Rating medical emergency teamwork performance: development of the Team Emergency Assessment Measure (TEAM). *Resuscitation* 2010;81:446–52.
17. Ochsmann EB, Zier U, Drexler H, *et al*. Well prepared for work? Junior doctors' self-assessment after medical education. *BMC Med Educ* 2011;11:99.
18. Hayes CW, Rhee A, Detsky ME, *et al*. Residents feel unprepared and unsupervised as leaders of cardiac arrest teams in teaching hospitals: a survey of internal medicine residents. *Crit Care Med* 2007;35:1668–72.
19. Brydges R, Hatala R, Zendejas B, *et al*. Linking simulation-based educational assessments and patient-related outcomes: a systematic review and meta-analysis. *Acad Med* 2015;90:246–56.
20. Morgan L, Hadi M, Pickering S, *et al*. The effect of teamwork training on team performance and clinical outcome in elective orthopaedic surgery: a controlled interrupted time series study. *BMJ Open* 2015;5:e006216.
21. Cheng A, Eppich W, Grant V, *et al*. Debriefing for technology-enhanced simulation: a systematic review and meta-analysis. *Med Educ* 2014;48:657–66.
22. Gaba D, Fish K, Howard S. *Crisis Management in Anesthesiology*. New York: Churchill Livingstone, 1994.
23. Morey JC, Simon R, Jay GD, *et al*. Error reduction and performance improvement in the emergency department through formal teamwork training: evaluation results of the MedTeams project. *Health Serv Res* 2002;37:1553–81.
24. DeVita MA, Schaefer J, Lutz J, *et al*. Improving medical emergency team (MET) performance using a novel curriculum and a computerized human patient simulator. *Qual Saf Health Care* 2005;14:326–31.
25. Theilen U, Fraser L, Jones P, *et al*. Regular in-situ simulation training of paediatric Medical Emergency Team leads to sustained improvements in hospital response to deteriorating patients, improved outcomes in intensive care and financial savings. *Resuscitation* 2017;115:61–7.
26. Sawyer T, Eppich W, Brett-Fleegler M, *et al*. More Than One Way to Debrief: A Critical Review of Healthcare Simulation Debriefing Methods. *Simul Healthc* 2016;11:209–17.
27. van de Ridder JM, Stokking KM, McGaghie WC, *et al*. What is feedback in clinical education? *Med Educ* 2008;42:189–97.
28. Cook DA, Hamstra SJ, Brydges R, *et al*. Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. *Med Teach* 2013;35:e867–e898.
29. Phrampus PE, O'Donnell JM. *et al*Debriefing Using a Structured and Supported Approach. In: Levine AL, DeMaria SJ, Schwartz AD, . eds. *The Comprehensive Textbook of Healthcare Simulation*. New York: springer, 2013:p.73–84.
30. Cheng A, Rodgers DL, van der Jagt ?lise, *et al*. Evolution of the Pediatric Advanced Life Support course. *Pediatric Critical Care Medicine* 2012;13:589–95.
31. Stanford School of Medicine. What Are Cognitive Aids? Cognitive Aids in Medicine . 2016 http://cogaids.stanford.edu/whatarecogaids.html (cited 2016 Dec 14).
32. Harrison TK, Manser T, Howard SK, *et al*. Use of cognitive aids in a simulated anesthetic crisis. *Anesth Analg* 2006;103:551–6.
33. Marshall S. The use of cognitive aids during emergencies in anesthesia: a review of the literature. *Anesth Analg* 2013;117:1162–71.
34. Haugen AS, Søfteland E, Almeland SK, *et al*. Effect of the World Health Organization checklist on patient outcomes: a stepped wedge cluster randomized controlled trial. *Ann Surg* 2015;261:821–8.
35. Cheng A, Hunt EA, Donoghue A, *et al*. Examining pediatric resuscitation education using simulation and scripted debriefing: a multicenter randomized trial. *JAMA Pediatr* 2013;167:528–36.

36. Stegmann K, Pilz F, Siebeck M, *et al*. Vicarious learning during simulations: is it more effective than hands-on training? *Med Educ* 2012;46:1001–8.
37. Weinman J. *Psychology as applied to medicine*. 2nd ed. Bristol, England: Wright, 1987.
38. Jenkins V, Fallowfield L. Can communication skills training alter physicians' beliefs and behavior in clinics? *J Clin Oncol* 2002;20:765–9.
39. Stroben F, Schröder T, Dannenberg KA, *et al*. A simulated night shift in the emergency room increases students' self-efficacy independent of role taking over during simulation. *BMC Med Educ* 2016;16:177.
40. Rall M, Oberfrank S. 'Human factors 'und 'crisis resource management'–Erhöhung der Patientensicherheit. *Unfallchirurg* 2013;116:892–9.
41. Couto TB, Kerrey BT, Taylor RG, *et al*. Teamwork Skills in Actual, In Situ, and In-Center Pediatric Emergencies. *Simulation in Healthcare* 2015;10:76–84.
42. Dorer B. ESS Round 7 translation guidelines [Internet]. *European Social Survey* 2014. http://www.europeansocialsurvey.org/docs/methodology/ESS_R7_Translation_Guidelines_FINAL.pdf (cited 2017 Jan 5).
43. Feldman M, Lazzara EH, Vanderbilt AA, *et al*. Rater training to support high-stakes simulation-based assessments. *J Contin Educ Health Prof* 2012;32:279–86.
44. Park HR, Kim CJ, Park JW, *et al*. Effects of team-based learning on perceived teamwork and academic performance in a health assessment subject. *Collegian* 2015;22:299–305.
45. Maignan M, Koch FX, Chaix J, *et al*. Team Emergency Assessment Measure (TEAM) for the assessment of non-technical skills during resuscitation: Validation of the French version. *Resuscitation* 2016;101:115–20.
46. Cant RP, Porter JE, Cooper SJ, *et al*. Improving the non-technical skills of hospital medical emergency teams: The Team Emergency Assessment Measure (TEAM™). *Emerg Med Australas* 2016;28:641–6.
47. Arora S, Ahmed M, Paige J, *et al*. Objective structured assessment of debriefing: bringing science to the art of debriefing in surgery. *Ann Surg* 2012;256:982–8.
48. Freytag J, Stroben F, Kämmer JE, *et al*. Teamwork-related debriefings in undergraduate emergency-medicine simulations : evaluation of the 'TeamTAG' using a prospective- randomized pilot-study design. *Paper Presented at 5th International Conference for Research in Medical Education (RIME 2017)*. Düsseldorf:German Medical Science GMS Publishing House, 2017

BMJ Open

Improving patient safety through better teamwork: how effective are different methods of simulation debriefing? Protocol for a pragmatic, prospective and randomised study

Julia Freytag, Fabian Stroben, Wolf E Hautz, Dorothea Eisenmann and Juliane E Kämmer

BMJ Open 2017 7:
doi: 10.1136/bmjopen-2017-015977

The future of practical skills in undergraduate medical education – an explorative Delphi-Study

Abstract

Background: 64% of young medical professionals in Germany do not feel adequately prepared for the practical requirements of the medical profession. The goal of "outcome-orientated training" is to structure medical curricula based on the skills needed when entering the workforce after completing undergraduate medical education, and thus to bridge the gap between the skills graduates have attained and those necessary for a career in the medical profession. Outcome frameworks (OFs) are used for this purpose. In preparation for developing the National Competence-Based Catalogue of Learning Objectives for Medicine (NKLM) – the German OF – the "Consensus Statement of Practical Skills in Undergraduate Medical Education" (which structures the teaching and acquisition of practical skills in Germany and which strongly influenced the "Clinical-Practical Skills" chapter of the NKLM) was published in 2011.
It is not uncommon for at least a decade to elapse between the definition and implementation of an OF and the students' graduation, which can further increase the gap between necessary and acquired skills. Thus, the purpose of this paper is to posit theses for future development in healthcare and to apply these theses to a current OF.
Methodology: Partially structured interviews with experts were used to generate theses pertaining to general, future development in healthcare. These theses were assessed by physician experts based on the likelihood of implementation by the year 2025. The 288 learning goals of the consensus statement were assessed for their relevance for medical education in the interim.
Results: 11 theses were generated for the development of medicine, and these theses were assessed and discussed by 738 experts. These theses include the increase in diseases associated with old age, the increasing significance of interprofessional cooperation, and the growing prevalence of telemedicine applications. Of the 288 learning goals of the consensus statement, 231 of the goals were assessed as relevant, and 57 were deemed irrelevant for the short-term future.
Discussion: The theses on the future of healthcare, which were generated in this study and which were validated by numerous experts, provide indications of future developments of overall requirements for medical school graduates. For example, when applied to the content of the "Clinical-Practical Skills" NKLM chapter, they largely validate the future relevance of developing practical skills while also providing indications for their further development as applied to the consensus statement.

Keywords: Skills, practical skills, clinical skills, medical training, consensus method, Delphi survey, learning goals, outcomes, competencies, NKLM

Katja Anne Dannenberg[1,2]
Fabian Stroben[1]
Therese Schröder[3]
Anke Thomas[3]
Wolf E. Hautz[4]

1 Charité – Universitätsmedizin Berlin, Lernzentrum (Skills Lab), Berlin, Germany

2 Charité – Universitätsmedizin Berlin, Department of Emergency Medicine at Campus Benjamin Franklin, Berlin, Germany

3 Charité – Universitätsmedizin Berlin, Gynecology and Obstetrics Clinic, Berlin, Germany

4 Inselspital Bern, University Emergency Center, Bern, Switzerland

Introduction

On the one hand, the significance of obtaining practical skills during undergraduate medical studies has increased significantly in recent years [1], [2]. On the other hand, 64.7% of young medical professionals in Germany state that they do not feel adequately prepared for the practical requirements of the medical profession [3], a figure which is startlingly high, even compared to international data [4], [5]. Possible causes identified by the graduate survey in Cologne (Stosch C et al., unpublished) and a national survey (partially published in [6]), were both the narrow scope of practical training and the inadequate or lacking integration of this training in curricula and examinations. In order to bridge the gap between education and training, medical curricula are increasingly oriented toward national framework curricula, known as "outcome frameworks" (OF) [7], [8], which – generally speaking – describe the skills and knowledge which should be obtained during a training period in a competence-oriented fashion. Various outcome definitions exist internationally [9], [10], [11]. The Tuning Project [12] in Europe is an attempt to synchronize the many national OFs currently in existence. The German "National Competence-Based Catalogue of Learning Objectives for Medicine" (NKLM) was developed by the medical faculty association in cooperation with the Society for Medical Education (GMA) [13] and was initially published in June of 2015 after a six-year period of development [14]. In preparation for developing the NKLM, the "Consensus Statement of Practical Skills in Undergraduate Medical Education" was developed by the committee for practical skills of the GMA in 2011 [15]. This consensus statement "can and should have a formative effect on faculties to adjust their curricula in accordance with guidelines" [15] and strongly influenced the "Clinical-Practical Skills" chapter of the NKLM. The recommendations of the consensus statement have been implemented and validated within at least one faculty department [16]. In addition, the statement serves to assist the simulator network – a merger of the DACH region Skillslabs – to structure its simulator database [17].

There are, however, notable differences in content and structure between different OFs [18], [19], which raises the question of which OF should reasonably be referenced for teaching proficiency. In addition, developing medical curricula is generally a lengthy process: the six stages of the Kern cycle as a widely taught model of curriculum development [20], for example, require a considerable period of time between the initial definition of requirements, implementation, evaluation, and adaptation. Furthermore, an average of 6.4 years [21] elapse between beginning undergraduate medical education and beginning to practice medicine [21]. This contrasts starkly with rapid developments in medicine and the use of new technologies which have become ubiquitous. Consequently, there is a risk that the contents of curricula developed based on current OFs are no longer up-to-date when the medical professionals educated accordingly enter the medical profession.

1. Object of the Study

The object of this study is to examine the "Consensus Statement of Practical Skills in Undergraduate Medical Education," and thus an important part of the NKLM, for medium-term sustainability. The results should, on the one hand, serve to provide details for the further development of the NKLM; and on the other hand, help enhance the future stability of OF and curricula by means of overarching trends in healthcare which must yet be identified. The applied explorative Delphi method, as well as its results, can also serve to further develop local and national curricula.

2. The explorative Delphi method

Originally developed in the 1950s as a technique for exploring technical developments in a military context [22], this method had been continually developed in the intervening decades [23] and is now considered an established method for analyzing uncertain developments and identifying strategic treatment options [22], [24], [25]. In principle, the Delphi method serves to collect group opinions and to focus group communications [24], as well as to qualitatively and quantitatively assess uncertain facts [24]. Although widely varied definitions of the Delphi method exist [24], certain common basic principles can be identified: anonymity of experts, multiple repetitions of the survey, statistical summary of group opinion, and controlled feedback [22]. The use of the Delphi method has been tested in various contexts, though here it is predominantly of interest to sufficiently documented applications in medical education research, such as for developing guidelines [26], [27], [28], [29].

Methods

The project was structured into preparatory and working phases. During preparations, literary research followed by partially structured stakeholder interviews was used to develop theses about developments in healthcare. These theses were then assessed by means of an expert survey. In addition, the same expert cohort assessed the 288 learning goals of the "Consensus Statement of Practical Skills in Undergraduate Medical Education" within the framework of a 2-level, explorative Delphi survey. The course of the study is depicted in Figure 1.

1. Preparatory Phase: Theses on healthcare development

Guidelines for partially structured interview with various healthcare practitioners were developed by means of selective literature research. The topics discussed in the interview included the following:

- The future development of healthcare
- Potential changes to care and to the disease spectrum
- Changes in medical technology and telemedicine
- Interdisciplinarity and cooperation with other occupational groups
- Future changes to undergraduate medical education
- Medical occupations in Germany and abroad
- Medical skills needed in the future

During the preparatory phase, 9 interviews were conducted with experts in the fields of public health, medical technology and pedagogy, clinical and outpatient, practical skills, and students of human medicine (cf. Table 1for details). Interview partners were chosen by means of a

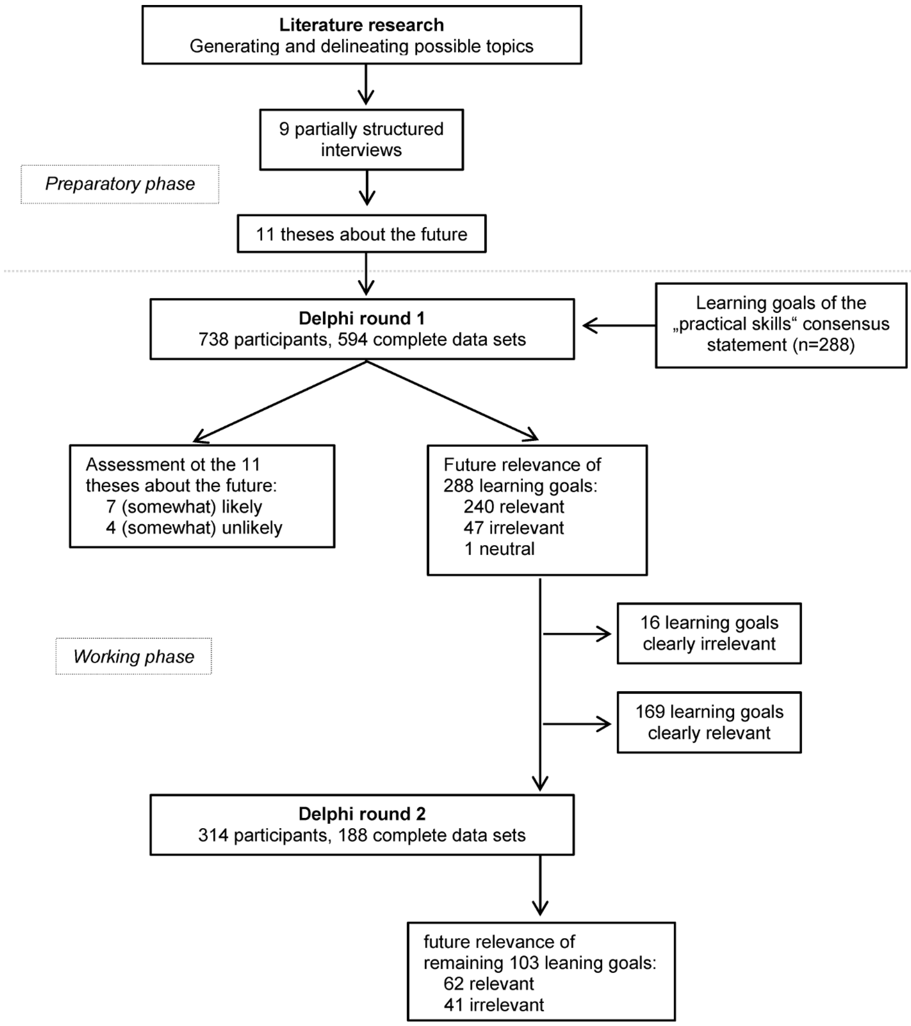

Figure 1: Study design and results overview.

"purposive sampling strategy" [30] with the goal of obtaining as broad a spectrum of perspectives on these topics as possible. All interviewed experts were prepared for the interview. Interviews lasted an average of 22 minutes.

The interviewees' answers were then grouped by topic using a qualitative text analysis and summarized as theses by means of inductive categorization per Mayring [31] by an interdisciplinary research group comprised of two students of futures studies, including one nurse and one student of human medicine with paramedic training, one practicing physician, and one computer scientist. The goal of the Mayring analysis is to systematically process the written communications at hand, and to identify similarities and differences [32]. The principles of categorization were a) category selectivity, and b) a high level of abstraction of the same.

2. Working Phase: Expert interviews on future theses generated, and on consensus statement learning goals

After generating theses, their probability of occurrence was assessed within the framework of an expert interview. The individual learning goals of the "Practical Skills in Undergraduate Medical Education" consensus statement were then assessed by the same experts within the framework of a two-level Delphi study. Physicians in all German medical university hospitals whose email addresses were available on the internet, as well as established physicians, were contacted via email to request their participation in the study. In addition to these 8,000 physicians contacted, others were approached at conferences (e.g., the Skills-Lab Symposium 2012) to request their participation in the study.

Each participant then assessed the probability of occurrence of these theses on the future of medical care (generated during the preparatory phase) using a 4-level Likert scale (1 – very likely to 4 – very unlikely). Afterward, each participant was then assigned randomly to a group of ten in order to assess the future relevance of the consensus statement learning goals. This statement defines 288 learning goals assigned to one of 16 organ systems. There is a statement for three different training stages ("clinical traineeship, practical year, advanced training") based on a three-tiered scale for each learning goal, to what extent this should be mastered ("seen demonstrated, performed under supervision, performed repeatedly"), and the survey further distinguishes between core and elective goals [15]. Each group was asked to assess a portion of the consensus statement learning goals (ca. 30/group) vis-à-vis their relevance for general medical training up to completing undergraduate medical education in the year 2025 using a 4-level Likert scale (1 – highly relevant to 4 – not at all relevant). The learning goal assessments were depicted based on the degree of mastery required by the advanced training stage, as stipulated in the consensus statement.

After round 1 assessments and individual review of the results by the research group, the round 2 learning goals to be assessed were determined by consensus. Selection criteria included a wide distribution of assessment and the estimated significance of each learning goal as assessed in round 1. The learning goals were then re-evaluated by physicians participating in round 2, who were provided with the result of round 1 evaluations. Registered participants were assigned randomly to two groups for this purpose. Each group re-evaluated circa 50 learning goals.

3. Data evaluation

The online interview was conducted using LimeSurvey (http://www.limesurvey.org). Data evaluation was conducted with Microsoft Excel (Microsoft Corporation, Redmond, WA, USA) and IBM SPSS Statistics 21.0 (SPSS Inc., Chicago, IL, USA).

Results

1. Participants

738 experts registered online for round 1 of the study, and 594 complete data sets (19.5% dropout rate) were usable. 314 experts were registered for round 2, and 188 complete data sets (40.1% dropout rate) were usable. Since the learning goals were assessed within the context of the organ system assigned to them, we only took complete data sets into consideration.

Participants were only asked to assess theses about the future during round 1. Partially-completed questionnaires were taken into consideration here, as the theses about the future can logically be interpreted on an individual basis. For this purpose, 651 expert opinions (11.8% dropout rate) were available.

Table 1: Overview of interview partners

Interview participant	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Work environment	Rural area	University	University	University	University	Professional association	University	Urban area	Rural area / Scandinavia
Profession	General medicine	Public Health/ Health Science	Physician, Master of Medical Education	Student of human medicine	Resident physician/ employed in a teaching organization	Medical technology	Nursing education	Public Health/ Health science	Emigrated specialist
Sex	M	M	M	F	F	M	M	M	F
Interview duration in min	18:44	28:11	24:34	24:14	17:20	18:04	28:31	19:59	16:45

A large portion of experts in round 1 had more than one year of work experience (96.0%), and 137 experts had been practicing for more than 15 years. The overwhelming majority of physicians worked on an inpatient basis in maximum-care hospitals (87.9%). Study participants represented a total of 26 disciplines. The number of experts in round 2 was smaller, though the characteristics of their working environments were similar. In comparison to German Medical Association (BÄK) statistics from 2014, the proportion of inpatient physicians is large (88.6% in round 1 and 79.3% in round 2, compared to 51.0% by BÄK figures), and the exact distribution of specialties also differed somewhat. Overall, 57.8% (round 1) and 53.7% (round 2) of experts work in disciplines such as surgery, internal medicine, or anesthesia, in addition to general medicine. This is consistent with BÄK data, which lists this figure at 48.8% [http://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/Statistik2014/Stat14AbbTab.pdf]. An exact analysis of the participant cohort in comparison to BÄK figures from 2014 is depicted in Table 2.

2. Theses on the future of healthcare

A total of 11 theses on the future of health care were derived from the partially structured interviews (cf. Table 3).

The theses on the future of healthcare trends generated in the preparatory phase are listed below. It was assumed that these theses were listed in accordance with expert assessment (cf. Table 3 and Figure 2):

Aspects of managing dementia are becoming significantly more important in physician communication. Increasingly balanced patient-physician relationships are emphasizing non-authoritarian forms of discussion and reasoning.

Increasing mechanization also poses a barrier to entry in the medical field: the relevance of purely manual skills is decreasing, yet IT technologies still cannot determine medical history or diagnoses, requiring the skills of doctors. Diagnostics and patient monitoring will lead to less physical contact, and patients prefer internet and smartphones for this purpose.

The physician remains the personal point of contact in established practice concepts. Duties once performed purely by physicians are, however, increasingly being delegated or substituted. Mobile treatment concepts of primary care are not catching on.

Business economic considerations are slowly moving into focus: while business economic and organizational aspects are included in training, patients' financial concerns also play a role in the type of care and treatment.

3. Assessment of learning goals

288 learning goals were assessed by experts in the 2 rounds of the study. The average of all expert assessments was used to determine whether a learning goals was deemed relevant or irrelevant. In round 1 of the Delphi study, 240 learning goals were assessed as relevant or highly relevant (average<2.5) and 57 were assessed as somewhat relevant or irrelevant (average>2.5), while 1 learning goal was assessed as neither (average=2.5).

After reviewing the results of round 1, 103 learning goals were selected for assessment in round 2 based on the distribution of round 1 results; of these, 71 were assessed as relevant, 31 were assessed as irrelevant, and 1 was assessed as neutral in round 1. In round 2, experts assessed 62 learning goals as relevant and 41 as irrelevant. In comparing the two rounds, 13 learning goals (12.6%) were rated less relevant and 4 (3.9%) were rated as more relevant in round 2. A total of 231 learning goals were considered relevant and 57 learning goals were considered irrelevant. Figure 1 depicts an overview of these results.

A more extensive review based on organ systems revealed that a large portion of learning goals for the sensory organ system (65.0%) were assessed as irrelevant for the future. Likewise, the future relevance of numerous learning goals associated with the skin, urogenital, and GI tract organs (≥30.0% each) was called into question.

31 of the 55 (56.4%) elective learning goals were assessed as irrelevant for the future – on the other hand, more than half (54.4%) of non-relevant learning goals are elective. Only 8% of the learning goals and skills which should be mastered upon completing advanced training were assessed as irrelevant for the future, yet 42% of skills were assessed at the lowest level of skill ("seen demonstrated"). The exact results are depicted in Table 4.

The online appendix of this study shows an overview of all consensus statement learning goals and their assessments in the two rounds of the Delphi study.

Discussion

This paper attempts, on the one hand, to anticipate future global requirements for medical school graduates; and on the other, to concretely examine the future relevance of practical medical skills as an example of a limited scope of competency in undergraduate medical studies. For the latter, the learning goals of the consensus statement of practical skills in undergraduate medical studies was assessed by means of an explorative Delphi study as preliminary work to the NKLM "clinical-practical skills" chapter.

The panel of experts in the Delphi study possesses many years of work experience and represents nearly all medical specialties. The majority of experts work in maximum-care hospitals, including university hospitals in which undergraduate training is primarily conducted in Germany. Though the experts' more intensive knowledge has influenced the contents and requirements of the study, outpatient physicians (who are needed primarily for contributing opinions on these theses about the future of medical care) are underrepresented. We can only speculate

Table 2: Work experience, environment, and specialties for participants in rounds 1 and 2.

	Delphi round 1		Delphi round 2		Practicing physicians in GER in 2014	
	n	%	n	%	n	%
total	594	100.0	188	100.0	365 247	100.0
Time elapsed since receiving license to practice medicine:						
< 1 year	19	3.2	4	2.1		
1–5 years	186	31.3	55	29.3		
6–10 years	147	24.7	41	21.8		
11–15 years	100	16.8	34	18.1		
> 15 years	137	23.1	46	24.5		
No response	5	0.8	8	4.3 *		
Primary working environment						
Inpatient/maximum care	522	87.9	146	77.7	186 329	51.0
Inpatient/standard and primary care	4	0.7	3	1.6 *		
Outpatient/urban	15	2.5	7	3.7	147 948	40.5 *
Outpatient/rural	10	1.7	7	3.7 *		
Medical education research	26	4.4	11	5.9	30 970	8.5
Other & no response	17	2.9	14	7.4 *		
Primary specialty					*N=255 020[a]*	
Anesthesia and intensive medicine	139	23.4	45	23.9	22 071	8.7 *
Internal medicine	118	19.9	32	17.0	60 697	23.8
Surgery	86	14.5	24	12.8	41 544	16.3
Pediatrics, gynecology, obstetrics	78	13.1	28	14.9	33 408	13.1
ENT & ophthalmology	55	9.3	10	5.3	13 239	5.2
Neurological disciplines	44	7.4	8	4.3	21 208	8.3
Radiological disciplines	24	4.0	6	3.2	10 099	4.0
General medicine	17	2.9	14	7.4 *	43 206	16.9 *
Other & no response	33	5.5	21	11.2 *	9 548	3.7

Notes: Percents based on total amount. Percents marked with * in round 2 or BÄK 2014 deviate by a minimum of 50% from those in round 1. [a] Physicians in advanced training are not listed in BÄK reviews with their desired training, and are thus not included here.
Surgery also includes urology, maxillofacial surgery; radiological disciplines also include radiology, radiation therapy, and nuclear medicine; neurological disciplines also include neurology, neurological medicine, neurosurgery, and psychiatry; pediatrics, gynecology, and obstetrics also include child and adolescent psychiatry; internal medicine also includes psychosomatic medicine, dermatology, physical medicine, and rehabilitation.

on the reasons for which so few outpatient physicians participated in the study.

At the beginning of this paper, 11 theses for relevant topics for future medical education were identified. We deliberately chose not to provide an OF as a basis for the interviews, as national OFs differ substantially in structure [18] as well as content [19].

An American group has already published a similar approach [33] which did not, however, take any further validation steps for its theses in comparison to our study. Below, we will discuss a few theses of this study in the context of their assessment by participating experts and derive implications for medical education:

The current demographic shift has caused an increase in diseases associated with advanced age, such as mild cognitive impairment and dementia [34], [35]. The learning goals catalog lists two learning goals which could be attributed to dementia illnesses. Determining the medical history of elderly patients and performing simple test procedures such as geriatric assessments or falling risk tests were both assessed as relevant for the future. However, emphasizing geriatric test procedures was listed in the catalog as an elective goal, though it should be a core learning goals according to the experts in this study.

The state of medical care, especially in rural areas with their own specific requirements [36], is in need of improvement [37]. Experience obtained in these places during voluntary training during undergraduate medical studies seemed to have a positive effect on students' learning and career choices [38] and could improve primary care. The mandatory primary care physician clinical traineeship [39] was recently introduced, and a GMA position paper emphasizes the significance of primary care during undergraduate medical studies [40].

In addition to these structural changes, working on an (interprofessional) team and using telemedicine or E-health applications will become more important in the future: delegating work to non-medical personnel increases the effectiveness of primary care [41], [42]. At the same time, learning goals which do not apply solely to physicians, such as applying plaster casts or demonstrating functional taping, have also been assessed as relevant for the future. Beginning to promote interprofessional cooperation during undergraduate medical studies, such as combined courses with trainees or students of other medical care professions, could be one method of implementing the interpersonal aspect of care more intensively during medical education.

Table 3: Expert responses (round 1) to the 11 stakeholder theses

Theses	N	Mean	SD	Very likely	Somewhat likely	Somewhat unlikely	Very unlikely	No response
1. Based on the demographic shift, special knowledge and skills in managing patients with dementia with be of great significance for all specialties in adult medicine in the year 2025.	626	1.65	0.69	278	273	46	11	18
2. Remote patient monitoring, video telephone services, and transmitting laboratory results using the internet and smartphones are accepted in the year 2025, and these services are used for the majority of patients.	626	1.85	0.83	244	232	112	21	17
3. In the year 2025, procedures performed previously only by physicians will also be performed and assessed by non-physician personnel.	651	1.85	0.91	277	232	83	48	11
4. A doctor in the year 2025 is a health manager whose training and education in basic knowledge of organization and business administration must be enhanced.	651	1.90	0.79	209	313	83	30	16
5. New diagnostic and treatment options due to innovative IT applications will lead to less physical contact between treating physicians and patients in the future.	651	2.13	0.84	146	303	137	46	19
6. The information gap between physician and patient will continue to decrease. Consequently, it is not the physician's authority that makes final decisions regarding treatment, but rather his or her ability to communicate and reason with the patient regarding diagnosis and treatment.	626	2.18	0.77	117	285	174	26	24
7. Financial concerns are the primary criteria for selecting supply and treatment options for patients in 2025.	651	2.25	0.88	142	243	199	50	17
8. The physician's most important tools in 2025 are his hands.	651	2.83	0.96	83	113	251	174	30
9. In 2025, primary medical care with primary care physicians and specialists is guaranteed overwhelmingly by means of mobile treatment plans such as daily practices, house visits, or buses, rather than in established local practices.	626	2.90	0.80	33	131	294	139	29
10. Medical history and diagnosis will be established automatically by certified IT systems. If needed, specially trained physicians will be consulted.	626	2.94	0.82	35	122	293	158	18
11. In 2025, the physician, in his person, is interchangeable, and is chosen by patients based on his function: access to treatment and diagnosis, rather than personal contact, is the deciding factor.	651	3.03	0.92	48	123	221	238	21

Notes: Assessed with 4-tier Likert scale (1 = very likely […] 4 = very unlikely)

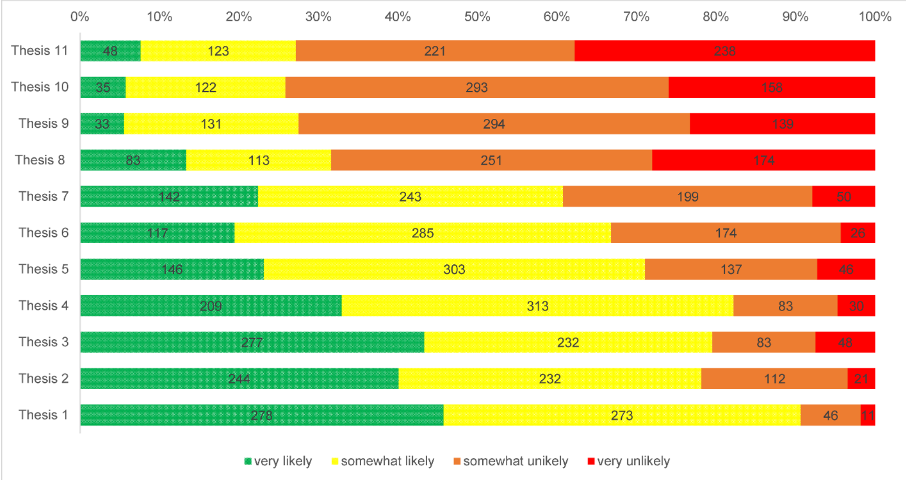

Figure 2: Graphical depiction of the experts' assessment of the 11 theses on the future of healthcare. See tab. 3 for allocation of these theses.

Table 4: Results of Delphi Rounds 1 + 2

	total	Relevant	Irrelevant	
	n	n	n	%
total	288	231	57	19.8
Organ Systems				
Sensory organs	40	14	26	65.0
Urogenital	23	15	8	34.8
GI tract	12	8	4	33.3
Skin	13	9	4	30.8
Respiratory	10	8	2	20.0
Nervous system	10	8	2	20.0
Musculoskeletal	11	9	2	18.2
Growth/aging	21	18	3	14.3
Blood/immune	10	9	1	10.0
Cardiovascular	22	20	2	9.1
Border area, emergency	36	34	2	5.6
Spanning multiple organ systems	36	35	1	2.8
Border area, soft skills	24	24	0	0.0
Border area, communication	14	14	0	0.0
Psyche	4	4	0	0.0
Endocrine system	2	2	0	0.0
Elective/core learning goals				
Elective learning goals	55	24	31	56.4
Core learning goals	232	206	26	11.2
Core/elective learning goals	1	1	0	0.0
Extent of mastery by completion of advanced training, per the consensus statement				
Seen demonstrated	62	36	26	41.9
Performed a few times	61	44	17	27.9
Performed routinely	165	151	14	8.5

Notes: Table organized by organ system, elective/core learning goals, and extent of mastery, as determined in consensus statement. Percent values calculated line by line.

In addition, electronic support system (health information technology) resources that are currently available could be used more effectively [42], [43]. Some positive effects of this technology, such as increased activity for COPD patients [44] or improved control of chronic asthma symptoms [45], have already been demonstrated. Integrating this growing field into training and education seems crucial, and could take place by means of telemedicine modules [46]. A dedicated learning goals catalog for e-health and telemedicine has already been published and can be consulted for future developments [47].

In consideration of this knowledge and the high relevance of soft skills and communication ability in this Delphi study, telephone- or internet-based physician-patient interaction could also grow more significant [48]. In order to do this field justice, training for communication skills (e.g., via telephone) should be intensified [49], as has been implemented in individual cases [50], [51]. Older patients in Germany see telemedicine methods more critically, however, and miss personal contact with and direct feedback from their physician [52].

Prioritization of core and elective learning goals in the OF original publication [15] (which were, in part, determined by means of the Delphi methodology and our results) mutually validate each other. More than 90% of the consensus statement learning goals assessed as needing to be mastered and nearly all border area learning goals were assessed by our participants as especially relevant for the future of the medical profession. On the other hand, more than 50% of skills listed in the consensus statement as elective were also assessed as less relevant by our study.

The experts assessed practical learning goals overwhelmingly as relevant for the future, primarily in the large categories of communication skills, soft skills, interinstitutional skills, cardiovascular, and emergency, but also in narrower disciplines such as mental health and the endocrine system. The portion of learning goals not relevant for the future is largest for the sensory organs and for the urogenital and GI tract systems. This could be related to the choice of experts, but could also be due to the fact that the learning goals were phrased in a very specific manner, and thus there are a great many of them. In other catalog categories, more learning goals tended to be summarized as one, which made it difficult for experts to provide a differentiated assessment. In addition, it cannot be determined whether rejection of a learning goal was due to a general lack of future relevance, or because experts considered the learning goal relevant for the future, but believed that the learning goal should be a part of specialty training rather than general medical training.

In a further step, the detailed results of this study (cf. online appendix) could be used, just like other validation studies [16], to re-assess the individual learning goals of the consensus statement and the NKLM, which would contribute to a review of the consensus statement and the NKLM.

Preparing future physicians to practice medicine should be done, in parallel, on as many levels as possible. Numerous OFs, unlike this study, currently name the "self-directed learning" method resulting from "self-assessment" [53] as a significant method of improving the results of undergraduate medical studies [54]. At the same time, there are significant doubts as to the accuracy of self-assessments [53], [55], [56]. Thus, "lifelong learning" based on self-assessment cannot be the only methods of anticipating and addressing future developments.

This should be done at the level of the OF. In addition to instructions on effective, self-directed, and lifelong learning, optimizing current OFs could make significant contributions. Anticipating future developments in conjunction with current research results could, on the one hand, provide important incentives for new content, and on the other hand, explorative Delphi studies could examine current learning goals and OFs with respect to their sustainability and possibly identify deficiencies. In concrete terms, the results of the Delphi study could serve to justify specific revisions to and implementation of the NKLM in different departments. Broad trend-setting decisions on possible future trends in undergraduate medical education can be derived from the assessment of these theses on the future.

1. Limitations

Cognitive bias must be accepted as a significant limitation of any expert survey. This is of particular importance for the application of the explorative Delphi method for assessing issues that are uncertain, per se, as they take place in the future, since the line between rational assessment and the experts' personal desires or fears could be blurred [57]. The structure of the survey could also have influenced expert opinions. After first assessing the likelihood of implementing theses in the future, experts were then asked to assess the future relevance of learning goals. This could have led to a bias. Though experts were asked to base their assessments on general education leading up to the medical exam, it cannot be determined whether this was actually done, and to what extent experts assessed general education as opposed to specialty training.

The study population of this paper is comprised primarily of inpatient physicians at maximum-care facilities. Consequently, a bias against outpatient treatment methods cannot be ruled out. In addition, colleagues practicing general medicine were underrepresented (just 2.9% of experts), while anesthesiology and intensive care – two highly specialized subjects – were overrepresented, which could explain the strong emphasis on technological trends in the theses generated. One possible explanation for the large number of anesthesiology practitioners represented in the study could be their disproportionate involvement in conveying practical skills. A follow-up survey with primarily outpatient physician seems reasonable.

Conclusions

The explorative Delphi method provides an adequate opportunity to allow a current outcome framework to be assessed by experts on the basis of its future relevance. In addition, future trends can be anticipated by means of generating and assessing theses. It is important to continually review and adapt current OF and curricula to future developments in order to provide optimal preparation for medical studies graduates for their future daily professional lives.

Acknowledgements

We would like to thank Sascha Dannenberg for his help and support with technical questions about futures studies, and for his assistance in carrying out the study. In addition, we would like to thank Nicole Ambacher and Daniel Knapp for their support in data collection. Finally, we would like to sincerely thank all of the experts who took part in this study.

Data

Data for this article are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.q4sc8 [58]

Competing interests

The authors declare that they have no competing interests.

Attachments

Available from
http://www.egms.de/en/journals/zma/2016-33/zma001061.shtml
1. Anhang.pdf (183 KB)
Online attachment – only in german

References

1. Nicky Hudson J, Tonkin AL. Clinical skills education: outcomes of relationships between junior medical students, senior peers and simulated patients. Med Educ. 2008;42(9):901-908. DOI: 10.1111/j.1365-2923.2008.03107.x

2. Yudkowsky R, Otaki J, Lowenstein T, Riddle J, Nishigori H, Bordage G. A hypothesis-driven physical examination learning and assessment procedure for medical students: initial validity evidence. Med Educ. 2009;43(8):729-740. DOI: 10.1111/j.1365-2923.2009.03379.x

3. Ochsmann EB, Zier U, Drexler H, Schmid K. Well prepared for work? Junior doctors' self-assessment after medical education. BMC Med Educ. 2011;11:99. DOI: 10.1186/1472-6920-11-99

4. Cave J, Woolf K, Jones A, Dacre J. Easing the transition from student to doctor: how can medical schools help prepare their graduates for starting work? Med Teach. 2009;31(5):403-408. DOI: 10.1080/01421590802348127

5. Goldacre MJ, Taylor K, Lambert TW. Views of junior doctors about whether their medical school prepared them well for work: questionnaire surveys. BMC Med Educ. 2010;10:78. DOI: 10.1186/1472-6920-10-78

6. Damanakis A. Ein kleiner Statusreport zu Skills Labs in der D-A-CH-Region. Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA). München, 05.-08.10.2011. Düsseldorf: German Medical Science GMS Publishing House; 2011. Doc11gma133. DOI: 10.3205/11gma133

7. Harris P, Snell L, Talbot M, Harden RM. Competency-based medical education: implications for undergraduate programs. Med Teach. 2010;32(8):646-650. DOI: 10.3109/0142159X.2010.500703

8. Busing N, Slade S, Rosenfield J, Gold I, Maskill S. In the spirit of Flexner: working toward a collective vision for the future of medical education in Canada. Acad Med. 2010 Feb;85(2):340-348. DOI: 10.1097/ACM.0b013e3181c8880d

9. Bloch R, Burgi H. The Swiss catalogue of learning objectives. Med Teach. 2002 Mar;24(2):144-150. DOI: 10.1080/01421590220120759

10. Frank JR, Danoff D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. Med Teach. 2007;29(7):642-647. DOI: 10.1080/01421590701746983

11. Metz JC. 'Blueprint 1994': common objectives of medical education in The Netherlands. Neth J Med. 1999;55(4):165-167. DOI: 10.1016/S0300-2977(99)00067-4

12. Cumming A, Ross M. The Tuning Project for Medicine-learning outcomes for undergraduate medical education in Europe. Med Teach. 2007;29(7):636-641. DOI: 10.1080/01421590701721721

13. Hahn EG, Fischer MR. Nationaler Kompetenzbasierter Lernzielkatalog Medizin (NKLM) für Deutschland: Zusammenarbeit der Gesellschaft für Medizinische Ausbildung (GMA) und des Medizinischen Fakultätentages (MFT). GMS Z Med Ausbild. 2009;26(3):Doc35. DOI: 10.3205/zma000627

14. Fischer MR, Bauer D, Mohn K, NKLM-Projektgruppe. Finally finished! National Competence Based Catalogues of Learning Objectives for Undergraduate Medical Education (NKLM) and Dental Education (NKLZ) ready for trial. GMS Z Med Ausbild. 2015;32(3):Doc35. DOI: 10.3205/zma000977

15. Schnabel KP, Boldt PD, Breuer G, Fichtner A, Karsten G, Kujumdshiev S, Schmidts M, Stosch C. A consensus statement on practical skills in medical school - a position paper by the GMA Committee on Practical Skills. GMS Z Med Ausbild. 2011;28(4):Doc58. DOI: 10.3205/zma000770

16. Blaum WE, Dannenberg KA, Friedrich T, Jarczewski A, Reinsch AK, Ahlers O. The practial use of the consensus statement on practical skills in medical school-a validation study. GMS Z Med Ausbild. 2012;29(4):Doc58. DOI: 10.3205/zma000828

17. Damanakis A, Blaum WE, Stosch C, Lauener H, Richter S, Schnabel KP. Simulator Network project report: a tool for improvement of teaching materials and targeted resource usage in Skills Labs. GMS Z Med Ausbild. 2013;30(1):Doc4. DOI: 10.3205/zma000847

18. Hautz SC, Hautz WE, Feufel MA, Spies CD. Comparability of outcome frameworks in medical education: Implications for framework development. Med Teach. 2015;20:1-9. DOI: 10.3109/0142159x.2015.1012490

19. Hautz SC, Hautz WE, Keller N, Feufel MA, Spies C. The scholar role in the National Competence Based Catalogues of Learning Objectives for Undergraduate Medical Education (NKLM) compared to other international frameworks. GMS Ger Med Sci. 2015;13:Doc20. DOI: 10.3205/000224

20. Kern DE. Curriculum development for medical education: a six-step approach. Baltimore u.a.: Johns Hopkins Univ. Press; 1998.

21. Wijnen-Meijer M, Burdick W, Alofs L, Burgers C, ten Cate O. Stages and transitions in medical education around the world: clarifying structures and terminology. Med Teach. 2013;35(4):301-307. DOI: 10.3109/0142159X.2012.746449

22. Rowe G, Wright G. The Delphi technique as a forecasting tool: issues and analysis. I J Forecast. 1999;15(4):353-375. DOI: 10.1016/S0169-2070(99)00018-7

23. Grunwald A. Modes of orientation provided by futures studies: making sense of diversity and divergence. Eur J Futures Res. 2013;2(1).

24. Häder M. Delphi-Befragungen: ein Arbeitsbuch. 2. Aufl. ed. Wiesbaden: VS; 2009. DOI: 10.1007/978-3-531-91926-3

25. Gönder D. KiTA 2030: Eine Delphi-Befragung zur Zukunft von Kindertageseinrichtungen in Deutschland. iF-Schriftenreihe Sozialwissenschaftliche Zukunftsforschung 04/2014. Berlin: Freie Universität Berlin, Institut Futur; 2014.

26. Williams PL, Webb C. The Delphi technique: a methodological discussion. J Adv Nurs. 1994;19(1):180-186. DOI: 10.1111/j.1365-2648.1994.tb01066.x

27. Houwink EJ, Henneman L, Westerneng M, van Luijk SJ, Cornel MC, Dinant JG, Vleuten Cv. Prioritization of future genetics education for general practitioners: a Delphi study. Genet Med. 2012;14(3):323-329. DOI: 10.1038/gim.2011.15

28. De Villiers MR, de Villiers PJ, Kent AP. The Delphi technique in health sciences education research. Med Teach. 2005;27(7):639-643. DOI: 10.1080/13611260500069947

29. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs. 2000;32(4):1008-1015.

30. Tavakol M, Sandars J. Quantitative and qualitative methods in medical education research: AMEE Guide No 90: Part II. Med Teach. 2014;36(10):838-848. DOI: 10.3109/0142159X.2014.915297

31. Mayring, P. Qualitative Inhaltsanalyse. Grundlagen und Techniken. 11., aktualisierte und überarbeitete Auflage, Weinheim, Basel: Beltz Verlag; 2010.

32. Flick U, von Kardorff E, Steinke I. Qualitative Forschung. Ein Handbuch. Orig.-Ausg., 8. Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verl; 2010.

33. Pershing S, Fuchs VR. Restructuring Medical Education to Meet Current and Future Health Care Needs. Acad Med. 2013;88(12):1798-1801. DOI: 10.1097/ACM.0000000000000020

34. Pater C. Mild cognitive impairment (MCI) - the novel trend of targeting Alzheimer's disease in its early stages - methodological considerations. Curr Alzheimer Res. 2011;8(7):798-807. DOI: 10.2174/156720511797633250

35. Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. JAMA. 201417;312(23):2551-2561.

36. Weeks WB, Wallace AE. Rural-urban differences in primary care physicians' practice patterns, characteristics, and incomes. J Rural Health. 2008;24(2):161-170. DOI: 10.1111/j.1748-0361.2008.00153.x

37. Rosenblatt RA, Andrilla CH, Curtin T, Hart LG. Shortages of medical personnel at community health centers: implications for planned expansion. JAMA. 2006;295(9):1042-1049. DOI: 10.1001/jama.295.9.1042

38. Barrett FA, Lipsky MS, Lutfiyya MN. The impact of rural training experiences on medical students: a critical review. Acad Med. 2011;86(2):259-263. DOI: 10.1097/ACM.0b013e3182046387

39. Bundesministerium für Gesundheit. Erste Verordnung zur Änderung der Approbationsordnung für Ärzte vom 17. Juli 2012. Bundesgesetzbl. 2012;I(34);1539. Zugänglich unter/available from: http://www.bgbl.de/xaver/bgbl/start.xav?start=%2F%2F*[%40attr_id%3D%27bgbl112s1539.pdf%27]#__bgbl__%2F%2F*[%40attr_id%3D%27bgbl112s1539.pdf%27]__1436264296102

40. Huenges B, Gulich M, Böhme K, Fehr F, Streitlein-Böhme I, Rüttermann V, Baum E, Niebling WB, Rusche H. Recommendations for undergraduate training in the primary care sector-position paper of the GMA-Primary Care Committee. GMS Z Med Ausbild. 2014;31(3):Doc35. DOI: 10.3205/zma000927

41. Altschuler J, Margolius D, Bodenheimer T, Grumbach K. Estimating a reasonable patient panel size for primary care physicians with team-based task delegation. Ann Fam Med. 2012;10(5):396-400. DOI: 10.1370/afm.1400

42. Chen PG, Mehrotra A, Auerbach DI. Do we really need more physicians? Responses to predicted primary care physician shortages. Med Care. 2014;52(2):95-96. DOI: 10.1097/MLR.0000000000000046

43. Urban E, Ose D, Joos S, Szecsenyi J, Miksch A. Technical support and delegation to practice staff - status quo and (possible) future perspectives for primary health care in Germany. BMC Med Inform Decis Mak. 2012;12:81. DOI: 10.1186/1472-6947-12-81

44. Lundell S, Holmner A, Rehn B, Nyberg A, Wadell K. Telehealthcare in COPD: a systematic review and meta-analysis on physical outcomes and dyspnea. Respir Med. 2015;109(1):11-26. DOI: 10.1016/j.rmed.2014.10.008

45. Fathima M, Peiris D, Naik-Panvelkar P, Saini B, Armour CL. Effectiveness of computerized clinical decision support systems for asthma and chronic obstructive pulmonary disease in primary care: a systematic review. BMC Pulm Med. 2014;14:189. DOI: 10.1186/1471-2466-14-189

46. Brockes MC, Wirth F, Schmidt-Weitmann S, Battegay E, Gerke W. Implementierung des Moduls Klinische Telemedizin/E-Health" in das Medizinstudium der Universität Zürich. GMS Z Med Ausbild. 2010;27(1):Doc14. DOI: 10.3205/zma000651

47. Rohrig R, Stausberg J, Dugas M, GMDS project group Medical Informatics Education in Medicine". Development of national competency-based learning objectives "Medical Informatics" for undergraduate medical education. Methods Inf Med. 2013;52(3):184-188. DOI: 10.3414/ME13-04-0001

48. Car J, Sheikh A. Telephone consultations. BMJ. 2003;326(7396):966-969. DOI: 10.1136/bmj.326.7396.966

49. Car J, Freeman GK, Partridge MR, Sheikh A. Improving quality and safety of telephone based delivery of care: teaching telephone consultation skills. Qual Saf Health Care. 2004;13(1):2-3. DOI: 10.1136/qshc.2003.009241

50. Fisher J, Martin R, Tate D. Hands on + hands free: simulated on-call interaction. Clin Teach. 2014;11(6):425-428. DOI: 10.1111/tct.12180

51. Dickinson M, Pimblett M, Hanson J, Davis M. Reflecting reality: pager simulations in undergraduate education. Clin Teach. 2014;11(6):421-424. DOI: 10.1111/tct.12185

52. Terschuren C, Mensing M, Mekel OC. Is telemonitoring an option against shortage of physicians in rural regions? Attitude towards telemedical devices in the North Rhine-Westphalian health survey, Germany. BMC Health Serv Res. 2012;12:95. DOI: 10.1186/1472-6963-12-95

53. Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. Acad Med. 2005;80(10 Suppl):S46-54. DOI: 10.1097/00001888-200510001-00015

54. Hautz SC, Hautz WE, Feufel MA, Spies CD. What makes a doctor a scholar: a systematic review and content analysis of outcome frameworks. BMC Med Educ. 2015. Under Review.

55. Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. JAMA. 2006;296(9):1094-1102. DOI: 10.1001/jama.296.9.1094

56. Li ST, Paterniti DA, Co JP, West DC. Successful self-directed lifelong learning in medicine: a conceptual model derived from qualitative analysis of a national survey of pediatric residents. Acad Med. 2010;85(7):1229-1236. DOI: 10.1097/ACM.0b013e3181e1931c

57. Gerhold L. Standards und Gütekriterien der Zukunftsforschung: ein Handbuch für Wissenschaft und Praxis. Wiesbaden: Springer VS; 2015. DOI: 10.1007/978-3-658-07363-3

58. Dannenberg KA, Stroben F, Schröder T, Thomas A, Hautz WE. Data from: The future of practical skills in undergraduate medical education - an explorative Delphi-Study. Dryad Digital Repository. 2016. DOI: 10.5061/dryad.q4sc8

Corresponding author:
Katja Anne Dannenberg
Charité – Universitätsmedizin Berlin, Lernzentrum (Skills Lab), Charitéplatz 1, D-10117 Berlin, Germany, Phone: +49 (0)30/450-576403, Fax: +49 (0)30/450-576922
katja-anne.dannenberg@charite.de

Please cite as
Dannenberg KA, Stroben F, Schröder T, Thomas A, Hautz WE. The future of practical skills in undergraduate medical education – an explorative Delphi-Study. GMS J Med Educ. 2016;33(4):Doc62.
DOI: 10.3205/zma001061, URN: urn:nbn:de:0183-zma0010614

This article is freely available from
http://www.egms.de/en/journals/zma/2016-33/zma001061.shtml

Received: *2015-06-25*
Revised: *2016-02-29*
Accepted: *2016-03-16*
Published: *2016-08-15*

11. Lebenslauf

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

12. Publikationsliste

Originalbeiträge:

Steinbart D, Dannenberg KA, Gerken JD, Eisenmann D, Stroben F: The "Lernzentrum" in the Course of Curricular Changes: How Perceive Students a Partly Curricular Implementantation of a Former Voluntary Peer Assisted Learning Program? GMS J Med Educ 2018. *Under Review.*

Schauber SK, Hautz SC, Kämmer JE, Stroben F, Hautz WE: "Do different response formats affect how test takers approach a clinical reasoning task? An experimental study on antecedents of diagnostic accuracy using a constructed response and a selected response format." Adv Health Sci Educ 2018. *Submitted for Review.*

Freytag J, Stroben F, Hautz WE, Schauber SK, Kämmer JE: "Rating the quality of teamwork—a comparison of novice and expert ratings using the Team Emergency Assessment Measure (TEAM) in simulated emergencies" Scand J Trauma Resusc Emerg Med 2018. *Accepted/ In press.*

Eisenmann D, Stroben F, Gerken JD, Machner M, Hautz WE: "Interprofessional emergency trainings lead to changes in the workplace" West J Emerg Med 2018, Jan;19(1):185-192.

Freytag J*, Stroben F*, Hautz WE, Eisenmann D, Kämmer JE: "Improving patient safety through better teamwork: How effective are different methods of simulation debriefing? A pragmatic, prospective and randomized study". BMJ Open 2017, 7:e015977

Stroben F*, Schroeder T, Dannenberg KA, Thomas A, Exadaktylos A, Hautz WE: „A Simulated Night Shift in the Emergency Room Increases Students' Self-efficacy Independent of Role Taking over during Simulation". BMC Med Educ 2016, 16:177

Dannenberg KA, Stroben F, Schroeder T, Thomas A, Hautz WE: „The future of practical skills in undergraduate medical education – an explorative Delphi-Study". GMS J Med Educ 2016, 33(4):Doc62

*(geteilte) Erstautorenschaft

Kongressbeiträge - Vorträge

Kämmer JE, Schauber SK, Hautz SC, Stroben F, Hautz WE: "Simple interventions can increase the diagnostic performance of young physicians." 59th Conference of Experimental Psychologists (TeaP 2017), Leipzig, Deutschland. 26.03.-29.03.2017

Freytag J & Stroben F: „Effektives Teamwork? – Implementierung eines Feedback-Konzeptes zur Teamzusammenarbeit bei einer Nachtdienst-Simulation für Studierende". Internationales Skills Lab Symposium. Essen, Deutschland. 18.-19.03.2016

Deselaers T, Vorwerk L, Stroben F: „Emergency Room – Beurteilung und Behandlung von Akutpatienten anhand des ABCED-Schemas". Internationales Skills Lab Symposium. Essen, Deutschland. 18.-19.03.2016

Stroben F, Vorwerk L, Eisenmann D: „Schaffung einer Krankenhausumgebung durch Telefonsysteme und Konsiliardienste in Simulationen – Erfahrungen vom "Nachtdienst: Bist du bereit?"" Skills Lab Symposium. Halle/Leipzig, Deutschland. 13.-14.03.2015

Stroben F: „Nachtdienst: Bist du bereit? – Eine Nacht in einer simulierten Notaufnahme im Lernzentrum Berlin". 9. Jahrestagung der deutschen Gesellschaft für Interdisziplinäre Notfall- und Akutmedizin (DGINA), Nürnberg, Deutschland. 6.-8.11.2014

Stroben F, Schröder T, Dannenberg KA, Blaum WE: „Nachtdienst: Einfluss von Simulation und Rolle während der Simulation auf die subjektive Handlungsfähigkeit PJ-Studierender". Internationales Skills Lab Symposium. Bern, Schweiz. 14.-15.03.2014

Ott M, Föhr P, Stroben F, Schmidt A, Dannenberg KA, Eisner R, Kuhner M, Dahlbock R, Orlob S, Wildner G: „ConECT – studentische Notfalllehre vernetzt – Eine Vernetzungs- und Kommunikationsplattform für engagierte Studierende in der Notfall-medizinischen Lehre". Internationales Skills Lab Symposium. Bern, Schweiz. 14.-15.03.2014

Kongressbeiträge - Poster

Freytag J & Stroben F: „Schriftliches Feedback nach Simulationsszenarien – eine Möglichkeit den Lerneffekt zu verstärken?". GMA-Jahrestagung 2018. Wien, Österreich. 19.-22.09.2018

Eberz P, Freytag J, Lück S, Hitzblech T, Stroben F, Eisenmann DE: „SimulationspatientInnen(SP)-Feedback und kommunikative Kompetenzen Studierender – eine Mixed-Methods-Studie zur Erfassung von Absichten zur Verhaltensänderungen nach SP-Gesprächen". GMA-Jahrestagung 2018. Wien, Österreich. 19.-22.09.2018

Freytag J & Stroben F: „Graphic feedback summaries after large simulation events – a possibility to deepen the learning experience?". AMEE 2018. Basel, Schweiz. 25.-29.08.2018

Freytag J, Stroben F, Hautz WE, Eisenmann D, Kämmer JE: "Validation of the German "Teamwork Emergency Assessment Measure" TEAM – a study with student and expert raters." AMEE 2017. Helsinki, Finnland. 26.-30.08.2017.

Eisenmann D, Gerken J, Stroben F, Machner M, Hautz WE: "What changes do students of emergency nursing, medicine and student paramedics commit to after an inter-professional simulation training?" AMEE 2017. Helsinki, Finnland. 26.-30.08.2017.

Gerken J, Steinbart D, Dannenberg KA, Eisenmann D, Stroben F: „Students perspective on curricular implementation of a PAL-programme – a qualitative approach" AMEE 2017. Helsinki, Finnland. 26.-30.08.2017.

Eisenmann D, Stroben F, Machner M: „Entwicklung eines interprofessionellen Simulationstrainings für Auszubildende in der präklinischen und klinischen Notfallversorgung." Internationales Skills Lab Symposium. Erlangen, Deutschland. 31.03.-01.04.2017

Steinbart D, Stroben F, Dannenberg KA: „Ist ein freier Fertigkeiten-Parcours definierten Workshops überlegen? – Vergleich der Peer-Teaching-Repetitorien „Neuro-Tag" und Notfall-Tag" im Lernzentrum der Charite". GMA-Jahrestagung 2016. Bern, Schweiz. 14.-17.09.2016

Freytag J & Stroben F: „Teaching CRM principles to undergraduates using a peer-teaching approach". AMEE 2016. Barcelona, Spain. 27.-31.08.2016

Stroben F, Schroeder T, Dannenberg KA, Thomas A, Exadaktylos A, Hautz WE: „Training to increase students' self-efficacy in the ER – a simulated night shift". 9th European Congress on Emergency Medicine (EuSEM 2015). Torino, Italy. 10.-14.10.2015

Stroben F, Schroeder T, Dannenberg KA, Hautz WE: „Simulation training for final year medical students: Design features and effects improving self-efficacy". AMEE 2015. Glasgow, United Kingdom. 06.-09.09.2015

Dannenberg KA, Steinbart D, Stroben F, Ziegeler K: „Implementation of an extracurricular, student-organised peer-teaching programme into a party curricular programme". AMEE 2015. Glasgow, United Kingdom. 06.-09.09.2015

Dannenberg KA, Steinbart D, Stroben F, Ziegeler K: „Implementation eines extracurricularen, studentischen Peer-Teaching-Programms in ein teilcurriculares Programm". Skills Lab Symposium. Halle/Leipzig, Deutschland. 13.-14.03.2015

Ott M, Föhr P, Stroben F, Dannenberg KA, Orlob S, Wildner G: „ConECT – studentische Notfalllehre vernetzt". 7. Kongress der Arbeitsgemeinschaft für Notfallmedizin. Graz, Österreich. 24.-26.06.2014

Kongressbeiträge - Sonstige

Disqspace: Toutaoui KJ, Stroben F, Eberz P, Eisenmann D: „„Private Practice – Fälle aus der Hausarztpraxis": Tutorien zur Vermittlung von Perspektiven der Primärversorgung." Internationales Skills Lab Symposium. Erlangen, Deutschland. 31.03.-01.04.2017

Workshop: Stroben F, Freytag J, Dannenberg KA: „Aspekte des Falldesign (notfall-) medizinischer Simulationen – Realismus vs. Kosteneffizienz." Internationales Skills Lab Symposium. Erlangen, Deutschland. 31.03.-01.04.2017

Workshop: Stroben F & Dannenberg KA: „Wie organisiere ich mein studentisches Projekt, sodass es jeden Studiendekan begeistert? – Projektmanagement für Anfänger". Internationales Skills Lab Symposium. Essen, Deutschland. 18.-19.03.2016

13. Danksagung

Ohne KollegInnen und UnterstützerInnen wäre diese Dissertation kaum möglich – weswegen diesen Personen gedankt sei:

Herzlich danke ich Prof. Dr. Anke Thomas für die Betreuung meiner Dissertation und die Unterstützung an allen Punkten des Prozesses.
Besonderer Dank geht an PD Dr. Wolf E. Hautz, MME, der mich für das Thema der Ausbildungsforschung begeistern konnte, mich durch die Promotion geleitet hat und mich auch sonst bei allen Fragen und Sorgen auf meinem Weg unterstützt.

Zutiefst dankbar bin ich meinen beiden Mitautorinnen Katja A. Dannenberg und Julia Freytag, die nicht nur gemeinsam mit mir forschen und promovieren, sondern auch abseits der Arbeit immer unterstützend für mich da sind. Weiterhin danke ich allen anderen MitautorInnen und den Personen, die mich wissenschaftlich unterstützen: Dr. Dorothea Eisenmann, Dr. Stefanie C. Hautz, Dr. Juliane E. Kämmer, Mareen Machner, Therese Schröder, Prof. Dr. Aristomeinis Exadaktylos, Jan D. Gerken und Dr. Stefan Schauber.

Gedankt sei allen aktiven und ehemaligen TutorInnen und Mitarbeitenden des Lernzentrums und des Simulationspatientenprogramms der Charité, insbesondere Dr. Katharina Ziegeler, Tobias Deselaers, Hanno Heuzeroth, David Steinbart und Leopold Vorwerk für ihre Unterstützung bei der Durchführung der Simulationen.

Meinen FreundInnen danke ich für die emotionale Unterstützung während des Anfertigens meiner Dissertation und meinen Eltern für das Ermöglichen des Studiums und den Glauben an ihren Sohn.

Ohne Euch wäre dies nicht möglich gewesen.