

Aus der Klinik für Allgemeine Pädiatrie und Neonatologie

Universitätsklinikum des Saarlandes und Medizinische Fakultät der Universität des Saarlandes

Direktor: Prof. Dr. med. Michael Zemlin

**Systematisches simulationsbasiertes Ultraschalltraining zum Erlernen und
Verbessern von Ultraschallkompetenzen im erweiterten Zweittrimesterscreening:
FESIM III**

Dissertation zur Erlangung des Grades eines Doktors der Medizin

der Medizinischen Fakultät

der UNIVERSITÄT DES SAARLANDES

2023

vorgelegt von: Paula Freundt

geboren am: 21. Mai 1996 in Haltern

Tag der Promotion: 13.12.2023
Dekan: Prof. Dr. M. Menger
Berichterstattende: Prof. Dr. M. Zemlin
Prof. Dr. A. Bücken
Prof. Dr. S. Becker

Teilergebnisse dieser Arbeit wurden auf folgenden Kongressen vorgestellt:

- 64. Kongress der Deutschen Gesellschaft für Gynäkologie und Geburtshilfe e.V., München, 12.–15. Oktober 2022
J. Steinhard, **P. Freundt**, P. Janzing, E. Tutdibi, N. Nourkami-Tutdibi
Effizientes Simulator-Training in der Pränatalmedizin - Erfahrungen aus den FESIM-Studien
- 31. ISUOG world congress: online, 15.10. - 17.10.2021
J. Steinhard, **P. Freundt**, P. Janzing, Th. von Ostrowski, M. Langer, M. Zemlin, E. Tutdibi, N. Nourkami-Tutdibi *Controlled prospective study on the use of systematic simulator-based training with a virtual, moving fetus for learning 2 trimester scan: FESIM III*, Ultrasound Obstet Gynecol., <https://doi.org/10.1002/uog.24309>

Teilergebnisse dieser Arbeit wurden in folgenden Fachzeitschriften zur Publikation angenommen:

- **P. Freundt** ‡, N. Nourkami-Tutdibi ‡, E. Tutdibi, P. Janzing, Th. von Ostrowski, M. Langer, M. Zemlin, J. Steinhard, *Controlled Prospective Study on the Use of Systematic Simulator-Based Training with a Virtual, Moving Fetus for Learning Second-Trimester Scan: FESIM III*, Ultraschall in Med (2023), Georg Thieme Verlag KG, <https://doi.org/10.1055/a-1984-8320>
- J. Steinhard, **P. Freundt**, P. Janzing, V. Popov, R. Menkhaus, L. Ross, *Künstliche Intelligenz und Simulation in der Pränatalmedizin – was wir von Maschinen lernen können*, Die Gynäkologie (2022), Springer, <https://doi.org/10.1007/s00129-022-04996-5>

Inhaltsverzeichnis

Inhaltsverzeichnis	4
Abkürzungsverzeichnis	6
Kurzzusammenfassung	7
Abstract: Systematic simulation-based ultrasound training to learn and improve ultrasound competences for second trimester screenings	9
1. Einleitung	11
1.1. Sonographie in der Geburtshilfe	11
1.2. Standards der pränatalen Ultraschalldiagnostik in Deutschland	12
1.2.1. Ultraschalluntersuchungen in der Schwangerschaftsvorsorge in Deutschland	12
1.2.2. Das Mehrstufenkonzept	13
1.2.3. Standardebenen und Dokumentation der Ultraaschallscreenings	14
1.2.4. Ausbildung und Akkreditierung im geburtshilflichen Ultraschall in Deutschland	16
1.3. Ultraschallsimulatoren	18
2. Ziele und Fragestellung der Arbeit	21
3. Methode	22
3.1. Studiendesign	22
3.3. Die Module der Studie	24
3.4. Ausgewählte Standardebenen	27
3.5. Studienteilnehmer*innen	29
3.6. Ablauf, Training, Tests	30
3.7. Testauswertung	32
3.7.1. Auswertung nach der Zeit	32
3.7.2. Auswertung nach Korrektheit in der Ebenendarstellung	33
3.8. Fragebogen zur Ebenen-Evaluation und Feedback	34
3.9. Statistische Analyse	35
4. Ergebnisse	36
4.1. Vorerfahrung im gynäkologisch-geburtshilflichen Ultraschall	36
4.2. Ergebnisse der Testungen am Simulator im Task-Modul	37
4.3. Ergebnisse der Testungen am Simulator im Case-Modul	37
4.4. Vergleich der Einzelebenen	40
4.4.1. Vergleich der Einzelebenen-Performanz zwischen den Gruppen	40
4.4.2. Vergleich der Einzelebenen innerhalb einer Gruppe	41
4.5. Zweitbewertung durch Expertengremium	42
4.6. Ebenen-Evaluation	42
4.6.1. Schwierigkeitsbewertung der Gruppen im Vergleich	42

4.6.2.	Schwierigkeitsbewertung der Ebenen im Vergleich.....	45
4.6.3.	Korrelation zwischen Testperformanz und Schwierigkeitsbewertung.....	46
4.7.	Feedback zum Training am Simulator	46
5.	Diskussion.....	47
5.1.	SIM-UT vs. Klinische Erfahrung	47
5.2.	Durchführbarkeit und Möglichkeiten von SIM-UT.....	50
5.4.	Limitationen durch den Simulator	55
5.5.	Limitationen im Studiendesign	56
5.6.	Weitere Möglichkeiten für den Einsatz von Ultraschallsimulatoren.....	56
5.7.	Schlussfolgerung und Ausblick	57
6.	Abbildungsverzeichnis	58
7.	Tabellenverzeichnis	59
8.	Literaturverzeichnis	60
9.	Anhang	66
9.1.	Anhang 1	66
9.2.	Anhang 2	70
10.	Publikation.....	71
11.	Danksagung.....	72
12.	Lebenslauf	73

Abkürzungsverzeichnis

<u>Abkürzung</u>	<u>Erläuterung</u>
FESIM	Fetal Simulation Study
SIM-UT	Simulationsbasiertes Ultraschalltraining
DEGUM	Deutsche Gesellschaft für Ultraschall in der Medizin
MuRL	Mutterschaftsrichtlinien
KV	Kassenärztliche Vereinigung
SSW	Schwangerschaftswoche
TTC	Time to Completion
RAI	Rate of Appropriate Images

Kurzzusammenfassung

Einleitung: Das Erlernen pränataler Ultraschalluntersuchungen ist schwierig und eine inadäquate Ausbildung auf diesem Gebiet kann weitreichende Folgen, insbesondere für das postnatale Outcome des Kindes haben. Ziel der Studie war zu untersuchen, ob Trainees mit minimaler Erfahrung im pränatalen Ultraschall ausführliche Untersuchungsprotokolle, wie das der Feindiagnostik im zweiten Trimester, das erweiterte Zweittrimesterscreening, mittels strukturiertem simulationsbasierten Ultraschalltraining erlernen können und inwiefern sie sich durch das Training an Kompetenzniveaus von (Assistenz-)Ärzt*innen und Expert*innen annähern können.

Methodik: In einer prospektiven kontrollierten Studie, FESIM III, wurden die Lernkurven von Trainees während eines sechswöchigen strukturierten simulationsbasierten Ultraschalltraining erhoben und analysiert. Gegenstand des Trainings war ein erweitertes Zweittrimesterscreening mit 23 Standardebenen mittels high-end Simulator. Zwei Mal pro Woche fand ein strukturiertes Einzeltraining statt. Die Traineegruppe bestand aus 11 Medizinstudierenden mit minimaler Vorerfahrung (<10h) im gynäkologisch-geburtshilflichen Ultraschall. Der Lernerfolg wurde alle zwei Wochen durch einen standardisierten, praktischen Test am Simulator überprüft. Die Lernkurven wurden mit zwei Referenzgruppen verglichen, welche den gleichen Test am Simulator durchführten. Referenzgruppe A bestand aus zehn Assistenz- und Fachärzt*innen der Frauenheilkunde und Geburtshilfe. Referenzgruppe B war eine Expert*innengruppe mit zehn mindestens DEGUM-Stufe-II-zertifizierten Pränataldiagnostiker*innen. Die Testungen wurden aufgezeichnet und nach zwei Kriterien bewertet: (1) der Zeit, die zum Aufsuchen der Standardebenen und der Durchführung der Gesamtuntersuchung benötigt wurde, und (2) der Rate korrekt dargestellter Standardebenen.

Zudem wurde in einem Fragebogen für die Trainee- und Referenzgruppen die subjektiv empfundenen Schwierigkeitsniveaus der verschiedenen Standardebenen abgefragt. Es erfolgte zusätzlich eine Evaluation des simulationsbasierten Ultraschalltrainings durch die Trainees.

Die zur statistischen Analyse verwendeten Tests sind der Mann-Whitney-U- und der Kruskal-Wallis-Test. Korrelationen wurden anhand des Pearson-Korrelation-Koeffizienten bestimmt.

Ergebnisse: Die Trainees steigerten die Geschwindigkeit in der Durchführung der Untersuchung im Laufe des Trainings signifikant ($p < 0,001$). Nach sechswöchigem Training führten sie die Untersuchung signifikant schneller durch als die Assistenz- und Fachärzt*innen der Referenzgruppe A ($p < 0,011$). Die Expert*innengruppe blieb in der Gesamtperformanz signifikant schneller ($p < 0,001$). Hinsichtlich der korrekten Darstellung der Standardebenen erreichten die Trainees das Fertigkeiten-Level von Assistenz- und Fachärzt*innen, während die Expert*innen mit einer Rate von 100% die Untersuchung besser durchführten.

Die Schwierigkeitsbewertung der Ebenen lässt sich in drei Kategorien einteilen: einfach – mittelschwierig - schwierig. Die Evaluation erfolgte basierend auf subjektiv empfundener Schwierigkeit. Alle Trainees gaben an, viel Freude am simulationsbasierten Training zu haben.

Schlussfolgerung: Simulationsbasiertes Ultraschalltraining erwies sich als hoch effektive Methode zum Erlernen des erweiterten Zweittrimesterscreenings. Der standardisierte Test am Simulator bildete verschiedene Kompetenzniveaus ab. Trainees mit minimaler Vorerfahrung im geburtshilflichen Ultraschall profitierten vom regelmäßigen Training und erreichten nach 8h das Fertigkeiten-Niveau der Referenzgruppe A. Nach Abschluss des Trainings führten die Trainees die Untersuchung signifikant schneller durch als Referenzgruppe A, während Referenzgruppe B mit zehn spezialisierten Pränataldiagnostiker*innen signifikant schneller blieb. Simulationsbasiertes Ultraschalltraining ist ein realistischer Ansatz zum Erlernen fortgeschrittener Ultraschall-Fertigkeiten. Es stellt eine Lösung für die lange geforderte Änderung in Ultraschallcurricula dar, da es in der Organisation einfach und langfristig kosteneffizient umsetzbar ist und überdies hinaus von Trainees als beliebte Lehr- und Lernmethode mit hoher Akzeptanz angenommen wird.

Abstract: Systematic simulation-based ultrasound training to learn and improve ultrasound competences for second trimester screenings

Introduction: Education and training of prenatal ultrasound screenings are challenging, and insufficient training can have serious consequences for the postnatal outcome of children. Objectives of this study were to assess if trainees with minimal experience in gynecologic and obstetrical ultrasound could learn how to conduct second trimester screenings via structured simulation-based ultrasound training, and how far they can approach competence levels of physicians and fetal medicine specialists.

Methods: FESIM III was a prospective controlled trial. Learning curves of trainees were analyzed during a six-week structured simulation-based ultrasound training. Subject of the training was a second trimester screening with 23 standard planes on a high-end ultrasound simulator. Structured individual training took place twice a week. The trainee group consisted of 11 medical students with minimal experience (<10h) in gynecologic-obstetric ultrasound. The learning progress was assessed every two weeks in a standardized, hands-on test on the simulator. Learning curves were compared with two reference groups performing the same test on the simulator. Reference group A consisted of ten residents and specialists in obstetrics and gynecology. Reference group B was a group of experts with ten at least DEGUM-II-certified fetal medicine specialists.

The tests were recorded and evaluated according to two criteria: (1) the time to completion needed to obtain all standard planes and to conduct the screening, and (2) the rate of appropriate-obtained images.

Furthermore, the subjectively perceived difficulty to obtain the different standard planes was assessed in questionnaires. The trainee-group was also asked to evaluate their experience of six-week simulation-based ultrasound training.

For statistical analysis Mann-Whitney-U-Test and Kruskal-Wallis-Test were used. Correlations were determined by Pearson-Correlation-Coefficient.

Results: Trainees shortened their needed time to completion significantly during the study ($p < 0,001$). After six weeks of training trainees performed significantly faster than the residents and specialists of reference group A ($p < 0,011$). The expert group stayed significantly faster in their overall performance ($p < 0,001$). Regarding the rate of appropriate-obtained images the trainees achieved the level of reference group A after four weeks of training, while the expert group remained better with a rate of 100% appropriate obtained images.

The evaluation of difficulty in obtaining standard planes can be divided into three categories: easy – intermediate – difficult. All trainees evaluated their experience during simulation-based training positively.

Conclusion: Simulation-based ultrasound training is a helpful and effective tool to learn skills for second trimester ultrasound screenings. The standardized test on the simulator discriminated between different competence levels. Trainees with minimal experience reached the standard of the resident- and specialist group,

while experts remained significantly faster. The method could be an answer to long requested changes in ultrasound curricula because of its uncomplicated organisation and its popularity among trainees.

1. Einleitung

1.1. Sonographie in der Geburtshilfe

In den letzten Jahrzehnten hat sich die Sonographie zu einem der wichtigsten diagnostischen Tools in der Geburtshilfe entwickelt und gilt in der Schwangerschaftsvorsorge als das zentrale bildgebende Verfahren [6]. Die Untersuchung ermöglicht nicht nur die Sicherung einer intrauterinen Schwangerschaft sowie der Feststellung von Plazentalage, Größe und Alter des Fetus, sondern dient auch der Erkennung von fetalen Anomalien und Fehlbildungen [6]. Im Vergleich zu anderen Bildgebungen können sonographische Untersuchungen direkt am Patient*innenbett durchgeführt werden und liefern Echtzeitbilder [4]. Weiterhin sind Ultraschallgeräte beinahe überall verfügbar und können somit schnell eingesetzt werden, sodass einerseits Bildgebungen sowohl im Notfall zügig durchführbar sind und andererseits Patient*innen für geplante Untersuchungen keine langen Wartezeiten für die Diagnostik in Kauf nehmen müssen.

In der Pränataldiagnostik ist der hohe Stellenwert einer möglichst frühzeitigen fetalen Diagnose unter anderem darin begründet, dass sie die perinatale Morbidität und Mortalität senken kann [20,27]. Die pränatale Detektion fetaler Fehlbildungen, wie z. B. Spina bifida aperta, hat zudem eine Bedeutung für fetale intrauterine Therapien, die das fetale Outcome verbessern können [1]. Ein breites, bevölkerungsbasiertes Screening ist unter anderem deshalb sinnvoll, weil sich 80% der fetalen Pathologien und Anomalien vor allem bei Schwangerschaften außerhalb der Risikogruppen zeigen [27]. Technischer Fortschritt und die Entwicklung von Screeningprogrammen erlauben eine pränatale Detektion und differenzierte Diagnostik [39,61]. Dennoch bleiben die Vorsorgeprogramme hinter ihren Möglichkeiten zurück: Trotz hochauflösender high-end Geräte und vermehrter Inanspruchnahme von Ultraschalleistungen hat sich die Rate der pränatalen Diagnosen bzw. pränatale Detektion von Fehlbildungen nur wenig erhöht [17,51].

Die Qualität sonographischer Untersuchungen wird maßgeblich durch die Fähigkeiten und Fertigkeiten-Niveaus der Untersucher*innen bestimmt [4]. Anders als bei Kernspin- oder Computertomographien ist das Ergebnis der Sonographie von der Schallkopfführung der Untersuchenden abhängig. Somit stellt das Erlernen pränataler Sonographie in der medizinischen Ausbildung eine Herausforderung dar. Die Lernkurven im gynäkologisch-geburtshilflichen Ultraschall sind lang. Die Erwartungen, denen sich Ärzt*innen in der Facharztausbildung Gynäkologie und Geburtshilfe in Bezug auf pränatale Sonographie im klinischen Alltag ausgesetzt sehen, übersteigen signifikant die Eigenwahrnehmung ihrer Kompetenz [56]. Eine besondere Schwierigkeit des pränatalen Ultraschalls besteht darin, dass der Fetus in seiner Position äußerlich nicht sichtbar ist und sich auch während der Untersuchung kontinuierlich bewegt. Es gibt weniger sogenannte anatomische Landmarken im Vergleich zur Ultraschalluntersuchung des Herzens oder des Abdomens. Auch das Training zur Erkennung fetaler Anomalien ist komplex, da viele der Anomalien und Fehlbildungen nur eine geringe Prävalenz zeigen und somit die Häufigkeit, in der sie Assistentenärzt*innen während der gynäkologischen Facharztweiterbildung begegnen gering ist. Zusätzlich kann die sonographische Darstellbarkeit einzelner Standardebenen durch die maternale Bauchdecke oder die fetale Position eingeschränkt sein. Es gibt Hinweise, dass unter den

gynäkologischen und geburtshilflichen Ultraschalluntersuchungen, das Erlernen transabdominaler Ultraschalluntersuchungen mehr Training in Anspruch nimmt, als das Erlernen transvaginalen Ultraschalls [57]. Aufgrund der Komplexität fetaler Untersuchungen und deren Bedeutung für das peri- und postnatale Management von Mutter und Kind sowie das neonatale Outcome, ist es wichtig, dass die Durchführung pränataler Ultraschalluntersuchungen strukturiert und nachhaltig gelehrt wird. Seit Jahren fordern Expert*innen standardisierte Curricula und Zulassungstests für gynäkologisch-geburtshilfliche Ultraschalluntersuchungen [4,49].

1.2. Standards der pränatalen Ultraschalldiagnostik in Deutschland

1.2.1. Ultraschalluntersuchungen in der Schwangerschaftsvorsorge in Deutschland

In Deutschland orientiert sich die Schwangerschaftsvorsorge an den Mutterschaftsrichtlinien des Gemeinsamen Bundesausschusses. Diese erklären: „Vorrangiges Ziel der ärztlichen Schwangerenvorsorge ist die frühzeitige Erkennung von Risikoschwangerschaften und Risikogeburten“ [6]. Dabei dient die Ultraschalldiagnostik als Standardmethode der Bildgebung, die routinemäßig einmal pro Trimester durchgeführt wird, bei Bedarf zur weiterführenden Diagnostik auch häufiger. Die Zeiträume für die sonographischen Trimesterscreenings nach Mutterschaftsrichtlinien (MuRL) sind jeweils:

- 8+0 – 11+6 SSW
- 18+0 – 21+6 SSW
- 28+0 – 31+6 SSW

Beim Ultraschallscreening im zweiten Trimester kann die Schwangere zudem zwischen einer rein biometrischen Messung von Kopf, Abdomen und Femur oder einem systematischen, erweiterten Basis-Ultraschall mit der genaueren Beurteilung von ausgewählten Organsystemen wählen.

Bei Vorliegen von Risikofaktoren für die Schwangerschaft ist die Überweisung für weiterführende sonographische Diagnostik indiziert, die durch spezialisierte Gynäkolog*innen durchgeführt wird. Mögliche Indikationen sind z. B. das Vorliegen von Krankheiten und Fehlbildungen in der familiären Anamnese, psychische Belastung der Schwangeren, Auffälligkeiten im Basis-Ultraschall oder die Überwachung von Mehrlingsschwangerschaften [6,38]. Die Basis-Screenings gemäß Mutterschaftsrichtlinien definieren den gesetzlichen Mindeststandard an Ultraschalluntersuchungen in der Schwangerschaft. Es existieren weitere, darüber hinaus gehende Untersuchungsprotokolle, wie zum Beispiel die weiterführende Differentialdiagnostik des Fetus (Fehlbildungsdiagnostik) und die Screenings nach den Qualitätsstandards der Deutschen Gesellschaft für Ultraschall in der Medizin (DEGUM). Die Qualitätsanforderungen an die Fehlbildungsdiagnostik in Deutschland unterscheiden sich sowohl im Umfang als auch in der Anzahl an geforderter Bilddokumentation je nach zuständiger Kassenärztlicher Vereinigung (KV). Die DEGUM hat als bundesweit vertretene

Gesellschaft einheitliche Anforderungen an die verschiedenen Untersuchungen des Mehrstufenkonzepts [27,38].

In den vergangenen zwei Jahrzehnten hat das Ersttrimesterscreening mehr und mehr an Bedeutung in der Früherkennung von komplexen Fehlbildungen gewonnen. So konnten ca. die Hälfte der strukturellen Anomalien und Chromosomenabberationen schon mit einem differenzierten Screening im ersten Trimester detektiert werden [15,37,53]. Dennoch stellt die Fehlbildungs- bzw. Feindiagnostik im zweiten Trimester weiterhin den Goldstandard in der Pränataldiagnostik dar. Das Erkennen fetaler Fehlbildungen und damit assoziierten Erkrankungen ist hier mit der größten Sensitivität und Spezifität möglich, da einige Anomalien aufgrund der Entwicklung des Fetus im ersten Trimester noch nicht detektierbar sind [38]. Zudem sind die Schallbedingungen in der 20. und 24. Gestationswoche am besten. Gegenstand dieser Arbeit wird das Organscreening im zweiten Trimester sein, welches im Folgenden näher erläutert wird.

1.2.2. Das Mehrstufenkonzept

Das in Deutschland etablierte Mehrstufenkonzept soll jeder Schwangeren eine an die individuelle Risikokonstellation der Schwangerschaft angepasste Vorsorge ermöglichen. Dieses Konzept sieht drei Stufen vor: ein flächendeckendes Stufe-I-Screening ist für alle Schwangeren empfohlen. Patient*innen mit auffälligem Screening und z.B. dem Verdacht auf Schwangerschaftsanomalien werden an eine*n Stufe-II-Untersuchende*n überwiesen. Die Stufe-II-Diagnostik wird von Schwerpunkt-Untersuchenden durchgeführt, die über eine entsprechende Qualifikation und höhere Gerätestandards verfügen. Sollten Unklarheiten nach einer Stufe-II-Diagnostik bestehen, ist eine Überweisung für eine Stufe-III-Diagnostik angezeigt [19]. Auf dieses Mehrstufenkonzept beruft sich die DEGUM [27].

Das Mehrstufenkonzept funktioniert in Deutschland nur bedingt. Eine Studie zur pränatalen Detektionsrate fetaler Fehlbildungen vor und nach Einführung des MuRL Iib-Screenings 2013 in Hessen zeigt, dass dieses als Stufe-I-Diagnostik nicht ausreicht [50]. Insgesamt lag die pränatale Erkennungsrate von Anomalien sowohl vor als auch nach Einführung des Screenings bei lediglich 25% und die fetalen Anomalien, die im Rahmen des Iib-Screenings erkennbar wären, blieben mit 45 % auf einem ähnlich niedrigen Niveau wie vor dessen Einführung [50]. Die Anforderungen für die Durchführung des Iib-Screenings sind deutlich niedriger als z.B. für die Stufe-I-Diagnostik der DEGUM. Im Folgenden werden die Anforderungen und Umfang der verschiedenen Ultraschallscreenings in Deutschland näher beleuchtet.

1.2.3. Standardebenen und Dokumentation der Ultraaschallscreenings

Zusammenfassend lassen sich die etablierten Untersuchungen für das Zweittrimesterscreening in der Reihenfolge aufsteigenden Umfangs sortieren: die Screenings nach MuRL Iia, MuRL Iib, die DEGUM-Stufe-I-Diagnostik, die weiterführende Differentialdiagnostik des Fetus (Fehlbildungsdiagnostik) und die DEGUM-Stufe-II- und -III-Diagnostik.

Im Folgenden soll ein kurzer Überblick über die Standardebenen der Untersuchungen gegeben werden: Die MuRL Iib umfasst zusätzlich zur rein biometrischen Untersuchung nach Iia das sonomorphologische Screening von Kleinhirn, Hals- und Rückenkontur, der Herzaktion und der Beurteilung des Vierkammer-Blicks, der Bauchwand, Magenblase und Harnblase. Diese durch die MuRL „definierten Mindestanforderungen entsprechen den Qualitätsanforderungen an die DEGUM-Stufe I von 2006 [16]“ [27].

Die gestiegenen Anforderungen an die DEGUM-Stufe-I-Diagnostik wird in den aktualisierten Qualitätsanforderungen folgendermaßen begründet: „gestiegene Erfahrung im Ultraschallscreening [...] sowie eine moderne Ultraschalltechnik selbst in den heutigen „Basisgeräten“ ermöglichen eine Steigerung unserer Qualitätskriterien in einem modernen Screening gemäß den Forderungen zur Stärkung der pränatalen Basisdiagnostik (IQWiG)“ [27]. Das erweiterte DEGUM-Stufe-I-Screening umfasst zusätzlich zu den Inhalten des Screenings nach MuRL Iib seit 2019 die Beurteilung der Nieren in zwei Ebenen, die Echogenität der Lungen und weiterer thorakoabdomineller Organe. Die Erweiterung um die Nierenebenen erscheint nicht zuletzt deshalb sinnvoll, da Anomalien von Nieren und Harnsystem zu den häufigsten kongenitalen Anomalien zählen [20,61]. Optionale Bestandteile der DEGUM-I-Untersuchung sind Standardebenen des Gesichts im Profil und der Lippenaufsicht, sowie die Ausflusstrakte von rechter und linker Herzkammer [27]. Diese optionalen Ebenen sind in der DEGUM-II-Untersuchung fester Bestandteil. Zusätzlich sind dort aber auch noch eine genauere Untersuchung der Extremitäten, von Händen, Füßen und weiterer Strukturen von Kopf, Gehirn, Thorax und Abdomen erforderlich. Bei unklaren Fällen in der DEGUM-II-Diagnostik ist eine Überweisung an eine*n DEGUM-III-Untersuchende*n angezeigt [38].

Die Anforderungen an die Fehlbildungsdiagnostik (Stufe-II-Diagnostik) sind in Deutschland weder zentral geregelt, noch gehen sie zwangsläufig mit einer Stufe-II-Zertifizierung durch die DEGUM einher, vielmehr unterscheiden sie sich je nach der zuständigen Kassenärztlichen Vereinigung (KV). Am Beispiel der KV Westfalen-Lippe erschließt sich aus dem Dokumentationsbogen für die Fehlbildungsdiagnostik, dass die Untersuchung dieselben Organe wie das DEGUM-II-Screening umfasst [29]. Allerdings müssen im Rahmen der Untersuchung weniger Strukturen beurteilt werden, insbesondere Hirnstrukturen wie z.B. das Cavum septum pellucidi oder die Cisterna magna werden nicht berücksichtigt.

Bilddokumentationen gelten als ein wichtiger Baustein in der Qualitätssicherung des Untersuchungsstandards [3,38]. Einerseits kann fehlende Dokumentation zu einer schlechteren Patient*innen-Versorgung führen, andererseits können dokumentierte Bilder die Grundlage zur Verbesserung und Qualitätskontrolle darstellen [3]. Außerdem stellen sie einen bedeutenden Teil für die Beantwortung forensischer Fragestellungen dar [38].

Die Pflicht zur Bilddokumentation der Untersuchungen und deren Befunde unterscheidet sich für die verschiedenen Screenings sowohl in der Anzahl der zu dokumentierenden Befunde als auch in der Genauigkeit der Definition dessen, was dokumentiert werden muss. Die Pflicht zur Bilddokumentation der Basisultraschalluntersuchungen gemäß IIa und IIb der MuRL fällt am geringsten aus: Hier sind lediglich die biometrischen Ebenen und auffällige Befunde zu dokumentieren [6]. Das Screening nach DEGUM I schreibt neben der Bilddokumentation der biometrischen Messungen noch die weiteren sechs Standardebenen vor [27]. Bei Indikation für weiterführende Diagnostik muss diese entsprechend umfassender dokumentiert werden. Die Anzahl der zu dokumentierenden Ebenen unterscheidet sich zwischen den Kassenärztlichen Vereinigungen. So fordert die KV Westfalen-Lippe bei der Fehlbildungsdiagnostik die Dokumentation von mindestens ca. 15 Ebenen [30], während in der DEGUM II Untersuchung 21 Ebenen dokumentiert werden müssen [38]. Die Mindestanforderungen an die Bilddokumentation der beiden Untersuchungen sind in Tabelle 1 gegenübergestellt. Die Unterschiede in der Dokumentation der beiden Stufe-II-Untersuchungen bestehen zum Beispiel in den Standardebenen des fetalen Herzens (s. Tab. 1). Herzfehler sind die am häufigsten isoliert auftretende kongenitale Fehlbildung mit einem Prävalenz von ca. 1% in Deutschland [13,32].

Körperregion	DEGUM-Stufe-II-Diagnostik	Anforderungen an die Fehlbildungsdiagnostik nach KVWL
Kopf:	Schädel mit Planum frontooccipitale	Kopfform
	Cerebellum	intrazerebrale Strukturen
	Gesichtsprofil mit Nasenbein	Gesichtsprofil
	Orbitae	Frontalschnitt
	Aufsicht Lippen/Nase	
Rumpf	Wirbelsäule mit Hautkontur und Sagittalschnitt	Nackenregion, gesamte Wirbelsäule
	Herz: 4-Kammerblick	Herzposition, 4-Kammerblick
	Herz: Linksventrikulärer Ausflusstrakt	
	Herz: Rechtsventrikulärer Ausflusstrakt	
	Zwerchfell im Sagittal- und Frontalschnitt	Lungen- und Leberstruktur
	Abdomenquerschnitt mit Magen	
	Fetaler Nabelschnuransatz	Nabelschnuransätze (plazentar und fetal), Nabelschnurgefäße
	Nieren beidseits	
	Harnblase	
Extremitäten:	Femur und Humerus	Extremitäten
	Tibia und Fibula	
	Radius und Ulna	
	Hände und Füße	
	Plazentasitz	

Tabelle 1: Gegenüberstellung der Mindestanforderungen an die Bilddokumentation der weiterführenden Differentialdiagnostik [30,38]

1.2.4. Ausbildung und Akkreditierung im geburtshilflichen Ultraschall in Deutschland

In Deutschland existieren weder im Medizinstudium noch in der Facharztweiterbildung Gynäkologie und Geburtshilfe standardisierte Curricula für die Ausbildung in der Sonographie [28,47]. Lehre in der Sonographie findet teilweise in freiwilligen Wahlfächern im Studium statt, in der Klinik durch Zuschauen und teilweise supervidierte Sonographie oder in externen, kostenpflichtigen Fortbildungen in der Weiterbildungszeit. Da standardisierte nationale Lehrformate fehlen, variiert hier die Qualität und das Angebot der sonographischen Lehre innerhalb Deutschlands erheblich. Praktisches Training und fachliche Supervision fallen auf vielen Fortbildungen in Deutschland kurz aus, während die theoretische Wissensvermittlung mehr Raum einnimmt [21]. Hofer et al. untersuchten über einem zweijährigen Zeitraum 34 Ultraschallkurse, die in verschiedenen Teilen Deutschlands stattfanden und von regionalen Ärztekammern sowie der Bundesärztekammer durchgeführt wurden. Während die Schallzeit pro Teilnehmer*in und Kurstag sich auf 16,3 +/- 7,0 min (SD) Minuten belief, nahm der theoretische Teil pro Tag 76 +/- 25 min (SD) ein [21]. In einer deutschlandweiten Studie wurden Medizinstudierende zur Ultraschalllehre an ihren Fakultäten befragt [47]. An der Umfrage nahmen >1% aller Medizinstudierenden in Deutschland teil (n=1040), die an 31 Fakultäten in Deutschland immatrikuliert waren. Die Ergebnisse zeigen, dass Medizinstudierende den Wunsch haben, Ultraschalllehre würde bereits in ihrem Studium beginnen: ca. 98% gaben Interesse an Ultraschallkursen im Studium an und etwa 94% sprachen sich für Ultraschallkurse als verpflichtenden Kurs im Curriculum aus, auch wenn dieser mehr Arbeit in bereits vollen Stundenplänen bedeuten würde. Gleichzeitig wurde nach den Gründen gefragt, warum die Studierenden aktuell keinen Ultraschallkurs an ihrer Universität belegen würden. Der häufigste Grund (65%), war dass ihr Stundenplan nicht genug Zeit für ein weiteres Fach freigegeben würde. Doch auch >29% der Teilnehmenden gaben an, dass an ihrer Fakultät kein Ultraschallkurs für Studierende angeboten würde [47].

In Deutschland wird pränataler Ultraschall ausschließlich von (Fach-)Ärzt*innen für Frauenheilkunde und Geburtshilfe durchgeführt. Der Erwerb zur fachlichen Befähigung ist durch die Ultraschallvereinbarungen der Kassenärztlichen Bundesvereinigung festgelegt und umfasst 300 attestierte geburtshilfliche B-Bild-Sonographien der utero-plazento-fetalen Einheit, sowie 100 Duplex-Sonographien ([28] Anhang I Spalte 3). Eine obligatorische praktische oder theoretische Basisausbildung darüber hinaus dagegen, ist – ähnlich wie in einigen anderen europäischen Ländern auch – nicht vorgeschrieben [49]. Das hat zur Folge, dass Ärzt*innen mit Bestehen der Facharztprüfung selbstständig die Basisultraschallscreenings in der Schwangerschaftsvorsorge durchführen können. Eine Ausnahme stellt der erweiterte Basis-Ultraschall nach MuRL IIb dar, dessen Durchführung nur von Fachärzt*innen vorgenommen werden darf, nachdem ein entsprechender Befähigungsnachweis gegenüber der KV erbracht wurde. Die Anforderungen an den Befähigungsnachweis sind durch die Anlage VI Ultraschallvereinbarung der KBV [28] geregelt und werden als einfache Online-Prüfung ohne hohen Anspruch von den Kassenärztlichen Vereinigungen durchgeführt.

Die weiterführende Differenzialdiagnostik des Fetus, der sogenannte Fehlbildungultraschall, erfordert ebenfalls eine Genehmigung durch die KV. Für die Genehmigung müssen 200 weiterführende differentialdiagnostische Sonographien nachgewiesen werden, wovon 30 eine Fehlbildung oder Entwicklungsstörung aufweisen sollen. Die Untersuchungen müssen unter Anleitung einer*s qualifizierten Ärztin*Arztes stattgefunden haben. Alternativ kann der Nachweis über einen Ultraschallkurs, ein Kolloquium oder die 18-monatige Tätigkeit in einem Fachzentrum erbracht werden [28].

Die flächendeckende und vollständige Stufe-I-Diagnostik ist für ein funktionierendes Mehrstufenkonzept unabdingbar und Voraussetzung für die pränatale Detektion kongenitaler Anomalien. Das 2013 eingeführte erweiterte Basisscreening gemäß Mutterschaftsrichtlinie, das Iib-Screening, kann laut einer kürzlich veröffentlichten Studie keine ausreichend hohen pränatalen Detektionsraten gewährleisten [50]. In Hessen wurden die pränatalen Detektionsraten vor und nach Einführung des MuRL Iib-Screenings im Jahre 2013 verglichen. Dafür wurden Daten aus den Zeiträumen von 2010-2013 und von 2013-2016 miteinander verglichen. Die Untersuchungen ergaben, dass sich durch die Einführung des Iib-Screenings die Detektionsrate fetaler Fehlbildungen nicht signifikant änderte, welche sowohl vor als auch nach Einführung des erweiterten Basisscreenings um die 25% lag. Dieses Defizit ist bei einer Abrechnung des Iib-Screenings in >90% der Zweittrimesterscreenings nicht darauf zurückzuführen, dass Schwangere sich gegen eine ausführliche Diagnostik und für das Iia-Screening entschieden. Interessant ist ebenfalls, dass in Hessen trotz der Vorgabe zur Bewertung der vier Herzkammern im Iib-Screening, nur 16,7% der Herzfehler, die im Vierkammer-Blick detektierbar wären, pränatal diagnostiziert wurden [50]. Die Arbeitsgruppe plädiert daher für die Stärkung der sonographischen Basisdiagnostik in der Schwangerschaftsvorsorge und den Ersatz des Iib-Screenings durch ein differenzierteres Ultraschallscreening, das von qualifizierten Ärzt*innen durchgeführt werden soll [50].

Die Notwendigkeit einer guten Ausbildung für die eigenständige Ausübung pränataler Ultraschalldiagnostik ist gegeben. Das Konzept, in der ärztlichen Ausbildung eigenständiges Schallen über das Erreichen einer Mindestanzahl bescheinigter sonographischer Untersuchungen zu regulieren, ist nicht unumstritten [26,55]. Die Qualitätssicherung für pränatale Basisultraschalluntersuchungen in Deutschland erscheint fragwürdig. 300 supervidierte Ultraschalluntersuchungen der utero-plazento-fetalen Einheit reichen für den Erwerb des selbstständigen Schallens aus [2] - doch Supervision fällt im klinischen Alltag aufgrund der hohen Arbeitsbelastung des Personals in deutschen Kliniken oft spärlich aus. Ebenfalls fragwürdig erscheint die Aussagekraft des Befähigungsnachweises für das erweiterte Basisultraschallscreening Iib gemäß MuRL. Die Richtlinien für die Online-Prüfung für den Befähigungsnachweis sind durch Anlage VI der Ultraschallvereinbarung der KVB geregelt und werden durch die jeweiligen Kassenärztlichen Vereinigungen durchgeführt [28]. Die Prüfung umfasst 30 zufällig generierte Fälle mit Bild- und Videomaterial aus einem Fragepool mit mindestens 200 Fällen. Die Prüfungsfragen sollen mit „Ja“, „Nein“ oder „Weiß ich nicht“ beantwortet werden. Außerhalb der obligaten DEGUM-Re-Zertifizierungen existieren keine regelmäßigen Qualitätskontrollen.

Die Relevanz einer guten sonographischen Ausbildung wird auch durch eine bereits 2008 veröffentlichte Analyse des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) unterstrichen, welche das Ultraschall-Screening in der Schwangerschaft im Hinblick auf die Detektionsrate fetaler Anomalien untersuchte. Die Ergebnisse der Analyse zeigten eine Assoziation zwischen hoher Qualifikation des*r Untersuchenden und höheren Detektionsraten fetaler Anomalien. Zudem werden im Fazit der IQWiG-Studie die Ergebnisse von Studien über das mehrstufige Screening-Programms in Deutschland wie folgt bewertet: „Die in diesen Studien erzielten Detektionsraten sind jedoch niedrig und stellen ein mehrstufiges Screeningprogramm mit Eingangsscreening auf Stufe I infrage“ [24]. Daraufhin wurden die Mutterschaftsrichtlinien 2012 neu aufgesetzt, doch trotzdem hat sich das Screening, vor allem im Vergleich mit der DEGUM-Stufe-I-Diagnostik, kaum verändert. Es wurde - bis auf den Onlinetest zur Befähigung zum MuRL Screening IIB - kein weiterer Qualifikationsnachweis für die Stufe-I-Diagnostik eingeführt und die pränatale Detektionsrate kongenitaler Fehlbildungen hat sich nicht verbessert [50].

1.3. Ultraschallsimulatoren

Ein Ansatz den hohen Anforderungen an die Lehre pränatalen Ultraschalls im klinischen bzw. universitären Alltag gerecht zu werden, stellt der Einsatz von Ultraschallsimulatoren dar. Simulation wurde bereits vor 2500 Jahren in die chirurgische Ausbildung aufgenommen, damit Anfänger*innen noch vor dem Praktizieren an Patient*innen vorbereitet werden konnten [43]. Auch in der Geburtshilfe existiert das Konzept des simulationsbasierten Lernens bereits seit Jahrhunderten [45]. Heutzutage existieren durch den technologischen Fortschritt und Einsatz neuer Simulatoren gute Evidenzen für die Lehre mit Ultraschallsimulatoren [40,54]. Durch den Einsatz von Ultraschallsimulatoren kann eine ruhige und Trainee-zentrierte Lernatmosphäre samt -umfeld geschaffen werden. Untersuchungen können durch Übung am Simulator beliebig oft wiederholt werden. So können sich Lernende auf sich und die Schallkopfführung und/oder Bildeinstellung konzentrieren, wohingegen im klinischen Alltag selbstverständlich Patient*innen anstelle von Lehre im Fokus der Untersuchung stehen. Durch den Simulator haben Lernende zudem die Möglichkeit zahlreiche und seltene Pathologien kennenzulernen, ohne dass sich tatsächlich kranke oder betroffene Patient*innen zu Übungszwecken finden müssen.

Das Prinzip des simulationsbasiertem Ultraschalltrainings (SIM-UT) wird auch für die sonographische Ausbildung von gynäkologisch-geburtshilflichen Ultraschalluntersuchungen untersucht. Ultraschallanfänger*innen können grundlegende Skills der fetalen Biometrie mithilfe von Ultraschallsimulatoren erlernen und verbessern. Direkte Trainingseffekte wurden schon nach einzelnen Trainingseinheiten am Simulator nachgewiesen [48]. So verbesserten Anfänger*innen durch die Möglichkeit zu repetitiven Schallübungen sowohl die Genauigkeit ihrer Messungen signifikant, als auch die Geschwindigkeit im Aufsuchen der Ebenen, die sich teilweise halbierte, ohne dass dabei die Untersuchung an

Genauigkeit verlor [8,11,31]. Hinzu kommt, dass Lernende selbst von dem Prinzip des SIM-UT überzeugt sind. In einer britischen Umfrage gaben 73% der befragten Assistenzärzt*innen SIM-UT als essenziellen Bestandteil ihrer gynäkologisch-geburtshilflichen Weiterbildung an und insgesamt 80% bewerteten es als ein nützliches Instrument ergänzend zum klinischen Training [44].

Die positiven Effekte von SIM-UT zeigen sich auch in der Patient*innenzufriedenheit. Tolsgaard et al befragten in einer multizentrischen, randomisierten Studie Patient*innen in der gynäkologischen Notaufnahme nach ihrer Einschätzung bezüglich verschiedener Qualitätsmerkmale der gynäkologisch-geburtshilflichen Untersuchung [59]. Im Rahmen der Studie wurden Ärzt*innen aus vier Zentren randomisiert in einer Kontroll- und eine Interventionsgruppe eingeteilt, wobei die Interventionsgruppe zusätzlich zum klinischen Training auch Training mittels SIM-UT absolvierte. Insgesamt bewerteten die Patient*innen die Untersuchungsqualität von Untersuchenden aus der Interventionsgruppe signifikant höher. In Untersuchungen durch Ärzt*innen, die zusätzlich SIM-UT zum klinischen Training durchlaufen hatten, gaben die Befragten weniger Unbehagen durch die Untersuchung an, während Sicherheitsempfinden und Vertrauen höher als in der Kontrollgruppe angegeben wurden. Des Weiteren arbeitete die Interventionsgruppe auch effizienter, wodurch sich Untersuchungszeiten verkürzten und seltener Oberärzt*innen den Untersuchungen hinzugezogen werden mussten.

Insgesamt erwies sich der Einsatz von Ultraschallsimulatoren zur Qualitätssicherung und zum Anheben des Standards aktueller sonographischer Untersuchungen bis dato als nützlich [7,10,33]. Evidenz gibt es sowohl für das Feststellen von praktischen Ultraschallkenntnissen von transvaginalen Ultraschalls [33], als auch für geburtshilfliche transabdominelle Untersuchungen [7,10].

Bei dem jährlichen französischen Staatsexamen für Ultraschall in Gynäkologie und Geburtshilfe im Jahr 2013, wurden ca. 20% der Teilnehmer*innen zufällig ausgewählt [10]. Diese sollten zusätzlich zu dem praktischen Prüfungsteil, bestehend aus einer Zweittrimester-Ultraschalluntersuchung an einer schwangeren Person, dieselbe Untersuchung an einem high-end Simulator wiederholen. Die Ergebnisse am Ultraschallsimulator unterschieden sich nicht signifikant von denen am Live-Modell. Chalouhi et al. schlussfolgerten, dass die Überprüfung praktischer, sonographischer Kenntnisse an Live-Modellen mit der an high-end Ultraschallsimulatoren verglichen werden kann, wobei die Prüfung am Simulator zusätzliche organisatorische und ökonomische Vorteile aufzeigt.

Ob sich auch unterschiedliche Erfahrungslevel von Ultraschaller*innen, Assistenzärzt*innen und Oberärzt*innen der Gynäkologie und Geburtshilfe durch einen praktischen Test an einem Ultraschallsimulator abbilden lassen, wurde von Burden et al. untersucht [7]. Die Überprüfung erfolgte durch die Durchführung zweier standardisierte Tests. Zuerst sollten die Teilnehmer*innen die Scheitel-Steiß-Länge eines Embryos bestimmen und anschließend in einem Dritttrimester-Screening die fetale Biometrie anhand sechs Parameter durchführen. Die Ergebnisse der Untersuchungen wurden auf Messgenauigkeit und Schnelligkeit analysiert. Insgesamt fanden sich alle Prüflinge mit der Handhabung des Simulators zurecht. Die Erfahrungslevel der Teilnehmer*innen wurden durch die Testergebnisse widerspiegelt. Die Assistenzärzt*innen benötigten

signifikant mehr Zeit für die Darstellung der Standardebenen, während die Expert*innen im Durchschnitt eine höhere Messgenauigkeit bewiesen. Daraus schlossen Burden et al., dass die benötigten Kenntnisse zur Durchführung einer simulationsbasierten Sonographie denen für klinische Sonographie entsprechen. Insgesamt resümierte die Arbeitsgruppe, dass verschiedene Kompetenzlevel durch simulationsbasierte Tests abgebildet werden können.

2. Ziele und Fragestellung der Arbeit

In der prospektiven, experimentell-didaktischen Studie „Fetal Simulation Study III“ FESIM III sollte der Lernfortschritt von Trainees mit minimalen Vorerfahrungen im pränatal-diagnostischen Ultraschall während des Trainings an einem high-end Ultraschallsimulator erfasst werden. Dabei sollte untersucht werden, inwiefern Anfänger*innen durch selbstständiges, strukturiertes simulationsbasiertes Ultraschalltraining (SIM-UT) sich den Kompetenzlevels von Ärzt*innen und DEGUM-II-zertifizierten Fachärzt*innen sowohl im Hinblick auf Korrektheit der Standardebenen eines umfangreichen Zweittrimesterscreenings als auch auf Geschwindigkeit der Untersuchung annähern können. Folgende Fragestellungen standen dabei im Fokus:

- Wie viel Training benötigen Anfänger*innen bis sie ein fachkompetentes, professionelles Level in der sonographischen Zweittrimester-Untersuchung erreichen?
- Wie weit können sich Trainees mit Hilfe von strukturiertem, selbstständigem SIM-UT dem Expert*innenlevel in Bezug auf Geschwindigkeit und Korrektheit der Standardebenen eines weiterführenden Zweittrimester-Ultraschallscreenings annähern?
- Könnte die Prüfung am Simulator eine sinnvolle Methode für Qualitätskontrollen im deutschen Mehrstufenkonzept der sonographischen Pränatalmedizin sein?
- Wie kann strukturiertes SIM-UT an deutschen Universitäten oder Krankenhäusern implementiert werden?
- Gibt es Standardebenen in der Zweittrimesterdiagnostik die schwieriger darzustellen sind als andere?

3. Methode

3.1. Studiendesign

In einer prospektiven, experimentell-didaktischen Studie wurden die Lernkurven von Medizinstudierenden an einem Sonographie-Simulator für pränatal-diagnostische Zweittrimester-Ultraschalluntersuchungen analysiert. Studienleiter ist DEGUM-Kursleiter, Ausbilder der ÄKWL und KVWL Dr. Johannes Steinhard. Die Studie wurde vom 30. August bis zum 11. Oktober 2020 im Sim-Zentrum des St.-Franziskus-Hospitals Münster durchgeführt. Über den Zeitraum von sechs Wochen absolvierten die Studierenden zwei Mal pro Woche selbstständiges, strukturiertes SIM-UT. Der Lernfortschritt der Studierenden wurde regelmäßig überprüft, aufgezeichnet und schließlich mit der Ausführung der Ultraschalluntersuchung, hier genannt Performanz, von (Assistenz-)Ärzt*innen und mindestens DEGUM-II-zertifizierten Fachärzt*innen für Frauenheilkunde und Geburtshilfe verglichen.

3.2. Der Sonographie-Simulator

Eingesetzt wurde der Sonographie-Simulator U/S Mentor von Symbionix (Beit Golan, Israel). Der U/S Mentor ist ein high-end Simulator zum Erlernen und Verbessern von Fähigkeiten für sonographische Untersuchungen und Interventionen. Der U/S Mentor verfügt über vielfältige Module aus verschiedenen medizinischen Fachbereichen. Integriert sind Module von Thorax-, Abdomen-, Schilddrüsen-, gynäkologischer und geburtshilflicher Sonographie. Dafür ist auch die Anwendung transösophagealen und transvaginalen Ultraschalls möglich.

Das System besteht aus einem Computer, zwei lebensgroßen Mannequins, vier verschiedenen Ultraschallsonden, einem Fußpedal, einem Transmitter und einer Umschaltbox, die die verschiedenen Teile des Simulators zusammenführt. Für ein realitätsnahes Schallen ist es möglich sowohl an einem weiblichen als auch an einem männlichen Mannequin-Torso Ultraschalltraining durchzuführen.

Die Bewegungen des Schallkopfes werden durch Kommunikation zwischen Schallkopf und dem sogenannten Transmitter registriert. Dieser besteht aus einer kleinen Box, die auf die Rückseite der Mannequins eingesetzt wird. Diese Box emittiert ein radiäres Magnetfeld, welches von Sensoren in den angeschlossenen Schallsonden detektiert wird. Dadurch kann der Simulator die Position der Schallköpfe am Torso bestimmen und das korrespondierende B-Bild auf dem Computerbildschirm wiedergeben. Aufgrund des Magnetfeldes können alle Strukturen aus jeglichem Winkel betrachtet werden, sofern dies anatomisch möglich ist.

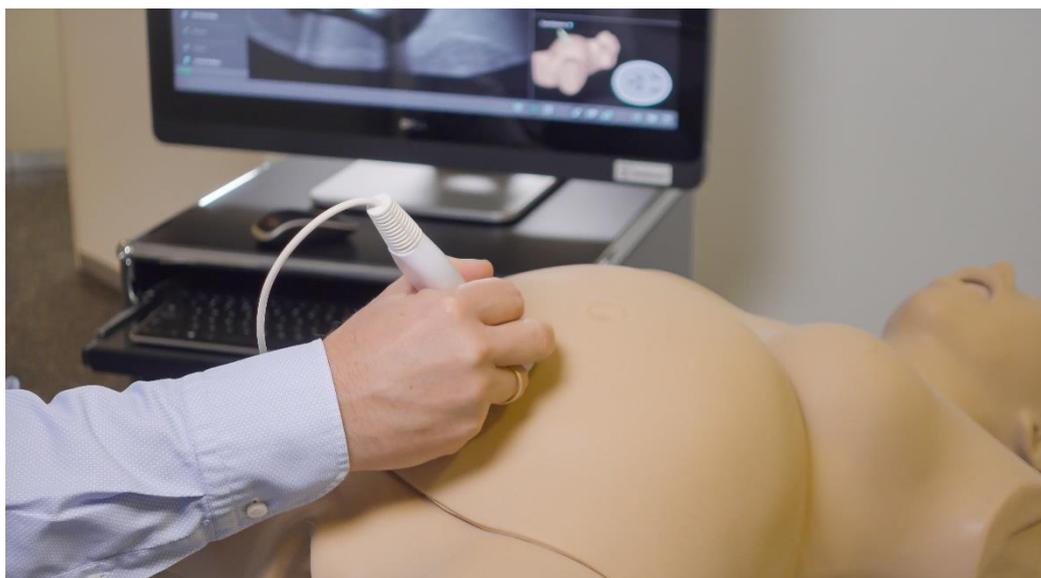


Abbildung 1: Der U/S Mentor: Mannequin für Zweittrimester-Screening mit Abdomensonde. Im Hintergrund: Computer des Simulators
Abbildung modifiziert nach [18]

Das Bildmaterial ist auf Basis von MRT-Befunden dreidimensional generiert. Dadurch werden die Organe dreidimensional simuliert, sodass der realistische Eindruck einer Sonographie am Menschen entsteht. Es sind außerdem Artefakte, wie z.B. Schallschatten durch ossäre Strukturen eingearbeitet. Dadurch entstehen zwei Vorteile gegenüber anderen Simulatoren. Die dreidimensionale Simulation des Bildes ermöglicht, dass eine anatomische Struktur wie im Rahmen einer echten sonographischen Untersuchung aus jeder Position und jedem Winkel betrachtet werden kann. Im Gegensatz dazu können Simulatoren, die real-generierte Schallbilder integriert haben, lediglich diese vorher aufgenommenen Bilder wiedergeben. Das flexible Herangehen von verschiedenen Seiten ist auf diesem Wege nicht möglich. Die Simulation des Bildmaterials ermöglicht weiterhin die fetalen Bewegungen in den pränatalen Ultraschallmodulen. Die Bewegungen eines simulierten Fetus sind eine weitere Errungenschaft dieser Methode.

Das Programm des Simulators ist auf Englisch. Am Beispiel der geburtshilflichen Sonographie werden die Anwendungen des U/S Mentor während einer laufenden Untersuchung beschrieben. Der Bildschirm des Computers ist zu Trainingszwecken in vier Teile geteilt, siehe Abbildung 2:



Abbildung 2: Bildschirm des Simulators während des Zweittrimestermoduls im Task-Modus bei aktivierter Hilfestellung. Abbildung modifiziert nach [18]

Die Symbole in der Spalte am linken Bildschirmrand dienen der Bildeinstellungen und -optimierung, sowie der biometrischen Messungen. Der Computer lässt sich per Mausklick, Tastatur und mit dem Touchfunktion des Bildschirms bedienen. Den größten und zentralen Teil des Bildschirms nimmt das B-Bild ein. Die typische Kaffeefilter-Form füllt sich mit einem Bild, sobald der Schallkopf aufgesetzt wird. Rechts im Bild befinden sich zwei Felder untereinander. Im oberen sieht man ein animiertes dynamisches Modell des Fetus ggf. mit Plazenta in Relation zur aktuellen Schallkopfposition. Das Modell lässt sich beliebig einstellen: So kann alternativ das animierte Modell in der jeweiligen Schnittbildebene durch den Schallstrahl geteilt, oder auch nur die inneren Organe des Fetus angezeigt werden. Im unteren Feld ist schließlich der Mannequin und mit dem darüberfahrenden Schallkopf in der jeweils aktuellen Position sichtbar. Hier gibt es eine weitere optionale Hilfestellung des Simulators: nähert sich der Schallkopf der richtigen Standardebene, färbt er sich von einem grauen Farbton zu einem intensiven Dunkelblau. Daneben befindet sich ein schematisches Piktogramm der einzustellenden Standardebene. Außerdem kann durch Aktivierung der kleinen Fetus-Symbole mittig in der obersten Bildzeile der Fetus zur Bewegung animiert oder Bewegungen gestoppt werden.

3.3. Die Module der Studie

Für die Durchführung der Studie wurde der weibliche Torso und die Konvexsonde für transabdominalen Ultraschall verwendet. Der U/S Mentor verfügt über ein Modul für Zweittrimesterscreenings, welches wir für die Studie genutzt haben. Das Zweittrimesterscreening wird am Simulator an einem virtuellen Fetus in der 20. SSW vorgenommen, die Standardebenen sind gemäß dem Standard der International Society of Ultrasound in Obstetrics and Gynecology (ISUOG) gewählt. Das Modul ist in zwei Aufgabentypen unterteilt: es enthält zwei

sogenannte Task-Modi und vier Case-Modi. Im Programm besteht ein Task darin, dass an einem sich nicht bewegendem Fetus Anatomie und Standardebenen trainiert werden können. Der Fetus ist im Task-Modus des Zweittrimesterscreenings physiologisch für die 20. SSW ohne jegliche Anomalien entwickelt und präsentiert sich in Beckenendlage. In dem von uns ausgewählten Task 1 lag der Fokus auf dem Aufsuchen der Ebenen und dem Erkennen der fetalen Anatomie. Dafür verfügt der Simulator über weitere Hilfestellungen. Am linken Bildschirmrand ist eine Spalte mit der Liste der darzustellenden Standardebenen aufgezeigt, deren Punkte bei erfolgreichem „Freeze“ der jeweiligen Ebene als erledigt gekennzeichnet werden. Außerdem können fetale Strukturen vom Simulator mit ihren anatomischen Bezeichnungen versehen werden. Die anatomischen Bezeichnungen können beliebig aktiviert und deaktiviert werden. Wenn die korrekte Ebene dargestellt ist, kann sie durch Aktivierung des Fußpedals „eingefroren“ bzw. gesichert werden. Ist die Ebene nicht richtig dargestellt, erlaubt der Task-Modus keine Sicherung des Bildes. Im Task-Modus geht es also darum, mit Hilfestellung und direktem Feedback durch den Simulator die entsprechende Standardebene aufzusuchen und durch „Freeze“ zu sichern. Als weitere Hilfestellung im Task-Modus steht unten im Bild eine knappe Beschreibung, wie die Ebene aufzufinden ist, zusammen mit den Strukturen, die in der Ebene abgebildet sein sollen.

Als Beispiel folgt die Anleitung zum Auffinden des Gesichts im Profil: *„Place the transabdominal transducer on the belly surface and locate the fetal head. Tilt the transducer to obtain a midsagittal plane of the fetal face.*

Note: This view is recommended when technically feasible.

Structures to be demonstrated: Nasal bone, mandible, maxilla.”

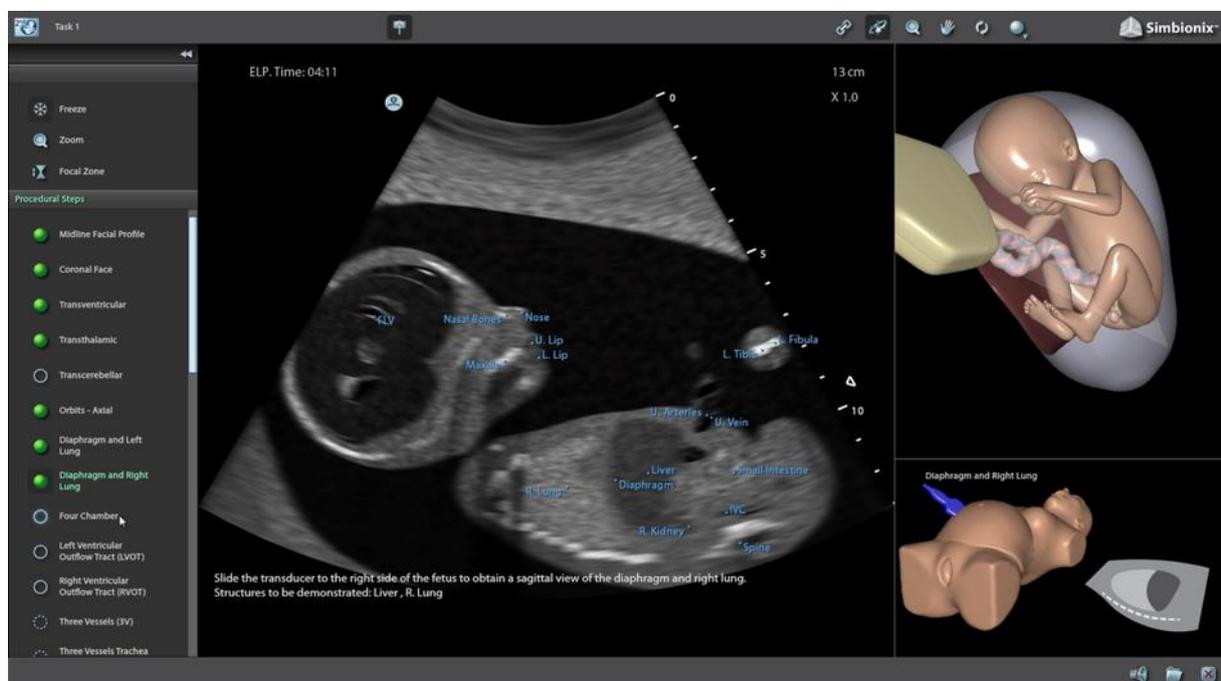


Abbildung 3: Task-Modus für Standardebene Diaphragma/rechte Lunge. Links im Bild ist die Liste der Standardebenen, anatomische Labels sind aktiviert. Der Schallkopf auf dem Mannequinmodell unten rechts ist tiefblau, da die Ebene korrekt dargestellt ist. Die Zeilen unten im Bild beschreiben die Anleitung des Simulators zum Auffinden der Ebene.

Weiterhin verfügt das Modul des Zweittrimesterscreenings über vier Cases. Diese Fälle bestehen aus einer kurzen Anamnese, von denen einer (Case 4) eine physiologische Schwangerschaft simuliert, die anderen Schwangerschaften weisen pathologische Aspekte auf. Der simulierte Fetus bewegt sich in jedem Case spontan hin und her, hier ist es vollkommen zufällig in welchen Zeitintervallen er sich bewegt und welche Position er danach einnimmt. Im Case der physiologischen Schwangerschaft präsentiert sich der simulierte Fetus stets in Beckenendlage, wobei die Positionen variieren. In den pathologischen Cases existieren sowohl Beckenend- als auch Schädellagen des Fetus. Für die Durchführung der einzelnen Fälle stehen einige der Hilfestellungen wie im Task zur Verfügung: das Fetus-Modell in Relation zum Schallkopf, der Mannequin mit Schallkopfbewegungen und die Skizzen der einzelnen Standardebenen. Im Unterschied zum Task-Modus sind diese Hilfestellungen hier allerdings fakultativ und jederzeit deaktivierbar. Der größte Unterschied zwischen der Durchführung eines Tasks und der Durchführung eines Case besteht jedoch darin, dass es beim „Freeze“ eines Bildes kein direktes Simulator-Feedback gibt. Die Aufgabe im Case besteht darin an einem sich bewegenden Fetus die Standardebenen aufzusuchen und selbst zu entscheiden, ob diese richtig dargestellt ist. Bilder können folglich zu jedem Zeitpunkt „eingefroren“ werden. Weiterhin besteht im Case-Modus keine Auflistung der Standardebenen in der Spalte am linken Bildschirmrand und keine Anleitung zum Auffinden der Ebenen. Stattdessen werden im Anschluss an den Fall die gesicherten Bilder ihren jeweiligen Standardebenen zugeordnet. Während der Untersuchung konnte man sich die bereits gesicherten Bilder anschauen, z.B. um zu überprüfen, welche Ebene bereits dargestellt und abgespeichert wurde. Die Hilfestellungen können unabhängig voneinander deaktiviert werden. So kann eine Überprüfung der praktischen Fähigkeiten ohne Verfälschung durch Simulator-bedingte Hilfestellungen gewährleistet werden. Außerdem besteht im Case-Modus die Möglichkeit sonographische Messungen, z. B. zur Fetometrie durchzuführen.

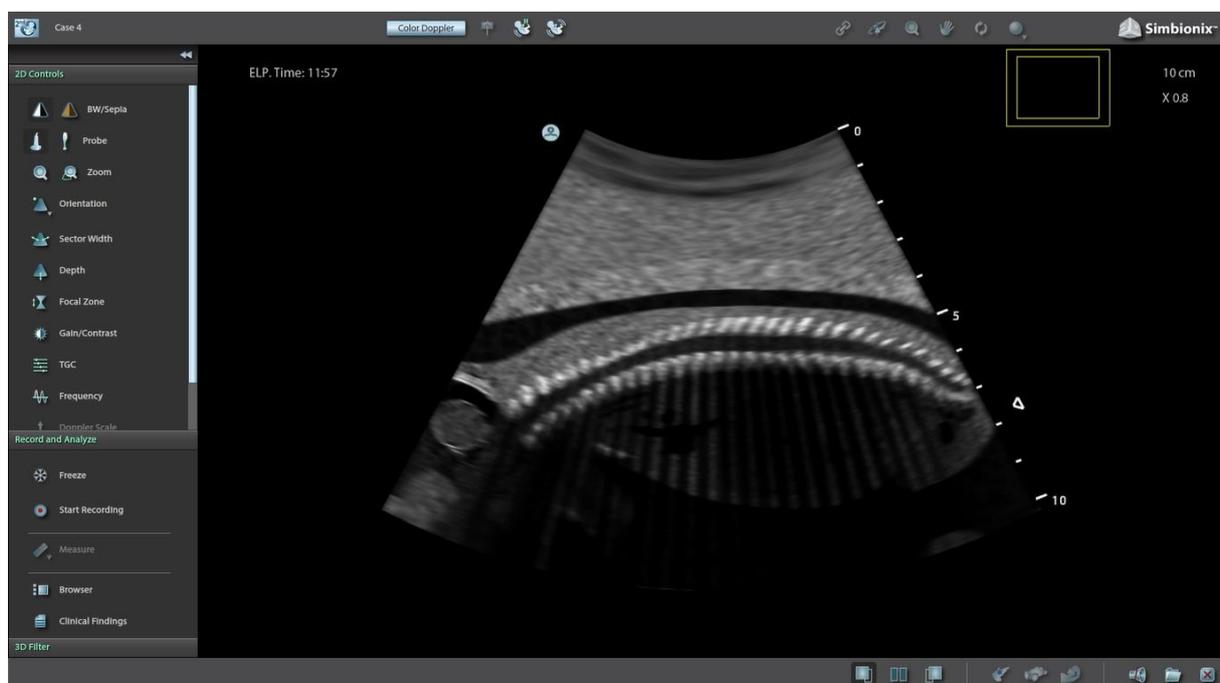


Abbildung 4: Case-Modus des Simulators bei deaktivierter Hilfestellung. Abbildung modifiziert nach [18]

Nach Abschluss eines Tasks oder Cases öffnet das Programm beim Schließen der Module automatisch eine Auswertung. Die Auswertung zeigt beim Task die Zeit, in der die Aufgabe ausgeführt wurde und vergleicht diese mit dem letzten Versuch. In den Cases stellt es sich komplexer dar. Da die Bilder dort flexibel, ohne vorheriges Auswählen aus einer Liste, „eingefroren“ und gesichert werden können, werden sie durch den Simulator automatisch der jeweiligen Standardebene zugeordnet. Diese können in der Auswertung durchgesehen und dahingehend kontrolliert werden, welcher Standardebene die Bilder durch den Simulator zugeordnet sind und ob es Bilder gibt, die keiner Standardebene zugeordnet werden können. Außerdem zeigt das Programm für die jeweilige Standardebene einen Goldstandard, mit dem das eigene Bild verglichen werden kann.



Abbildung 5: Auswertung des Case-Modus für die Abdomenstandardebene. Das linke Bild wurde durch die Teilnehmerin gemacht, das rechte Bild zeigt den Goldstandard des Simulators. Unter den Bildern werden sowohl abgebildeten Strukturen aufgelistet als auch eine Angabe der für den Goldstandard fehlenden Strukturen.

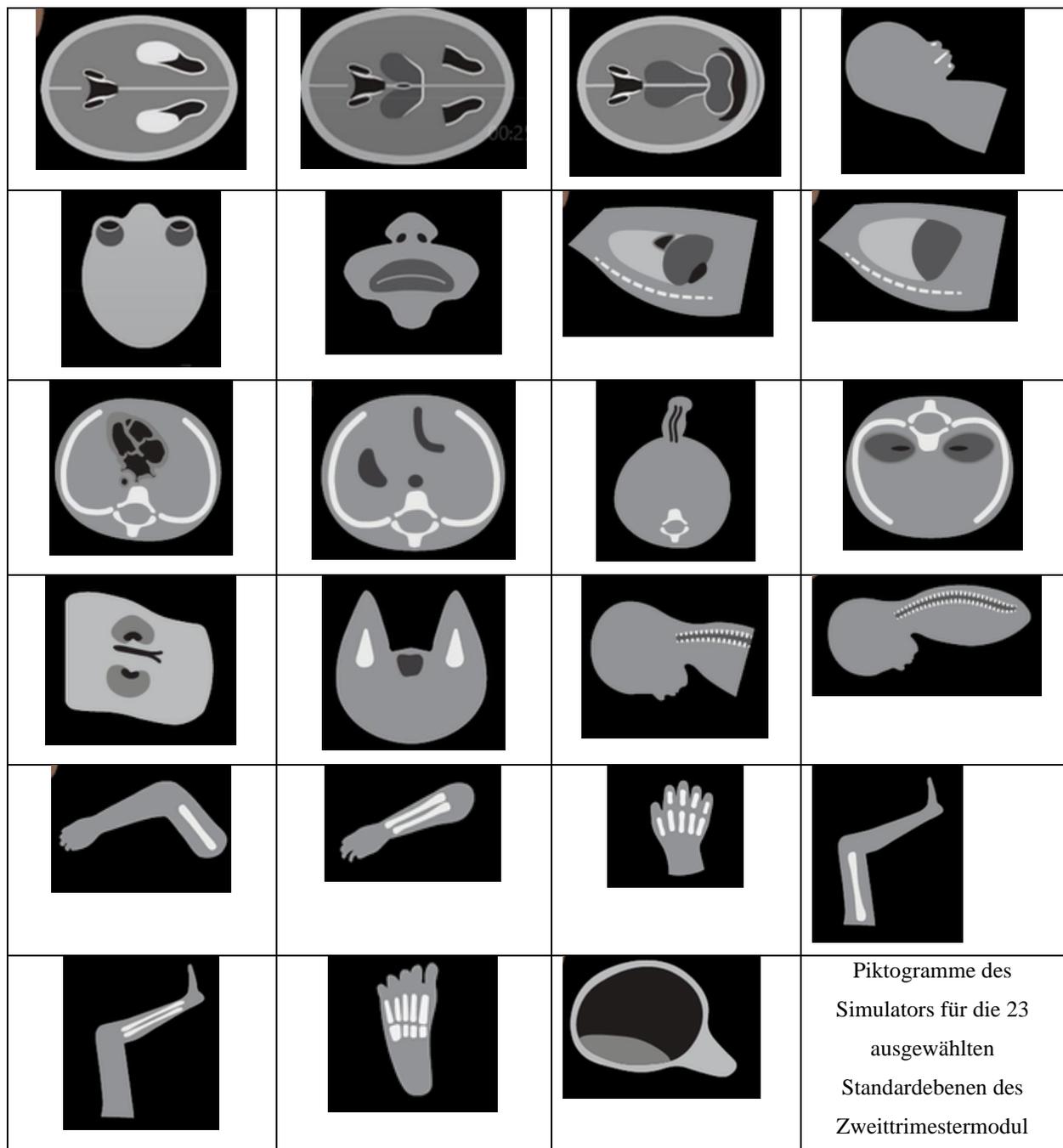
3.4. Ausgewählte Standardebenen

Vor Beginn der Studie testeten mein Co-Doktorand (Thema Fetale Echokardiografie) und ich das Zweittrimestermodul des U/S Mentor. Da wir beide keine Vorerfahrung im gynäkologisch-geburtshilflichen Ultraschall hatten und auch sonst alle Einschlusskriterien für die Interventionsgruppe auf uns zutrafen, waren die Ausgangsbedingungen identisch zu denen der Trainees. Das Zeittrimestermodul des U/S Mentor enthält 35 Standardebenen. Während der Testung des Moduls war das Aufsuchen vieler Standardebenen frustan. Ohne Vorerfahrung und Einführung dauerte es sehr lange bis Ebenen korrekt aufgefunden waren und teilweise war die richtige Darstellung während der ersten Übungen am U/S Mentor gar nicht möglich. Da vor Beginn der

Studie unklar war, in welchem Maße sich die Trainees hinsichtlich der Geschwindigkeit im Ablauf der Untersuchung verbessern würden, wurde – in Zusammenschau mit der eigenen Erfahrung aus den Übungen am Simulator – die Anzahl der eingeschlossenen Ebenen auf 23 festgelegt. Diese sind in Tabelle 2 aufgeführt. Die Extremitätenebenen kürzten wir um die Hälfte der angedachten zwölf Ebenen auf sechs, indem festgelegt wurde, dass die Teilnehmenden nur jeweils eine Extremität darstellen und „einfrieren“ müssen. Ob sie die linke oder rechte Seite auswählten, blieb den Teilnehmenden selbst überlassen. Eine Mischung aus beiden Seiten, z.B. rechtem Humerus und linker Hand, war ebenfalls zulässig. Die Intention dahinter war, dass die Trainees die Extremitätenebenen aufzusuchen lernen, diese überprüft wurden und gleichzeitig Zeit in den Tests eingespart wurde. Außerdem wurde die Genitalebene herausgenommen, da diese kein obligater Bestandteil der DEGUM-Stufe-II-Diagnostik ist. Schließlich entschieden wir uns noch dazu, die Herzebenen lediglich auf den Vier-Kammer-Blick zu reduzieren. Dies war einerseits dadurch begründet, dass der Simulator über ein eigenes Modul für die fetale Echokardiographie mit deutlich besserer Auflösung des Herzens verfügt, und andererseits dadurch, dass eine andere Trainee-Gruppe in einer parallelaufenden Studie die fetale Echokardiographie gesondert und differenzierter übte, sodass der Simulator hinsichtlich dieser Ebenen separat untersucht wurde.

Region	Ebene	Region	Ebene
Gehirn:	Transventrikulär	Abdomen:	Abdomen mit Magen
	Transthalamisch		Nabelschnurinsertion
	Transzerebellär		Nieren sagittal
Gesicht:	Profil mit Nasenbein		Nieren transversal
	Koronares Gesicht		Blase transversal
	Orbitae	Extremitäten:	Humerus
Wirbelsäule:	Nacken sagittal		Radius + Ulna
	Wirbelsäule sagittal		Hand
Thorax:	Vier-Kammerblick		Femur
	Zwerchfell und Lunge links		Tibia + Fibula
	Zwerchfell und Lunge rechts		Fuß
		Plazenta:	Plazenta

Tabelle 2: Ausgewählte Standardebenen zur Übung und Testung des Zweittrimesterscreenings



Ultraschalluntersuchungen nur minimale Vorerfahrung hatten. Deshalb galten folgende Kriterien zum Einschluss von Studierenden in die Interventionsgruppe:

- a. Immatrikulation als Medizinstudent*in in Deutschland
- b. Erfolgreicher Abschluss des Anatomiekurses
- c. <10h Vorerfahrung im pränatal-diagnostischen Ultraschall
- d. 2-3h Zeitkapazität pro Woche über den gesamten Studienzeitraum für SIM-UT und Tests

Mit interessierten Studierenden, die auf die Einladungsmail antworteten, wurde in einem Telefongespräch der Ablauf der Studie näher erläutert und erfragt, ob alle Einschlusskriterien zuträfen. Die Interventionsgruppe bestand aus 11 Medizinstudierenden im 2.-6. Studienjahr. Des Weiteren gab es zwei Vergleichsgruppen mit unterschiedlichem Kompetenzniveaus. Die erste Vergleichsgruppe, Referenzgruppe A, setzte sich aus 10 Assistenz- und Fachärzt*innen mit unterschiedlich langer Arbeitserfahrung aus der Klinik für Gynäkologie und Geburtshilfe des Teams des St-Franziskus-Hospitals Münster zusammen. Dafür luden wir alle Ärzt*innen der Abteilung ein, an der Studie teilzunehmen. Interessierte Ärzt*innen konnten anschließend persönlich oder per E-Mail einen Termin für ihre Testungen vereinbaren. Die geburtshilfliche Abteilung des Hospitals ist ein Perinatalzentrum-Level I. Alle Ärzt*innen hatten zum Zeitpunkt der Studie praktische Erfahrung sowohl in gynäkologischen als auch in geburtshilflichen Ultraschalluntersuchungen. Eine teilnehmende Ärztin war an einer anderen gynäkologischen Klinik mit perinatalem Schwerpunkt, dem Clemens-Hospital in Münster, angestellt. Die zweite Gruppe, Referenzgruppe B, fungierte als Expert*innengruppe. Sie zählte 10 hochspezialisierte mindestens DEGUM-II-zertifizierte Fachärzt*innen für Gynäkologie. Diese wurden durch den Studienleiter eingeladen an der Studie teilzunehmen, Terminvereinbarungen erfolgten anschließend über uns.

3.6. Ablauf, Training, Tests

Die Studie gliedert sich in drei Phasen: Einführung, Training und Tests. Am 30.08.2020 fand eine Einführungsveranstaltung für die Studierenden statt. Ziel war es, einen Überblick über Richtlinien zu ultraschallgestützten Schwangerschaftsvorsorgeuntersuchungen in Deutschland zu vermitteln. Dabei wurden sowohl ausgewählte Studien zur pränatalen Detektion fetaler Fehlbildungen präsentiert als auch eine Studie zu simulationsbasiertem pränatal-diagnostischen Ultraschalltraining. Am Folgetag startete die Trainingsphase, für die sich die Studierenden flexibel über einen Online-Terminplaner (Doodle) in Trainingsslots über den Tag verteilt anmelden konnten. Das SIM-UT wurde als regelmäßiges selbstständiges Einzeltraining über eine Stunde zwei Mal wöchentlich durchgeführt. Lediglich in der ersten Woche wurden die Studierenden von uns begleitet, um ihnen den Simulator, die Programme und Standardebenen zu erklären. Zu diesem Zweck umfasste der erste

Trainingstermin 1,5h, alle folgenden Termine umfassten 1h. Der Zugang zum Simulator-Raum war für die Trainees über die Pforte der Klinik frei zugänglich.

Da das SIM-UT eigenständig durchgeführt wurde und die Trainees sich unterschiedlich schnell in die Module der Pränatalsonographie einfanden, wurde das Training flexibel gestaltet. Grobe Vorgabe war in den ersten zwei Wochen intensiv die Standardebenen mithilfe der beiden Task-Module zu üben und gleichzeitig pro Trainingssession auch mindestens einmal das Case 4-Modul zu üben. Auf diese Weise sollten die Trainees die Anforderungen der Standardebenen des Zweittrimesterscreenings gut kennenlernen und sich in der Pränatalsonographie, insbesondere mit dem Umstand eines sich bewegenden und die Position wechselnden Objektes, dem Fetus, vertraut machen. Mit zunehmender Übung hatten die Trainees die Möglichkeit zum Schallen der Fälle mit fetalen Fehlbildungen. Zusätzlich wurden sie dazu angehalten so früh wie möglich die Hilfestellungen im Case-Modus auszuschalten und ohne Hilfestellung zu schallen.

Alle zwei Wochen wurde der Lernfortschritt der Interventionsgruppe durch standardisierte Tests überprüft. Dafür sollten sowohl im Task 1 am unbewegten Fetus als auch im Case 4 am sich zufällig bewegenden Fetus innerhalb einer Stunde alle ausgewählten 23 Standardebenen aufgefunden und „eingefroren“ werden. Die Zielsetzung, die Standardebenen so genau wie möglich in möglichst kurzer Zeit einzustellen, wurde den Studierenden vor der Prüfung mitgeteilt. Die Reihenfolge der Ebenen spielte dabei keine Rolle. Ein Test vor Beginn des SIM-UT, im Sinne eines Pre-Tests, fand aufgrund der Erfahrungen der Doktorand*innen in der Testung des Zweittrimestermoduls vor Beginn der Studie nicht statt. Die Durchführung des Tests wäre ohne Vorerfahrung im pränatalen Ultraschall und SIM-UT nicht praktikabel.

Derselbe Test wurde auch für die Erhebung von Geschwindigkeit und Genauigkeit der Referenzgruppen eingesetzt. Diese führten den Test einmalig im Studienverlauf durch. Die Testtermine wurden individuell vereinbart. Vor der jeweiligen Testung gaben wir allen teilnehmenden Ärzt*innen eine 30-minütige Einführung in die Bedienung des Simulators und die zu schallenden Module. Während der Einführung stand es den Teilnehmenden frei, Fragen zu stellen und alle vorgeschriebenen Standardebenen auszutesten. Alle Tests wurden von mir oder dem Co-Doktoranden der Partnerstudie durchgeführt und aufgezeichnet. Die Aufzeichnung erfolgte zur Absicherung vor Programmausfällen doppelt. Der U/S Mentor speichert alle Untersuchungen automatisch in einem Video ab. Da es teilweise zu Abstürzen des Programms kam, wurden die Test-Performanzen mit dem Screen Recorder Elgato Game Capture HD60 S+ aufgenommen. Gleichzeitig wurde Protokoll geführt, zu welchem Zeitpunkt welche Ebene „eingefroren“ wurde, damit im Falle eines uneindeutigen Bildes, die Zuordnung zur Standardebene anhand des Protokolls möglich war.

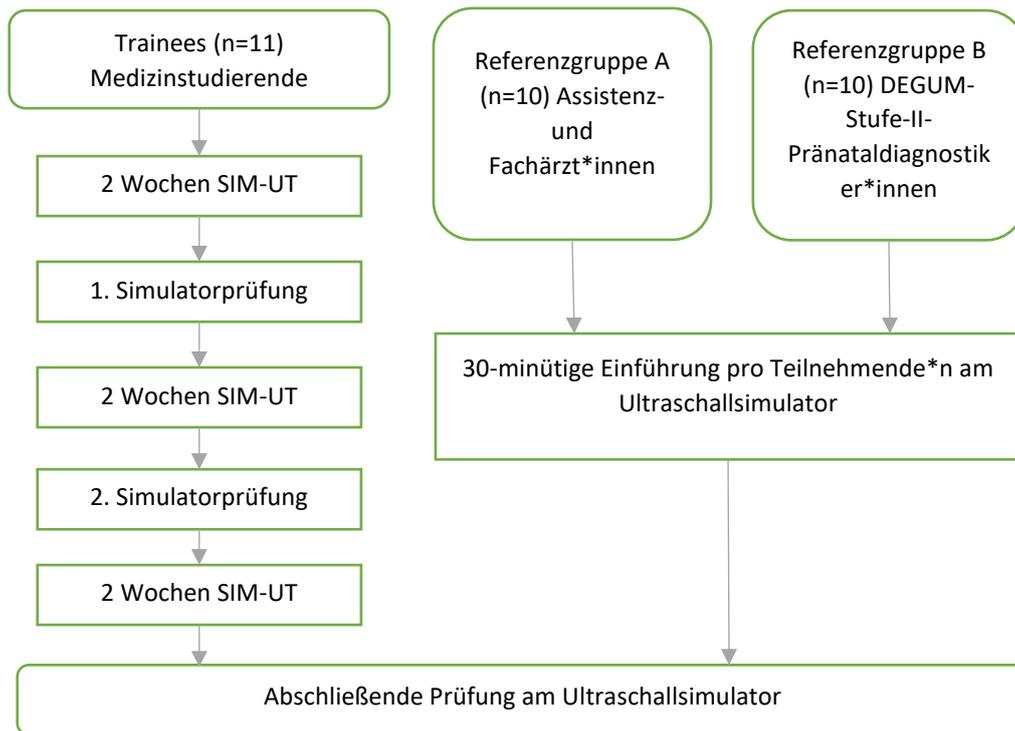


Abbildung 7: Flussdiagramm über Ablauf von Training und Testungen

3.7. Testauswertung

Die Performanz aller Teilnehmenden wurde anhand desselben standardisierten Tests erhoben. Dabei unterschied sich die Bewertung des Task 1 von der des Case 4. Die Kriterien, nach denen die Untersuchung am Simulator durchgeführt werden sollte, waren (a) die Ebene korrekt darzustellen und „einzufrieren“ und (b) die Geschwindigkeit in der dies durchgeführt wurde.

3.7.1. Auswertung nach der Zeit

Die Zeit als Parameter für die Überprüfung der Kompetenz ist im pränataldiagnostischen Ultraschall nicht ungewöhnlich und eine signifikant kürzere Untersuchungsdauer ist ein Merkmal für erfahrene Untersucher*innen [7,8,11,14,31]. Die Geschwindigkeit im Aufsuchen der Ebenen dient damit als Indikator sowohl für die Kenntnis der aufzusuchenden Ebene als auch für eine gute Orientierung im pränatalen Ultraschall. Die Geschwindigkeit spielt zudem eine zentrale Rolle in der Beurteilung von Effizienz und Effektivität.

Um die benötigte Zeit der Teilnehmenden in der Darstellung der Ebenen zu ermitteln, wurde das gesamte Videomaterial sekundengenau ausgewertet. Die Gesamtzeit, die alle Teilnehmenden zum Darstellen aller 23

Standardebenen benötigten, wurde als Total Time to Completion (TTC) definiert und mit Microsoft Excel ermittelt. Die Zeit wurde auf die Sekunde gestoppt: Die Zählung begann mit dem Beginn des Schallens und endete jeweils mit dem „Freeze“ des Bildes, ausgelöst durch das Fußpedal. Im Task-Modus gab es die eindeutige Zuordnung zur Standardebene, indem die Standardebene auf der Liste als erledigt gekennzeichnet wurde. Im Case-Modus war der Zeitpunkt des „Freeze“ in der Videoaufzeichnung durch einen Ton und den Schriftzug „Image saved“ kenntlich. Wurde eine Ebene mehrmals dargestellt und gespeichert, wurden alle für die Ebene aufgewandten Zeitintervalle summiert. Start- und „Freeze“-Zeitpunkt wurden sekundengenau in eine Exceltabelle eingetragen und dort in die entsprechende Dauer zur Darstellung einer Ebene ausgerechnet. Die Zeit, in der ein gesichertes Bild betrachtet wurde bzw. während der Bildschirm im „Freeze“-Modus war, wurde nicht gewertet.

3.7.2. Auswertung nach Korrektheit in der Ebenendarstellung

Die Korrektheit der Ebene wurde durch die im Modul des Simulators integrierte Auswertung bestimmt. Bei der Auswertung listet der Simulator auf, welche Strukturen für eine korrekte Standardebene abgebildet sein sollen. Fehlte eine Struktur, so wurde diese in der Auswertung angegeben. Da der Simulator allerdings teilweise vermeintlich richtige Ebenen nicht zuordnete, wurde eine Zweitbewertung für alle vom Simulator abgelehnten Ebenen vorgenommen.

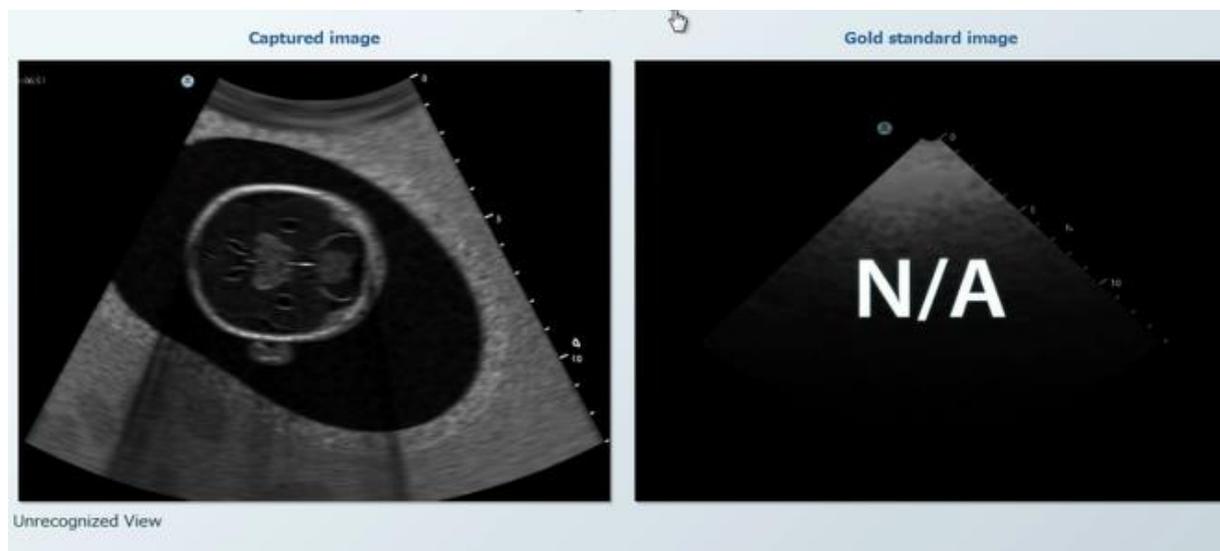


Abbildung 8: Eine im Case-Modus generierte Kleinhirnebene, die der Simulator keiner Standardebene zuordnete. Abbildung modifiziert nach [18]

Im Task-Modus des Simulators, in dem der Fetus in einer Position unbewegt verbleibt, ist die Funktion „Freeze“ nur möglich, wenn die Ebene nach Simulatorbewertung korrekt eingestellt ist (siehe 3.3). Im Case-Modus, in dem die Teilnehmenden frei untersuchen und die Ebenen darstellen sollten, war die direkte Feedback-Funktion deaktiviert. Ebenen, die unklar eingestellt oder gespeichert wurden, wurden nicht erkannt. Da die fehlende Zuordnung zu Standardebenen gehäuft bei augenscheinlich fehlerfreien Ebenen auftrat, entschieden wir, diese abgelehnten Ebenen von einem unabhängigen Expertengremium verblindet bewerten zu lassen. Das Gremium bestand aus drei DEGUM-II-zertifizierten Fachärzten der Gynäkologie, die zusätzlich alle Mitglieder der Qualitätssicherungskommission für gynäkologisch-geburtshilflichen Ultraschall der Kassenärztlichen Vereinigung Westfalen-Lippe sind. Alle drei Gremium-Mitglieder bekamen eine verblindete Version der vom Simulator abgelehnten Ebenen, das bedeutet, dass bei der zu bewertenden Ebene weder Name noch Gruppenzugehörigkeit der untersuchenden Person kenntlich war. Diese sollten sie analog zum deutschen Schulnotensystem bewerten („1 - sehr gut“ bis „6 - mangelhaft“). Vorab wurden dem Gremium die 23 Standardebenen des Simulators erläutert und auf Eigenheiten des Simulators hingewiesen, die wir im Laufe der Studie in der Simulatorbewertung beobachtet hatten. Dies waren z. B., dass die Kleinhirnebene so angeschnitten werden muss, dass Thalamus und Cavum septum pellucidum ebenfalls mit auf dem Bild sind oder dass z.B. in die Orbitalebene so transversal dargestellt werden muss, dass das Kleinhirn okzipital ebenfalls dargestellt ist. Die Bewertung der Ebenen wurde gemäß dem Standard der Qualitätssicherungskommission der KV Westfalen-Lippe durchgeführt, welche sich am DEGUM-Standard orientiert. Abschließend wurde der Durchschnitt der drei Noten ermittelt und alle bis „4 - ausreichend“ bewerteten Ebenen gingen als korrekt dargestellt in die Endwertung mit ein.

3.8. Fragebogen zur Ebenen-Evaluation und Feedback

Alle Teilnehmenden erhielten einen Fragebogen zur subjektiven Bewertung der Schwierigkeit die verschiedenen Ebenen in der Untersuchung darzustellen. Auf einer 10 cm langen Skala von „sehr leicht“ bis „sehr schwierig“ waren die Ebenen zu evaluieren. Der Fragebogen unterschied sich in einer Version für Studierende und einer für Ärzt*innen. Es wurde 21 der 23 Ebenen des angewandten Screenings abgefragt. Die beiden Zwerchfellebenen wurden für die Schwierigkeitsevaluation als eine Zwerchfellebene zusammen erfasst und die Ebene des Vierkammerblicks wurde in der parallellaufenden Studie zum fetalen Echo evaluiert.

Die Evaluationsbögen wurden nach Abschluss des SIM-UT an die Interventionsgruppe ausgeteilt. Die Referenzgruppen erhielten die Evaluationsbögen am Tag ihres Tests am Simulator. Falls einzelne Ebenen noch nie vor der Testung am Simulator eingestellt wurden, hatten die Teilnehmenden die Möglichkeit die Ebene in der Evaluation zu streichen. Mit der Befragung sollte untersucht werden, ob bestimmte Ebenen von einer Mehrheit als schwierig oder einfach darzustellen bewertet werden. Zur Auswertung wurde die angegebene Schwierigkeit millimetergenau ausgemessen und in Microsoft Excel zusammengeführt.

Der Fragebogen für Studierende enthielt nur Fragen zur Schwierigkeitsbewertung für die Darstellung der Ebenen, der nach sechswöchigem SIM-UT auszufüllen war. Die Version für Ärzt*innen fragte zusätzlich nach Eckdaten der ärztlichen Ausbildung, wie z.B. Jahr der Approbation und ggf. Facharztprüfung oder DEGUM-Level. Zusätzlich sollten die Ärzt*innen ihre Ultraschallerfahrung quantifizieren, insbesondere wie viele pränatal-diagnostische Ultraschalluntersuchungen sie bereits durchgeführt haben (siehe Anhang 1).

Nach Ablauf der Studie sollten die Studierenden außerdem Feedback zur Studienteilnahme geben (siehe Anhang 2). Mittels offener Fragen wurde nach ihrer Erfahrung mit dem Simulator gefragt und z.B. danach, ob sie ein simulationsbasiertes Ultraschallwahlfach an ihrer Universität belegen würden, wenn dies in dieser Form angeboten würde.

3.9. Statistische Analyse

Die Performanz der Teilnehmenden wurde für beide Aufgaben des Tests, dem Task- und dem Case-Modus, nach zwei Variablen bewertet. Einerseits nach der zeitlichen Dauer, die zum Darstellen und „Freeze“ einer Ebene benötigt wurde und andererseits nach der Korrektheit der Ebene, d.h. ob sie alle nötigen Strukturen der Standardebene korrekt abbildet.

Die Total Time to Completion (TTC) wurde aus den einzelnen Zeitintervallen zum Darstellen aller 23 Standardebenen in Microsoft Excel berechnet. Die Prozentzahl korrekt dargestellter Ebenen wurde als Rate of Appropriate Images (RAI) definiert. Für die statistische Analyse wurde IBM SPSS Statistics 27 verwendet. Statistische Signifikanz wurde ab einem p-Wert $<0,05$ gewertet. Für den Vergleich der RAI und TTC der Interventionsgruppe über die Zeit wurden paarige, nicht-parametrische Tests (Friedman und Wilcoxon) angewandt. Die Performanz- und Schwierigkeitsbewertungen der Gruppen untereinander wurden mit nicht-parametrischen, unpaarigen Tests (Kruskal-Wallis und Mann-Whitney U) verglichen. Ebenso die Analyse, ob innerhalb einer Gruppe die Teilnehmenden für eine Ebene signifikant mehr Zeit benötigten als für eine andere Ebene.

Korrelationen von Schwierigkeitsbewertung der Ebenen und Testperformanz wurden mit dem Korrelationskoeffizient nach Pearson ermittelt.

4. Ergebnisse

In die Interventionsgruppe der Studie wurden 11 Medizinstudierende als Trainees eingeschlossen. Die Ergebnisse aus den zweiwöchentlichen, standardisierten Tests wurden mit denen der beiden Referenzgruppen à 10 Ärzt*innen verglichen. Es kam weder in der Interventionsgruppe noch in den Referenzgruppen zu einem drop-out von Teilnehmer*innen.

Alle Teilnehmenden nahmen an der Evaluation über die subjektive Schwierigkeit der Ebenen teil. Die Referenzgruppen gaben in einem weiteren Fragebogen Eckdaten ihrer ärztlichen Laufbahn, sowie Auskunft über Erfahrung in der Sonographie mit ggf. Zertifizierungen an. Zusätzlich evaluierten die Studierenden in einem Feedbackbogen ihre Erfahrung und Zufriedenheit mit dem absolvierten SIM-UT.

4.1. Vorerfahrung im gynäkologisch-geburtshilflichen Ultraschall

Die Interventionsgruppe bestand aus 11 Medizinstudierenden, die an der Westfälischen-Wilhelms-Universität Münster eingeschrieben waren. Die Einschlusskriterien für die Trainees der Interventionsgruppe war die minimale Vorerfahrung in gynäkologisch-geburtshilflichen Ultraschall von <10h. Alle Trainees nahmen über den gesamten Studienzeitraum bis zum Abschluss an der Studie teil.

Die Referenzgruppe A setzte sich aus acht Assistenzärzt*innen und zwei Fachärzt*innen zusammen, die im Median 750 geburtshilflich-gynäkologische Sonographien durchgeführt hatten. Die Gruppe war mit Ausnahme einer Assistenzärztin in einem Level-I-Perinatalzentrum tätig. Die Ausnahme bildete eine Assistenzärztin aus der gynäkologisch-geburtshilflichen Abteilung mit perinatalem Schwerpunkt des Clemenshospitals in Münster. Unter den zwei Fachärzt*innen war einer DEGUM-Stufe-I-zertifiziert. In der Expert*innengruppe waren alle Teilnehmenden Fachärzt*innen für Frauenheilkunde und Geburtshilfe und mindestens DEGUM-Stufe-II-zertifiziert. Der Median der durchgeführten geburtshilflich-gynäkologischen Sonographien betrug für diese Gruppe 15.000. Die Vorerfahrung der Referenzgruppen divergierte wie in Tabelle 3 dargestellt:

	Fachärzt*innen	Assistenzärzt*innen	Median der durchgeführten geb.-gyn. Sonographien	Mindestanzahl geb.-gyn. Sonographien	Maximale Anzahl geb.-gyn. Sonographien	Median DEGUM-II-Untersuchungen	Gesamtzahl geb.-gyn. Sonographien
Referenzgruppe A, n=10	2	8	750	130	40.000	0	48.130
Referenzgruppe B, n=10	10	0	15.000	8.000	80.000	6.000	264.000

Tabelle 3: Vorerfahrung im geburtshilflich-gynäkologischen (geb.-gyn.) Ultraschall der Referenzgruppen A (Assistenz- und Fachärzt*innen) und B (DEGUM-Expert*innen)

4.2. Ergebnisse der Testungen am Simulator im Task-Modul

Im Task-Modus gibt der Simulator bei der Sicherung eines Bildes bzw. einer Ebene ein direktes Feedback und ein „Einfrieren“ eines Bildes ist nur möglich, wenn die Ebene vom Simulator als richtig erkannt wird. Dementsprechend war die Rate der korrekt eingestellten Ebenen unter den gesicherten Ebenen bei 100%. Für Untersuchungen, die im Task-Modus nicht innerhalb der vorgegebenen 30 Minuten abgeschlossen wurden, gingen Ebenen als fehlend in die Wertung mit ein. Das war bei zwei Trainees in der ersten Prüfung nach zwei Wochen SIM-UT und einer Ärztin der Referenzgruppe A der Fall (siehe Tab. 4). Den Trainees fehlten zusammen vier, der Ärztin 2 Ebenen:

Gruppe	Vollständig abgeschlossene Task 1-Untersuchungen	Anzahl fehlender Ebenen für die Gruppe insgesamt
Prüfung 1 Trainees n=11	9 (81,81%)	4
Prüfung 2 Trainees n=11	11 (100%)	0
Prüfung 3 Trainees n=11	11 (100%)	0
Referenzgruppe A n=10	9 (90%)	2
Referenzgruppe B n=10	10 (100%)	0

Tabelle 4: Vollständig abgelegte Prüfung für die Task 1-Untersuchung der Gruppen im Vergleich

4.3. Ergebnisse der Testungen am Simulator im Case-Modul

Die Interventionsgruppe profitierte über den gesamten Studienzeitraum signifikant vom simulationsbasierten Training. Die Gesamtzeit zum Darstellen aller 23 Standardebenen (Total Time to Completion - TTC) verkürzte sich zwischen den Tests nach 2, 4 und 6 Wochen jeweils signifikant: Die TTC von 1238 ± 283 Sek. nach 2 Wochen, betrug nach 4 Wochen SIM-UT nur noch 868 ± 186 Sek. und hatte sich am Ende des Trainingszeitraums fast halbiert auf 620 ± 189 Sek. ($p < 0,001$). Folglich schallten die Trainees bereits nach 4 Wochen so schnell wie die Referenzgruppe A, der Assistenz- und Fachärzt*innen und hatte sie nach 6 Wochen SIM-UT signifikant überholt (620 ± 189 vs. 1036 ± 389 Sek., $p < 0,001$). Die Expert*innengruppe, Referenzgruppe B, hingegen erzielte eine signifikant schnellere TTC als die anderen Gruppen (443 ± 243 Sek., $p < 0,001$, Abb. 9).

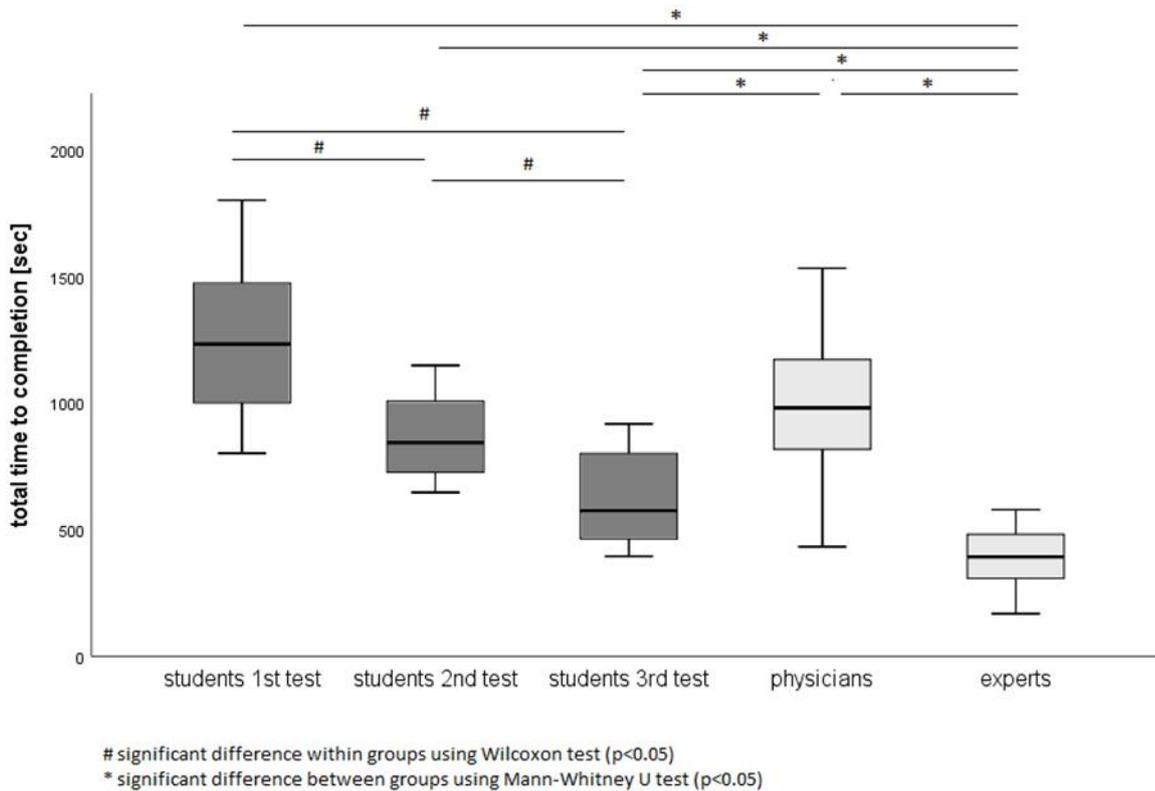


Abbildung 9: TTC aller 23 Ebenen der verschiedenen Gruppen mit Signifikanztests.

Legende: students 1st, 2nd und 3rd test: Tests der Interventionsgruppe nach 2, 4 und 6 Wochen. Physicians: Referenzgruppe A der Assistenz- und Fachärzt*innen, experts: Referenzgruppe B der DEGUM-Expert*innen.

Abbildung aus [18]

In der Betrachtung der TTC für die alleinige Darstellung eines biometrischen Basisschalls mit vier Standardebenen (Transthalämische Ebene für biparietalen Durchmesser und Kopfumfang, Abdomenebene für den Bauchumfang, Femurebene für die Femurlänge und die Darstellung der Plazenta), sind ähnliche Ergebnisse zu beobachten. Allerdings erreichten die Trainees kein signifikant schnelleres Ergebnis als die Referenzgruppe A in der Darstellung dieser vier, im klinischen Alltag häufig angewandten Standardebenen (78,18 Sek. vs. 135 Sek., $p=0,098$). Hier blieb die Expert*innengruppe signifikant schneller als die Trainees und Assistenz- und Fachärzt*innen (49,40 Sek., $p < 0,004$, vgl. Abbildung 10).

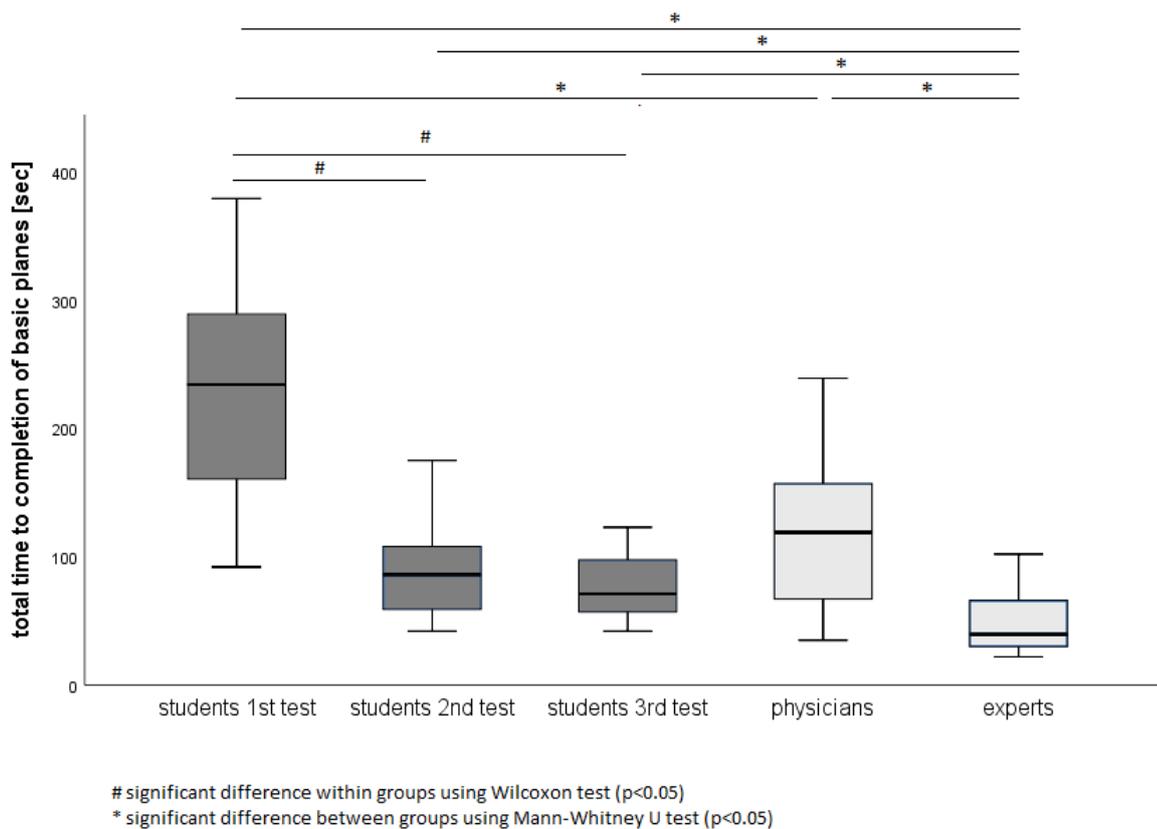


Abbildung 10: TTC der vier Basisebenen mit Signifikanztests. Legende: students 1st, 2nd und 3rd test: Tests der Interventionsgruppe nach 2, 4 und 6 Wochen. Physicians: Referenzgruppe A der Assistenz- und Fachärzt*innen, experts: Referenzgruppe B. Abbildung aus [18]

Im Hinblick auf die Rate korrekt dargestellter Ebenen erreichte die Interventionsgruppe bereits nach 2 Wochen eine RAI von >94%. Damit wies die Rate korrekt dargestellter Ebenen zwischen der Interventionsgruppe und den Assistenz- und Fachärztinnen der Referenzgruppe A bereits nach 2 Wochen SIM-UT keinen signifikanten Unterschied mehr auf und war für die weitere Studiendauer auf einem sehr ähnlichen Niveau 96,8% (Interventionsgruppe) vs. 96,5% (Referenzgruppe A). Die Referenzgruppe B der DEGUM-Expert*innen erreichte mit einer Rate von 100% korrekten Ebenen ein besseres Ergebnis als die beiden Vergleichsgruppen, wenngleich der Unterschied statistisch nicht signifikant war ($p=0,062$). Einen statistisch signifikant höheren Wert erzielte somit nur die Referenzgruppe B gegenüber der Interventionsgruppe nach 2 Wochen SIM-UT ($p < 0,008$). Insgesamt erreichten alle drei Gruppen eine hohe Rate an korrekt dargestellten Ebenen mit >96%.

4.4. Vergleich der Einzelebenen

4.4.1. Vergleich der Einzelebenen-Performanz zwischen den Gruppen

Nach Abschluss der sechs Wochen SIM-UT wiesen nur 8 der 23 Standardebenen signifikante Unterschiede im Aufsuchen der Einzelebene auf (Abbildung 11). Für vier der acht Ebenen erzielte die Interventionsgruppe einen signifikant schnelleren Wert als Referenzgruppe A der Assistenz- und Fachärzt*innengruppe. Die Trainees unterschieden sich nach Abschluss des Trainings im Aufsuchen von 20 der 23 Standardebenen nicht mehr signifikant von den DEGUM-II-Expert*innen der Referenzgruppe B. Zwei der drei Ebenen, in der die Expert*innengruppe signifikant schnellere Zeiten erzielte, gehören dem Basisultraschall an (Transthalamische- und Femur Ebene). Während die Interventionsgruppe nur für drei Ebenen signifikant mehr Zeit benötigte, war die Referenzgruppe A in sieben der 23 Standardebenen signifikant langsamer als Referenzgruppe B, die Expert*innengruppe.

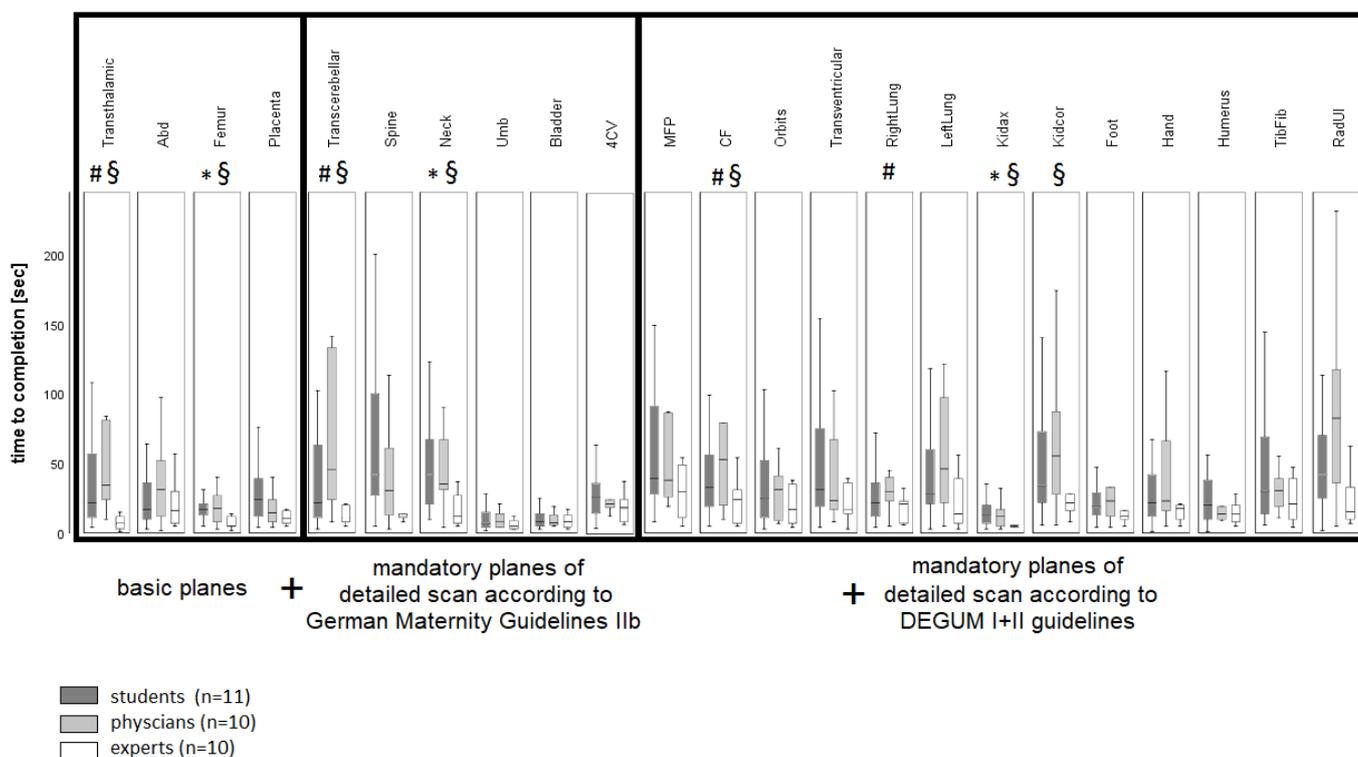


Abbildung 11: Boxplot der TTC für die Einzelebenen, aufgeteilt nach Untersuchungsprotokollen (Basisfetometrie, sonomorphologischer Schall nach MuRL I Ib, erweitertes Zweittrimesterscreening mit Ebenen der DEGUM I+II-Diagnostik)

Legende: #signifikanter Unterschied zwischen Interventionsgruppe (students) und Referenzgruppe A Assistenz- und Fachärzt*innen (physicians)

*signifikanter Unterschied zwischen Interventionsgruppe (students) und Referenzgruppe B DEGUM-II-Expert*innen (experts)

§signifikanter Unterschied zwischen Referenzgruppe A und B (experts und physicians)

Abkürzungen Standardebenen: Abd – Abdomen mit Magen, Umb – Nabelschnurinsertion, 4CV – Vierkammerblick, MFP – Gesichtsprofil mit Nasenbein, CF – koronares Gesicht, Right-/LeftLung – Diaphragma mir rechter/linker Lunge, Kidax – Nieren axial, Kidcor – Nieren koronar, TibFib – Tibia und Fibula, RadUI – Radius und Ulna

Abbildung aus [18]

4.4.2. Vergleich der Einzelebenen innerhalb einer Gruppe

Signifikante Unterschiede hinsichtlich der benötigten Zeit zur Standardebenendarstellung sind teilweise innerhalb einer Gruppe für verschiedene Ebenen zu beobachten. Es fällt auf, dass bei einem signifikanten Unterschied immer mindestens eine axiale Standardebene involviert ist, zumeist die der Nabelschnurinsertion oder der Harnblase. Zudem lässt sich beobachten, dass in der Interventionsgruppe mit dem jüngsten Erfahrungsalter am meisten signifikante Unterschiede in der Ebenendarstellung zu verzeichnen sind, während in der Expert*innengruppe nur vier Vergleiche einen signifikanten Unterschied in der benötigten Zeit ergeben. Gemessen an allen Standardebenen, ergaben sich für die einzelnen Gruppen folgende signifikante Unterschiede im Ebenenvergleich:

<u>Interventionsgruppe</u>			
Ebenen im Vergleich	Teststatistik [sec]	Std.-Fehler	Angepasste Signifikanz
Nabelschnurinsertion - Gesichtsprofil	132,864	31,187	0,005
Nabelschnurinsertion - Nieren koronar	130,227	31,187	0,008
Nabelschnurinsertion - Nacken	123,364	31,187	0,019
Nabelschnurinsertion - Wirbelsäule	122,818	31,187	0,021
Blase - Gesichtsprofil	-120,955	31,187	0,027
Nabelschnurinsertion - Radius, Ulna	120,909	31,187	0,027
Blase - Nieren koronar	-118,318	31,187	0,038
<u>Referenzgruppe A</u>			
Ebenen im Vergleich	Teststatistik [sec]	Std.-Fehler	Angepasste Signifikanz
Nabelschnurinsertion - Radius, Ulna	119,65	29,749	0,015
Nabelschnurinsertion - Nieren koronar	119,65	29,749	0,015
Blase - Nieren koronar	-114,7	29,749	0,029
Blase - Radius, Ulna	-114,7	29,749	0,029
Nabelschnurinsertion - Gesichtsprofil	112,6	29,749	0,039
Nabelschnurinsertion - koronares Gesicht	110,9	29,749	0,049
<u>Referenzgruppe B</u>			
Ebenen im Vergleich	Teststatistik [sec]	Std.-Fehler	Angepasste Signifikanz
Nieren axial - Nieren koronar	-120	29,731	0,014
Nieren axial - Gesichtsprofil	114,8	29,731	0,029
Nabelschnurinsertion - Nieren koronar	113,05	29,731	0,036
Nieren axial - transventrikuläre Ebene	-110,95	29,731	0,048

Tabelle 5: Signifikante Unterschiede zwischen den einzelnen Standardebenen in der time to completion innerhalb der Testgruppen

Signifikante Unterschiede treten folglich nur im Vergleich von axialen Rumpfebenen mit anderen Ebenen auf. Alle Vergleiche, bei denen ein Unterschied statistisch signifikant war, sind in Tabelle 5 aufgeführt. Zwischen dem Großteil der Ebenen besteht innerhalb einer Testgruppe kein signifikanter Unterschied in der benötigten Zeit zur Standardebenendarstellung.

4.5. Zweitbewertung durch Expertengremium

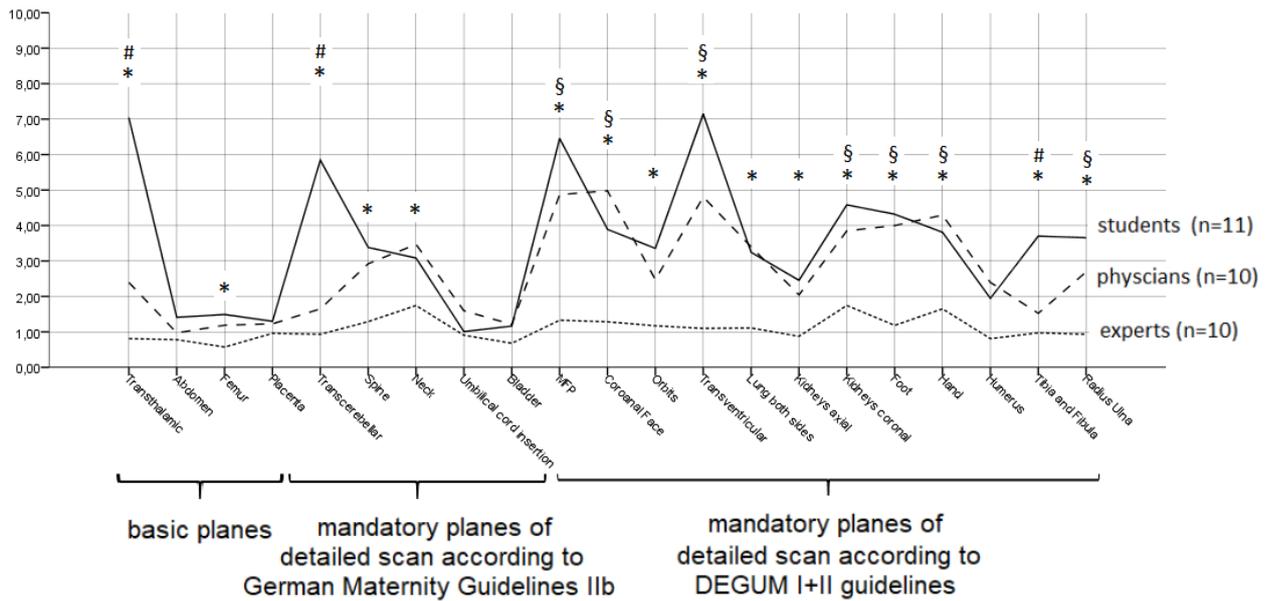
Die Ebenen, die von Teilnehmenden in den Prüfungen dargestellt und vom Simulator zu keiner Standardebene zugeordnet werden konnten, wurden von einem Expertengremium gemäß KV-Standard reevaluiert. Insgesamt wurden 268 Ebenen nach dem deutschen Schulnotensystem von jeweils allen drei Gremiumsmitgliedern einzeln bewertet. Die Re-Evaluation fand verblindet und in zufälliger Reihenfolge statt, sodass das Gremium weder zuordnen konnte von welchen Teilnehmenden die Ebenen gespeichert wurde noch welcher Gruppe sie angehörten. Aus den drei Einzelnoten wurde eine Durchschnittsnote gemittelt. Durchschnittsnoten von 1 bis 4,5 galten als bestanden und somit korrekt dargestellt. Durchschnittsnoten ab 4,5 galten als nicht bestanden und wurden somit als nicht korrekt dargestellt gewertet. Unter den 268 Ebenen, wurden 32, also 11,9%, als nicht korrekt bewertet.

4.6. Ebenen-Evaluation

Für die Analyse der Ultraschalluntersuchung interessierte uns gleichermaßen die subjektiv empfundene Schwierigkeit in der Darstellung der Standardebenen. Diese wurden im Evaluationsbogen jeweils auf einer 10 cm langen Skala von „sehr einfach“ bis „sehr schwierig“ bewertet (siehe Anhang 1). Die Studierenden evaluierten die Schwierigkeit der Standardebenen nach Abschluss des sechswöchigen SIM-UT, bezogen auf die praktischen Erfahrungen am Simulator. Die Ärzt*innen der Referenzgruppen sollten sich in der Evaluation auf die praktische Erfahrung im Ultraschall im klinischen Alltag beziehen und nicht auf die Testung am Simulator. Wenn Assistenzärzt*innen abgefragte Standardebenen außerhalb der Testung noch nicht angewandt hatten, strichen sie die betreffenden Ebenen auf dem Evaluationsbogen. Insgesamt wurde 19-mal eine Ebene gestrichen. Dies betraf nur Standardebenen erweiterter Untersuchungsprotokolle, wie z.B. die Gesichtsebenen oder die Darstellung fetaler Hand und Fuß.

4.6.1. Schwierigkeitsbewertung der Gruppen im Vergleich

Zwischen den Gruppen gab es signifikante Unterschiede in der Schwierigkeitsevaluation der Standardebenen. Es lässt sich eine nach Erfahrungsalter abfallende subjektiv empfundene Schwierigkeit in der Ebenendarstellung beobachten: Folglich bewerteten die Trainees mit minimaler Ultraschall-Vorerfahrung nach Abschluss des sechswöchigen SIM-UT die Darstellung im Mittel als schwieriger im Vergleich zu den Referenzgruppen.



#significant difference between students and physicians, *significant difference between students and experts, §significant difference between physicians and experts. Groups were compared by Mann-Whitney-U-test. A $p < 0.05$ was considered as significant

Abbildung 12: Schwierigkeitsbewertung der Ebenen. 0 = sehr einfach, 10 = sehr schwer. Mittelwerte der drei Gruppen.

students = Interventionsgruppe, physicians = Referenzgruppe A, experts = Referenzgruppe B

Abkürzungen Standardebenen: Abd – Abdomen mit Magen, Umb – Nabelschnurinsektion, 4CV – Vierkammerblick, MFP – Gesichtsprofil mit Nasenbein, CF – koronares Gesicht, Right-/LeftLung – Diaphragma mir rechter/linker Lunge, Kidax – Nieren axial, Kidcor – Nieren koronar, TibFib – Tibia und Fibula, RadUl – Radius und Ulna

Die Expert*innengruppe bewertete alle Ebenen als einfach darzustellen bei Mittelwerten zwischen 0,5 und 1,7 (siehe Tab. 6). Dadurch unterscheidet sich die Schwierigkeitseinschätzungen der Expert*innen für viele Ebenen signifikant von der der Interventionsgruppe, die 16 Ebenen als signifikant schwieriger bewertete. Referenzgruppe A bewertete ebenfalls sieben Ebenen als signifikant schwieriger als die Expert*innengruppe.

In Abbildung 12 ist gut zu erkennen, dass die 16 Ebenen, welche die Trainees signifikant schwieriger als die Referenzgruppe B evaluierten, sich sowohl aus Basisebenen zusammensetzen als auch aus spezifischeren Ebenen, die Teil der Feindiagnostik und des DEGUM-II-Standards sind. Dagegen finden sich die signifikanten Unterschiede in der Schwierigkeitsbewertung zwischen Referenzgruppe A und B nur für Ebenen, die über das Basisscreening und die Ebenen des Iib-Screenings nach Mutterschaftsrichtlinien hinausgehen. Abbildung 12 zeigt ebenfalls, dass die Interventionsgruppe und Referenzgruppe A ein ähnliches Verteilungsmuster für schwierige und leichte Ebenen vorweisen. Signifikante Unterschiede in den Schwierigkeitseinschätzungen der beiden Gruppen gibt es nur für drei der 21 bewerteten Ebenen (transthalamische-, transzerebelläre- und Tibia-Fibula-Ebene).

Die fünf Ebenen, für die kein signifikanter Unterschied zwischen Trainees und Expert*innen vorlag, waren drei axiale Standardebene (Abdomen-, Nabelschnurinsertion- und Harnblase-Ebene), sowie Plazenta- und Humerus-Standardebene.

Schwierigkeits- evaluation	Trainees		Referenz- Gruppe A		Referenz- Gruppe B	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
Transthalamisch	7,0455	1,57630	2,3944	2,53247	0,8100	1,14547
Abdomen	1,4091	1,30112	0,9722	0,87110	0,7810	1,11870
Femur	1,4909	1,01139	1,1889	1,02432	0,5730	0,83803
Plazenta	1,3000	1,35868	1,2278	0,96731	0,9600	1,00022
Transzerebellär	5,8455	2,67408	1,6500	1,44539	0,9333	1,13248
Wirbelsäule	3,3818	1,76228	2,9222	2,45956	1,2900	1,26377
Nacken	3,0818	1,54455	3,4688	2,38791	1,7450	1,83885
Nabelschnur- insertion	1,0091	0,62202	1,5944	1,25086	0,9050	1,24865
Blase	1,1636	2,05877	1,2111	0,98608	0,6800	1,11833
Gesichtsprofil	6,4545	2,96424	4,8714	1,39010	1,3300	1,65734
Koronares Gesicht	3,8909	2,31666	4,9786	1,51213	1,2850	1,49370
Orbitae	3,3545	2,16904	2,4714	1,63296	1,1722	1,21272
Transventrikulär	7,1455	1,77559	4,8063	3,03885	1,1000	1,38002
Lungen/Dia- phragma bds.	3,2444	1,80285	3,4000	2,25579	1,1110	1,09188
Nieren axial	2,4545	1,97148	2,0438	1,44505	0,8750	1,10183
Nieren koronar	4,5818	2,56664	3,8500	2,23788	1,7450	1,53992
Fuß	4,3182	2,41446	4,0000	2,25071	1,1850	1,05726
Hand	3,8091	2,13563	4,2929	2,50673	1,6500	1,71756
Humerus	1,9455	1,80353	2,3944	2,22380	0,8100	1,05483
Radius, Ulna	3,6545	1,84465	2,6938	1,44233	0,9330	1,18458
Tibia, Fibula	3,7000	1,73321	1,5250	1,11580	0,9730	1,16722

Tabelle 6: Mittelwerte der Schwierigkeitsbewertungen der drei Gruppen.

4.6.2. Schwierigkeitsbewertung der Ebenen im Vergleich

Eine Unterteilung der Ebenen nach Schwierigkeitsbewertung wurde in drei Kategorien vorgenommen: *einfach* – Bewertungen zwischen 0 und 1,9, *mittelschwer* – Bewertungen zwischen 2,0 und 4,9, sowie *schwer* – Bewertungen >5,0. Ausgehend von der Interventionsgruppe, welche als Trainees nach sechs Wochen SIM-UT die Ebenen bewerteten, ist die Unterteilung in die Kategorien wie folgt:

Schwierigkeitsbewertung der Interventionsgruppe	
Schwierige Ebenen	Transventrikuläre Ebene
	Transthalämische Ebene
	Gesichtsprofil
Mittelschwierige Ebenen	Transzerebelläre Ebene
	Nieren koronar
	Fuß
	Koronares Gesicht
	Hand
	Tibia und Fibula
	Radius und Ulna
	Wirbelsäule
	Orbitae
	Lungen/Diaphragma beidseits
Einfache Ebenen	Nacken
	Nieren axial
	Humerus
	Femur
	Abdomen
	Placenta
	Blase
Nabelschnurinsertion	

Tabelle 7: Einteilung der Ebenen in Kategorien nach Einschätzung der Interventionsgruppe

Nach Evaluation der Interventionsgruppe sind fünf Standardebenen einfach darzustellen. Für diese fünf Ebenen besteht kein signifikanter Unterschied zwischen der Evaluation der Trainees und der der Expert*innen aus Referenzgruppe B. Die Evaluation von Trainees und Referenzgruppe A ähnelt sich, insbesondere für die leicht darzustellenden Ebenen. Der Unterschied in der Bewertung besteht zwischen den beiden Gruppen hauptsächlich darin, dass die Ärzt*innen der Referenzgruppe A die durch die Trainees als schwierig evaluierten Ebenen, als mittelschwerig einschätzten bei Mittelwerten unter 5,0 (Transventrikuläre- transthalämische Ebene und Gesichtsprofil, s. Tabelle 6).

4.6.3. Korrelation zwischen Testperformanz und Schwierigkeitsbewertung

Alle drei Gruppen wiesen eine signifikante Korrelation zwischen ihrer Testperformanz, gemessen an der TTC, und der allgemeinen Schwierigkeitsbewertung der Standardebenen des Screenings auf:

Korrelationen				
Gruppe			Case	SCase
Trainees (Interventionsgruppe)	Case	Korrelation nach Pearson	1	,281**
		Signifikanz (2-seitig)		,000
		N	253	229
	SCase	Korrelation nach Pearson	,281**	1
		Signifikanz (2-seitig)	,000	
		N	229	229
Referenzgruppe A	Case	Korrelation nach Pearson	1	,260**
		Signifikanz (2-seitig)		,001
		N	230	172
	SCase	Korrelation nach Pearson	,260**	1
		Signifikanz (2-seitig)	,001	
		N	172	172
Referenzgruppe B	Case	Korrelation nach Pearson	1	,154*
		Signifikanz (2-seitig)		,026
		N	230	208
	SCase	Korrelation nach Pearson	,154*	1
		Signifikanz (2-seitig)	,026	
		N	208	208
**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.				
*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.				

Tabelle 8: Korrelation nach Pearson zwischen Testperformanz (Case) und der Schwierigkeitsbewertung (SCase).

4.7. Feedback zum Training am Simulator

Das Training am Simulator wurde von den Trainees sehr gut angenommen. Im Feedbackbogen, der mittels offener Fragen nach der subjektiven Erfahrung von SIM-UT im Rahmen der Studie fragte, gaben alle Trainees an, dass ihnen das SIM-UT gefallen habe. Die Frage, ob die Studierenden ein simulationsbasiertes Ultraschall-Wahlfach an ihrer Universität belegen würden, wenn diese ein anböte, beantworteten alle Trainees mit „Ja“.

5. Diskussion

In der Studie wurden die Lernkurven von Trainees mit minimaler Vorerfahrung im pränatalen Ultraschall anhand eines erweiterten Zweittrimesterscreenings an einem high-end Simulator untersucht. Die Lernfortschritte wurden nach 4h, 8h und 12h SIM-UT untersucht und mit der Performanz zweier Referenzgruppen am Ultraschallsimulator verglichen. Die erste Referenzgruppe A bestand aus 10 Assistenz- und Fachärzt*innen der Gynäkologie und Geburtshilfe mit unterschiedlich langer ärztlicher Tätigkeit, die im Median 750 gynäkologisch-geburtshilfliche Ultraschalluntersuchungen durchgeführten hatten. Die zweite Referenzgruppe B war die Expert*innengruppe mit 10 DEGUM-Stufe-II-zertifizierten Fachärzt*innen, der Median dieser Gruppe für durchgeführte gynäkologisch-geburtshilfliche Sonographien lag bei 15.000.

5.1. SIM-UT vs. Klinische Erfahrung

Die Traineegruppe, welche erst zu Anfang der Studie die Standardebene des erweiterten Zweittrimesterscreenings kennenlernte, verbesserte sich in ihrer Testperformanz am Simulator zwischen den drei Testungen nach zwei, vier und sechs Wochen SIM-UT signifikant. Die Dauer der Untersuchung verkürzte sich im Case-Modus zwischen dem ersten und dem letzten Test um über die Hälfte der benötigten Zeit (1238 ± 283 Sek. vs. 620 ± 189 Sek., $p < 0,001$). Damit holte sie nach bereits vier Wochen, also 4 h SIM-UT zeitlich die Referenzgruppe A der Assistenz- und Fachärzt*innen ein und war nach Abschluss des SIM-UT signifikant schneller in der Durchführung der Module ($p < 0,001$). Die Expert*innengruppe blieb im Vergleich zu beiden Gruppen signifikant schneller in der Durchführung des Zweittrimesterscreening, was für eine Aussagekraft der DEGUM-Stufe-II-Zertifizierung spricht ($p < 0,001$).

Referenzgruppe A verfügte über eine Erfahrung von im Median 750 gynäkologisch-geburtshilfliche Ultraschalluntersuchungen und eine Gesamtzahl von 48.130. Die durchschnittliche Erfahrung liegt etwa auf der Minimalebene zum „Facharztstandard“. Referenzgruppe B kam auf insgesamt 264.000 Sonographien, bei einem Median von 15.000 Untersuchungen pro Expert*in. Den Trainees gelang mittels SIM-UT die klinische Erfahrung der Assistenz- und Fachärzt*innengruppe aufzuholen und den Abstand zwischen Expert*innengruppe und sich deutlich zu reduzieren. Inwiefern die Anzahl durchgeführter Ultraschalluntersuchungen ein Prädiktor für die tatsächliche diagnostische Genauigkeit sein kann, untersuchten Tolsgaard et al. in einer Studie [55]. 101 Trainees wurden nach erfolgreichem Ablegen der französischen Staatsprüfung für geburtshilflichen Ultraschall eingeladen, fünf Module für transvaginalen Ultraschall an einem Ultraschallsimulator durchzuführen. Anschließend wurde eine Umfrage hinsichtlich ihrer sonographischen Vorerfahrung erhoben. Die Ergebnisse zeigen, dass die vorherige Durchführung einer bestimmten Anzahl an Ultraschallscans durchaus sinnvoll ist, um überhaupt ein gewisses Level zu erreichen, sie jedoch als Prädiktor allein nur eine sehr geringe Aussagekraft hat: Eine signifikante Korrelation zwischen der Anzahl bereits durchgeführter Sonographien und der diagnostischen Genauigkeit in den Simulatormodulen wurde zwar

nachgewiesen, es gab allerdings keine Assoziation zwischen der Dauer der postgraduellen Weiterbildung der Ärzt*innen und der Performanz am Simulator. Insgesamt ergaben die Untersuchungen von Tolsgaard et al., dass die Anzahl durchgeführter Scans ein schlechter Prädiktor für die sonographisch-diagnostische Genauigkeit darstellt und dass anstelle einer Mindestanzahl durchgeführter Untersuchungen vor eigenständiger klinischer Durchführung ein Test am Simulator die Fähigkeiten dazu überprüfen könne.

Die Ergebnisse der im Rahmen unserer prospektiven Studie durchgeführten Tests lassen ähnliche Rückschlüsse für das Zweittrimesterscreening zu. Die Referenzgruppe B, bestehend aus DEGUM-Stufe-II-Expert*innen, die signifikant mehr Erfahrung im geburtshilflichen Ultraschall aufwiesen, stellte alle 23 Standardebenen korrekt dar. Auch hier lässt sich eine Assoziation zwischen der Anzahl der durchgeführten Untersuchungen und der Genauigkeit und Schnelligkeit in der Diagnostik feststellen. Jedoch wurde das Niveau der Referenzgruppe A bereits nach vier Wochen von den Trainees erreicht und nach sechs Wochen bereits überholt, obwohl deren minimale Vorerfahrung im gynäkologisch-geburtshilflichen Ultraschall nur durch das SIM-UT erweitert wurde. Betrachtet man also lediglich die Anzahl der durchgeführten Ultraschalluntersuchungen als Prädiktor für Kompetenz im Ultraschall, ist diese hier nicht repräsentativ.

Der Median der durchgeführten geburtshilflichen Ultraschalluntersuchungen der Referenzgruppe A von 750 Scans überstieg den vorgesehenen Grenzwert von 300 attestierten geburtshilflichen B-Bild-Sonographien der utero-plazento-fetalen Einheit aus dem Weiterbildungskatalog für Gynäkologie und Geburtshilfe. Interessant ist, wie schnell die Interventionsgruppe die Vorerfahrung der Referenzgruppe erreichte und sich durch das SIM-UT sogar weiter steigerte. In der Schwierigkeitsevaluation der Ebenen muss beachtet werden, dass einige Teilnehmende der Assistenz- und Fachärzt*innengruppe Ebenen in der Evaluation gestrichen hatten, da diese außerhalb der Teilnahme an der Studie teilweise noch nicht von den Ärzt*innen angewendet worden waren. Dieser Fall trat 19-mal für Ebenen des erweiterten Zweittrimesterscreenings ein. Der Median für durchgeführte DEGUM-Stufe-II-Screenings lag für die Referenzgruppe A bei null und die Testung am Simulator beinhaltete viele Ebenen, die erst in der Feindiagnostik bzw. einem DEGUM-Stufe-II-Screening untersucht werden. Für eine differenziertere Betrachtung hinsichtlich der Aussagekraft sollten folglich die Zeit und Korrektheit der Ebenen aus den Basisscreenings Ia gemäß MuRL zwischen den Gruppen verglichen werden: Die Assistenz- und Fachärzt*innen der Referenzgruppe A erreichten – sowohl für alle 23 Ebenen als auch isoliert für die vier Basisebenen – einen niedrigeren Mittelwert als die Interventionsgruppe nach den ersten zwei Wochen des SIM-UT. Statistisch signifikant war der Unterschied allerdings nur für das Basisscreening, in welchem die Ärzt*innen mehr Erfahrung hatten ($p=0,026$). Nach weiteren zwei Wochen SIM-UT hatte die Interventionsgruppe die Referenzgruppe eingeholt und für den restlichen Studienzeitraum unterschieden sich die Mittelwerte für die TTC der Basisebenen nicht mehr signifikant (nach 4 Wochen: $p=0,26$, nach 6 Wochen: $p=0,98$). Dagegen fällt im Vergleich der TTC für die gesamte Untersuchung aller 23 Ebenen auf, dass die Interventionsgruppe nach sechs Wochen eine signifikant kürzere TTC als Referenzgruppe A erreichte ($p=0,011$).

In dieser Hinsicht profitierten die Ärzt*innen der Referenzgruppe A offenbar von der klinisch gesammelten Erfahrung im Vergleich zur Interventionsgruppe. Trotzdem war es unerwartet, dass die Studierenden mit fehlender Vorerfahrung bereits nach 4 Trainingsstunden die klinische Erfahrung der Ärzt*innen aufholten und sie im Gesamtscreening sogar signifikant überholten. Das ist umso alarmierender, da Referenzgruppe A bereits die Anzahl der pränatalen Ultraschalluntersuchungen laut Anforderungen des Weiterbildungskatalogs für die Facharztprüfung erreicht hatte. Insgesamt können wir die Ergebnisse von Tolsgaard et al. [55] auch für transabdominale geburtshilfliche Ultraschalluntersuchungen dahingehend bestätigen, dass die alleinige Anzahl von durchgeführten Untersuchungen ein unzureichender Prädiktor für die tatsächliche Ultraschallkompetenz ist.

Die Anforderungen an die von uns für das SIM-UT ausgesuchte Untersuchung schließt alle Ebenen der DEGUM-Stufe-I-Diagnostik mit ein. Darüber hinaus wurden bis auf die Herzebenen alle Ebenen der DEGUM-II-Diagnostik ebenfalls trainiert und verglichen. Der Ausschluss der Herzebenen erfolgte trotz ihrer immensen Relevanz bis auf den Vierkammer-Blick, da in einer parallellaufenden Studie die Durchführbarkeit von SIM-UT für die fetale Echokardiographie am Simulator detailliert und differenziert untersucht wurde (FESIM II). Dadurch war es uns jedoch möglich, weitere sehr relevante und schwierige Ebenen der DEGUM-Stufe-II-Diagnostik miteinzuschließen, ohne den Rahmen des Trainings zu sprengen. Im Hinblick auf häufige kongenitale Fehlbildungen, wie z.B. die des Harnsystems, hielten wir es für sinnvoll, die Lernkurven von SIM-UT über die reinen Biometrie-Ebenen und die des sonomorphologischen Screenings gemäß MuRL IIb hinaus zu analysieren.

Immer mehr Studien belegen die Erlernbarkeit von Grundlagen-Skills und biometrischen Screenings im gynäkologisch-geburtshilflichen Ultraschall mithilfe von Simulation [8,31]. Unsere Studie unterscheidet sich im Hinblick auf die bereits vorhandenen Erkenntnisse insbesondere durch die Untersuchung der Durchführbarkeit von SIM-UT für ein umfangreiches Zweittrimesterscreening mit 23 Ebenen. Durch die Ergebnisse kann nun erstmals gezeigt werden, dass auch komplexere Untersuchungsprotokolle von Trainees mittels SIM-UT erlernt werden können. Des Weiteren stellt der Zeitraum des Trainings mit über sechs Wochen und die Regelmäßigkeit von zwei Trainingseinheiten pro Woche ein neues Modell dar. Eine Studie, die die Lernkurve von Ultraschallanfänger*innen für eine so ausführliche Untersuchung in einem regelmäßig stattfindenden Training analysiert, ist nach unserem Kenntnisstand noch nicht für pränatale Sonographie durchgeführt worden. Auch die hohe Zertifizierung aller Mitglieder der Expert*innengruppe mit mindestens DEGUM-Stufe-II stellt eine Stärke unserer Studie dar.

5.2. Durchführbarkeit und Möglichkeiten von SIM-UT

Die Anforderungen an pränatale Ultraschalluntersuchungen sind hoch und die Aufgabe, Lehre in Facharztweiterbildungen und Studium den Anforderungen anzupassen, komplex. Eine Möglichkeit erfolgreiche Lehre unkompliziert in die Weiterbildung zu implementieren ist der Einsatz von Ultraschallsimulatoren. Die Ergebnisse von Burden et al., dass SIM-UT selbstständig und mit minimaler Supervision für Trainees durchführbar ist [8], können wir durch unsere Ergebnisse bestätigen. Der Zugang zum Ultraschallsimulator wurde durch die Pforte des St.-Franziskus-Hospitals mittels protokollierter Schlüsselausgabe reguliert. Wir kontrollierten via Online-Terminplaner (Doodle) die Einteilung der Termine sowie stichprobenartig das Training vor Ort. Insgesamt wurde das SIM-UT durch die Trainees jedoch selbstständig über eine Online-Liste (Doodle-Liste) eingeteilt und der organisatorische Aufwand des Trainings war gering. Ähnliche Modelle könnten folglich in Kliniken für Ärzt*innen oder an Universitäten für Studierende eingeführt werden. Die Betreuung des Trainings durch studentische Tutor*innen, wie von uns durchgeführt, könnte die Implementierung von SIM-UT an Universitäten und deren Kliniken erleichtern. Studien belegen bereits den Erfolg für studentisch geleitete Ultraschalllehre. Dies gilt sowohl für Peer-Teaching, also Lehre von Studierenden für Studierende, als auch Ärzt*innen-Training durch speziell geschulte Studierende, welches ebenso effektiv ist, wie Training durch Ärzt*innen [22,41].

Ein Hindernis der Implementierung von Ultraschallsimulatoren in bestehende Curricula werden die Anschaffungskosten der Systeme sein. Dennoch stellt SIM-UT summa summarum eine kostengünstige und effektive Methode dar, denn selbstständiges Training am Simulator erfordert weder die Einstellung zusätzlichen Lehrpersonals noch die Organisation und Bezahlung von Schallmodellen [9]. Fortbildungen im Bereich der Pränatal-Sonographie und Unterricht erforderten stets schwangere Freiwillige. SIM-UT kommt dagegen ohne Live-Modelle aus. Seit dem 01.01.2021 wird pränataler Ultraschall zu nicht-medizinischen Zwecken durch die Verordnung zum Schutz vor schädlichen Wirkungen nichtionisierender Strahlung bei der Anwendung am Menschen (NiSV) strenger limitiert [5]. Aus diesem Grund ist beispielsweise auch das sogenannte „Babyfernsehen“ als IGeL-Leistung nicht mehr möglich. Aus ethischer Sicht ist es fragwürdig ausschließlich an Menschen, noch dazu Schwangeren oder Patient*innen, zu üben, wenn ein gleichwertig effektives wie realistisches Training an Simulatoren durchgeführt werden kann [62]. Da Simulation die Ausbildung von Ärzt*innen vor der Durchführung diagnostischer und therapeutischer Maßnahmen an Patient*innen ermöglicht, können Fehlerquoten verringert und die Patient*innensicherheit gefördert werden [40].

Da SIM-UT die Effektivität der Arbeitsabläufe im klinischen Alltag steigern kann, werden somit ebenfalls Ressourcen eingespart [58,59]. Die Kosten reduzieren und amortisieren sich stärker, wird der Simulator von mehreren Fachrichtungen eines Krankenhauses oder einer Universität benutzt. Es zeigt sich bereits für viele Fachbereiche, dass simulationsbasierte Ultraschalllehre eine effektive, praktische und schonende Methode darstellt [23,26,42].

Die Annahme von SIM-UT durch die Trainees erwies sich als sehr positiv. Alle Teilnehmenden des Trainings beschrieben ihre subjektive Erfahrung mit SIM-UT positiv und gaben darüber hinaus an, dass sie simulationsbasierte Ultraschall-Wahlfächer an ihrer Universität annehmen und belegen würden, würden diese angeboten werden. Zusätzlich schrieben einige Studierende, dass sie sich simulationsbasierte Ultraschall-Wahlfächer auch für andere Fachrichtungen wünschen würden. Zu ähnlichen Ergebnissen kam eine Studie von Cook et al., in welcher Studierende in einer einzigen Trainingseinheit transvaginalen gynäkologischen Ultraschall am U/S Mentor von Symbionix übten [12]. Auch das von Cook et al. durchgeführten SIM-UT wurde von allen teilnehmenden Studierenden positiv bewertet, indem sie es hilfreich und angenehm wahrnahmen. Dass die Teilnehmenden unserer Studie auch das regelmäßige Training über sechs Wochen positiv evaluierten, ist eine wichtige Information für zukünftige Curricula, die sich den Herausforderungen sonographischer Lehre stellen.

Im Zeitalter von virtual-reality Technologien, in dem mit Hilfe von Simulatoren Patient*innenversorgung realitätsnah imitiert wird, erscheint der Umstand, praktische Lehre ausschließlich im klinischen Alltag durchzuführen als nicht mehr zeitgemäß und inadäquat [35,36,62]. Befürworter*innen von alternativen Lehransätzen geben auch die Aspekte der Patient*innensicherheit und Qualität medizinischer Diagnostik bzw. Versorgung an. Für Ultraschalluntersuchungen gibt es bereits viele Belege dafür, dass nicht allein die Anzahl durchgeführter Untersuchungen Aufschluss über die Kompetenz der Untersucher*innen gibt [26,55]. Wie schnell Ärzt*innen und Studierende Ultraschallkompetenzen erlernen, ist individuell verschieden. Ein Ansatz, der die interindividuellen Schwankungen im Lernen miteinbezieht, ist das Konzept des sogenannten Mastery- oder Proficiency-based Learning. Der Ansatz folgt der Strategie, dass auf ein festgelegtes Kompetenzniveau in einem sicheren und simulierten Umfeld trainiert wird. Mastery Learning berücksichtigt, dass das Erreichen des Niveaus individuell unterschiedlich viel Zeit in Anspruch nehmen kann und vor Durchführung der Untersuchungsmethode in der klinischen Praxis überprüft wird [26,35]. So bestand auch für die Trainees in unserer Studie die Möglichkeit, individuell an den eigenen Schwächen in der Durchführung der Untersuchung zu üben. Dies geschah ebenfalls in einem sicheren Umfeld, in dem der Fokus auf den Lernenden lag und nicht auf Patient*innen, wodurch wenig Distraktoren auf die Lernenden einwirkten. Da die Trainees das Kompetenzniveau von Assistenz- und Fachärzt*innen im praktischen Test am Simulator erreichten und im Hinblick auf die Geschwindigkeit sogar signifikant überstiegen, ist anzunehmen, dass sie die notwendige Kompetenz für klinisches Training an Patient*innen erreichten. Die Beobachtung, dass die Rate an korrekt dargestellten Ebenen bereits nach vier Wochen die der Referenzgruppe A erreichte, die Trainees sich nach weiteren zwei Wochen strukturiertem SIM-UT aber immer noch in der Geschwindigkeit zur Durchführung der Untersuchung steigerten, ist eine Beobachtung, die ebenfalls im Mastery Learning gemacht wurde. Daher wird empfohlen, das SIM-UT nicht nach Erreichen des festgelegten Kompetenzniveau einzustellen, sondern dieses fortzuführen, um Routine und Effektivität in Untersuchungsabläufe zu bringen [26].

Mastery Learning stellt somit eine Methode in der Gewährleistung eines hohen Ultraschallstandards dar, indem Untersuchende erst nachdem sie ein festgelegtes Kompetenzniveau im simulierten Setting erreicht haben, im

klinischen Alltag Ultraschall-Untersuchungen durchführen. Dass die Dauer der Facharztweiterbildung und Anzahl durchgeführter Ultraschalluntersuchungen nur wenig Aussagekraft für den Kompetenzerwerb haben, wurde bereits diskutiert. Trotzdem wird die Zulassung zum eigenverantwortlichen und selbstständigen Sonographieren in Deutschland nach diesen Kriterien sowie durch Abschließen der Facharztprüfung reguliert. Für Basisultraschalluntersuchungen gibt es weder einen Zulassungstest, noch werden Ultraschallfähigkeiten und -kenntnisse nach einmaliger Online-Prüfung für den sonomorphologischen Schall nach MurRL IIB reevaluiert. Aufgrund der niedrigen pränataler Detektionsraten fetaler Fehlbildungen und deren teils schweren Konsequenzen, werden Zulassungstests und Qualifikationsnachweise für sonographische Pränataldiagnostik zunehmend gefordert [4,50]. Eine Möglichkeit die Kompetenz von Untersuchenden zu überprüfen, sind standardisierte Tests an Ultraschallsimulatoren. So könnten auch Untersuchende akkreditiert oder reevaluiert werden, die bereits im klinischen Setting praktizieren. Der im Rahmen unserer prospektiven Studie durchgeführten Test scheint zwischen verschiedenen Kompetenzniveaus zu differenzieren. Die Lernkurve der Trainees zeigt, dass die nach zweiwöchigem SIM-UT durchgeführten Test zu signifikant schlechteren Ergebnissen hinsichtlich der Genauigkeit der Ebenen führten als bei beiden Referenzgruppen. Der Test attestierte ihnen folglich das Kompetenzniveau von Anfänger*innen. Über die Dauer des Trainings näherten sie sich dem Level der Referenzgruppen immer weiter an. Zuerst schlossen sie schnell zum Level der Referenzgruppe A auf und wurden im Hinblick auf die Geschwindigkeit in der Durchführung des Tests signifikant schneller. Gleichzeitig erzielte die Expert*innengruppe bessere Ergebnisse und signifikant kürzere Schallzeiten. Würden Studierende durch alleiniges SIM-UT auch das Level hochspezialisierter und zertifizierter Pränataldiagnostiker*innen einholen, würde die Validität des verwendeten Simulators, im Sinne der Übertragbarkeit auf realen Ultraschall, als Testinstrument herabgesetzt werden. Es gibt bereits ähnliche Beobachtungen für gynäkologisch-geburtshilflichen Ultraschall, hier unterscheiden sich die Genauigkeit fetaler biometrischer Messungen am Ultraschallsimulator signifikant zwischen Trainees und Spezialist*innen, ebenso wie die Geschwindigkeit in der Durchführung von simulationsbasierter Untersuchungen [7].

Madsen et al. untersuchten einen Ultraschallsimulator hinsichtlich seiner Validität und Reliabilität zur Überprüfung von Ultraschallkompetenzen mittels eines Tests für transvaginalen Ultraschall [34]. In derselben Studie wurden die Lernkurven von Trainees beim Erlernen dieser Kompetenzen durch SIM-UT analysiert. Dafür erstellte die Arbeitsgruppe einen umfassenden Test aus sieben Modulen mit 156 Aufgaben, die z.B. in der Darstellung einzelner Strukturen in Standardebenen, Messungen von Strukturen, Bildoptimierungen und Schallkopfführung bestanden. Insgesamt ergaben diese Untersuchungen, dass der Simulator zwischen Untersuchungsergebnissen von spezialisierten Oberärzt*innen und Trainees differenzieren und Kompetenzlevel abbilden konnte. Allerdings kritisierte die Arbeitsgruppe die Auswertung des Simulators: ein Teil (105/153) der Aufgaben im Simulatortest bildete die Ultraschallkompetenz der Teilnehmenden nicht repräsentativ ab, weshalb nur 56 der 153 Aufgaben, die durch die vorherige Untersuchung der Aufgaben als valide erfasst wurden, für Training und Feedback angewandt wurde. Ähnliche Erfahrungen machten wir auch im Rahmen unserer Studie mit der automatisierten Auswertung des Simulators. Dessen Feedback und Auswertungen sich nicht mit der

Zweitbewertung durch das Expertengremium deckten, da von den durch den Simulator als inkorrekt befundene Ebenen im Rahmen der Studie lediglich rund 12% in der Zweitbewertung des Expertengremiums als inkorrekt bewertet wurden. So war zwar die Auswertung durch den Simulator nicht ausreichend, jedoch ist die reine Durchführung eines Tests am Simulator zur Überprüfung von Ultraschallkompetenzen eine geeignete Methode, die ohne Schallmodelle auskommt.

In früheren Studien wurde die Abwesenheit fetaler Bewegungen an Ultraschallsimulatoren kritisiert, da diese die Realität des Trainings limitierte [31]. Neuere high-end Simulatoren, wie der in der vorliegenden Studie eingesetzte Simulator, simulieren nun auch die fetalen Bewegungen und der Fetus wechselt während der Untersuchung zufällig zwischen verschiedenen Positionen hin und her. Der reale Eindruck der Simulation wird zusätzlich durch Schallartefakte wie Schallschatten verstärkt.

5.3. Subjektive Schwierigkeitsbewertung der Standardebenen

Für die Erlernbarkeit des erweiterten Zweittrimesterscreenings interessierte uns auch die Evaluation der Standardebenen hinsichtlich ihrer Schwierigkeit in der Darstellung. Mit der subjektiven Schwierigkeitsbewertung sollte untersucht werden, ob es Ebenen gibt, die mehrheitlich als schwieriger oder einfacher in der Darstellung wahrgenommen werden. Außerdem interessierten uns die Unterschiede in der Evaluation zwischen den drei Gruppen, z.B. ob die Schwierigkeit bei wenig Erfahrung auch als höher empfunden wird. Basierend auf den Ergebnissen erfolgte eine Einteilung der Ebenen in einfach-mittelschwierig-schwierig. Die korrekte Darstellung von Standardebenen gilt als Schlüsselkompetenz für pränataldiagnostische Ultraschalldiagnostik und die pränatale Erkennung kongenitaler Fehlbildungen [60]. Eine systematische Einteilung von Ebenen des erweiterten Zweittrimesterscreenings nach Schwierigkeit existiert nach unserem Kenntnisstand bisher nicht. Die Unterscheidung zwischen verschiedenen Schwierigkeitsniveaus in der Darstellung der Standardebenen ist aus didaktischen Gründen besonders interessant. Curricula von Ultraschallkursen und Fortbildungen könnten sich so zunächst auf das Erlernen von Basisebenen und einfach darstellbaren Standardebenen fokussieren und sich im Verlauf den anspruchsvolleren Ebenen widmen.

Die subjektiv empfundene Schwierigkeit nahm mit dem Erfahrungsalter der Teilnehmenden ab. Zwar bewerteten die Trainees deutlich mehr Ebenen als schwierig als Referenzgruppe A, deren Mittelwerte die Ebenen vor allem in mittelschwierig und einfach unterteilten, signifikant waren die Unterschiede zwischen den beiden Gruppen allerdings nur für drei Ebenen. In Abbildung 12 lassen sich Ähnlichkeiten in der Evaluation der beiden Gruppen ablesen. Referenzgruppe B dagegen bewertete alle Ebenen als einfach darzustellen, was für ihre hohe Expertise und Zertifizierung spricht und gleichzeitig auf die Erlernbarkeit der Ebenen hindeutet, wenn man die hohe Diskrepanz in der Einschätzung zwischen Trainees und Expert*innen betrachtet. Deshalb werden zur Beurteilung der Schwierigkeitsbewertung hauptsächlich die Bewertungen der Interventionsgruppe und Referenzgruppe A miteinander verglichen.

In der Untergruppe der einfach darzustellenden Ebenen sind vor allem die axialen Rumpfebenen wie Nabelschnur-Insertion-, Abdomen- und Harnblasenebene, sowie Ebenen des Basisscreenings wie Femur- und Plazentaebene zu finden. Weiterhin fällt auf, dass Kopf- und Gesichtsebenen mehrheitlich als schwierig und mittelschwerig darzustellen eingeschätzt wurden. Diese Einschätzung spiegelt sich auch in der Test-Performanz der Teilnehmenden wider. So fand sowohl die Interventionsgruppe als auch Referenzgruppe A die einfache Ebene der Nabelschnurinsertion signifikant schneller auf als die schwierig evaluierte Ebene des Gesichtsprofils. Signifikante Unterschiede in der Schnelligkeit der Ebenendarstellung innerhalb einer Gruppe fanden sich ausschließlich in Verbindung mit mindestens einer axialen Rumpfebene. Dieser positive Zusammenhang zwischen subjektiver Schwierigkeitsbewertung und der tatsächlichen hands-on Performanz der Teilnehmenden deutet darauf hin, dass die subjektive Schwierigkeitsbewertung weniger vom individuellen Empfinden, sondern vielmehr von erlebten Herausforderungen während des Schallens und der Kompetenz der verschiedenen Gruppen abhängt. Dieser Zusammenhang wird besonders deutlich im Vergleich der allgemeinen Schwierigkeitsbewertung der Ebenen einer Gruppe mit ihrer TTC, welche signifikant korreliert.

Ein interessanter Aspekt in der subjektiven Evaluation ist die abnehmende Schwierigkeitsbewertung mit dem Erfahrungsalter. Referenzgruppe A bewertete die Standardebenen als einfacher darzustellen als die Trainees, obwohl diese nach Abschluss des SIM-UT die Untersuchung signifikant schneller abschlossen. Hier muss allerdings berücksichtigt werden, dass nicht alle Ärzt*innen der Referenzgruppe A alle Ebenen evaluierten, da einzelne Teilnehmende aufgrund mangelnder Erfahrung Standardebenen aus der Bewertung ausschlossen.

Die eingeschränkte Reliabilität der subjektiven Schwierigkeitsbewertung liegt in der Natur der Umfrage. Dennoch zeigen sich signifikante Korrelationen zwischen subjektiver Evaluation und tatsächlicher Performanz. Außerdem konnte für unterschiedliche Erfahrungsgruppen voneinander unterscheidbares Schwierigkeitsempfinden in der Darstellung von verschiedenen Standardebenen analysiert und abgebildet werden.

Zusammenfassend kann festgehalten werden, dass viele Standardebenen des erweiterten Zweittrimesterscreenings durch SIM-UT gut erlernbar sind und Lernende subjektiv die Darstellung dieser größtenteils als einfach bis moderat einschätzt. Dennoch wurden wenige Ebenen durch Interventions- und Referenzgruppe A als schwierig darzustellen bewertet. Die subjektive Einschätzung spiegelt sich in ihrer Testperformanz und dem signifikanten Leistungsfortschritt der Interventionsgruppe wider. Gleichzeitig wurde für schwierig einzustellende Ebenen signifikant mehr Zeit in der Darstellung dieser benötigt. Die fünf schwierigen Standardebenen sind ausschließlich Gesichts- und Kopfebenen (das Gesichtsprofil und koronare Gesicht, die transventrikuläre-, transthalämische- und transzerebelläre Ebene). Ultraschallcurricula sollten diese Schwierigkeiten in der Ebenendarstellung miteinbeziehen, um den bestmöglichen Standard des pränataldiagnostischen Ultraschalls zu gewährleisten. Die transthalämische Ebene als Basisebene ist Bestandteil jeder Fetometrie für das fetale Schätzwicht. Eine falsches Schätzwicht, was z. B. zu einem Verdacht auf fetale Makrosomie führen kann, kann weitreichende Folgen für den Entbindungsmodus bedeuten [46].

Angepasstes Basisultraschalltraining kann zu besseren Ergebnissen und schnelleren Lernkurven führen [56]. Selbstsicherheit und -einschätzung sind wichtige Aspekte für die Patient*innenbetreuung und den Skill-Transfer in das klinische Umfeld; Tolsgaard et al. erkannten Bilderkennung und -interpretation als wichtige Faktoren für Selbstsicherheit von Ultraschalltrainees [56].

5.4. Limitationen durch den Simulator

Während der Positionswechsel des Fetus durch die zufälligen Bewegungen in unterschiedlichen Zeitabständen ein Vorteil des Simulators darstellt, ist die beschränkte Lagevarianz des Fetus als Nachteil zu werten. Für die Untersuchung eines physiologisch entwickelten Fetus im Zweittrimester-Screening gab es lediglich Module, in denen sich der Fetus in Beckenendlage befand. Zweittrimester-Screenings in Schädellage waren nur möglich bei der Untersuchung der Pathologiefälle, wie z.B. der Case 5 in welchem sich der Fetus mit Ventrikulomegalie und Lippen-Kiefer-Gaumenspalte präsentierte. Eine andere Limitation des Simulators ist die fehlende Diversität der Schallbedingungen durch maternale und fetale Eigenschaften wie z.B. die maternale Bauchdecke oder Plazentaverkalkungen, welche die Schallbedingungen verändern können.

Ein Kritikpunkt am U/S Mentor von Symbionix war die Zuordnung der im Case-Modus generierten Ebene zur zugehörigen Standardebene durch das Simulator-Feedback, welches sich nach Abschluss der Untersuchung automatisch öffnete. Die Kriterien, wann der Simulator gesicherte Bilder als nicht zuzuordnen bewertete, war für uns teilweise nicht ersichtlich. Die Zweitbewertung der Ebenen durch das unabhängige Expertengremium zeigt ebenfalls, dass die Bewertung des Simulators in vielen Fällen nicht nachvollziehbar war. Die Trennschärfe in der Bewertung des Simulators, ob eine Ebene ausreichend oder mangelhaft dargestellt ist, muss zuverlässiger werden, mindestens so lange wie Trainees im SIM-UT auf dessen Feedbackfunktion angewiesen sind. Trotzdem konnte Basistraining für die Darstellung von Standardebenen und den Ablauf der Untersuchung schon jetzt mit zuverlässigem Feedback im Task-Modus am Simulator erlernt werden.

Weiterhin wurden am Simulator keine Video-Loops beim „Freeze“ gespeichert – was ein Zurückspulen bis zur richtigen Ebene erlaubt -, wie es die Ärzt*innen der Referenzgruppen aus der Benutzung herkömmlicher Ultraschallgeräte gewohnt sind. Dies limitiert die Simulation einer realistischen Ultraschalluntersuchung. Auch das gummiartige Oberflächenmaterial der Mannequins und die fehlende Anwendbarkeit von Schallgel, bewirkten, dass das Gleiten des Schallkopfs nur eingeschränkt möglich war. Während der Durchführung der Studie konnten wir dies durch das Abdecken des Torsos mit Frischhaltefolie verbessern, sodass der Schallkopf leichter über das Abdomen zu bewegen war. Trotz dieser Limitationen in der Bedienung des Simulators, fanden sich alle Teilnehmenden schnell in die Handhabung des U/S Mentors ein und konnten die Module intuitiv bedienen.

5.5. Limitationen im Studiendesign

Das Studiendesign hat zwei Limitationen. Zum einen wären größere Gruppegrößen wünschenswert gewesen. Die Gruppengrößen mit 10-11 Teilnehmenden waren hauptsächlich durch die Verfügbarkeit des Simulators limitiert. Durch die parallellaufende Studie FESIM II trainierten montags-freitags 22 Studierende zwei Mal wöchentlich für je eine Stunde am Simulator. An jedem zweiten Wochenende erfolgte samstags und sonntags die Überprüfung des Lernfortschritts von 22 Studierenden à eine Stunde. Zusätzlich mussten die Termine zur Testung der Referenzgruppen eingeplant und zeitlich blockiert werden. Folglich war der Simulator täglich über viele Stunden belegt und die Vergrößerung der Interventions- und Referenzgruppen aus organisatorischen Gründen kaum möglich.

Es wäre interessant gewesen nach Abschluss des sechswöchigen SIM-UT den Skill-Transfer im Zweittrimesterscreening vom Simulator auf die Untersuchung eines echten Fetus zu überprüfen. Leider konnte dies bei Ende des Trainingszeitraums Mitte Oktober 2020, aufgrund verschärfter Kontaktbeschränkungen und Hygienemaßnahmen in der Corona-Pandemie, nicht mehr realisiert werden. Der Skill-Transfer von komplexen Untersuchungsprotokollen nach regelmäßigen und strukturierten SIM-UT sollte daher in zukünftigen Studien überprüft werden.

5.6. Weitere Möglichkeiten für den Einsatz von Ultraschallsimulatoren

Für die Steigerung der pränatalen Detektionsrate kongenitaler Fehlbildungen ist die Kenntnis über die sonographische Präsentation der Fehlbildungen unabdingbar. In einer deutschlandweiten Studie testeten Staboulidou et al. die Effektivität von SIM-UT hinsichtlich der Erkennung fetaler Pathologien bei Fachärzt*innen für Frauenheilkunde und Geburtshilfe [52]. Sie führten in jedem deutschen Bundesland standardisierte Ultraschallkurse durch, die ihren Schwerpunkt auf der sonographischen Darstellung fetaler Anomalien hatten. Dabei nutzten sie eine Kombination aus theoretischem Input über häufige kongenitale Fehlbildungen und SIM-UT. Vor und nach dem Kurs wurde in Fragebögen im Multiple Choice-Format mit Ultraschallbildern von häufigen Fehlbildungen das Wissen der Teilnehmenden erfasst. Die Ärzt*innen verbesserten ihre Ergebnisse durch den Kurs signifikant um nahezu 40%. Daraus schloss die Arbeitsgruppe, dass SIM-UT ein mögliches Instrument ist, die Detektionsrate bei Fachärzt*innen zu steigern.

Der Vorteil liegt hier auf der Hand; wenn Pathologien an Ultraschallsimulatoren realistisch dargestellt werden, können Studierende und Ärzt*innen die Detektion dieser lernen und üben, bevor sie auf eine tatsächlich betroffene Schwangere im klinischen Alltag treffen. Dies kann so weit reichen, dass auf risikobehaftete, invasive Pränataldiagnostik durch gut und vollständig durchgeführte Ultraschalldiagnostik teilweise verzichtet werden könnte.

5.7. Schlussfolgerung und Ausblick

Akkurate Ultraschalldiagnostik in der Schwangerschaft ist unabdingbar für die Detektion fetaler Malformationen und eine angemessenen Geburtsplanung. Um das peri- und postnatale Outcome von Neugeborenen zu verbessern, muss auch der Ultraschallstandard in Deutschland angehoben werden. Für erfolgreiche pränataldiagnostische Ultraschallscreenings ist die genaue Kenntnis der Standardebenen und deren korrekte Darstellung eine zentrale Voraussetzung [60]. Aus unseren Ergebnissen geht eine gute Erlernbarkeit von umfangreichen Untersuchungsprotokollen im pränatalen Ultraschall durch den Einsatz eines Ultraschallsimulators hervor. Die Interventionsgruppe erlernte durch SIM-UT die sichere Darstellung der 23 Standardebenen des erweiterten Zweittrimesterscreenings. Um sich dem Ziel, strukturierte und standardisierte Ultraschalllehre in der Gynäkologie und Geburtshilfe zu etablieren, anzunähern, ist dies eine wichtige Feststellung. Ultraschallsimulatoren bieten die Chance in einem sicheren Lernumfeld notwendige Ultraschallkompetenzen zu erlernen und zu vertiefen, ohne Patient*innen zu schaden oder die Qualität der Diagnostik durch mangelnde Vorkenntnisse oder parallel erfolgende Lehre herabzusetzen. Zudem geht der Einsatz von Simulatoren mit dem erheblichen Vorteil des eigenständigen Trainings und geringem organisatorischen Aufwand einher. Trotz der Limitationen des Simulators zeigen unsere Ergebnisse, dass der U/S Mentor ein hilfreiches Instrument ist, um Trainees für klinisches Training vorzubereiten oder sich auch noch während des klinischen Trainings weiter zu verbessern. Trainees ohne oder mit minimaler Vorerfahrung im pränatal-diagnostischen Ultraschall schlossen im Laufe des SIM-UT im Hinblick auf die Geschwindigkeit in der Ebenendarstellung mit Assistenz- und Fachärzt*innen auf und zeigten selbst im Vergleich mit hochzertifizierten Pränataldiagnostiker*innen in 87% der Ebenen keine signifikanten Unterschiede mehr. Nach 12 h SIM-UT stellten die Trainees über 96% der Ebenen korrekt dar und befanden sich damit auf Assistenz- und Fachärzt*innen Niveau. Die Expert*innen blieben in der Gesamtperformanz der TTC signifikant besser und stellten alle Standardebenen korrekt dar (RAI=100%), doch auch an deren Niveau konnten sich die Trainees mittels SIM-UT annähern. Des Weiteren stellen simulationsbasierte, praktische Tests eine Möglichkeit zur Akkreditierung oder Re-evaluation von Ultraschallkompetenzen von Ärzt*innen dar. Unsere Ergebnisse legen nahe, dass der von uns eingesetzte Simulator ebenfalls für die Testung von Ultraschallkompetenzen geeignet wäre. Für zukünftige Studien wäre es interessant die Übertragbarkeit von Fähigkeiten am Simulator auf Patient*innen und Live-Modelle zu überprüfen. Insgesamt können wir uns der Forderung, SIM-UT in Ultraschallcurricula zu etablieren, aufgrund der sehr guten Erlernbarkeit, der einfachen Umsetzung und dem damit einhergehenden Erfolg der Methode, nur anschließen [25,40].

6. Abbildungsverzeichnis

Abbildung 1: Der U/S Mentor: Mannequin für Zweittrimester-Screening, Abdomensonde.	23
Abbildung 2: Bildschirm des Simulators während des Zweittrimestermoduls bei aktivierter Hilfestellung. ...	24
Abbildung 3: Task-Modus für Standardebene Diaphragma/rechte Lunge.....	25
Abbildung 4: Case-Modus des Simulators bei deaktivierter Hilfestellung.	26
Abbildung 5: Auswertung des Case-Modus für die Abdomenstandardebene.....	27
Abbildung 6: Piktogramme des Simulators für die Standardebenen (deaktivierbare Hilfestellung).....	29
Abbildung 7: Flussdiagramm über Ablauf von Training und Testungen.....	32
Abbildung 8: Eine im Case-Modus generierte Kleinhirnebene.....	33
Abbildung 9: TTC aller 23 Ebenen der verschiedenen Gruppen mit Signifikanztests.	38
Abbildung 10: TTC der vier Basisebenen mit Signifikanztests.	39
Abbildung 11: Boxplot der TTC für die Einzelebenen, aufgeteilt nach Untersuchungsprotokollen.....	40
Abbildung 12: Schwierigkeitsbewertung der Ebenen.	43

7. Tabellenverzeichnis

Tabelle 1: Gegenüberstellung der Mindestanforderungen an die Bilddokumentation der weiterführenden Differentialdiagnostik.....	15
Tabelle 2: Ausgewählte Standardebenen zur Übung und Testung des Zweittrimesterscreenings	28
Tabelle 3: Vorerfahrung im geburtshilflich-gynäkologischen Ultraschall der Referenzgruppen A und B.....	36
Tabelle 4: Vollständig abgelegte Prüfung für die Task 1-Untersuchung der Gruppen im Vergleich	37
Tabelle 5: Signifikante Unterschiede zwischen den einzelnen Standardebenen in der TTC innerhalb der Testgruppen	41
Tabelle 6: Mittelwerte der Schwierigkeitsbewertungen der drei Gruppen.....	44
Tabelle 7: Einteilung der Ebenen in Kategorien nach Einschätzung der Interventionsgruppe	45
Tabelle 8: Korrelation nach Pearson zwischen Testperformanz und der Schwierigkeitsbewertung.....	46

8. Literaturverzeichnis

1. Adzick NS, Thom EA, Spong CY, Brock JW, Burrows PK, Johnson MP, Howell LJ, Farrell JA, Dabrowiak ME, Sutton LN, Gupta N, Tulipan NB, D'Alton ME, Farmer DL (2011) A Randomized Trial of Prenatal versus Postnatal Repair of Myelomeningocele. *N Engl J Med* 364:993–1004
2. Ärztekammer Westfalen-Lippe (2020) Allgemeine Inhalte der Weiterbildung für Abschnitt B.
3. Aspler A, Wu A, Chiu S, Mohindra R, Hannam P (2021) Towards quality assurance: implementation of a POCUS image archiving system in a high-volume community emergency department. *Can J Emerg Med*
4. Benacerraf BR, Minton KK, Benson CB, Bromley BS, Coley BD, Doubilet PM, Lee W, Maslak SH, Pellerito JS, Perez JJ, Savitsky E, Scarborough NA, Wax J, Abuhamad AZ (2018) Proceedings: Beyond ultrasound first forum on improving the quality of ultrasound imaging in obstetrics and gynecology. *J Ultrasound Med* 37:7–18
5. BGB (2018) Verordnung zum Schutz vor schädlichen Wirkungen nichtionisierender Strahlung bei der Anwendung am Menschen. URL: <http://www.gesetze-im-internet.de/nisv/BJNR218700018.html>
6. Bundesausschuss der Ärzte und Krankenkassen (2013) Richtlinien des Bundesausschusses der Ärzte und Krankenkassen über die ärztliche Betreuung während der Schwangerschaft und nach der Entbindung (Mutterschafts-Richtlinien). 1–32
7. Burden C, Preshaw J, White P, Draycott TJ, Grant S, Fox R (2012) Validation of virtual reality simulation for obstetric ultrasonography: A prospective cross-sectional study. *Simul Healthc* 7:269–273
8. Burden C, Preshaw J, White P, Draycott TJ, Grant S, Fox R (2013) Usability of virtual-reality simulation training in obstetric ultrasonography: A prospective cohort study. *Ultrasound Obstet Gynecol* 42:213–217
9. Carolan-Rees G, Ray AF (2015) The ScanTrainer obstetrics and gynaecology ultrasound virtual reality training simulator: A cost model to determine the cost viability of replacing clinical training with simulation training. *Ultrasound* 23:110–115
10. Chalouhi GE, Bernardi V, Gueneuc A, Houssin I, Stirnemann JJ, Ville Y (2016) Evaluation of trainees' ability to perform obstetrical ultrasound using simulation: Challenges and opportunities. *Am J Obstet Gynecol* 214:525.e1-525.e8
11. Clark AE, Shaw CJ, Bello F, Chalouhi GE, Lees CC (2020) Quantitating skill acquisition with optical ultrasound simulation. *Australas J Ultrasound Med* 23:183–193

12. Cook J, Rao V V., Bell F, Durkin M, Cone J, Lane-Cordova A, Castleberry L (2020) Simulation-based clinical learning for the third year medical student: Effectiveness of transabdominal and transvaginal ultrasound for elucidation of OB/GYN scenarios. *J Clin Ultrasound* 48:457–461
13. Dastgiri S, Stone DH, Le-Ha C (2002) Prevalence and secular trend of congenital anomalies in Glasgow, UK. *Arch Dis Child* 86:257–263
14. Dromey BP, Ahmed S, Vasconcelos F, Mazomenos E, Kunpalin Y, Ourselin S, Deprest J, David AL, Stoyanov D, Peebles DM (2021) Dimensionless squared jerk: An objective differential to assess experienced and novice probe movement in obstetric ultrasound. *Prenat Diagn* 41:271
15. Edwards L, Hui L (2018) First and second trimester screening for fetal structural anomalies. *Semin Fetal Neonatal Med* 23:102–111
16. Eichhorn KH, Schramm T, Bald R, Hansmann M, Gembruch U (2006) Qualitätsanforderungen an die DEGUM-stufe I bei der geburtshilflichen ultraschalldiagnostik im zeitraum 19 bis 22 schwangerschaftswochen. *Ultraschall der Medizin* 27:185–187
17. Ferrier C, Dhombres F, Khoshnood B, Randrianaivo H, Perthus I, Guilbaut L, Durand-Zaleski I, Jouannic JM (2019) Trends in resource use and effectiveness of ultrasound detection of fetal structural anomalies in France: A multiple registry-based study. *BMJ Open* 9:1–6
18. Freundt P, Nourkami-Tutdibi N, Tutdibi E, Janzing P, Von Ostrowski T, Langer M, Zemlin M, Steinhard J (2023) Controlled Prospective Study on the Use of Systematic Simulator-Based Training with a Virtual, Moving Fetus for Learning Second-Trimester Scan: FESIM III. *Ultraschall der Medizin* 44:E199–E205
19. Hansmann M (1981) Identification and exclusion of disturbances of foetal development via ultrasonic screening and on-target examination - A multistage concept. *Ultraschall der Medizin* 2:206–220
20. Hindryckx A, De Catte L (2011) Prenatal diagnosis of congenital renal and urinary tract malformations. *Facts, views Vis ObGyn* 3:165–74
21. Hofer M, Mey N, Metten J, Hartwig HG, Mödder U (2002) Qualitätssicherung von ultraschallkursen in der ärztlichen fort- und weiterbildung: Status-quo-analyse und verbesserungspotential. *Ultraschall der Medizin* 23:189–197
22. Hofer M, Kamper L, Heussen N, Martin O, Heverhagen J (2022) Influence of Clinical Expertise between Clinician-Instructors Versus Student-Instructors on the Effectiveness of Ultrasound Courses. *Ultraschall der Medizin* 43:58–63
23. Huang GS, Sheehan FH, Gill EA (2022) Transesophageal echocardiography simulation: A review of current technology. *Echocardiography* 39:89–100

24. IQWiG (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen) (2008) Ultraschallscreening in der Schwangerschaft: Testgüte hinsichtlich der Entdeckungsrate fetaler Anomalien.
25. Issenberg SB, McGaghie WC, Petrusa ER, Gordon DL, Scalese RJ (2005) Features and uses of high-fidelity medical simulations that lead to effective learning: A BEME systematic review. *Med Teach* 27:10–28
26. Jensen JK, Dyre L, Jørgensen ME, Andreasen LA, Tolsgaard MG (2018) Simulation-based point-of-care ultrasound training: a matter of competency rather than volume. *Acta Anaesthesiol Scand* 62:811–819
27. Kähler C, Schramm T, Bald R, Gembruch U, Merz E, Eichhorn KH (2020) Updated DEGUM quality requirements for the basic prenatal screening ultrasound examination (DEGUM Level I) between 18+0 and 21+6 weeks of gestation. *Ultraschall der Medizin* 41:499–503
28. Kassenärztliche Bundesvereinigung (2021) Qualitätssicherungsmaßnahmen nach § 135 Abs. 2 SGB V zur Ultraschalldiagnostik (Ultraschallvereinbarung) (Anlage 3 BMV-Ä). *Dtsch Arztebl Int* 118:A745
29. KVWL (2019) Ultraschall: Musterdokumentation weiterführende Differentialdiagnostik des Feten (GOP 01773).
30. KVWL (2019) 01773 Weiterführende sonographische Diagnostik II.
31. Le Lous M, Tsatsaris V, Tesnière A, Grangé G (2017) Improving students' ability to perform a standardized foetal biometry plane using ultrasound simulators. *J Gynecol Obstet Hum Reprod* 46:439–443
32. Lindinger A, Schwedler G, Hense HW (2010) Prevalence of congenital heart defects in newborns in Germany: Results of the first registration year of the PAN study (July 2006 to June 2007). *Klin Padiatr* 222:321–326
33. Madsen ME, Konge L, Nørgaard LN, Tabor A, Ringsted C, Klemmensen ° A K, Ottesen B, Tolsgaard MG (2014) Assessment of performance measures and learning curves for use of a virtual-reality ultrasound simulator in transvaginal ultrasound examination. *Ultrasound Obs Gynecol* 44:693–699
34. Madsen ME, Konge L, Nørgaard LN, Tabor A, Ringsted C, Klemmensen AK, Ottesen B, Tolsgaard MG (2014) Assessment of performance measures and learning curves for use of a virtual-reality ultrasound simulator in transvaginal ultrasound examination. *Ultrasound Obstet Gynecol* 44:693–699
35. McGaghie WC, Siddall VJ, Mazmanian PE, Myer J (2009) Lessons for continuing medical education from simulation research in undergraduate and graduate medical education. *Chest* 135:62S–68S
36. McGaghie WC (2015) Mastery learning: It is time for medical education to join the 21st century. *Acad Med* 90:1438–1441

37. Merz E, Thode C, Eiben B, Faber R, Hackelöer BJ, Huesgen G, Pruggmaier M, Wellek S (2011) Kalkulation des Risikos für chromosomale Anomalien im Ersttrimester-Screening unter Berücksichtigung einer individuellen gewichtsadaptierten Korrektur. *Ultraschall der Medizin* 32:33–39
38. Merz E, Eichhorn K-H, von Kaisenberg C, Schramm T (2012) Aktualisierte Qualitätsanforderungen an die weiterführende differenzierte Ultraschalluntersuchung in der pränatalen Diagnostik (= DEGUM-Stufe II) im Zeitraum von 18 + 0 bis 21 + 6 Schwangerschaftswochen. *Ultraschall der Medizin - Eur J Ultrasound* 33:593–596
39. Morris JK, Springett AL, Greenlees R, Loane M, Addor MC, Arriola L, Barisic I, Bergman JEH, Csaky-Szunyogh M, Dias C, Draper ES, Garne E, Gatt M, Khoshnood B, Klungsoyr K, Lynch C, McDonnell R, Nelen V, Neville AJ, O'Mahony M, Pierini A, Queisser-Luft A, Randrianaivo H, Rankin J, Rissmann A, Kurinczuk J, Tucker D, Verellen-Dumoulin C, Wellesley D, Dolk H (2018) Trends in congenital anomalies in Europe from 1980 to 2012. *PLoS One* 13:10–28
40. Nayahangan LJ, Dietrich CF, Nielsen MB (2021) Simulation-based training in ultrasound-where are we now? *Ultraschall der Medizin* 42:240–244
41. Nourkami-Tutdibi N, Tutdibi E, Schmidt S, Zemlin M, Abdul-Khaliq H, Hofer M (2020) Long-Term Knowledge Retention after Peer-Assisted Abdominal Ultrasound Teaching: Is PAL a Successful Model for Achieving Knowledge Retention? *Ultraschall der Medizin* 41:36–43
42. Orr KE, Hamilton SC, Clarke R, Adi MY, Gutteridge C, Suresh P, Freeman SJ (2018) The integration of transabdominal ultrasound simulators into an ultrasound curriculum: <https://doi.org/10.1177/1742271X18762251> 27:20–30
43. Owen H (2012) Early use of simulation in medical education. *Simul Healthc* 7:102–116
44. Patel H, Chandrasekaran D, Myriokefalitaki E, Gebeh A, Jones K, Jevic YB (2016) The Role of Ultrasound Simulation in Obstetrics and Gynecology Training. *Simul Healthc J Soc Simul Healthc* 11:340–344
45. Peachey GC (1924) Note upon the Provision for Lying-in Women in London up to the middle of the Eighteenth Century. *Proc R Soc Med* 17:72
46. Pretscher J, Kehl S, Stelzl P, Stumpfe FM, Mayr A, Schmid M, Staerk C, Schild R, Beckmann MW, Faschingbauer F (2020) Influence of Sonographic Fetal Weight Estimation Inaccuracies in Macrosomia on Perinatal Outcome. *Ultraschall der Medizin* 43:e56–e64
47. Recker F, Barth G, Lo H, Haverkamp N, Nürnberg D, Kravchenko D, Raupach T, Schäfer VS (2021) Students' Perspectives on Curricular Ultrasound Education at German Medical Schools. *Front Med* 8:758255

48. Rosen H, Windrim R, Lee YM, Gotha L, Perelman V, Ronzoni S (2017) Simulator Based Obstetric Ultrasound Training: A Prospective, Randomized Single-Blinded Study. *J Obstet Gynaecol Canada* 39:166–173
49. Salvesen KA, Lees C, Tutschek B (2010) Basic European ultrasound training in obstetrics and gynecology: Where are we and where do we go from here? *Ultrasound Obstet Gynecol* 36:525–529
50. Schmand C, Misselwitz B, Hudel H, Bedei I, Wolter A, Schenk J, Keil C, Köhler S, Axt-Flidner R (2022) Analysis of the Results of Sonographic Screening Examinations According to the Maternity Guidelines Before and After the Introduction of the Extended Basic Screening (Iib Screening) in Hesse. *Ultraschall der Medizin - Eur J Ultrasound*
51. Schwedler G, Lindinger A, Lange PE, Sax U, Olchvary J, Peters B, Bauer U, Hense HW (2011) Frequency and spectrum of congenital heart defects among live births in Germany: A study of the competence network for congenital heart defects. *Clin Res Cardiol* 100:1111–1117
52. Staboulidou I, Wstemann M, Vaske B, Elssser M, Hillemanns P, Scharf A (2010) Quality assured ultrasound simulator training for the detection of fetal malformations. *Acta Obstet Gynecol Scand* 89:350–354
53. Syngelaki A, Chelemen T, Dagklis T, Allan L, Nicolaides KH (2011) Challenges in the diagnosis of fetal non-chromosomal abnormalities at 11-13 weeks. *Prenat Diagn* 31:90–102
54. Taksoe-Vester C, Dyre L, Schroll J, Tabor A, Tolsgaard M (2021) Simulation-Based Ultrasound Training in Obstetrics and Gynecology: A Systematic Review and Meta-Analysis. *Ultraschall der Medizin* 42:E42–E54
55. Tolsgaard M, Velupillai C, Gueneuc A, Taksoe-Vester C, Hajal N, Levailant JM, Ville Y, Tabor A, Chalouhi G (2019) When Are Trainees Ready to Perform Transvaginal Ultrasound? An Observational Study. *Ultraschall der Medizin* 40:366–373
56. Tolsgaard MG, Rasmussen MB, Tappert C, Sundler M, Sorensen JL, Ottesen B, Ringsted C, Tabor A (2014) Which factors are associated with trainees' confidence in performing obstetric and gynecological ultrasound examinations? *Ultrasound Obstet Gynecol* 43:444–451
57. Tolsgaard MG, Ringsted C, Dreisler E, Klemmensen A, Loft A, Sorensen JL, Ottesen B, Tabor A (2014) Reliable and valid assessment of ultrasound operator competence in obstetrics and gynecology. *Ultrasound Obstet Gynecol* 43:437–443
58. Tolsgaard MG, Tabor A, Madsen ME, Wulff CB, Dyre L, Ringsted C, Nørgaard LN (2015) Linking quality of care and training costs: Cost-effectiveness in health professions education. *Med Educ* 49:1263–1271

59. Tolsgaard MG, Ringsted C, Rosthøj S, Nørgaard L, Møller L, Freiesleben NLC, Dyre L, Tabor A (2017) The Effects of Simulation-based Transvaginal Ultrasound Training on Quality and Efficiency of Care. *Ann Surg* 265:630–637
60. Van Nisselrooij AEL, Teunissen AKK, Clur SA, Rozendaal L, Pajkrt E, Linskens IH, Rammeloo L, Van Lith JMM, Blom NA, Haak MC (2020) Why are congenital heart defects being missed? *Ultrasound Obs Gynecol* 55:747–757
61. Wiesel A, Queisser-Luft A, Clementi M, Bianca S, Stoll C (2005) Prenatal Detection of Congenital Renal Malformations by Fetal Ultrasonographic Examination: An Analysis of 709,030 Births in 12 European Countries. *Eur J Med Genet* 48:131–144
62. Ziv A, Root Wolpe P, Small SD, Glick S (2003) Simulation-based medical education: An ethical imperative. *Acad Med* 78:783–788

9. Anhang

9.1. Anhang 1

GynSonoSim-Studie: Fragebogen zur Selbsteinschätzung

NAME:

VORNAME:

GEBDAT:

1. Approbation seit: _____ (Monat) _____ (Jahr)
2. Facharzt Frauenheilkunde/Geburtshilfe seit: _____ // kein Facharzt
3. Welche DEGUM-Qualifikation haben Sie:
 Keine DEGUM I DEGUM II DEGUM III
4. Ultraschallerfahrung
 - Wie viele US-Untersuchungen haben Sie insgesamt ca. _____ durchgeführt?
 - Wie viele gynäkologische/geburtshilfliche US-Untersuchungen haben Sie insgesamt ca. durchgeführt? _____
 - Fetale Biometrien nach MuRL im 2. und 3. Trimenon? _____
 - Organultraschalls nach MuRL 2B / DEGUM I? _____
 - Fehlbildungultraschall / Feindiagnostik (Level DEGUM II) _____
 - Fetale Echokardiographien (mit Ausflusstraktbeurteilung) _____
5. Wie viele Herzfehler haben Sie im Ultraschall gesehen? _____
6. Wie viele Herzfehler haben Sie selbstständig erkannt? _____
7. Welche der folgenden Herzfehler haben Sie schon selbst diagnostiziert:
 - ASD
 - VSD
 - AVSD
 - TOF (Fallot)
 - HLHS
 - Aortenisthmusstenose
 - DORV
 - Pulmonalstenose
 - Aortenstenose
 - Partielle Lungenvenenfehleimündung

8. Welche der folgenden Fehlbildungen haben Sie schon selbst diagnostiziert:

- Pyelektasie
- Omphalozele
- Gastroschisis
- Spina bifida
- Skelettfehlbildungen
- LKG
- Duodenalatresie
- Lungenfehlbildungen
- Nierenagenesie (auch einseitig)

Subjektive Bewertung der Schwierigkeit von Standardebenen

Bitte auf der Achse ankreuzen: Die folgende Ebene einzustellen finde ich...

sehr leicht ←—————→ **sehr schwierig**

Falls Sie eine Ebene bisher noch nicht eingestellt haben oder sie Ihnen unbekannt ist, streichen Sie diese bitte durch.

Ebenen des Zweittrimesterschalls (Fetus)

Medianes Gesichtsprofil

←—————→

Koronares Gesicht (Nasenspitze, Ober- und Unterlippe)

←—————→

Transventrikulär (Seitenventrikel, Vorderhörner)

←—————→

Transthalamisch (BIP/KU-Ebene)

←—————→

Kleinhirn

←—————→

Augenhöhlen axial

←—————→

Zwerchfell

←—————→

Abdomen (für Abdomenumfangsmessung)

←—————→

Fetaler Nabelschnuransatz

←—————→

Nieren Axial



Nieren koronal (Aorta längs dazwischen)



Blase



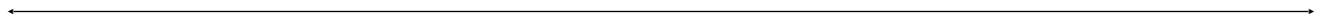
Nacken (Ebene für Nackentransparenzmessung)



Wirbelsäule im Längsverlauf (Ausschluss Spina bifida)



Humerus längs



Radius und Ulna längs



Hand



Femur längs



Tibia und Fibula längs



Fuß



Äußeres Genital



Plazenta inklusive Cervix (Ausschluss P. praevia)



Ebenen der fetalen Echokardiographie

Vierkammerblick



Linksventrikulärer Ausflusstrakt



Rechtsventrikulärer Ausflusstrakt



Dreigefäßblick – A. pulmonalis, Aorta ascendens und VCS als drei separaten Gefäße nebeneinander



Dreigefäßblick – mit Darstellung des Ductus und Aortenbogens



Bikaval – Rechtes Atrium mit Einmündung beider Vv. cavae



Aortenbogen



Duktaler Bogen



Hohe kurze Achse (auf Ebene der Aortenklappe)



Tiefe kurze Achse (Darstellung des Ventrikelseptums unterhalb der AV-Klappenebene)



9.2. Anhang 2

Feedbackbogen für die Simulatorstudie

Wie hat euch eure Teilnahme an der Studie insgesamt gefallen?

Wie gefällt euch der Simulator zum Erlernen von geburtshilflichem Ultraschall?

Hättet ihr euch noch mehr inhaltlichen / theoretischen Input gewünscht? Wenn ja, welche Art von Input (z.B. Auffinden der Standardebenen, Pathologien, Klinisches)? Und könnte dieser von Seiten des Simulators übernommen werden?

Wenn es an der Uni ein Wahlfach mit Simulator-basiertem Ultraschalltraining gäbe, würdet ihr daran teilnehmen?

Wie fandet ihr das Maß an tutorialer Begleitung?

Habt ihr Verbesserungsvorschläge zum Simulator? Was wären Kritikpunkte inhaltlicher oder technischer Art?

Was hat euch im Verlauf der Studie genervt, was fandet ihr gut?

Hier ist noch Platz für sonstige Anregungen / Kritik / Lob / o.ä.

10. Publikation

Teilergebnisse dieser Arbeit wurden auf folgenden Kongressen vorgestellt:

- 64. Kongress der Deutschen Gesellschaft für Gynäkologie und Geburtshilfe e.V., München, 12.–15. Oktober 2022
J. Steinhard, **P. Freundt**, P. Janzing, E. Tutdibi, N. Nourkami-Tutdibi
Effizientes Simulator-Training in der Pränatalmedizin - Erfahrungen aus den FESIM-Studien
- 31. ISUOG world congress: online, 15.10. - 17.10.2021
J. Steinhard, **P. Freundt**, P. Janzing, Th. von Ostrowski, M. Langer, M. Zemlin, E. Tutdibi, N. Nourkami-Tutdibi *Controlled prospective study on the use of systematic simulator-based training with a virtual, moving fetus for learning 2 trimester scan: FESIM III*, Ultrasound Obstet Gynecol., <https://doi.org/10.1002/uog.24309>

Teilergebnisse dieser Arbeit wurden in folgenden Fachzeitschriften zur Publikation angenommen:

- **P. Freundt** ‡, N. Nourkami-Tutdibi ‡, E. Tutdibi, P. Janzing, Th. von Ostrowski, M. Langer, M. Zemlin, J. Steinhard, *Controlled Prospective Study on the Use of Systematic Simulator-Based Training with a Virtual, Moving Fetus for Learning Second-Trimester Scan: FESIM III*, Ultraschall in Med (2023), Georg Thieme Verlag KG, <https://doi.org/10.1055/a-1984-8320>
- J. Steinhard, **P. Freundt**, P. Janzing, V. Popov, R. Menkhaus, L. Ross, *Künstliche Intelligenz und Simulation in der Pränatalmedizin – was wir von Maschinen lernen können*, Die Gynäkologie (2022), Springer, <https://doi.org/10.1007/s00129-022-04996-5>

11. Danksagung

Zunächst bedanke mich sehr bei Herrn Prof. Dr. med. Zemlin für die Supervision meiner Dissertation und die großzügige Ermöglichung der Open-Access Publikation.

In besonderem Maße bedanke ich mich bei Herrn Dr. med. Johannes Steinhard für die Einführung in die sonographische Pränataldiagnostik, die spannenden Hospitationen in seiner Sprechstunde, die Überlassung dieses wichtigen Themas und die Studienleitung. Seine zuverlässigen und schnellen Antworten, die fachlichen Anregungen und Telefonate waren mir eine große Hilfe.

Großer Dank gilt meiner Betreuerin Frau Dr. med. Nasenien Nourkami-Tutdibi für die wissenschaftliche und herzliche Unterstützung in Publikation und Dissertation. Durch ihre Betreuung war es möglich, an meiner Heimatuniversität zu promovieren.

Des Weiteren möchte ich mich herzlich bei Herrn Dr. med. Erol Tutdibi für die Unterstützung in den statistischen Analysen, die Ausgestaltung von Graphiken sowie die fortlaufende und freundliche Beratung bedanken.

Ganz besonders danke ich meinem besten Freund und Co-Doktoranden Paul Benjamin Janzing für die gemeinsame Umsetzung der Studie und den steten Austausch während des Schreibprozess.

Ich bedanke mich bei allen Studierenden, die sich neben Examenslernplänen, Famulaturen und PJ für das Training am Simulator motiviert und kontinuierlich teilgenommen haben. Ebenso danke ich den Ärzt*innen, die durch ihre Studienteilnahme diese Arbeit möglich gemacht haben.

Vielen Dank an meine lieben Freund*innen für eure Geduld und offenen Ohren. Vielen Dank auch an Burgis Janzing für den Hinweis auf die Doktorand*innenstelle.

Meiner Familie danke ich für ihre große und liebevolle Unterstützung im Studium und während der Doktorarbeit. Ich möchte mich herzlich bei meiner Mutter Sabine und meiner Schwester Hanna für die Durchsicht des Textes bedanken.

12. Lebenslauf