EVALUATING VIRTUAL REALITY SIMULATION FOR TRANSVAGINAL SCAN TRAINING IN GYNAECOLOGY

by

NATALIE JANE WOODHEAD

A thesis submitted to The University of Birmingham for the degree

DOCTOR OF MEDICINE

School of Clinical and Experimental Medicine

University of Birmingham

March 2018

ABSTRACT

Ultrasound has a broad range of clinical applications and is increasingly being used at the point of care. With growing use comes an increasing need to train health professionals to perform ultrasound proficiently. This thesis appraises the evidence for the use of simulation in gynaecological ultrasound scan training. The thesis explores the role of simulation in training, the validity of a virtual reality simulator for transvaginal ultrasound, and its use in real clinical practice. As transvaginal ultrasound is an intimate examination the opinions of women on having a scan and training health professionals were sought.

A systematic review showed that use of simulation training in gynaecological ultrasound when compared with standard theoretical and clinical approaches is associated with significant improvements for both learners and patients. Outcomes such as acquisition of skill, confidence and competence favour incorporating simulation into training, as does reduction in patient discomfort. An attempted pilot feasibility study, challenged the notion that simulation using the Medaphor Transvaginal ScanTrainer® is effective as a self-directed tool. However, the simulator did demonstrate construct validity and was felt to be useful in teaching a systematic approach to performing gynaecological ultrasound. Women were supportive of simulation to reduce discomfort and increase trainees confidence in clinical practice.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABREVIATIONS

2D	Two Dimensional
3D	Three Dimensional
4D	Four Dimensional
BMUS	British Medical Ultrasound Society
BNI	British Nursing Index
CINHAL	Cumulative Index of Nursing and Allied Health
CRL	Crown Rump Length
CT	Computed Tomography
DUC	Deanery Ultrasound Co-ordinator
EAT	Error avoidance training
EMBASE	Excerpta Medica dataBASE
EMT	Error management training
EPAU	Early Pregnancy Assessment Unit
ERIC	Institute of Education Sciences
FAST	Focussed Abdominal Sonography for Trauma
GMC	General Medical Council
HEE	Health Education England
HyCoSy	Hystero Contrast Sonography
MRI	Magnetic Resonance Imaging
NHS	National Health Service
O&G	Obstetrics and Gyanecology
OSAT	Objective Structured Assessments of Technical Skill
OSAUS	Objective Structured Assessment of Ultrasound Skills
RCOG	Royal College of Obstetricians and Gynaecologists
RCR	Royal College of Radiologists
SBME	Simulation Based Medial Education
SCoR	Society and College of Radiographers
TA	Transabdominal
TV	Transvaginal
TVU	Transvaginal Ultrasound
UK	United Kingdom
UoB	University of Birmingham
US	Ultrasound
USB	Universal Serial Bus
VR	Virtual reality

CHAPTER 1

Introduction and Objectives

This thesis will evaluate virtual reality simulation for teaching gynaecological ultrasound to trainee obstetricians and gynaecologists.

Importance of transvaginal ultrasound in modern gynaecology practice

Transvaginal ultrasound (TVU) scanning involves the insertion of an ultrasound probe into the vagina to give high-resolution images of the female pelvic organs. Traditionally pelvic examination was done blindly, with a vaginal speculum and bimanual palpation. The value of such examination in asymptomatic women has been questioned [(1)] and ultrasound is being increasingly applied as a particularly useful method for the gynaecologist to image the pelvis. The first practical endovaginal scan probes came into practice in 1985 and gained popularity in the late 1990s,[(2)] use has become widespread since then. The transvaginal route allows the probe to be placed nearer to the organs of interest and, coupled with the use of higher frequency ultrasound, results in better image quality and resolution compared with transabdominal imaging [(3)] and does not require a full bladder. Ultrasound has obvious advantages over other imaging modalities such as magnetic resonance imaging (MRI) and computed tomography (CT) in availability, cost, examination time and avoidance of exposure to potentially carcinogenic ionising radiation. [(4)]

Applications of transvaginal ultrasound to clinical practice

TVU is ideal to assess the uterus, endometrium, ovaries, bladder and adnexa.[5] It has uses in early pregnancy, acute gynaecology, outpatient gynaecology and assisted reproduction. Ultrasound has transformed the care and management of first trimester pregnancy problems where it is increasingly used with no data to suggest harmful effects at diagnostic levels. [6] In particular TVU can accurately identify a pregnancy from as early as five to six weeks' gestation. It is used to investigate women who present with early pregnancy complaints such as pain and bleeding. As it identifies the location of a pregnancy at such early gestations it can allow rapid diagnosis of early intrauterine pregnancy and potential ectopic pregnancy. The reported sensitivity and specificity in diagnosis of ectopic pregnancy with TVU are 87-99% and 94-99%.[7] Ectopic pregnancy has an incidence of approximately 11 per 1,000 pregnant women, meaning each year in the UK nearly 12,000 women are diagnosed. Unfortunately women still die from ectopic pregnancy, with four maternal deaths reported between 2013 and 2015, however, the case fatality rate has decreased a factor in this reduction could be earlier diagnosis.[8] The earlier diagnosis of ectopic pregnancy, before tubal rupture and its associated life threatening complications, has allowed a shift to medical and conservative approaches to the management of ectopic pregnancy, although surgery remains the mainstay of treatment.[9]

In the office or gynaecology outpatient setting TVU can be used to evaluate women with menstrual disorders to detect uterine fibroids and polyps and help to guide appropriate therapy. [5] By allowing examination of endometrial thickness TVU is also the most effective way of evaluating postmenopausal bleeding [10] and can be used to effectively triage postmenopausal women into low and high risk groups, allowing targeted investigation of only the high risk group and avoidance of unnecessary intervention in the low risk. Ultrasound, in particular TVU can be used to

assess ovarian cysts and adnexal masses to determine their dimensions and internal architecture. This morphological assessment of ovarian masses, combined with information about menopausal status, can allow a gynaecologist to differentiate the likelihood of benign and malignant ovarian cysts. In expert hands, the morphological description of an ovarian cyst outperforms malignancy risk scoring algorithms that use biochemical markers.[11] Ultrasound has also played a major role in the development of modern investigation and monitoring of treatment in women with infertility.[5] It allows investigation of uterine, endometrial, ovarian and tubal morphology. Ultrasound combined with contrast agents can be used to assess tubal patency (Hystero Contrast Sonography HyCoSy) with advantages over traditional tubal patency tests, which involve risks of x-ray irradiation or surgical risks from laparoscopy. During fertility treatment, TVU is used to track follicular development and ultrasound guided oocyte retrieval is a standard procedure for in vitro fertilisation treatment.

Growth in point of care ultrasound practice

When ultrasound was first introduced to clinical practice it was largely carried out in the radiology department, overseen by radiologists and done by sonographers. Over subsequent years ultrasound machines have become more compact, higher quality and less expensive. The clinical application of ultrasound has expanded and with that has come increasing demand. This has led to growth in point of care ultrasound practice: ultrasound carried out and interpreted by the clinician at the bedside.[12] It involves a specialist, who may not have had extensive training in imaging, performing targeted investigations or procedures. Given its clinical value, its ease of use, and relative safety, ultrasound is changing the way gynaecologists, and doctors from many other

specialities, examine and diagnose their patients. The concept of an 'ultrasound stethoscope' is said to be moving from the theoretical to the possible. [13, 14] The advent of point of care ultrasound has allowed development of new models of care in women's health, with 'one-stop' see and treat practice increasingly been used and favoured by women. [15]

As a consequence of the expansion in application of ultrasound and movement to point of care practice there is an increasing demand for ultrasound education and training.[16] Ultrasonography is a user dependent technology and there are concerns that as usage spreads, competence needs to be ensured and use appropriate to limit unnecessary imaging and its consequences.[12, 17] Point of care scanning is difficult to regulate, ensuring that only certified practitioners perform scans within a proper governance system, with agreed protocols, good quality written reports, archived images, monitored training and supervision is a real challenge.[14] A practitioner may perceive it is safe and beneficial for patients to have a scan but fail to recognise that poor quality scanning can put patients at risk, through missed or incorrect diagnosis. A recent illustrative example of this point came in 2011 when a series of papers provided evidence that guidelines on the diagnosis of miscarriage were unsafe.[18] One study showed significant intra and interobserver reproducibility using transvaginal ultrasound to measure gestational sac and crown rump length.[19] This led to rapid amendments to Royal College of Obstetricians and Gynaecologists (RCOG) guidance to more conservative and safer diagnostic values in 2011.

In the context of early pregnancy scanning variations are seen between centres in the reported sensitivity and specificity of ultrasound to diagnose ectopic pregnancy [7] and the proportion of non-diagnostic scans depending on level of expertise in

ultrasound.[20] Similarly subjective assessment by ultrasound is valuable in discriminating malignant from benign ovarian cysts by pattern recognition in expert hands but it is recognised this may not be universally achievable in all clinical settings without such expertise.[11, 21] Diagnostic error can only be overcome by proper training, understanding both the limitations and the potential of ultrasound equipment.[17]

Provision of ultrasound training in gynaecology

Provision of training in the UK

High quality training of operators in the use of ultrasound is key to its safe and effective use in clinical practice. Gynaecologists were historically self-trained in ultrasound; this is no longer acceptable [22] and a move to include ultrasound scanning as part of the curriculum for trainee obstetricians and gynaecologists reflects this. Pelvic ultrasound scanning is part of the RCOG curriculum in the 'intermediate ultrasound in gynaecology' and 'intermediate ultrasound in early pregnancy' modules. These are optional modules usually done following a compulsory basic ultrasound module. Trainees will choose them if they have a particular interest in developing advanced training skills where ultrasound is beneficial. The modules use both the transabdominal (TA) and transvaginal [23] route, however, for many undertaking these modules this will be their first experience of performing TVU. Included within the curriculum are knowledge requirements for safety, machine set-up and operation, principles of report writing as well as recognition of relevant pathology. The modules are assessed using a competency-based curriculum and logbook with work-based assessments (objective structured assessments of technical

skill [OSATS] and case-based discussion) as evidence of training rather than a numbers-based logbook.[24]

Radiologists are trained through the Royal College of Radiologists (RCR). Similar to the RCOG system trainees with a particular interest in gynaecology would choose to undertake an optional uro-gynaecological module. This can be to level one or level two competences, with level two producing experts who focus the majority of their caseload in gynaecology scanning. Again similar to the RCOG training the module is competency based and the methods of assessment prescribed are both formative (workplace-based assessment) and summative (formal professional examinations).

In the UK, trained sonographers should have a Postgraduate Certificate or Diploma in Medical or Clinical Ultrasound (or the older Diploma of Medical Ultrasound awarded by the College of Radiographers). Training for a Postgraduate Certificate or Diploma in Ultrasound takes between 12-18 months to complete and includes modules in gynaecology. Short focussed courses are available to allow other professionals, such as nurses working in early pregnancy, to train in ultrasound. A Postgraduate Certificate or Diploma usually requires a logbook of cases to be kept and a specified number of hours to be spent scanning. In practice sonographers perform the majority of the workload therefore quickly gain experience so often have advanced technical skill and departments generally have a robust system of supervision of practice by senior sonographers and radiologists.[25]

For doctors registered with the RCR, the training they receive in gynaecological ultrasound is delivered mostly by sonographers; there are relatively few consultant radiologists with this subspecialty interest.[25] The technical training they receive is to a high standard and they can often do more general imaging, for example they could

scan the abdomen if the pelvic findings dictate that this is necessary. As the modules in gynaecological ultrasound for obstetricians-gynaecologists and radiologists are optional, uptake of training may be influenced by availability of training and role models with the expertise to train. Within the specialties in the forthcoming years who is more likely to provide high-level diagnostic ultrasound services, radiologists or gynecologists, may depend on availability and uptake of training.[25]

European training

There is a generally held belief that it is beneficial for all obstetricians and gynaecologists to be able to scan to a basic level in modern practice. [17] In the UK this is demonstrated by incorporation of scan training into a competence-based curriculum. However there are separate, independent ultrasound training approaches across individual countries in Europe, and this creates a problem of inconsistency in training standards. The UK is the only country to have a competence based system for training of speciality doctors in gynaecological ultrasound, other countries require a logbook or specified number of scans for certification, this number can range from 50 scans in Norway to 500 in Germany. [17] In Denmark a logbook of 90 scans is required and this is achieved through focussed clinical supervision, during specialist ultrasound training over a 3-week period, which has been shown to be associated with confidence in performing gynaecological scans independently.[26] There exists not only variation in the requirements of training but also the approach to ultrasound training and the delivery. In European countries such as Denmark, Sweden and Norway it seems trainees are commonly allowed and indeed expected to perform more TV scans early in their gynaecology training than TA scans which are done by more experienced practitioners.[26] In contrast in the UK trainees are first exposed to

TA scanning, although are expected to do only a very basic level scan independently (for example third trimester presentation and placental localisation) then training in TV scanning generally starts at an intermediate training level once in year 3 or above of specialty training.[24] Another stark contrast between different European countries is the estimated percentages of obstetrics and gynaecology scans performed by sonographers and midwives. In the UK it is estimated over 90% of scans are done by sonographers and midwives and less than 5% by obstetricians and gynaecologists, where as in Germany and France the pattern is the complete opposite with 70-85% of scans performed by obstetricians and gynaecologists and less than 10% done by nurses and sonographers.[17] These findings have implications on the training opportunities available for obstetricians and gynaecologists in an apprentice style-learning model and who will be delivering the training. It also has implications for the necessity of training obstetricians and gynaecologists to scan, and allowing them to maintain their skills once trained in a system designed around speciality doctors doing only a small proportion of the imaging.

Ultrasound training challenges

There is some worry about delivery of ultrasound training to obstetricians and gynaecologists in the UK. In 2010 when the existing ultrasound-training programme was launched, the RCOG released a joint statement with the Society and College of Radiographers (SCoR) and the British Medical Ultrasound Society (BMUS). [27] In this statement sonographers' contribution to training was recognised as substantial and that this involvement would need to be sustained if training programs are to be delivered effectively. Furthermore, the widespread application of ultrasound in

obstetrics and gynaecology means that local and regional training programs need to accommodate a variety of trainees in addition to obstetricians and gynaecologists. These include other doctors such as trainee radiologists as well as sonographers, nurses and midwives. The statement appreciated that the viability of training programmes was dependent upon the number of qualified trainers and the departmental capacity to deliver training. A problem arises when you do not have enough trained operators in the workforce to deliver the training. It is estimated by the SCoR that there is at least a 12% vacancy rate across the UK, which cannot be filled with qualified sonographers.[28] Compounding the problem of workforce shortage is the rise in demand for ultrasound. The Centre for Workforce Intelligence, commissioned by the Department of Health, projects a 45% increase in demand for ultrasonography by 2025.[29] This means departments are under strain to meet demand for ultrasonography. The demand of training adds further departmental strain as having a trainee scanning under supervising increases the time taken to complete the examination,[30] meaning there is a real problem in training the workforce of the future.

TVU is constantly gaining popularity amongst obstetricians and gynaecologists as a valuable aid to their clinical practice given its applications but this is a relatively new skill, compared to other traditional competencies in the curriculum. Hence, the apprenticeship approach to learning used in other areas is not as applicable to ultrasound,[31] due in part to the paucity in experienced gynaecologists that are already doing it. This leaves a situation in which there is a bottle neck effect on training, with a large number of trainees wanting to learn; as they recognise ultrasound as a valuable skill, but limited opportunities to train; given the limited number of operators and the demands on ultrasound departments. Ideally you want to be able to get trainees

quickly up skilled in ultrasound, but ultrasound is a skill that takes time to acquire and relies on pattern recognition for identification of pathology. Training also need to be ensured to a good standard, as ultrasound can be dangerous with risk of progression of unrecognised disease or inappropriate intervention.[32] Given the challenges of delivering ultrasound scan training in the clinical setting, moving training away from the clinical setting and allowing trainees to practice on a simulator is very appealing as a potential solution.

Simulation in healthcare

Simulation is a developing area of healthcare and medical education.[33] The history of simulation stretches back over centuries. In the modern day simulation is used in many high-risk industries. The aviation, military and nuclear power industries adopted the use of simulation training for non-technical skills training of crew and resource management in an emergency; given the alternative of real world training would be dangerous and expensive. These industries have demonstrated simulation training to be highly effective and a major contributor to improved safety.[34, 35] Parallels can be drawn between the high stakes emergency scenarios faced in the aviation industry and the health industry; therefore it is unsurprising simulation based medial education (SBME) is gaining popularity in healthcare.

The Society of Simulation in Healthcare defines simulation as the 'imitation or representation of one act or system by another'. They go on to state that healthcare simulations can be said to have four main purposes – education, assessment, research and health system integration in facilitating patient safety.[36] Simulation is often referred to as 'high' or 'low', fidelity in this referring to the degree of realism, and although a commonly used term in the SBME literature it is not well defined or

straightforward to explain. Maran and Glavi [37] have suggested, fidelity can be assessed on two levels: 'engineering fidelity' or authenticity (does the simulation look realistic?), and 'psychological fidelity' (does the simulator contain the critical elements to accurately simulate the specific behaviors required to complete the task?).

The modern era of medical simulation from the 1950s onwards is described by Bradley [35] to have three distinct movements. The first is the resuscitation movement; starting with the production of Resusci-Anne [38] a low cost, yet effective training model for resuscitation. The second is the development of more sophisticated simulators to reproduce aspects of the human patient. Sim One [39] was a simulator produced in the 1960s, which could breath, had a heartbeat and would show a physiological response to drug and anaesthetic gases. As the technology became more sophisticated these models were superseded. In the 1980s, David Gaba, who is considered a pioneer in medical simulation,[40] along with his colleagues developed the comprehensive anaesthetic simulation environment. [41, 42] They focussed on training clinical teams in simulated environments that closely resembled the real workplace. Part of the curriculum copied elements from aviation models of simulated training to manage 'crisis' situations. [43] This novel approach led the way in the more widespread provision of higher fidelity simulation in healthcare. The third movement is described as medical education reform. In the later part of the 20th century Bradley [35] explains the recognition that theory alone was not enough to prepare undergraduate students for the real world of being a doctor, they needed to learn clinical skills. At the same time there were also changes to postgraduate education with the need for continuing medical education and drive to revalidation. This has seen a rise in simulation methodologies in both undergraduate and postgraduate education.

There has been a cumulative growth in SBME literature since the 1980s.[35] In 2005 Issenberg et al. conducted a review of the features and uses of high fidelity medical simulations that lead to effective learning and noted that few journal articles were published in the decades of 1970 and 1980 however beginning in the early 1990s the growth became exponential.[44] The growing body of interest and research is further evidenced by the development and growth of societies dedicated to simulation in medical education.[45] Simulation is being used because it provides a safe and supportive educational environment.[46] It allows users of all levels from novice to expert to develop skills without the pressure and risk of the real clinical environment and encourages the acquisition of skills through experience.[47] Simulation is a broad term, which covers a huge range of different scenarios. Medical simulation can be divided into four main areas: [35, 40, 48]

Simulated patients and environments

These have become commonplace in the last few decades, particularly used for undergraduate education to assess communication skills. Simulated patients may be role players or real patients trained as 'professional' patients. Simulated environments could be recreated or real environments not in clinical use; a realistic environment can increase learner's engagement.

Partial task simulation

These are models that usually represent a single part of the human body. They are used for procedural based learning of skills such as venepuncture and suturing. They allow the learner to focus on a single task; they could also be cleverly incorporated into a scenario with a real patient so communication as well as procedural skill could be assessed.

Computer based simulation or virtual reality (VR) simulation

This covers any form of simulation using a computer display and user interaction. It incorporates (i) multimedia programmes, which use images, audio and video as adjuncts to learning; (ii) interactive computer systems, which depending on the users actions will manipulate scenarios and provide feedback on decisions made and actions taken; (iii) sophisticated computer technology in virtual reality (VR). VR refers to the production of computerised images of real things or environments and these are often combined with haptic systems which provide the sensations of touch and feel when undertaking a procedure. In endoscopic surgery such VR systems are often combined with a training model.

Integrated simulation

Refers to computer-enhanced mannequins, which are human mannequins under sophisticated computer control. The simulator response can be manipulated by the trainer or vary according to the user's actions. This technology ranges from low to moderate or high fidelity simulators. A well-known example of such a model is SimMan.

SBME is used in multiple medical specialties including anaesthetics, emergency care, cardiology, obstetrics and surgery.[48] Simulation has been used for skills learning and non-technical skills learning. It lends its self to rehearsing rarely encountered emergency scenarios as well as technical skills in areas such as surgery, particularly minimal access surgery. Surgical skill requires the development of psychomotor competencies, a process that requires regular practice. During the last two decades the use of VR simulators has become a key element in many surgical training programs, and considerable amounts of time and monetary resources are now invested in SBME

for technical skills training.[49] Parallels can be drawn between ultrasound skills learning and minimal access surgical skills in the complexity and need for a combination of motor skills and visual-cognitive skills. VR simulation training for surgical trainees in laparoscopic surgery appears to reduce the operating time and improve the operative performance. [50] Given the parallel skills needed for laparoscopy and ultrasound it is possible that the same effect would be seen for ultrasound training on a VR simulator.

Ultrasound simulators

Ultrasound is a technical skill that is acquired through experience. Simulation for ultrasound is appealing as some of the hours needed to train could be done away from the pressured clinical environment. There are various types of simulator for ultrasound training. These will be discussed and classified according to the methods used to simulate the ultrasound (US) image, the different user interfaces simulators have and the output device used. [40, 51]

Phantom mannequin models

For ultrasound training using simulation physical phantoms can be used. These are models that mimic usually one area of the human body, such as a single organ or can be more complex models. They are made of different materials such as foam, gelatin and rubber.[52] These materials mimic the acoustic properties and therefore ultrasound appearance of real tissue. Phantom models made using these materials can be used with commonplace ultrasound probes and machines. Some examples are the plastic fetal model, 'Space Fan ST', proposed for training on routine second-trimester screening by the Japanese company Kyoto Kagaku [53] and the 'Blue phantom'

models produced by CAE Healthcare of which there are models for all specialties, including a transvaginal simulator for gynaecology.[54]

Computer models

Computer based simulators do not use a real ultrasound probe, they use a mock probe and simulate the ultrasound image in the computer. Many different methods for computer based ultrasound simulation exist. They can be classified according to the methods used to simulate the ultrasound images.[40, 51]

Methods used to simulate ultrasound images:

i) Interpolative simulation

This common method simulates 2-dimensional (2D) ultrasound images by interpolation usually from 3-dimensional (3D) volumes that have been acquired from real patients examinations. The image obtained is realistic when they are simulated from a position similar to the position used to acquire the 3D volume. [50] However the quality and resolution of obtained images can be compromised with this method of simulating ultrasound because the images are acquired in a single plane and the other planes are then reconstructed. Furthermore, the probes movements during simulation may not resemble those utilised during the real examination. It can also be difficult to simulate image view dependent artifacts or remove existing artifacts.

ii) Generative image based

This method uses technology to reconstruct a 2D image from other images such as CT or MRI volumes. The image quality depends on simulation methods but in general is lower quality than interpolative methods. View dependent artifacts can be

considered. Although large numbers of CT and MRI images are available image based simulation has limitations, in particular for ultrasound of small moving anatomy such as the fetus and the heart.

iii) Generative model based

For this method a 2D image is reconstructed from computer software while the user scans a 3D or 4-dimensional (4D) model, this overcomes some of the problems of generative image based simulation. Models have been made of the heart mimicking the movement of the valves; extracting a 2D slice from the model and re-texturing it creates the 2D ultrasound image. Its quality is dependent largely upon the angle applied between the mock probe and the phantom. It has been said that such a model offers 'cartoon-like' images, with a certain degree of simplification, an issue that has yet to be overcome. This model does not yet fully reproduce the difficulties encountered in real scanning; such as posterior shadowing arising from bone structures and this method cannot offer a range of cases. It is usually based around a single case; a lot of work would go into modeling various pathologies.

User interface

The user interface refers to the methods simulators use for tracking the position of the mock ultrasound probe, the simulation of haptics and the output device, which is usually a standard computer monitor.

Mock probe

Many systems use a patient mannequin and mock-probe that is connected to a computer; the computer screen displays US images depending on the probes position and movements. The use of electromagnetic tracking to define the

probes position is the commonest, particularly where the ultrasound transducer is inserted into the patient. The mock probe usually contains a 3D sensor, capable of acquiring virtual position data instantaneously. As the data is transferred to the computer it evaluates location and position of the probe on the mannequin and displays correlating 2D ultrasound images. An alternative to electromagnetic tracking is optical tracking, however electromagnetic devices have the advantage, as they do not require a free line of sight to the tracked objects. [51]

Haptic force-feedback devices

A haptic device can be used instead of a mannequin and a mock probe, allowing measurement of the pressure applied to the probe and providing realistic feedback on this force. With certain techniques this can be used to simulate the deformation of the image as a result of the pressure applied to a patient. However haptic devices can have limited working range of movement and need more abstraction from the user because there is no mannequin representing the patient.[51]

Output devices

Most systems display the 2D ultrasound image on a standard computer monitor. Systems not using a patient mannequin have to visualise the relative position of the probe and the patient. Most of the haptic systems provide a 3D animated illustration of the anatomy surrounding the probe and its position and relations with the different organ systems.

Medical applications of ultrasound simulators

Simulators for ultrasound are used in many areas of medicine for diagnostic ultrasound as well as interventional procedures. Many simulators have been developed to teach the basic skills of cardiac ultrasound examination.[55] Ultrasound simulators have been used in transthoracic echocardiography and the more challenging invasive procedure to learn, transoesophageal echocardiography. Several studies have investigated the effectiveness of simulation-based-echocardiography training compared with conventional methods such as theoretical lectures and hands on training. Studies have been critised for being underpowered due to limited sample size but they do show high levels of compliance and satisfaction amongst users and suggest significant improvements amongst trainees exposed to simulation in terms of anxiety levels, performance, efficiency, competency and recognition of pathology.[55]

Simulation has been used for scanning of the abdomen for general diagnostic ultrasound and focussed abdominal sonography for trauma (FAST) scanning. Terkamp et al.[56] showed validity for a simulator that covers different pathologies in the right upper quadrant of the abdomen. In trauma assessment FAST scanning has gained popularity as a point of care ultrasound diagnosis.[57] The UltraSim simulator has found to be a convenient and objective method of introducing surgery residents to ultrasound that compared favourably to hands on experience.[58] Studies are now being published demonstrating validity of simulators for FAST scanning.[59, 60] Presumably given the more intimate nature of the examination ultrasound simulators for mammography have been developed, there are mannequin models that allow practice of ultrasound as well as needle guided procedures. There are also more sophisticated models being developed using VR simulation.[61].

One of the main areas ultrasound simulators have been studied is in obstetrics and gynaecology[40] given the importance of ultrasound scanning in the speciality and perhaps the more intimate and invasive nature of a transvaginal ultrasound examination.

The focus of the thesis is on simulation for gynaecological ultrasound. The commercially available ultrasound simulators for gynaecology are presented below and the evidence for their effectiveness is presented in chapter two as a systematic review of the literature.

Overview of available ultrasound simulator systems

There are now a number of commercially available simulator systems for ultrasound.[40, 51] Most commercial systems provide multiple cases for training often with varying difficulty levels. The basic system is often purchased then additional modules and cases can be purchased separately or as packages.

Much of the literature pertaining to ultrasound simulation deals with the "higher-stakes" interventional or transvaginal procedures in which patient safety or issues with consent for intimate examinations may limit trainee performed procedures.[62] Transabdominal ultrasound training that is relatively safe, not particularly uncomfortable to women, and may even be appealing for routine obstetric examinations when the baby is viewed is less likely to attract either the attention of training programme directors or the interest of commercial simulator developers. Therefore it is not surprising that there are numerous transvaginal simulator models

commercially available. The table below, Table 1, details a summary of the available ultrasound simulator models for transvaginal ultrasound scanning.

Bluephantom™ produce ultrasound training models using the patented 'SimulexUS tissue' which is designed to match the acoustic properties of human tissue. The models can be used with any ultrasound machine; you simply scan the model to get the image on screen. For transvaginal gynaecology there are models available for general pathology, ectopic pregnancy, intrauterine pregnancy and sonohysterography/sonosalpingography.

Unlike many of the other models, which use a tracked probe, Transvaginal ScanTrainer® uses a haptic probe with force feedback technology to replicate the feel of a real ultrasound scan performed on a patient. The images displayed are real patient scans from 3D data. The simulator software system has a learning management system with allows customisation of modular learning, and means users and tutors can track their progress.

The Schallware US simulator® allows practice at ultrasound examinations using a tracked probe, which is moved over a model, for example of the female pelvis. The system uses data from 3D volumes to produce 2D B mode images. The core system sold by the company comes with a dummy torso, three modules (each with up to 12 cases) and the corresponding probes. The user scans a model torso with the provided probe, which is plugged into a computer model. This system is also designed, like ScanTrainer® around cases to allow an element of self directed learning, there are medical histories, questions and descriptions of pathology built into the modules.

Sonosim® like other models on the market uses a tracked probe and data from 3D volumes to produce 2D B mode images. Sonosim® probes rather than being used with a large mannequin can be used on a small pad. In fact SonoSim® can be used from the comfort of your own home, they offer a package for personal use where you can purchase the SonoSim® probe, SonoSim® drive on a USB and a scan pad. The USB is plugged into a personal computer, allowing you to work through the library of cases and test your knowledge while practicing ultrasound skills.

UltraSim was one of the first simulators on the market. The Ultrasim® simulator uses a full size mannequin model alongside probes which use 3D position sensors to tell the UltrasSim® computer the position and angle at which the body volume is being scanned and the corresponding image is displayed. The UltraSim® compuer display is made to look most like a real ultrasound machine and the controls mimic those of the real machine.

U/S Mentor® VR and Vimedix® use a mannequin and a tracked probe to send information to a computer to display corresponding software generated images. These simulator companies are now also moving into immersive virtual reality (VR) immersion where by the user wears lenses to see the anatomy come to life.

Most of the above simulators are designed in a way that they have in built instructions and work packages so that users can interact with them without necessarily the presence of a tutor allowing unlimited practice on a variety of cases. Companies are also making the technology more compact and mobile meaning it could be used anywhere.

Table 1 Summary of the key aspects of commercially available simulator systems for transvaginal scanning

System	Developer and manufacturer	Type of simulation	Data Acquisition	User interface
Bluephantom™	CAE Healthcare Sarasota, USA and Ville St Laurent, Quebec, Canada	Mannequin	3D data from phantom	Mannequin with real ultrasound probe
Transvaginal ScanTrainer®	Medaphor Ltd, Cardiff Medicentre, Wales, UK	Computer model	3D data from real exam	Haptic device and probe
Schallware US simulator®	Schallware GmbH, Berlin, Germany	Computer model	3D data from real exam	Mannequin and tracked probe
Sonosim®	SonoSim Inc. California USA	Computer model	3D data from real exam	Tracked probe
Ultrasim®	Medsim Ltd, Kfar Sava, Israel	Computer model	3d data from real exam	Mannequin and tracked probe
U/S Mentor® VR	3D Systems (formally Simbionix), USA	Computer model	Software generated images	Mannequin and tracked probe
Vimedix®	CAE Healthcare Sarasota, USA and Ville St Laurent, Quebec, Canada	Computer model	Software generated images	Mannequin and tracked probe

SUMMARY

Transvaginal ultrasound is an extremely useful imaging modality in gynaecology with many applications. There is an increasing demand for ultrasound imaging and point of care ultrasound practice is expanding. To meet the demands for imaging in modern gynaecology many practitioners need to be trained in the safe and effective use of ultrasound, however training is time consuming and challenging to deliver. Simulation in medical education is a growing field and simulation based medical education (SBME) is shaping the future of medical education. During the past two decades the use of virtual reality simulators has become a key element of many surgical training programmes for learning technical skills. Ultrasound is similarly a complex skill to master. The last decade has seen expansion in the development of virtual reality simulators for ultrasound training. There are many systems commercially available. At the time of developing and embarking upon the work that forms this thesis, although commercial systems were available and some investment had been made in the systems, there was a scarce evidence base for their evaluation or use in training, which was the motivation for the work presented in this thesis.

OBJECTIVES OF THE THESIS

1. To summarise and critically appraise the available research on the effectiveness of simulators for training in gynaecological ultrasound.

2. To explore the national status of simulation availability and use in ultrasound training for obstetricians and gynaecologists

3. To assess the efficacy and cost-efficiency of VR simulation for transvaginal ultrasound scan training

4. To validate virtual reality simulation in assessing transvaginal-scanning skills in gynaecology by establishing face, content and construct validity.

5. To explore women's views and experiences of transvaginal ultrasound examination, with a particular focus on training professionals in transvaginal ultrasound.

CHAPTER 2

Ultrasound simulation for gynaecological ultrasound training: a systematic review

ABSTRACT

Introduction

Simulation is a potentially promising solution to some of the challenges in delivering gynaecological ultrasound training. Simulators are available for gynaecology, and in the last decade research in simulation-based education has rapidly expanded. However there have been no prior systematic reviews on the effectiveness of simulation to teach gynaecological ultrasound. The aim of this review was to summarise and critically appraise the available research on the effectiveness of simulators for training in gynaecological ultrasound.

Methods

Medline, EMBASE, CINHAL, BNI, PubMed, ERIC, Cochrane library and ISRCTN Register of Clinical Trials were searched using selected terminology. No language restrictions were applied. The selection criteria were studies of all types that investigated the use of a simulator to teach students or health-related professionals gynaecological ultrasound. Data evaluating study outcomes, along with methodological details, were extracted in duplicate. The outcomes were classified according to the Kirkpatrick hierarchy and included: patient related outcomes, such as

discomfort; organisational efficiency outcomes, such as examination times; and trainee related outcomes, such as competence, skill and confidence.

Results

Fourteen studies with 573 participants were included: seven randomised controlled trials, three controlled observational studies and four uncontrolled observational studies. The use of simulation practice in gynaecological ultrasound in addition to standard approaches for teaching is associated with significant improvements in competence, skill and confidence of trainees and reduced patient discomfort, increased perceived patient safety and enhanced confidence of the trainee.

Conclusion

Our findings suggest that simulation in the early stages of training in gynaecological ultrasound is beneficial for the trainee, patients and has organisational benefits through reduced examinations times and need for supervision. However future research needs to confirm these findings in different settings.

INTRODUCTION

Pelvic ultrasound is a valuable diagnostic tool but ensuring high quality training faces the generic challenges of teaching any craft skill requiring patient contact, namely time, supervision and willing participants. However, achieving proficiency in transvaginal ultrasound through instruction on patients is even more problematic given its intimate nature. The advent of simulation technology for ultrasound offers a promising solution to some of those challenges. Technology enhanced simulation has widespread appeal and many assert its educational utility [33] but some question the evidence base for these beliefs.[63]

Over the past three decades, but particularly in the last ten years, there has been rapidly growing interest in simulation based medical education (SBME). In 2005 Issenberg et al.[44] carried out a systematic review of high fidelity medical simulators for effective learning and identified 109 studies for inclusion, then by 2011 Cook et al.[63] in their review added a further 500 studies, illustrating the rapid expansion of published literature on the subject.[44] The extensive systematic review and meta-analysis by Cook et al.[63] on technology-enhanced simulation for health professions education demonstrated large effects of SBME on outcomes of knowledge, skills, and behaviours and moderate effects for patient-related outcomes when compared to no intervention.[63] However in this well conducted review there were no included studies relating to simulation technology for ultrasound.

In 2012 Sidhu et al.[62] published a systematic review aiming to know whether ultrasound or ultrasound procedural simulation training of postgraduate physicians led to improvements competence in ultrasound measured in a clinical or in a simulator setting. Owing to the paucity of evidence relating to skills transfer from ultrasound

simulation to clinical practice at that time the review considered articles that measured performance in ultrasound practice in general. The scope of the review was wide, exploring diagnostic and interventional ultrasound studies across a range of medical specialities. The authors concluded that the literature reflected a general enthusiasm for ultrasound SBME, but there was little compelling evidence to support the widespread adoption to improve competence in clinical ultrasound. The review noted the scarcity of studies, as well as the heterogeneity and poor general quality of the assessed studies. At the time of conducting their review in 2012 the authors recognised the evidence base for ultrasound simulation was in its infancy compared to other simulation based procedural education such as laparoscopy.[64]

There have been two narrative reviews published about training with ultrasound simulators since 2012 reflecting the interest and possible expansion of published evidence on the subject. Blum et al.[51] presented an overview of available ultrasound simulators, the technology they use, their possible applications and briefly discuss evidence for evaluation. Chalouhi et al.[40] focussed on obstetrics and gynaecology presenting an editorial review and overview of publications on ultrasound simulators within the speciality. Chalouhi et al. highlighted[40] the findings from two studies published in 2014, one of which found the performance of novices improved with practice on a simulator[65] and the other reported that simulators did not perform as well as live models for the training of novices concluding they should be used as an adjunct to, rather than substitute for live models.[40, 66]

Systematic literature reviews enable an understanding of the available, relevant evidence in a particular field. These data summaries can also identify and then explore through subgroup analyses any variations in the size and direction of results.

Areas where data are lacking or biased can be highlighted thereby informing the need and design of future studies. We undertook a systematic review of the effectiveness of simulation in gynaecological ultrasound training in light of the increasing adoption of simulation in ultrasound training and the anticipated parallel expansion in the primary literature. Using this methodology we hoped to examine the effectiveness of ultrasound simulators in teaching gynaecological ultrasound examination in order to address current uncertainties in the absence of prior focused systematic reviews in this field.

Aim

The aim of this systematic review was to summarise and critically appraise the available research on the effectiveness of simulators for training in gynaecological ultrasound.

METHODS

The systematic review followed a prospectively developed protocol and used widely recommended and comprehensive methodology. [67]

Framing the question

A structured research ('PICO' – population, intervention, comparator, outcome(s)) question was framed; the population comprised of health professionals learning gynaecological ultrasound, the intervention was use of a simulator to teach gynaecological ultrasound against a comparator and the outcomes were some type of educational outcome. Gynaecological ultrasound was defined as up to the end of the first trimester and therefore included early pregnancy These outcomes were

subsequently classified according to the Kirkpatrick hierarchy, Table 3[68]. The detailed PICO framed research question is shown in Table 2 and Appendix 1.

Table 2 Framing the review question (PICO)

Population or participants of interest	Health professions learner: a student, a postgraduate trainee or practitioner in a profession directly related to human health.[63] This could be students and healthcare professionals from different disciplines such as medicine, nursing, midwifery, radiography and sonography.
Interventions or exposures	Use of a simulator to teach gynaecological US This could be computer-based virtual reality simulators, high fidelity and static mannequins.
Comparisons or controls	None, conventional training, training on one simulator compared to another or an alternative instructional method.
Outcomes of interest	Educational outcomes are classified into levels according to a hierarchy, called the modified Kirkpatrick hierarchy, frequently used in medical education. This attempts to capture the impact of educational interventions using the following levels: participation or completion, modification of attitudes or perceptions, modification of knowledge or skills, change in participants' behavior and change in delivery of care and health outcomes. This review collected information on any reported outcome demonstrating an effect on health professional learning as a result of using an ultrasound simulator.
Setting	Any
Study designs	Any, primary research studies of all types

Identification of relevant studies

A comprehensive search of the relevant published literature was carried out by forming a search strategy for the effectiveness of ultrasound simulators in obstetrics and gynaecology. The search was completed on 28th November 2017, and was performed using the electronic databases Medline [1946 to November 2017], EMBASE [1974 to November 2017], CINHAL [1981 to November 2017], BNI [1992 to November 2017], PubMed [1946 to November 2017], ERIC [1964 to November 2017], Cochrane library [1999 to November 2017], ISRCTN Register of Clinical Trials.

The search strategy was developed for Medline and included free text and MeSH terms. Subsequently the search was adapted to the other databases. The search was built according to the population and intervention by including search terms pertaining to 'gynaecology and obstetrics' 'ultrasonography' and 'simulation' in addition searching for all known manufactured simulated scan systems for gynaecology by brand name. The search strategy used is detailed in Appendix 2. Hand searching of journals dedicated to simulation was also undertaken and these included 'Simulation in Healthcare' and 'Simulation in Nursing' and 'Advances in Simulation' from inception to December 2017. The reference lists of all included studies were checked for any further relevant studies.

Inclusion and exclusion criteria

Included studies were those that investigated the use of a simulator, this could be a mannequin or computer based model, to teach health-related professionals (doctors,

midwives, nurses, sonographers, radiographers) or students in any of these disciplines how to perform a gynaecological ultrasound scan, in comparison with no intervention, conventional training, or alternative forms of simulation training.

Initially obstetrics was included in the search strategy to ensure we did not miss papers that may pertain to early pregnancy scanning which would be considered within the remit of a gynaecologist in the UK.

Studies were excluded if they:

- Used simulation to teach ultrasound that was not gynaecological (gynaecological defined as up to the end of the first trimester and therefore included early pregnancy)

- Studies comparing one group's performance to another all on a simulator to assess validity of the system or studies including expert performance as the comparator group (validity studies)

Conference papers were included when full text versions could not be identified through electronic bibliographic database searches if they met the inclusion/exclusion criteria. No language, country or study type restriction was applied and unpublished studies were sought.

Study selection

The study selection was a two-stage process conducted by two independent reviewers (NW and AM). The first stage entailed screening title and abstracts against the predefined selection criteria to identify relevant papers. The second stage was to read the manuscripts of the citations that appeared to fulfil the predefined selection criteria.

Where there was a difference of opinion on which papers to include a consensus opinion was made after discussion involving a third party arbitrator (TJC).

Data extraction

A predesigned form was used to extract data from studies fulfilling the inclusion criteria (Appendix 3). Data were extracted independently and in duplicate, with discussion and consensus used to overcome disagreements. Data were extracted for country, study design, population, the simulator(s) used, sample size, intervention and comparator. Outcomes and assessment were coded according to Kirkpatrick hierarchy; details of Kirkpatrick's levels are represented in Table 3. [44] We coded the level of impact being studied and summarised any results of the intervention at the appropriate level. However, some studies presented multiple outcomes so these outcomes were coded according to the relevant Kirkpatrick level, which sometimes led to reporting of study outcomes across more than one level. When an outcome assessment relating to skill was done on a simulator this was classified as level 2b, when an outcome assessment relating to skill was done on a real subject (professional patient educator or real patient) this was classified as level 3 because it showed transfer of learning and ability to apply skill. Effect sizes were extracted from study data such as mean values, standard deviations or statistical test results when means and standard deviations were not available.

Table 3 Kirkpatrick hierarchy levels

Level 1 Participation: covers learners' views on the learning experience, its organisation, presentation, content, teaching methods, and aspects of the instructional organisation, materials, quality of instruction.
Level 2a Modification of attitudes/perceptions: outcomes relate to changes in the reciprocal attitudes or perceptions between participant groups towards the intervention/simulation.
Level 2b Modification of knowledge/skills: for knowledge, this relates to the acquisition of concepts, procedures and principles; for skills this relates to the acquisition of thinking/problem-solving, psychomotor and social skills.
Level 3 Behavioral change: documents the transfer of learning to the workplace or willingness of learners to apply new knowledge and skills.
Level 4a Change in organisational practice: wider changes in the organisation or delivery of care, attributable to an educational programme.
Level 4b Benefits to patient/clients: any improvement in the health and well being of patients/clients as a direct result of an educational programme.

Assessing the quality of the literature

The quality of randomised control trials was graded using the Cochrane risk of bias tool for randomised control trials[69] and the modified Newcastle Ottawa scale for observational studies.[70]

Data synthesis

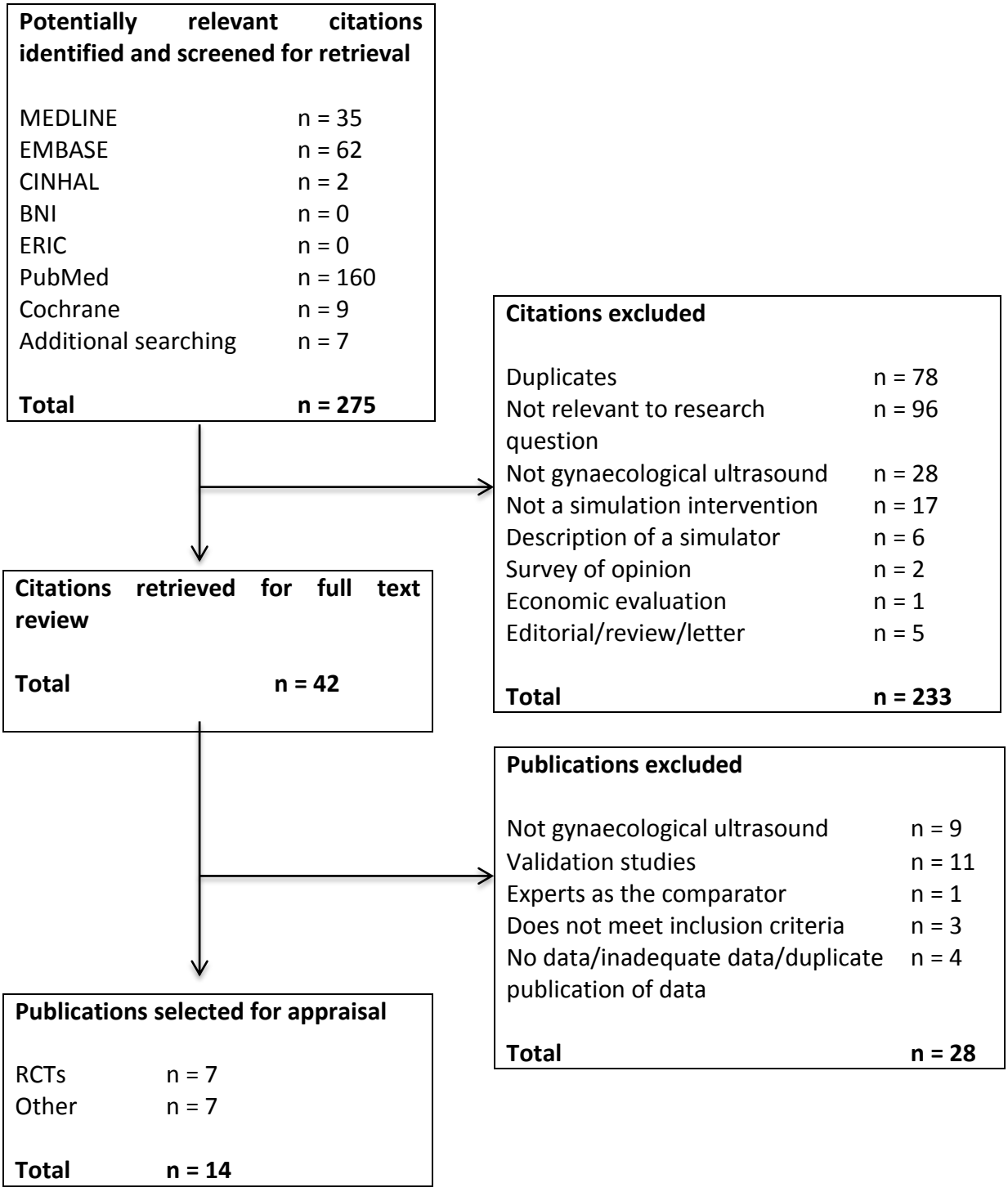
Data synthesis was undertaken according to the outcome assessed and the corresponding classification in the four levels of the Kirkpatrick hierarchy. Where there were groups of ≥ 3 studies reporting the same outcome we planned to performed subgroup analysis if possible by study design. Relevant outcome data from eligible studies that could not be meta-analysed were reported and synthesised within the review narrative.

RESULTS

Study selection

Figure 1 illustrates the flow diagram of the data collection process. The initial search of electronic databases identified 268 potentially relevant studies, 7 were added from additional searching. The predefined inclusion and exclusion criteria and screening of titles and abstracts led to the exclusion of 233 records. Eligibility was assessed for 42 records, including 10 conference abstracts. A total of 28 articles were subsequently excluded; 9 did not relate to gynaecological or early pregnancy ultrasound; 11 were validity studies, 1 used experts as the comparator group; 2 did not use a simulator and 1 was an assessment of the scoring system used to assess trainees; 4 presented no data, inadequate data or data in an abstract which later appeared in a published in a study.

Figure 1 PRISMA flow diagram for systematic review of ultrasound simulation for gynaecological ultrasound training



Study characteristics

The 14 studies that met our inclusion criteria were published between 2004 and 2017. Studies were published as either full-text articles,[66, 71-81] conference abstract[82] or data within a published thesis.[83] The included studies recruited a total of 573 participants, the number of participants per study ranged from 11 to 145. All included studies were published in English. The majority of the selected studies were performed in Europe (71%) the remainder in the USA (29%). Of the European studies four of these were from one group in Denmark, three from the UK, two from France and one from Germany. There were seven randomised controlled trials, three controlled observational studies, and four non-controlled observational studies. Twelve studies used doctors as the study population; these were predominately junior doctors (this term is used as a blanket term to encompass residents and doctors in speciality training). Of the junior doctors these were mostly obstetricians and gynaecologists in training, making up 39% of the total number of participants (n=226/573), other doctors were specialising in emergency medicine, general practice or radiology. Two studies used medical students[66, 74] and one study used midwives exclusively;[80] two studies had a mixed population of doctors, nurses and sonographers.[78, 79]

There were seven randomised controlled trials four were multicentre,[72-74, 83] with centre in one region of a country, three were single centre. [66, 71, 82] Six studies compared simulation to conventional teaching or training, one compared simulator practice in pairs to individual simulator practice.[74]

Amongst the RCTs the simulator interventions were varied and not always well described. Variation existed in the amount of time spent on the simulator, whether use

of the simulator was individual practice or small group teaching and whether practice was supervised or unsupervised. When it was stated, the time the intervention group spent on the simulator ranged from 40 minutes to approximately 3-4 hours. The comparators were also varied; three studies compared simulation to clinical teaching or training,[83] [72, 73] for two studies this was simulation prior to clinical training[72, 73] and the other was alongside.[83] One study compared a simulator session to conventional lecture based teaching,[71] two compared simulation teaching to practice on a live model.[66, 82] All the RCTs, with the exception of one in which the type of mannequin was not stated,[66] used the either the transvaginal ScanTrainer® (Medaphor, UK), BluePhantom (CAE Healthcare, USA) or a combination of the two models.

Of the three controlled observational studies one compared simulation training to clinical training,[77] one compared simulation followed by clinical training to clinical training alone,[75] one compared theoretical plus simulator teaching to theoretical teaching plus teaching on a live model.[66] They all used different simulators.

Of the four uncontrolled observational studies two cohorts used a simulator as part of course comprising of workshops with tasks and assessments,[78, 81] one used simulators to practice a specific skill[80] one study was of very poor quality and minimal description was given of the intervention.[79] All four studies used different combinations of simulators.

The time points for outcome assessment were varied. Half of the included studies 7/14 used an immediate assessment of outcome, within the first three days following the intervention.[66, 71, 74, 78-81] Four studies used a delayed outcome assessment, in two studies this was at 2-3 months[73, 75] and in two RCTs it was at six months.[72, 83] For

the remaining three studies the time of outcome assessment was not stated or not made clear. [76, 77, 82] Williams et al.[77] assessed outcome after 10 hours of either simulation or clinical practice but it is not clear over how many days, weeks or months the intervention or comparator group received this training.

Types of simulator

The studies used different types of gynaecological ultrasound simulators, high fidelity virtual reality (VR) simulators and low fidelity pelvic mannequin models. The predominant VR model was ScanTrainer® that was used in 7/14 (50%) of studies either alone, in four of the seven studies,[66, 71, 74, 83] or with a pelvic mannequin model the BluePhantom™ in three studies from Denmark.[72, 73, 80] Other virtual reality models were Sonotrainer,[76] Vimedix,[78, 81] Ultrasim[78, 79] and Simbionix US mentor.[78] Five studies used a combination of simulator types, the study by Arya et al.[78] used a combination of three virtual reality simulators, Vallabh-Patel[81] used two types of VR simulators and a mannequin and the Danish studies used a single VR model with a mannequin. Three studies used only a mannequin model as the intervention, this was BluePhantom™[54] in two studies[66, 75] and not specified in the other.[82] The exact version of software the simulator used or the modules accessed on the simulator was not always clearly stated and data to this level of detail was not extracted.

Study quality

The study quality assessments are shown in Figure 2 and Appendix 4 for randomised control trials and Table 4 for observational studies. The nature of the intervention made it impossible to blind the participants but eight of the studies did have blind

assessors. Overall the follow up was good, with only two studies [80, 83] losing more than 20% of the enrolled participants prior to the primary outcome assessment.

The included study designs and characteristics are summarised in Table 5.

Figure 2 Cochrane risk of bias summary for included randomised controlled trials

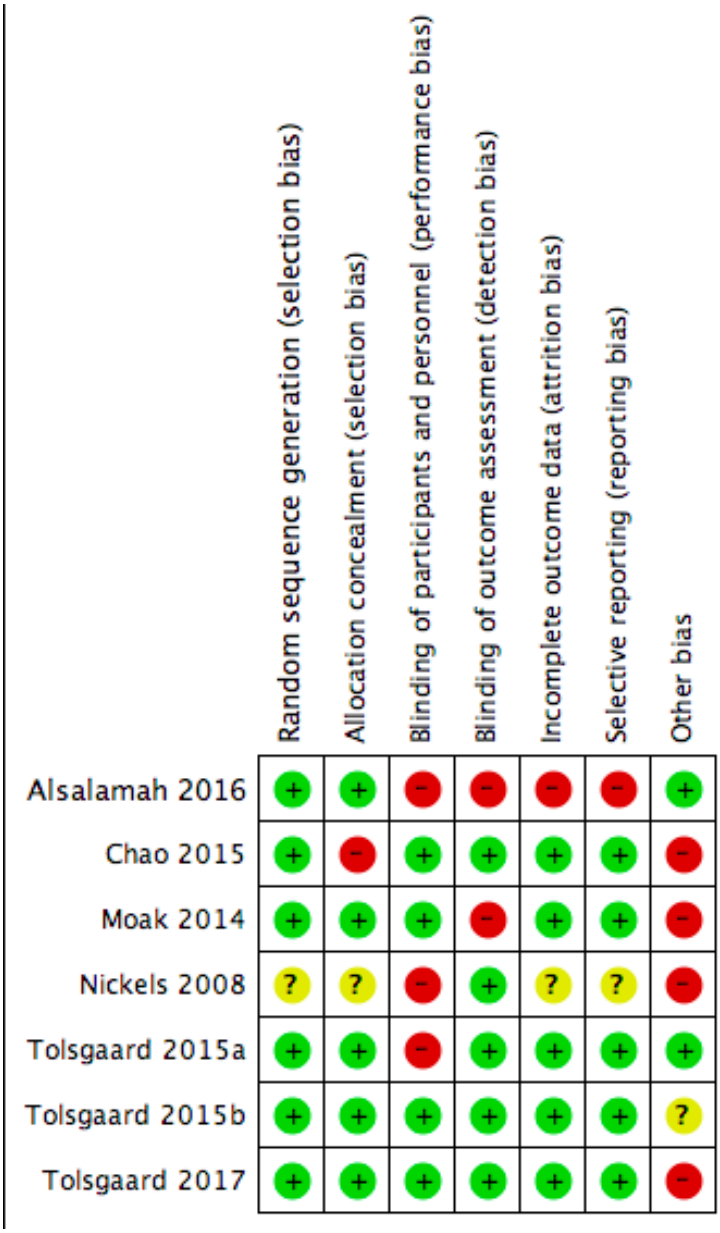


Table 4 Modified Newcastle Ottawa scale for observational studies

	Selection				**Comparability** (max 2 stars available)	**Outcome**			
	Representativeness of the intervention cohort[1]	Selection of the non-intervention cohort[1]	Ascertainment of the intervention[2]	Demonstration of baseline assessment prior to simulation intervention[3]	Comparability of cohorts on the basis of the design/analysis[4]	Assessment of outcome[5]	Was follow up long enough for outcomes to occur[6]	Adequacy of follow up of cohorts[7]	Total number of ★ Maximum score = 9
Lous (2017)	★	★	★	★	★	★	★	★	8
Maul (2004)	NS	NS	★	NA	NS	NS	NS	★	2
Williams (2013)	★	★	★	★	★	-	★	★	7
Arya (2017)	★	NA	★	★	NA	NA	★	★	5
Bahl (2006)	★	NA	NS	NA	NA	NA	-	NA	1
Madsen (2017)	★	NA	★	NA	NA	NA	★	-	3
Patel (2014)	★	NA	★	★	NA	★	★	★	6

Key: ★ each star represents if individual criterion within the subsection was fulfilled. NS represents data that was not specified in the study. NA represents not applicable - represents not done

1) Studies received a ★ if the sample included health education learners: a student, a postgraduate trainee or practioner in a profession directly related to human health. This could be students and healthcare professionals from different disciplines such as medicine, nursing, midwifery, radiography and sonography.
2) Studies received a ★ if the ascertainment of intervention was from a secure record, structured interview or written self report
3) Studies received a ★ if there were baseline measurements done prior to the intervention
4) Studies received a ★ if the participants were matched by experience. Studies received an additional ★ if they were matched for additional factors (eg age, professions etc)

5) Studies received a ★ if there were independent blinded assessments

6) Studies received a ★ if the follow up after the simulator is implicitly or explicitly known

7) Studies received a ★ if the follow up of the cohort is 80% or above

Table 5 Details of study design, population and intervention

Author Year	Country	Population	Sample size	Simulator	Intervention	Comparator
Randomised controlled trials						
Alsalamah [83] 2016	UK	Junior doctors n=77	**T =77** I = 41 C = 36	Transvaginal ScanTrainer® (Medaphor, UK)	Introduction to the simulator then access to a simulator, in **assessed** practice mode, alongside clinical training.	Introduction to the simulator then access to the simulator in **unassessed** practice mode, alongside clinical training
Chao[71] 2015	France	Junior doctors n=34 (O&G n=30, GP n=4)	**T = 34** I = 16 C = 18	Transvaginal ScanTrainer® (Medaphor, UK)	40 minutes on the simulator, working in small groups of 2-3 students per simulator with supervision from an experienced tutor	40 minutes of conventional teaching comprising of a 30 minute lecture with 10 minutes of questions
Moak[66] 2014	USA	Medical students n=145 (3rd year)	**T = 145** I = 70 C = 75	BluePhantom™ (CAE Healthcare, USA)	Standardised 1 hour training session on the simulator	Standardised 1 hour training session on a live model
Nickels[82] 2008	USA	Doctors n=19 (Emergency medicine)	**T = 19** I = 10 C = 9	Mannequin, type not specified	Didactic presentation then practice on simulator	Didactic presentation then practice on a live model
Tolsgaard[72] 2017	Denmark	Junior doctors n=54 (O&G)	**T= 54** I = 26 C = 28	Transvaginal ScanTrainer® (Medaphor, UK) BluePhantom™ (CAE Healthcare, USA)	30-minute introduction then independent practice on ScanTrainer with verbal feedback on request from an instructor. Practice continued on simulator until a pre-defined mastery level (88.4%) achieved. Then practice was continued on BluePhantom until proficiency achieved. Followed by clinical training as described for the comparator group for 6 months.	Clinical training alone, which comprised of an introductory lecture then scanning patients with supervision on request for 6 months.
Tolsgaard[73] 2015 a	Denmark	Junior doctors n=33 (O&G)	**T = 33** I = 18 C = 15	Transvaginal ScanTrainer® (Medaphor, UK)	Intervention described as above. Followed by clinical training as described for the comparator group for 2 months.	Clinical training, which comprised of an introductory lecture then scanning patients with supervision on request for

				BluePhantom™ (CAE Healthcare, USA)		2 months.
Tolsgaard[74] 2015 b	Denmark	Medical students n=30 (final year)	**T = 30** I = 16 C = 14	Transvaginal ScanTrainer® (Medaphor, UK)	10-minute introduction then 2 hours practice on simulator working in pairs (dyad practice), taking turns on the simulator.	10-minute introduction then 2 hours practice on simulator working individually.
Controlled observational studies						
Lous[75] 2017	France	Junior doctors n=36 (GP)	**T = 36** I = 26 C = 10	BluePhantom™ (CAE Healthcare, USA)	Theoretical lecture then a session of supervised learning on the simulator followed by clinical training as described for the comparator group for 2 months.	Clinical training, which comprised of scanning patients with supervision on request for 2 months.
Maul[76] 2004	Germany	Doctors n=45 (Obstetricians)	**T = 45** I = 21 C = 24	Tansvaginal SonoTrainer (Sonofit GmbH, Germany)	Theoretical and simulator training in NT and CRL measurements	Theoretical training in NT and CRL measurements
Williams[77] 2013	UK	Junior doctors n=11 (O&G n=10, range of grades, radiology n=1)	**T = 11** I = 6 C = 5	Transvaginal ScanTrainer® (Medaphor, UK)	Introduction to scan simulation then 10 hours unsupervised practice on a simulator.	Standard clinical training, which was 10 hours of supervised clinical training on patients.
Uncontrolled observational studies						
Arya[78] 2017	USA	Doctors n= 4 Junior doctors n=17 Nurses n=1 Sonographers n=6	**T= 28**	VIMEDIX (CAE Healthcare, USA) Ultrasim (MedSim, USA) Simbionix US mentor (3D systems Healthcare, USA)	A 4-hour workshop comprised of a 15min lecture and then 40 minutes of supervised training on simulators (2x20min sessions) done in small groups	
Bahl[79] 2006	UK	Doctors n= 2 Nurses n= 27 Sonographers n=2	**T=31**	Ultrasim (MedSim, USA)	No description	
Madsen[80] 2017	Denmark	Midwives n = 20	**T = 20**	Transvaginal ScanTrainer® (Medaphor, UK) BluePhantom (CAE Healthcare, USA)	30-minute introduction then independent practice on ScanTrainer with feedback from an instructor on failed metrics. Practice continued until simulator until a pre-defined mastery level was achieved. Then practice was continued on BluePhantom until proficiency level reached	
Patel[81] 2014	USA	Junior doctors n=24 (O&G, 1st year n=12,	**T = 24**	Mannequin (limbs and things, USA)	Introductory lecture followed by 2 x 45 min simulation stations done in groups of 2-3	

2nd year n=12)	VIMEDIX (CAE Healthcare, USA) Ultrasim (MedSim, USA)	residents and 1 transabdominal scanning station using live models.

O&G Obstetrics and gynaecology, GP General Practice, T= Total, I = Intervention, C = Control

Results synthesis

The outcomes and methods of assessment varied considerably across studies with many reporting multiple outcomes. Studies evaluating the same outcomes often presented the data in different ways. To aid interpretation, the results are presented according to the corresponding Kirkpatrick hierarchy level, noting that some studies report outcomes at multiple levels. Table 6 provides an overview of the studies and the corresponding applicable Kirkpatrick levels for outcomes reported. Table 7 provides detail of study outcomes, assessment of outcomes and results according to corresponding Kirkpatrick level.

Table 6 Overview of study outcomes according to Kirkpatrick hierarchy levels

	Kirkpatrick level				
Author, year	**4**	**3**	**2b**	**2a**	**1**
Randomised controlled trials					
Alsalamah 2016			■		
Chao 2015			■	■	
Moak 2014		■		■	
Nickels 2008		■	■		
Tolsgaard 2017	■				
Tolsgaard 2015 a		■			
Tolsgaard 2015 b		■	■		
Controlled observational studies					
Lous 2017		■	■		
Maul 2004	■	■			
Williams 2013			■	■	
Uncontrolled observational studies					
Arya 2017			■	■	
Bahl 2006					■
Madsen 2017		■	■		
Vallabh -Patel 2014			■		

Table 7 Detail of study outcomes, assessment of outcomes and results according to Kirkpatrick framework

Kirkpatrick Level 4: Organisational practice and benefit to patients					
Study	**Outcome**	**Time (Completeness)**	**Assessor**	**Assessment**	**Results**
Tolsgaard 2017 RCT	1° Patient reported discomfort	During first 6 months of clinical practice (96.3% n=52/54)	Blinded, patient assessor	Ratings on a 10-point Likert scale (no. of ratings n=939)	Relative difference 18.5% lower in intervention group (95% CI -25.5 - 10.7) (p <0.001)*
	2° Patient perceived safety in provider			Ratings on a 10-point Likert scale (no. of ratings n=1109)	Relative difference 7.9% higher in intervention group (95% CI 0.5 – 14.7) (p =0.04)*
	2° Patient perceived confidence in provider			Ratings on a 10-point Likert scale (no. of ratings n=1096)	Relative difference 11.1% higher in intervention group (95% CI 2.5 – 18.9) (p = 0.01)*
	2° Patient satisfaction			Ratings on a 10-point Likert scale (no. of ratings n=1097)	Relative difference 2.1% higher in intervention group (95%CI -6.3 - 9.8) (p = 0.61)
	2° Time taken to perform scan			Recorded time taken for examination	I: 6m 2s (95% CI 5m 47s – 6m 17s) C:7m 11s (95% CI 6m 50s – 7m 30s) Relative difference 1m 32s less in intervention group (CI 7s to 3m 6s) (p=0.03)*
	2° Need for supervision or repeat examination according to the clinical training time and type of diagnosis (e.g normal pregnancy, fetal demise, ectopic)		Blinded, nurse assessor	Recorded need for supervision or repeat examination by supervisor who would often taken over if asked for help	Interaction between days of clinical training and the intervention (p=0.005)* When the duration of clinical training was doubled, the odds for trainee supervision/repeated examinations by a supervisor reduced by 45.3% (95% CI 33.5 – 55.1%) in the intervention group and by 19.8% (95% CI 4.1 – 32.9%) in the control Significant effect of type of diagnosis on need for supervision/repeat examination by supervisor (p=<0.001) in both groups
Maul 2004	Time taken to preform scan	NS	NS	Recorded time taken for examination	I: 7m 30s C:14m 55s (p= <0.05)*

Kirkpatrick level 3: Behavioral change: documents the transfer of learning to the workplace, assessments on patients					
Study	**Outcome**	**Time (completeness)**	**Assessor**	**Assessment**	**Results**
Moak 2014 RCT	Competence	2-3 days after intervention (92.4% n=134/145)	Non- blinded live model was scanned and assessed professionalism Blinded assessor assessed scanning technique and another later reviewed and assessed images	A 40 point assessment score (Score made up of professionalism 10, scanning technique 18, image acquisition 12)	Mean test scores presented (%) Overall score I: 56% C: 69% p = <0.001^ Professionalism I: 71% C: 75% p = <0.45 Scanning technique I: 60% C: 73% p = <0.001^ Image acquisition I: 37% C: 59% p = <0.001^
Nickels 2008 RCT	Competence	NS	Two blinded assessors, performance assessed on a live model	A 17 point assessment score	Difference in scores pre/post Scanning I: Pre 47% (38-55%) post 85% (78 – 92%) C: Pre 51% (32-70%) post 88% (77 – 97%)
Tolsgaard 2015 a RCT	Competence	After two months of clinical practice (78.8% n=26/33)	Two blinded assessors, reviewed video recordings of scans and patient records	OSAUS score, maximum score 30	Adjusted absolute difference in OSAUS scores between the two groups 20.1% (95% CI 11.1 – 29.1) (p <0.001)*
Tolsgaard 2015 b RCT	Competence	The day after the intervention (80% n= 24/30)	Blinded assessor observed scan	OSAUS score, maximum score 25	Differences in OSAUS between groups I: 56.3 ± 14.8% C: 48.4 ± 11.8% (95% CI -3.8 – 19.6) Not significant
	Stress levels	Baseline and 2-	Self rating by	10 point Likert scale	Baseline I = 4.8 C = 5.7 p=0.46

Lous 2017		3 months after the intervention (100% n=36/36)	doctors		End point I = 4.8 C = 6 p=0.4 No significant difference between groups
Maul 2004	Accuracy of measurements	Following intervention (100% n=45/45)	NS	Mean absolute deviations from pre-defined standards	I: NT 0.31 ± 0.14 C:NT 0.62 ± 0.23 p = <0.05* I: CRL 1.48 ± 2.0 C:CRL 3.27 ± 2.48 p = 0.05*
Madsen 2017	Competence	Following intervention (30% n=6/20)	Blinded assessor observed scan	OSAUS score maximum score 25	Pearson correlation coefficient 0.81 between VR performance and clinical performance p=0.049*

Kirkpatrick level 2b: Modification of knowledge/skills, assessments on simulator

Study	Outcome	Time (completeness)	Assessor	Assessment	Results
Alsalamah 2016 RCT	Speed of acquisition of skill	At 6 months (Variation in reporting)	Non-blinded assessor	A pass/fail scoring checklist	Repeated measure no significant difference between groups at the final plateau Kaplan-Meier estimate no significant difference between group
Chao[71] 2015 RCT	Image quality 1° Individual image quality score	Images taken immediately after intervention, (100% n=34/34)	Blinded assessors later reviewed and assessed images taken on simulator	1° Score from 0-19 based on image quality criteria	
	2° Rate of adequate images per group			2° Rate of adequate images per group (defined as this was the number of images with 1 or zero criteria scored as zero divided by total number of images produced)	Mean scores intervention group 12 (SEM, 0.8) comparison 9 (SEM 1.0) p = 0.0302* Rate of satisfactory images intervention group 42% (n=27/64) comparison group 28% (n=20 of 72) Not statistically significant
Nickels 2008 RCT	Image interpretation	NS	Computer based quiz	24 point quiz for interpretation	I: Pre 62% (53-71%) post 81% (73 – 89%) C: Pre 52% (34-70%) post 85% (75 – 95%)
Tolsgaard[74] 2015 b RCT	Skill	The day after the intervention (80% n= 24/30)	Two observer blinded assessors	Pre and post test OSAUS scores, maximum score 25 points, presented as percentage mean scores	I: pre 28.5 ± 4.2% post 55± 6.3% C: pre 24.9 ± 3.7% post 49.3 ± 6.0%
	Performance and efficiency during training assessed on simulator	The day after the intervention (80% n= 24/30)	Validated simulator metrics	Simulator metric score	I: 82.1 ± 9.6% C: 72.1 ± 21.3% p=0.22
				Simulator efficiency score	I: 5.88 ± 1.13 points/attempt C: 2.79 ± 0.92points/attempt (p= <0.01)
Lous[75] 2017	Image quality	2-3 months after the intervention (100% n=36/36)	Blinded assessor scored images obtained on real patients, one case per participant with requirement of 4 images	Image quality score, maximum score 11.	Differences in satisfactory global score (score >10) I: 70% vs C: 20% p = 0.027*
Williams[77] 2013	Skills, accuracy of measurements, image quality and interpretation of findings	Following completion of intervention (82% n=9/11)	Combination of simulator scores, non-blinded researcher observing use of simulator and image quality later assessed by a sonographer	Difference in mean scores between groups as percentages	Overall percentage difference between the mean scores of the two groups 8.34% Mann-Whitney U value 13 p=0.0556
Arya 2017	Knowledge	Immediately after intervention (90.3% n=28/31)	MCQ	A comparison of means from a composite score of dichotomous items with a max score of 17. Scores were transformed into percentages	Pre-test 66.2% Post test 94.1% p <0.0001*

Madsen 2017	Skill on different simulator types (VR vs. Mannequin)	Following intervention (timing not explicitly stated) (100% n=20/20	VR performance scored on simulator metrics Mannequin performance scored by an instructor blinded to previous performance on an OSAUS scale	No. of attempts Total training time A comparison of mean performance scores (total performance scores divided by number of attempts).	No. of attempts Pearson correlation coefficient 0.39 between the two groups p= <0.093 Total training time Pearson correlation coefficient 0.78 between the two groups p= <0.001 Mean performance score Pearson correlation coefficient 0.28 between the two groups p= <0.23
Vallabh-Patel 2014	Knowledge	Immediately prior, during and after intervention (100% n=24/24)	MCQ for pre and post course assessment. Station assessment scores assessed by 3 physicians	Combined score presented as percentage combining percentage score for MCQ with ultrasound station score which was done on a standardised checklist, max score of 10, percentage scores presented	Year 1: Pre 78.1% Post 82.4% Mean score 81.1% Year 2: Pre 84.4% Post 84.6% Mean score 88.9%

Kirkpatrick Level 2a: Modification of attitudes/perceptions

Study	Outcome	Time (completeness)	Assessor	Assessment	Results
Chao[71] 2015 RCT	Confidence, speed, capacity to produce adequate images, capacity to get orientated	Immediately after intervention (85% n=29/34)	Survey	No detail on questions	100% improved confidence, capacity to produce adequate images and get orientated. 94% of control group and 86% of intervention group thought it would increase speed
Moak 2014 RCT	Preparedness for scanning on real patients	2-3 days after intervention (92.4% n=134/145)	Survey	10 point Likert scale	Mean scores I: 4.4 C: 6.2 p = <0.001^
Williams 2013	Confidence	Following completion of intervention (82% n=9/11)	Participant questionnaires followed by semi-structured interviews	Yes/No question and Interviews analysed by author	89% (n=8/9) reported intervention could help increase confidence 11% (n=1/9) reported depends on trainee
Arya 2017	Relevance of US to clinical learning and usefulness to improve performance and interpretation skill	Immediately after intervention (90.3% n=28/31)	Survey	10 statements rated on a 5 point scale	100% (n=28) agreed/strongly agreed that pelvic US simulation applies to their clinical practice. 100% (n=28) agreed/strongly agreed that performance and interpreting skills would be improved. 100% (n=28) believed simulation would improve care

Kirkpatrick Level 1: Participation

Study	Outcome	Time (completeness)	Assessor	Assessment	Results
Bahl 2006	Participants opinions on the simulator realism and place in teaching	Immediately after course (Varied response rate)	Survey	Statements on a 5-point rating	100% (n=31) Reported mannequin easy to scan and suitable for training n=29/31 mannequin and transducer handling satisfactory

* Indicates a significant result in favour of the intervention group (p <0.05)
^ Indicates a significant result in favour of the comparator (p <0.05)

NS = Not specified, CI = Confidence Intervals, OSAUS = Objective Structured Assessment of Ultrasound Skills, VR = Virtual reality

Kirkpatrick level 4: Organisational practice and benefit to patients

Two studies reported outcomes at Kirkpatrick level 4 demonstrating benefit to patients and/or the organisation in efficiency of care, through reduction in time taken to perform scans. The outcomes at level 4 come from a well designed RCT[72] with a low risk of bias and a controlled observational study which did not report the intervention or assessment in detail.[76] Tolsgaard et al.[72] found significantly lower patient discomfort scores in the group receiving simulation training prior to standard clinical training; the relative difference in patient discomfort scores was 18.5% (95% CI 10.7% to 25.5%). There were also significant differences in favour of the intervention group in patient perceived safety and confidence in the trainee scanning ability (Table 7). However, there was no significant difference in overall patient satisfaction. The time taken to perform a scan was reduced by approximately 20% ($p=0.03$) over the first six months of training [72] Maul et al. [76] report significantly faster recorded examination times for the intervention group compared to the control with reductions of almost 50% ($p<0.05$) in scans performed to measure CRL and NT. However the effect size should be interpreted with caution in the later study because of inadequate reporting of the assessment of this outcome including the timing of the evaluation relative to the amount of training received in each group. The need for supervision or repeat examination by a supervisor was reported by Tolsgaard et al.,[72] according to the amount of clinical training time i.e. scanning on real patients and also on the primary diagnosis, which were all early pregnancy or acute gynaecological problems. A significant interaction between days of clinical training and the intervention ($p=0.005$) was reported. When the duration of clinical training was doubled, the odds for trainee supervision or need for the supervisor to repeat the examination reduced in both groups. However, the magnitude of this effect

significantly favoured the simulation plus clinical training group compared with the clinical training alone control group (45.3%, 95% CI 33.5 – 55.1% versus 19.8%, 95% CI 4.1 – 32.9% respectively). There was a significant effect of type of diagnosis e.g normal pregnancy, fetal demise or ectopic pregnancy, on the need for supervision or need for repeated examination by a supervisor (p=<0.001), who would often take over to repeat the examination if asked for help, implying that the more complex the diagnosis the more likely supervision was sought.

Kirkpatrick level 3: Behavioural change, documents the transfer of learning to the workplace

Seven studies reported outcomes at Kirkpatrick level 3 demonstrating transfer of learning on a simulator to the workplace. The outcomes were defined at level 3 if skills transfer from the simulated environment to real patients was demonstrated, including on a professional patient educator. Four studies reporting outcomes at level 3 were RCTs,[66, 73, 74, 82] two controlled observational studies[75, 76] and one uncontrolled observational study[80] in which only 30% of those who received the intervention were assessed in clinical practice.

Five studies reported on competence demonstrated on patients using an assessment score. Three studies used the Objective Structured Assessment of Ultrasound Skills scores (OSAUS),[73, 74, 80] these were studies from the same research group in Denmark who has also conducted work to demonstrate validity of this assessment tool.[84] The other two studies used an alternative assessment score; one was a 40 point score with some similarities to the OSAUS score with inclusion of scanning technique and image acquisition as criteria,[66] however it also included a score of professionalism assessed by a live model[66] and the remaining study did not provide

sufficient detail of the assessment for comparison.[82] The full OSAUS is a 35 point score made up of seven criteria rated 0-5,[85] although the reported primary studies excluded one[73] or two[74] [80] of the seven criteria relating to non-technical skills, indication for the examination and/or medical decision-making, to give a 30 or 25 point total score. The randomised controlled study by Tolsgaard et al. reported competence assessed by OSAUS score and showed an adjusted absolute difference in OSAUS scores in favour of the group receiving simulation training of 20.1% (CI 11.1-29.1%) (P<0.001).[73] In contrast the study by Tolsgaard et al.[74] did not demonstrate any difference between the intervention and control groups but this study is not comparable as the intervention was dyad practice (i.e. training in pairs) on a simulator compared to individual practice and there was not a further control group without exposure to simulation. In the RCT by Madsen study using OSAUS performance scores, the VR simulator was shown to correlate well with clinical performance scores (Pearson correlation coefficient of 0.81 (P=0.049)) [80]. The two studies using their own assessment score for competence were studies comparing training on a simulator as the intervention with training on a live model. One of these RCTs showed an improvement in performance from baseline in both groups with no differences between simulation or live model training.[82] In contrast the RCT by Moak reported significantly better overall mean test scores, scanning technique and image acquisition in favour of training on a live model rather than a simulator. However, whilst the evaluation of these outcomes were blinded the assessment was performed on the same live model upon which the comparator group had practised on which may have biased the performance of the intervention group using the simulator.[66]

Lous et al.[75] is the only study to examine the level of stress reported by doctors at baseline and 2-3 months after scan training once the doctors had commenced scanning in clinical practice. No significant difference was identified between levels of reported stress as measured on a 10 point Likert scale between simulation with clinical training and clinical training alone.

Maul et al.[76] was the only study to report on accuracy of measurements with transfer of skill from training to real patients. Greater precision in measuring early fetal anatomy (CRL and NT) was seen in the intervention group who received simulation training in addition to theoretical instruction (Table 7).

Kirkpatrick level 2b: modification of knowledge and skills

The most frequently reported outcomes were at Kirkpatrick level 2b, with nine of the total 14 studies reporting an outcome at this level that demonstrated modification of knowledge or skills on a simulator. Outcomes were reported from four RCTs,[71, 74, 82, 83] two controlled observational studies[75, 77] and three uncontrolled observational studies.[78, 80, 81]

In addition to being the most numerous, chosen outcomes at Kirkpatrick level 2b were the most varied making comparisons difficult. The most frequently reported outcomes related to the quality and interpretation of images and the knowledge of, and skill in, scanning. Interpretation was further compounded by heterogeneity in the method and timing of assessment.

Four studies reported on image quality [71, 75, 77] or interpretation.[82] These were images taken either immediately after the intervention,[71] 2-3 months after the intervention[75] or not specified.[77, 82] Although two of these studies [71, 75] reference

the same quality criteria scoring scale, direct comparison is hindered because they assessed different elements included in the scoring scale such that one study assigns a maximum score of 11[75] and the other of 19. [71] Furthermore Lous et al.[75] assessed real patient images, the participants were required to take four pre-specified images from a real patient scan and return these to the researcher for assessment, whereas in the study by Chao et al.[71] images assessed were taken on the simulator. In both studies, the reported results on image quality showed a significant difference (p=0.03) [71] and (p=0.027) [75] in favour of the simulation intervention. However, like Chao et al., a study by Williams et al. also assessed image quality from images taken on the simulator but did not find trainees having simulation training obtained images of any better quality. These conflicting findings may reflect a type II statistical error because of the smaller sample in the latter study[77] or bias in favour of the simulation group in the study by Chao et al.[71] where assessment in both groups was undertaken on the simulator rather than a real patient.

Two uncontrolled observational studies [78, 81] reported on knowledge as the outcome, which was assessed through multiple-choice questions in both studies with the addition of a score based on performance in assessed ultrasound stations in the study by Patel.[81] Both studies showed an increase in knowledge between pre-and post-intervention scores; this was significant in the study by Arya et al. (P<0.0001)[78] but improvement was restricted to year one O&G trainees as opposed to year two in the other study.[81]

The only included study, which had the second largest number of recruited participants of all primary studies included in this review, to report on speed of skill acquisition was conducted by Alsalamah.[83] In this study both groups of junior

doctors, had access to a simulator alongside clinical training but the intervention group could use the simulator in 'assessed practice mode' where they had to pass certain tasks and could self-assess. The other group did not have access to this mode and could use the simulator in 'unassessed mode'. Despite the benefits of a larger sample not all research subjects received the intended intervention or the required clinical training resulting in multiple small subgroups for analysis. The study showed no significant difference between groups in skill acquisition at the six-month assessment.

The RCT by Tolsgaard et al. b [74] compared dyad practice on a simulator to individual practice and this showed similar skills improvement in both groups assessed by OSAUS scores and simulator metrics but significantly higher training efficiency in the dyad training group in terms of simulator score per number of attempts compared with the single student group ($p=0.01$).[74] In an uncontrolled observational study from the same research group, Madsen et al.[80] reported good correlation between time needed to achieve a predefined performance levels on the VR simulator and the mannequin used ($p<0.001$).

Kirkpatrick level 2a: modification of attitudes and perceptions

Three studies reported data at Kirkpatrick level 2a demonstrating modification of attitudes and perceptions after simulation, two were RCTs that also reported outcomes at higher levels,[66, 71] and one controlled observational study, which was the only study to use mixed methods and report qualitative data.[77] Two studies[71, 77] reported on confidence, which was assessed though a survey and in the study by Williams et al.[77] followed up with semi-structured qualitative interviews. No quality assessment was made on the qualitative component of the study. These studies report high

positive agreement in the simulator improving confidence although sample size is small. Moak et al.[66] reported that the comparator was more likely to prepare them for scanning on real patients but in this study the comparator was a live model. Arya et al.[78] report high levels of agreement in support of relevance of ultrasound to clinical learning, usefulness to improve performance and interpretation based on responses to a post-encounter survey with questions rated on a 5-point Likert scale.

Kirkpatrick level 1: participation

One uncontrolled observational study from 2006 reported outcome data at the level of participation. The study has a modified Newcastle Ottawa score of one and is lacking detail in the data presented precluding any meaningful inferences being drawn.[79]

DISCUSSION

Principle findings

The use of simulation training in gynaecological ultrasound, when compared with standard approaches for teaching gynaecological ultrasound in clinical practice is associated with significant improvements in both important outcomes for the learner and patient. Improved learner outcomes included competence, skill and confidence, and patient-centred outcomes included as reduced discomfort, increased perceived safety and enhanced confidence in the provider. These outcomes were invariably improved with the addition of ultrasound simulation in both randomised and observational study designs. However, the type(s) of simulator to use and how best to incorporate them into gynaecological ultrasound training programmes remains unclear. This is because many of the primary study designs were of poor

methodological quality and the nature of the interventions evaluated varied widely with different simulators being used or employed in different ways within differing clinical training schedules. The assessment and reporting of clinical outcomes were also heterogeneous precluding quantitative aggregation of data and limiting educational and clinical inferences.

Strengths and limitations

This is the only review focused on simulation for gynaecological ultrasound training. A reproducible, thorough and ordered approach to literature searching was carried out without language restrictions. The selection of studies and subsequent data abstraction was carried out in duplicate with disagreements resolved by consensus. We included both observational and randomised controlled studies to increase the data pool, and assessed the presence and magnitude of methodological bias within selected studies by applying appropriate, validated quality assessment tools. As with any research synthesis, the conclusions are bound by the published evidence. As the studies were so heterogeneous with variable quality no meta-analysis was deemed appropriate. The studies used a range of simulators, including high fidelity VR models as well as low fidelity mannequin models. No exclusions were made based upon the type of simulator as fidelity has been shown to show no significant advantage of one type over another.[86]

The studies reported on a range of outcomes across the Kirkpatrick hierarchy levels. It was not only outcomes that were varied but also outcome assessment; variation was seen in setting, assessor, and assessment tool. Some studies used validated outcome measures with the OSAUS[84] to assess competence,[73, 74, 80] or a previously published rating score[87] to assess image quality but even then used differing elements of the

tool or score.[71, 75] There was a potential source of bias unique to these studies linked to how the training was delivered and then how competence was assessed. When trainees had practiced, whether it be on a simulator [71, 77] or live model [66, 82] and then they went on to be assessed on the same modality with which they had practiced they did better than the comparator group who hadn't had the same level of practice.

For the purpose of this review outcomes were classified according to the Kirkpatrick hierarchy, the use of which has been critiqued for its suitability for appraising interventions in medical education.[88] For this review Kirkpatrick's levels were deemed applicable to grade the measured outcomes and give structure to the review with a clear distinction in its application as a framework and not as a critical appraisal tool.[88] However one of the problems with the Kirkpatrick framework is some outcomes may be subjective for example emotions could be classified at multiple levels. Stress was allocated at level 3 in this review as a behavioural change, as it was assessed after 2-3 months rather than immediately and stress could impact on workplace behaviour. However stress could also be classed perhaps at lower levels: level 1 (reaction) or level 2a (attitude or perception), which is where we placed learner perceived confidence. In this review distinction was made between skills demonstrated on a simulator (level 2b) and patients (level 3). For the latter classification level we include assessment on 'real' patients as well as live human models but the two are different entities: real patients may be bleeding, may be in pain or under psychological stress due to the clinical context for the scan.[89] In the included studies using live models[66, 82] participants were familiar with them and by virtue of their employment for training and willingness to give feedback they are not comparable to real patients or inanimate models.

In this review the research question related to the effectiveness of ultrasound simulation in gynaecology according to the Kirkpatrick hierarchy and so validation studies were excluded, because the outcomes were related to system – does it measure what it supports to purports to measure[90] this resulted in the exclusion of eleven citations. The application of this exclusion criterion could be seen to result in underrepresentation of the number of published studies pertaining to use of ultrasound simulators for gynaecological ultrasound (see chapter 5 for a review of the relevant validation studies). A systematic review on hysteroscopy simulators, [91] which simply aimed to identify studies relating to a simulator intervention, used the demonstration of different types of validity as a framework for the review. Skills transfer, Kirkpatrick level 2b could be said to represent concurrent validity,[44] and inclusion of validity studies examining face and construct validity would have given more level 1 and 2a Kirkpatrick level outcomes than are presented.

For the studies included in this review validity was presumed for the intervention and validation is important for effective learning[92] whereas in actual fact Tolsgaard et al. and Alsalamah were the only authors to use an intervention which they had first studied the validity of and published their findings.[65, 83]

Tolsgaard et al.[72] used a validated intervention and reported outcomes at the highest level of Kirkpatrick hierarchy demonstrating benefit to patients but the study did have some limitations. Nearly twice as many patients rated outcomes came from the intervention group than the control group possibly reflecting differences in levels of commitment to data collection in the two groups. This differential reporting is likely to have introduced some bias; for example, patients with considerable discomfort may have made informed consent difficult to obtain and thereby could have limited the

amount of patient reported outcomes in the comparator, non-simulation group. Such selection bias may have led to under estimation of effects; the authors did attempt to overcome this by a stratified analysis by type of diagnosis.[72] The other unpublished RCT with a validated intervention was the study by Alsalamah.[83] This RCT was fraught with challenges and an example of real world research in an uncontrolled clinical setting. The sample for analysis was too small as there was incomplete compliance with the intervention and follow up. Limiting factors were cited as access to simulation location, lack of protected time given to trainees to practice on the simulator and lack of clinical training sessions had an impact on time taken for skill acquisition.

Comparison with other studies

This review is the only focussed systematic review on ultrasound simulation for SBME in gynaecology. Prior systematic reviews and meta-analysis of SBME have either not included ultrasound,[63] or have broadly examined simulation based ultrasound education, including studies from different specialities and diagnostic as well as interventional procedures.[62] In obstetrics and gynaecology, one systematic review focussed on high fidelity obstetric ultrasound [93] finding a small body of moderate level evidence with positive results for the use of high fidelity simulators in obstetric ultrasound but with potential risk of bias throughout. Other reviews including gynaecological ultrasound simulation being have been narrative and non-systematic.[51, 71]

In addition to the review by Cook et. al[63] examining technology enhanced simulation for health professionals education this review adds focus on simulation for ultrasound and also included studies that compared simulator training to an alternative

instructional method. In keeping with the systematic quantitative review examining undertaken by Cook et al.[63], the current review also found overall benefits of technology enhanced simulation over standard, non-simulation based approaches to training on acquisition of knowledge, skills, behaviours and patient outcomes. However, interpretation of these findings were restricted in both reviews by the observation of substantial heterogeneity between studies. A prior review by Sidhu et al.[62] conducted in 2012 focussed on simulation-based education in ultrasound practice training and included six studies relating to diagnostic ultrasound. Only one primary study by Maul et. al.[76] was included in both reviews, the other five studies being excluded from the current review because three were validity studies, one concerned abdominal scanning and one was simulated emergency scenarios with ultrasound image interpretation as one part of the scenario. Since 2012 there has been increased interest and development of simulation-based training and so our review includes recently published data with a greater emphasis on gynaecological ultrasound, namely transvaginal scanning. Osbourne et al.[93] focussed their review on high fidelity simulators for obstetrics but like the current review in gynaecology, they identified only a small number of studies of variable quality and highlighted the gap in the evidence base for SBME in obstetrics.

Meaning of the study and implications for education and clinical practice

The use of SBME for training in gynaecological ultrasound appears to be effective in improving learner-based and patient-based outcomes of relevance. The best evidence comes from a single RCT that shows simulation learning to a pre-determined 'mastery level' prior to undertaking clinical training is associated with a reduction in patient discomfort and patient reported perceived safety and confidence in their

provider.[72] There was also a reduction in the need for repeated patient examinations and trainee supervision. Other RCTs and observational studies identified in this review also generally found a benefit of simulation on top of clinical training in acquisition of knowledge and skills such as competency in completing pre-set tasks, obtaining good quality images and precision of measurement. Thus, post-graduate training programmes for teaching gynaecological ultrasound should look at incorporating ultrasound simulation. In this way standard approaches for teaching gynaecological ultrasound in clinical practice can be improved for all stakeholders; students, trainers, patients and the wider health services. This is because investment in SBME has the potential to increase both the effectiveness of training in both learner-based and patient-based outcomes and also the efficiency of training with quicker acquisition of proficiency. Although no studies in the review conducted a formal cost-effectiveness analysis, this needs to be considered to determine if investing in modern simulation is economically viable.

Simulation is being adopted across medical education for doctors, nurses and allied health professionals to acquire craft skills such as diagnostic imaging and therapeutic interventions such as conducting surgical procedures. Simulation should be seen as an adjunct to clinical training and not as a replacement. However realistic the simulation may be, the authenticity cannot always stretch to the particular nuances of working in the real clinical environment such as dealing with stress, time pressures, noise and the variation in types of patients and pathologies to name just a few variables. Arguably, simulation training is even more attractive in gynaecological ultrasound as it helps overcome some of the barriers to acquiring experience and competencies in undertaking intimate procedures such as vaginal examination. The very nature of such examinations makes training more challenging and so SBME should help push

trainees along ‘the learning curve’, helping them acquire confidence thereby enhancing the educational value of such encounters for trainees and the experience of patients. We identified more published work evaluating simulation in gynaecological ultrasound as opposed to obstetric ultrasound, and this observation may reflect the greater appeal of simulated solutions to scanning because of the intimate nature of ultrasound transvaginal examination.

Unanswered questions and future research

In the last five years, advances in simulation technology supported by an increased recognition for the potential of SBME, has resulted in an expansion in the evidence-base relating to the effectiveness of ultrasound simulation in gynaecology. However, even with the inclusion of these newly published studies the evidence-base remains small with only 14 studies of variable quality included in this contemporaneous review. Furthermore, the studies varied widely in their design and were conducted in different health care systems across Europe and North America. To optimise SBME in gynaecological ultrasound, we need data from head to head comparisons between types of simulation and how best to employ them within training programmes. This was difficult in the current review because the included primary studies evaluated a wide variety of simulators ranging from high fidelity virtual reality (VR) simulators to low fidelity pelvic mannequin models and where the same simulator was used they were often utilised during training in different ways. In addition, heterogeneity between studies precludes rigorous comparisons and data aggregation. Thus, future studies of effectiveness should be randomised comparing not just methods of simulation, but also how simulation is optimally integrated into training programmes to supplement or replace current clinical training. The interventions should be

explicitly described in terms of instructional design, tasks set, chosen outcomes and methods of assessment. Identification and standardisation of key educational 'core' outcomes should be undertaken by a consensus group of expert clinicians and other stakeholders to aid comparison between studies. Research into the development of validated assessment tools would help this process.

The studies in this review included a high proportion of doctors training in obstetrics and gynaecology, this being the target population for focussed gynaecological ultrasound training in modern practice training for point of care ultrasound. However, future studies should include mixed populations such as trainee sonographers, nurse practitioners and radiologists to enhance the generalisability of findings. Health care systems and post-graduate training programmes vary across regions and national boundaries[17] such that studies in medical education are frequently context specific and a single study from one area is rarely sufficient to definitively answer a question.[94] The only high quality RCT evaluating SBME at Kirkpatrick level 4[72] needs to be replicated in different settings such as the UK National Health Service. Indeed, if the findings are to be applicable to RCOG trained gynaecologists in the UK training setting this is essential. More trials are needed at this higher level of utility, to demonstrate the impact of SBME on the way organisations practice and how interventions benefit patients and the wider health services. Economic evaluations to look at cost and utility should be undertaken alongside such studies.

CONCLUSION

Delivering high quality training in gynaecological ultrasound scanning is challenging, as it is a complex skill that takes time to acquire. Simulators have been developed and

introduced to help train professionals in pelvic ultrasound examination across the developed world. Evidence from individual studies evaluating the effectiveness of simulation instruction in gynaecological ultrasound compared with conventional supervised clinical training is associated with significant improvements in learner outcomes, such as competence, skill and confidence as well as patient outcomes such as reduced discomfort and increased safety and confidence in provider. Simulation training also appears to make clinical training more efficient in the reduced need for supervision. Future research needs to confirm these findings in different settings.

CHAPTER 3

National survey of deanery ultrasound co-ordinators on the status of simulation in ultrasound training for obstetricians and gynaecologists.

ABSTRACT

Introduction

The Royal College of Obstetricians and Gynaecologists ultrasound-training curriculum is delivered across the UK to postgraduate doctors, training in geographical areas known as deaneries. Deanery ultrasound coordinators oversee ultrasound education to approximately 1,850 trainees nationally. This study aimed to explore the national status of simulation availability and use in ultrasound training for obstetricians and gynaecologists.

Methods

All deanery ultrasound coordinators across the UK were asked to complete a web-based survey in April 2014.

Results

Of 26-deanery ultrasound co-ordinators 24/26 (92%) responded to the survey. Formal programmes to deliver ultrasound training existed in up to 10/22 (45%) of deaneries for certain scan modules. Mostly training was done through “hands-on” supervised scanning. Ultrasound simulation models were available, the ScanTrainer

transvaginal® model being the predominant type. 10/16 (62.5%) deaneries had at least one ScanTrainer transvaginal®, the next most popular type was the Bluephantom™ mannequin simulator 6/16 (37.5%), three deaneries had both. Not all deaneries had simulation equipment for ultrasound; at least 3/16 (18.7%) reported no simulator equipment and the status was unknown in the non-responders to this question 8/26 (30%). When simulation was used in delivery of ultrasound training it was most often for the early pregnancy and gynaecology modules. Use of simulators was on an ad-hoc basis 16/16 (100%) of the time and only in a formal programme by 1/16 (6%) deaneries, its use was often unsupervised 10/16 (62%) and it was reported that simulation equipment where it was available was underutilised (12/15; 80%). Deanery ultrasound coordinators were divided in opinion about whether simulation has significant role in the acquisition of scanning skills.

Conclusion

There has been some investment in ultrasound simulator equipment, predominantly for gynaecological scan training. However the ratio of trainees to simulators is high, simulators are not being used in a structured way as part of formal programmes; use is ad-hoc and largely unsupervised. Moreover those responsible overseeing the delivery of ultrasound training are unsure of its role in skill acquisition. Research needs to be done to inform and direct future training programmes involving simulation.

IINTRODUCTION

In the United Kingdom (UK) for postgraduate doctors specializing in obstetrics and gynaecology (O&G) the ultrasound curriculum is set by the Royal College of Obstetricians and Gynaecologists (RCOG) and approved by the General Medical Council (GMC). Health Education England (HEE) delivers the curriculum through postgraduate schools of obstetrics and gynaecology. The UK is divided geographically into regions and each region has a postgraduate school responsible for training doctors. Doctors registered with a training number from the RCOG undertake a seven-year competency based training programme which is delivered within these regional post graduate schools, known as deaneries.

The RCOG requires that all trainees are able to scan at least at a basic level. Ultrasound is an integral part of training and the basic ultrasound competencies must be achieved before a trainee can advance from basic training, year 1 and 2 to intermediate training in year 3. Basic ultrasound modules are seen as a springboard for further supervised training and the acquisition of more advanced ultrasound competencies. Beyond basic level training the RCOG offers three intermediate scan modules. These modules are optional on completion of the basic modules. Trainees will choose them if they have a particular interest in developing an advanced training skill or a career where ultrasound is beneficial, the content of these modules is outlined in Appendix 5 and 6.

The RCOG oversees and supports the ultrasound training of approximately 1856 specialty trainees in obstetrics and gynaecology,[95] through a national network of deanery ultrasound coordinators (DUCs) who sit on a committee chaired by the RCOG appointed ultrasound officer. The ultrasound officer and the DUCs will

usually be obstetricians and gynaecologists with expertise in scanning and interest in training. The geographical deaneries are large areas containing multiple teaching hospitals; there will be at least one DUC for each region, sometimes more for example if the region is divided further into north and south or one oversees gynaecology, and one oversees obstetric scanning. The DUC should have oversight of training which is then facilitated by a network of local ultrasound leads within individual hospitals; and a further network of clinical supervisors to deliver the training in ultrasound, these may be nurses, sonographers, midwives, radiographers and consultant obstetricians and gynaecologists.

Study hypothesis

The hypotheses that led to the conception of this study were:

- Simulation is not part of a structured programme for ultrasound training for trainee obstetricians and gynaecologists and
- Support for simulation training in ultrasound is poor on the basis of the scant evidence base for its use and the cost of the equipment.

The aims and objectives of the study were thus:

Aim

The main aim of this survey was to explore the national status of simulation availability and use in ultrasound training for obstetricians and gynaecologists.

Objectives

The objectives of this survey were to:

1. Determine the existing structure for delivery of the RCOG basic and intermediate scan training modules.

2. Assess if simulation was being used in delivery of ultrasound training.

3. Know which simulators were available and how they were used.

4. Gauge opinion from the people responsible for the delivery of ultrasound training in the UK on the role of simulation in trainees acquiring scanning skills.

METHODS

An electronic questionnaire survey containing 10 questions and 24-points (Appendix 7) was developed to assess the use and availability of ultrasound simulators for ultrasound training, and understand opinion on the role of simulation in modern ultrasound training.

The survey was developed with input from the RCOG ultrasound officer, the RCOG simulation advisory network lead at that time and a consultant in O&G with an interest in ultrasound training. The survey was kept intentionally concise for ease of completion. The survey was piloted on a group of four local ultrasound trainers to ensure clarity of questions and unambiguous language. The survey was then uploaded onto the web-based survey tool SurveyMonkey™. The individuals involved in survey development checked the formatting and ease of use of the electronic version. The survey link and explanatory cover letter were sent out via an administrator within the

RCOG to all DUCs. The data was collated and responses tracked using the online survey software.

The survey was carried out between February and April 2014. There were 26 DUCs in post at the time; we purposely sampled all 26 to get coverage of the picture across the whole of the UK. Email addresses were obtained from the RCOG ultrasound officer and a test email sent with an introduction to the survey. This was followed by an email invitation to participate, endorsed by the RCOG ultrasound officer, containing a link to the survey. The survey was accessible for eight weeks and two reminders to participate were sent by email. Participation was voluntary and no incentives were offered for participation.

The questions were a mix of open ended and closed questions. Closed questions were structured as single answer, multiple option or Likert scale. Descriptive statistics were calculated with the number of responses as the denominator. Respondents could skip questions that were considered not applicable or relevant to them.

RESULTS

The response rate was 24/26 (92%), with 22 of the respondents completing all the relevant survey questions. Respondents' demographics were not collected, however, by virtue of their roles as deanery ultrasound coordinators, all 26 potential participants were consultants in obstetrics and/or gynaecology.

Delivery of RCOG scan modules

A summary of the RCOG ultrasound modules offered in the UK and details of the structure of training is outlined in Table 8. All responding DUCs reported that their deaneries delivered basic scan modules. 9/22 (41%) had a formal programme for delivery of basic early pregnancy and 10/22 (45%) had a formal programme for basic fetal assessment. Not all deaneries were able to deliver the RCOG intermediate scan modules. The numbers of deaneries offering intermediate scan modules 3, 4 and 5, which form the intermediate scan curriculum was 18-19/22 (82-86%) depending on the module. Free text comments were used to expand upon training and organisational issues. Free text responses cited reasons for not being able to deliver intermediate modules as not all hospitals in the region having capacity to do so, for example module 3: ultrasound of normal fetal anatomy could only be offered in some deaneries by fetal medicine specialists in certain units. There were also examples of only 2-3 hospitals in a region being able to offer scan training therefore there would be an interview process to allocate training if needed. For the three intermediate modules, 6-7/19 (31-37%) of respondents reported that their deanery offered a formal programme for these modules. It was commented by some that this was delivered on a national level by RCOG and the numbers locally were not sufficient to demand a local course.

Table 8 Summary of the RCOG ultrasound modules offered in the UK and details of the structure of training

	Module 1 Basic early pregnancy n (%)	**Module 2 Basic fetal assessment** n (%)	**Module 3 Intermediate normal fetal anatomy** n (%)	**Module 4 Intermediate gynaecology** n (%)	**Module 5 Early pregnancy complications** n (%)
Deliver the module	24/24 (100)	22/24 (96)	19/22 (86)	19/22 (86)	18/22 (82)
Formal programme	9 /22 (41)	10/22 (45)	6/19 (31)	7/19 (37)	6/17 (35)
Hands on scanning	22/22 (100)	22/22(100)	18/18 (100)	19/19 (100)	18/18 (100)
Simulation	10 /22 (45)	6/22 (27)	5/18 (28)	10/19 (53)	8 /18 (44)
Theory	20/22 (91)	18/22(82)	10/18 (56)	13/19 (68)	14/18 (78)

Use of simulation in the delivery of ultrasound training

Supervised 'hands on' scanning of patients was the predominant method by which training was delivered across all modules and deaneries. Not all deaneries had access to simulators for ultrasound; at least 3/16 (18.7%) explicitly stated they had no simulation equipment, and for other non-responders the situation was unknown 8/26 (30%) in total. Simulation was used in delivery of training across all ultrasound modules to varying degrees. Simulation was used most frequently for module 1: basic early pregnancy scanning 10/22 (45%), module 4: intermediate gynaecology scanning 10/19 (53%) and module 5: early pregnancy complications 8/18 (44%). This could be explained by the availability of equipment (Table 9). The types of simulator available were surveyed, 16 DUCs responded; the most frequently available were the transvaginal ScanTrainer® (Medaphor, Cardiff, UK) reported by 6/16 (62.5%) and BluePhantom™ (CAE Healthcare, Redmond, USA) pelvic mannequin, reported by 6/16 (37.5%) (Table 9). When it was identified deaneries had simulation equipment

they were asked how many models they had, this ranged from 1-3 models within deaneries for the ScanTrainer®

Table 9 Types of simulator available in UK training deaneries

Type of simulator	Number of responses (n=16) n (%)
Blue Phantom™	6 (37.5)
Ultrasim®	1 (6.25)
Medaphor® Transabdominal	2 (12.5)
Medaphor® Transvaginal	10 (62.5)
Blue Phantom™ and Medaphor® Transvaginal	3 (18.75)
None	3 (18.75)

When simulators were available for training it was reported that the equipment was used on an ad hoc basis 16/16 (100%) of the time and only in a formal programme by 1/16 (6%) (Figure 3). The training was a mixture of supervised, 56% (n=9/16) and unsupervised, 62% (n=10/16) training (Figure 4).

Figure 3 DUCS responses with regards to if the use of simulation equipment is ad-hoc or part of a formal programme (n=16)

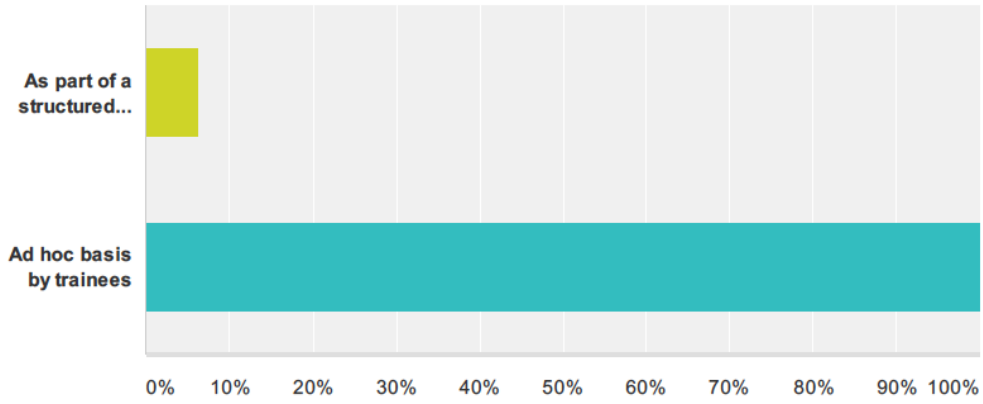
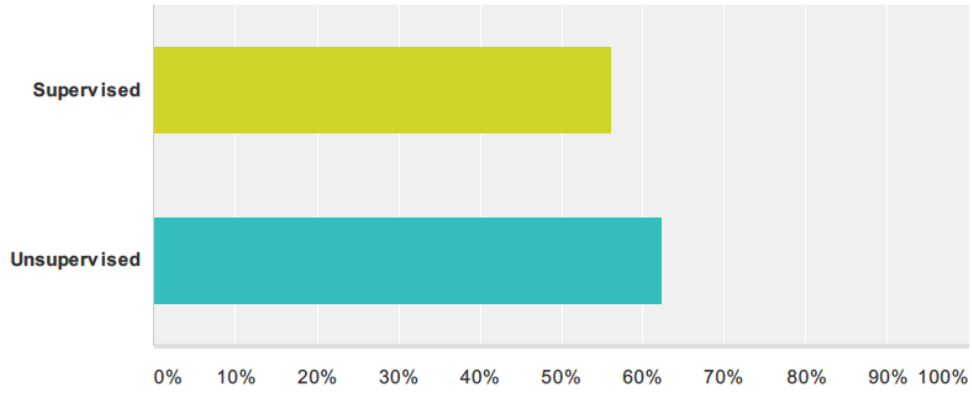
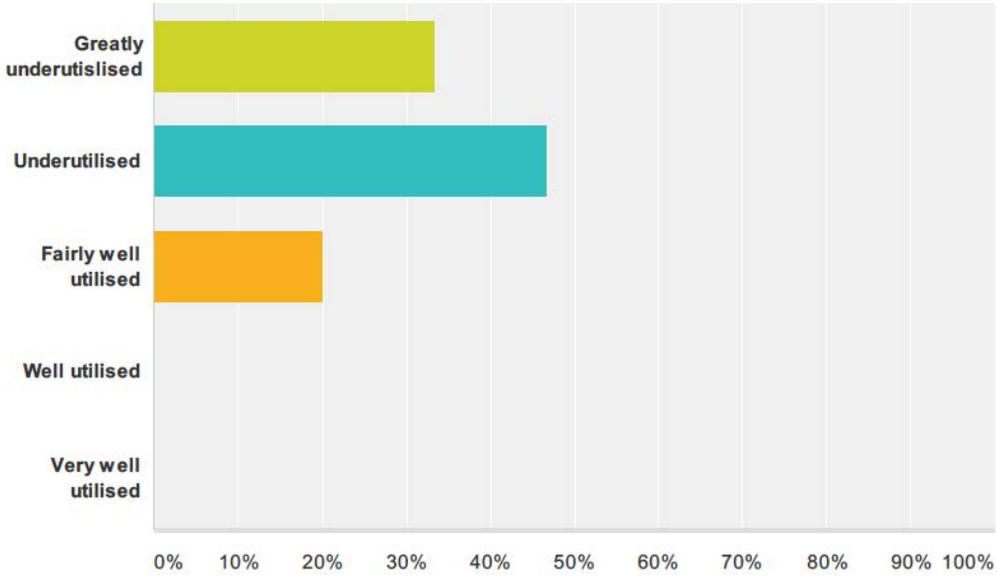


Figure 4 DUCs responses with regards to if simulation equipment is supervised or unsupervised (n=16)



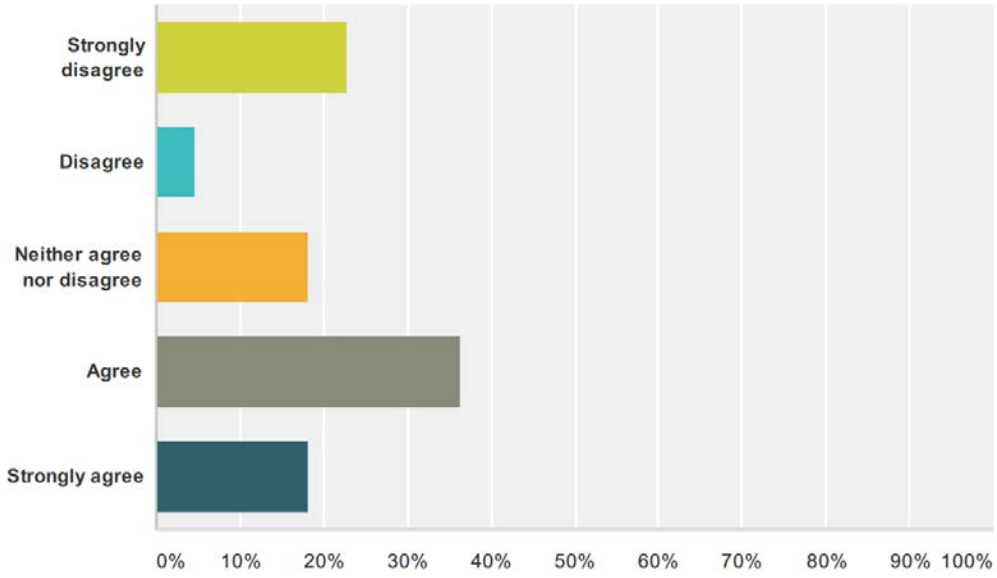
Where simulation equipment was available DUCs were asked their opinion on how well it was utilised, 12/15 (80%) reported that the equipment was underutilised or greatly underutilised, with only 3/15 (20%) reporting it was fairly well utilised (Figure 5).

Figure 5 DUCs views on the utilisation of simulation equipment (n=15)



DUCs were asked whether they thought simulation had a significant role in the acquisition of scanning skills for trainees in obstetrics in gynaecology. The response to this question showed a spread of opinion (Figure 6), 12/22 (54.5%) agreed or strongly agreed it had a role, 6/22 (27.2%) disagreed or strongly disagreed it had a role and 4/22 (18.2%) were undecided.

Figure 6 DUCS views as to whether they consider simulation has a significant role in acquiring scanning skills for O&G trainees (n=15)



DISCUSSION

Principle findings

This was a survey to explore the national status of simulation in ultrasound training available for trainee obstetricians and gynaecologists. The leading clinicians responsible for ultrasound training in the UK were approached. This survey found that not all deaneries were able to offer the three intermediate scan modules and use of formal programmes to deliver ultrasound training were limited. There were ultrasound

simulation models available nationally, with the transvaginal ScanTrainer® model being the predominant type followed by the Bluephantom™ mannequin simulator, however not all deaneries had simulation equipment for ultrasound. There is not uniform access to simulators for trainees and as the number of trainees across the county is estimated 1850 the trainee to simulator ratios are high. Simulation use in the delivery of ultrasound training in deaneries ranged from 28-53%; this variation was according to the different modules. When simulation was used in delivery of ultrasound training it was most often for the early pregnancy and gynaecology modules (module 1, 4 and 5), likely due to the predominance of the gynaecological simulator models. Use of simulators was on an-hoc basis and often unsupervised. DUCs reported that simulation equipment where it was available was underutilised. There was divided opinion about whether simulation has a significant role to play in the acquisition of scanning skills for trainees.

Strengths and limitations of the study

The study has several strengths, which increase its validity. The response rate of 92% was high increasing the precision of the findings, reducing selection bias and hence representative of practice across the UK, however there was a drop off rate through the survey. Purposive sampling enabled me to capture the views of consultants who had overall responsibility for developing and delivering ultrasound training regionally and nationally. The participants were chosen because of their position as a regional representative on a national committee, their role on that committee was to be well informed about the delivery of ultrasound training in their deanery. However, the study has some limitations, firstly, the survey was piloted on a small group of people with knowledge of the subject matter; they found the terminology self-explanatory

however terms such as formal programme where not explicitly defined and therefore open to interpretation. Nevertheless, because of the nature of survey participants and their active involvement in delivering aspects of the RCOG curriculum they would have had a high level of knowledge and understanding of the subject matter. Secondly, not all respondents that started the survey finished it; two respondents were lost after the first question and 15/26 completed the survey in full. I only surveyed the DUCS perspective; no other personnel involved in the delivery of ultrasound training such as the local ultrasound coordinators, clinical supervisors of trainees views were captured.

Comparison with other studies

The perspective of the trainee in regards to ultrasound training has been explored in the UK, Europe and Canada. Patel et al.[31] surveyed trainees (n=70) in the East Midlands deanery UK, on the provision of clinical and simulation training in ultrasound. They found that a relatively small number of trainees had completed RCOG intermediate scan modules, only between 6-18% of respondents, depending on the module, highlighting as in my survey, that there does seem to be issues with trainees accessing training in intermediate level scanning. Similarly in Canada[96] it is reported training in gynaecology scanning is lacking, with only one third of those surveyed believing they would be sufficiently trained in gynaecology ultrasound. Patel et al. [31] explored some of the barriers to training in the UK and the most common limiting factor in obtaining and maintaining competency in ultrasound was lack of time because of service commitment according to 58/70 (83%) trainees, lack of dedicated ultrasound training sessions reported by 55/70 (55%) and lack of trainer engagement reported 31/70 (44%). They found virtual reality simulators were

available to 50% of trainees and a formal simulation-teaching programme was available to 14%. Interestingly and in contrast to my national survey of trainers Patel et al.[31] found that trainees held very positive views of simulation for ultrasound training, 73% considered simulation essential in ultrasound training. There seems to be a discrepancy in opinion on the role of simulation in ultrasound training between trainers and trainees, with the trainees more positive about the role simulation could play in training. In Europe a large sample of 621 trainees from Denmark, Norway and Sweden were surveyed to explore the association between clinical training characteristics and trainees' level of confidence in performing ultrasound. This survey reported low levels of satisfaction with transabdominal and transvaginal ultrasound training, opinions on the role of simulation and attitudes towards simulation were not explored however the author does conclude that perhaps simulation based training can be used to address some of the challenges.[26] In the United States in 2010 a survey was sent to program directors for critical care medicine fellowships exploring barriers to US training. [97] The respondents (n=90) recognised the importance of ultrasound training in critical care medicine and 80% wanted to get their fellows trained but barriers to training were identified as fellow turnover, insufficient faculty training and perceived length of time required for echocardiography training.[97]

Implications for clinicians and policy makers

There is a curriculum and a structure to ultrasound training from the RCOG but this survey has highlighted issues exist with the uniform delivery of the training across the UK and in Europe.[17] There are low numbers of trainees completing RCOG intermediate scan modules. From this survey and that of Patel et al.[31] we do not know if it was lack of demand for the modules or if training could not meet the

demand. Given that not all deaneries were able to offer intermediate modules, the low satisfaction in training reported and the very positive response of trainees towards simulation as an opportunity to do some scanning you can infer it is the situation of training not meeting demand.

There has been investment in ultrasound simulators, at least 70% of deaneries have a simulator(s), the investment in these is perhaps in response to difficulties in training delivery, but access is not uniform with some deaneries having no equipment. In the UK there are an estimated 1856 trainees [95] so with just 10 transvaginal ScanTrainers® this is not sufficient for universal access and where simulators are available they are being utlised on an ad-hoc and largely unsupervised, not as part of formal programmes. This survey highlighted a spread of opinion about whether simulation has a significant role in the acquisition of scanning skills for trainees from those who have overall responsibility for developing and delivering ultrasound training regionally and nationally. This has implications for educational policy makers, as these trainers are key stakeholders in delivering training. If simulation is going to have a role in ultrasound training, the existing equipment needs to be better utilised or investment is going to be made in simulation in the future then this group will be a key group to convince of its effectiveness.

Unanswered questions and future research

This survey has shown that amongst key stakeholders in training, there is a lack of consensus on the role simulation has to play in the acquisition of ultrasound scanning skills. The survey was carried out in 2014 just as good quality and randomised control trials in this area were starting to be published. Repeating the survey now could possibly reveal a shift in opinion more in favour of simulation, as there is an emerging

evidence base for the use in learning (Chapter 2). It would also allow up to date knowledge of which simulators are available and if any further investment has been made in simulation.

This survey was a short quantitative survey and barriers to the effective delivery of ultrasound training, and in particular the RCOG intermediate scan curriculum were not explored, but should be in future surveys or with qualitative methods. This survey only explored the opinions and knowledge about delivery of ultrasound training from the DUCs, there would be value in also exploring the opinions of the local ultrasound coordinators and ultrasound trainers within individual hospitals of different sizes and where perhaps different models of ultrasound training exist to get a richer picture of delivery of ultrasound training in the UK. As well as the perspectives of the trainers the opinions of trainees on a national scale would be the other stakeholder perspective to be explored in future research.

CONCLUSION

In the UK in 2014 basic ultrasound scan training was being achieved but not all deaneries were able to offer the intermediate scan modules and use of formal programmes to deliver ultrasound training were limited. Some investment has been made in ultrasound simulators, in the main Scantrainer transvaginal® and Bluephantom™ manikins but the trainee to simulator ratio is high and access is not universal, some deaneries have no simulation models. Use of simulators is largely in an ad-hoc and unsupervised manner and thought to be underutlised. Moreover those responsible overseeing the delivery of ultrasound training are divided in opinion on the role simulation has in skill acquisition. The survey needs to be updated, the scope

broadened to include other stakeholders in training, local ultrasound co-ordinators, educational supervisors and trainees. Qualitative methods could be considered to gain richer insights. Alongside mapping current availability and access there needs to be research into learning and assessment of ultrasound training skills and the role simulation can play in order to inform and direct development of best practice training programmes involving simulation.

CHAPTER 4

Medaphor® Assessment in Specialist Training and its Economic Recompense Study (The MASTERS Study) - an assessment of study design and report of a failed observational study

ABSTRACT

This chapter describes the design of a prospective multicentre controlled observational study, done in real practice aiming to evaluate the efficacy and cost effectiveness of virtual reality (VR) simulation training with transvaginal ScanTrainer® (Medaphor, Cardiff, Wales, United Kingdom). A trial was designed in which two groups of obstetrics and gynaecology trainees were observed through their intermediate ultrasound training. Trainees in the control group had already started the Royal College of Obstetricians and Gynaecologists (RCOG) intermediate scan modules and were training by means of supervised clinical practice (conventional method). The intervention group were trainees planning to start an intermediate scan module, before starting supervised clinical practice they were asked to first practice on a high fidelity VR simulator. This trial faced several organisational and individual barriers that posed challenges to the recruitment process including variation in trainee engagement, accessibility to simulators and lack of available time limiting opportunities for clinical and simulator ultrasound training. Despite pragmatic adaptations to the study design to overcome the unanticipated challenges, it became

apparent that the intended outcomes and endpoints would not be reached and so the study was closed on the grounds of futility.

A descriptive analysis of the available data from this pilot feasibility trial is presented and the reasons for failure of the study explored with a view to informing the design and execution of alternative, feasible future trials. Training future doctors effectively and efficiently remains an important goal, because advances in imaging technology have increased the accessibility and portability of gynaecological ultrasound, such that it is an integral part of contemporary gynaecological practice in the National Health Service (NHS). In view of the lack of data in a UK NHS setting about the effectiveness and cost effectiveness of simulation for gynaecological ultrasound scanning future studies are urgently needed.

INTRODUCTION

Transvaginal ultrasound (TVU) is an invaluable diagnostic tool in obstetrics and gynaecology, which is considered safe, but its use is highly operator dependant.[17] Lack of knowledge into the limitations and applications of ultrasound and lack of skill could lead to diagnostic error and compromise patient safety. To overcome these risks ultrasound training needs to be robust and assured. However training in basic ultrasound for obstetrics and gynaecology trainees is varied and remains a challenge.[17] Attaining proficiency in obstetric and gynaecological ultrasound is thought to be associated with long learning curves [98] meaning that training times can be prolonged and is as a consequence frequently challenging to deliver, placing a substantial burden on scarce health care resources. More specifically, these barriers to delivering training in ultrasound include the availability of appropriate trainers and the impact of training on service delivery because training reduces efficiency of scan lists.[30] Furthermore, particularly in the case of junior doctors, ring-fenced training in ultrasound comes with an opportunity cost in that doctors are invariably taken away from other clinical duties, which need backfilling. There are time constraints on trainees due to the reduction in training hours in general as an effect of the European Training Working Time directive [99] and pressures on service delivery limiting opportunities for supervised training time due to service commitments.

Given the challenges of delivering ultrasound training balanced with the demand for training, recent advances in the development of haptic-feedback VR scan simulators, offer a potential solution for more effective and efficient training, reducing time taken to attain proficiency and freeing up human and other health care resources. The transvaginal ScanTrainer® is driven by computer generated software and has the two

main components; (i) *ultrasound training software* which provides real ultrasound images and virtual anatomy illustrations for orientation, (ii) *a learning management system* which provides assignments, tasks and feedback therefore allowing self-directed learning. To improve fidelity, it has a replica transvaginal probe that is attached to a haptic device, which provides force feedback as though scanning a real person, and a high-resolution, touch screen with the conventional ultrasound controls and a roller-ball mouse to navigate these, see Appendix 8. The launch of the ScanTrainer® was heralded as 'revolutionary,'[100] the technology benefiting trainees by allowing them to develop cognitive and hand-eye co-ordination skills in a flexible non-clinical learning environment and benefiting hospitals by releasing clinical equipment and expert supervision from the constraints of training to focus on service delivery. In this way it was hoped that scan training would be quicker and cheaper.[100] Despite the novelty of the technology and its potential for providing feasible, safe, effective and cost-effective training only a minority of hospitals providing training in gynaecological ultrasound invested in the pioneering system, as detailed in Chapter 3 it was the most prevalent simulator system in UK training deaneries. According to information obtained from the manufacturers in December 2012 there were ten transvaginal ScanTrainer® models installed in the UK.

At the time of designing and planning this study in 2012, there was little evidence to support the manufacturers claims of benefit and there was the suggestion that the equipment, although available, was being underutilised, see Figure 5, Chapter 3. The potential utility of simulation based medical education (SBME) as an adjunct to clinical ultrasound training had been recognised [17, 76, 101, 102] but there was limited evidence of skills transfer from simulation to clinical learning. We therefore aimed to fill this evidence gap designing and conducting a study to evaluate the effectiveness

of VR simulation for transvaginal ultrasound scan training using the ScanTrainer®. ScanTrainer® was chosen as the most common simulator available across the country and two models had already been purchased in the West Midlands.

Study hypothesis

The hypothesis that led to the conception of this study was that the use of an ultrasound simulator could reduce the amount of clinically supervised scanning a trainee needed to achieve competence, which could be demonstrated by transfer of skills to real patients and could improve efficiency of training and reduce costs.

Aim

To assess the efficacy and cost-efficiency of VR simulation for transvaginal ultrasound scan training.

Objectives

The study objectives were to:

- determine if use of an ultrasound simulator prior to conventional training on patients can reduce the amount of clinical training needed to reach competency
- demonstrate transfer of skills from a simulator to a scan performed on a live patient
- determine the cost efficiency of VR simulation for transvaginal ultrasound scan training

METHODS

The study was performed between January 2013 and October 2013. It was a multicentre, controlled observational study with two groups: a control group who were training using conventional 'hands-on' ultrasound training methods through supervised practice on patients and an intervention group, who were recruited to practice on a VR simulator; the transvaginal ScanTrainer®, prior to commencing their planned conventional 'hands-on' training.

Population

The population studied was obstetrics and gynaecology trainees undertaking RCOG intermediate level scan modules. Intermediate ultrasound modules in gynaecology and early pregnancy complications are two, of the three, optional modules available after completion of basic ultrasound training, see Appendix 6. The conventional training group trainees had already started their intermediate scan training in early pregnancy complications or gynaecology. For most it was assumed this would begin at the start of the academic year, which was August 2012, with rotation in hospital placements. Participants in the conventional group should be training in clinical practice doing supervised scanning and working towards completion of the modules. The intervention group comprised of trainees who had expressed an interest in doing the intermediate scan modules, had not yet started clinical training, but would be likely to do so within six months of recruitment into the study. The inclusion and exclusion criteria were kept broad and can be seen in Table 10.

Table 10 Inclusion and exclusion criteria

	Inclusion	Exclusion
Conventional group	O&G trainee already	

	undertaking intermediate scan module 4 or 5	
Intervention group	Eligible to start intermediate scan training modules within 6 months of recruitment into the study	Prior knowledge that they will be out of training within 6 months of entry into the study

Setting

It was a multicentre trial with collaboration across five deaneries in the UK: West Midlands, East Midlands, Mersey, Wessex and Severn. These deaneries were invited to participate in the study as they had access to the transvaginal ScanTrainer® and had an identified ultrasound trainer committed to volunteering time to run the study locally.

Sample

It was not known how many trainees at the time were doing intermediate scan modules. All trainees across the deaneries who may have been undertaking training or planning to in the near future undertaken an intermediate module were approached regarding participation. We planned to recruit those doing scan training to the control group and at least an equal number to the intervention group.

Recruitment

The deanery ultrasound co-ordinator (DUC) identified eligible participants, and the researcher sent information about the study via email and asked potential participants to register their interest with an electronic reply. The trainees' current status in ultrasound training dictated eligibility for either the control or intervention group. Those who had already started an RCOG intermediate scan module were recruited

into the control group, those who expressed an interest in starting an intermediate scan module within six months were enrolled as the intervention group.

Study welcome packs were sent electronically to participants and the local study coordinator then arranged to meet the participant at a time convenient to them to go over the study requirements and ensure for the intervention group that the participants could access the simulator and required modules.

Interventions

It was anticipated the conventional training control group would continue with their usual method of gynaecological ultrasound training during the first six months of the study. At three time points during the six months, (at two monthly intervals) they would be asked to supply data using a logbook (Appendix 9) relating to their training over a representative 2-week period.

The simulation training group participants were each given a unique user name to log on to the transvaginal ScanTrainer®. The intervention group participants were asked to complete two core training modules and one advanced module on the transvaginal ScanTrainer®, the content of which mapped to the RCOG intermediate training modules in early pregnancy complications (module 4) and gynaecology (module 5), (Appendix 6). The local study coordinator arranged to meet the trainee and showed them the location of the ScanTrainer® and how to log-on. With the assistance of the sales representative from the company the required modules had been linked to login details. The basic functions of the ScanTrainer® were explained, supported by the introductory video on the system. The trainee was expected to undertake self-directed learning on the simulator during the first six months of the trial. At three time points during the six months (at two monthly intervals) they were asked to supply data using

the logbook format relating to their training over a representative 2-week period. Data on their progress in simulation could be accessed from tutor area of the ScanTrainer® which gave an overview of all users progress.

After six months an assessment workshop was planned at a time point when the conventional training group was expected to have completed their intermediate scan module(s) and the intervention group were expected to have completed the modules on the simulator and be ready to commence their clinical training. A further six months of follow up was planned when it was assumed both groups would be doing supervised and unsupervised clinical practice before another assessment workshop at 12 months. Figure 7 shows the trial design.

Outcomes

The primary outcome was defined as the time taken to reach competence in clinical practice. This was measured in the reported time spent undertaking clinically supervised practice in order to attain RCOG module sign off for the intermediate ultrasound modules in gynaecology and early pregnancy complications, measured in actual training time (days, weeks, and months). In the UK competency in obstetric and gynaecological ultrasound is assessed and evidenced through completion of a series of work-based placed assessments; the tool for assessment of technical skill is the Objective Structured Assessment of Technical skill form. [(103)] A summative OSAT gives a judgement about the ability of the trainee to perform a specified procedure independently. There are OSATs for various conditions specified in the RCOG curriculum requirements, such as ultrasound assessment of the normal female pelvis and ultrasound assessment of uterine fibroids.[(104)] A trainee needs to show

evidence of satisfactory completion of three summative OSATs for each specified procedure in the curriculum, involving two assessors in the assessment process.[104]

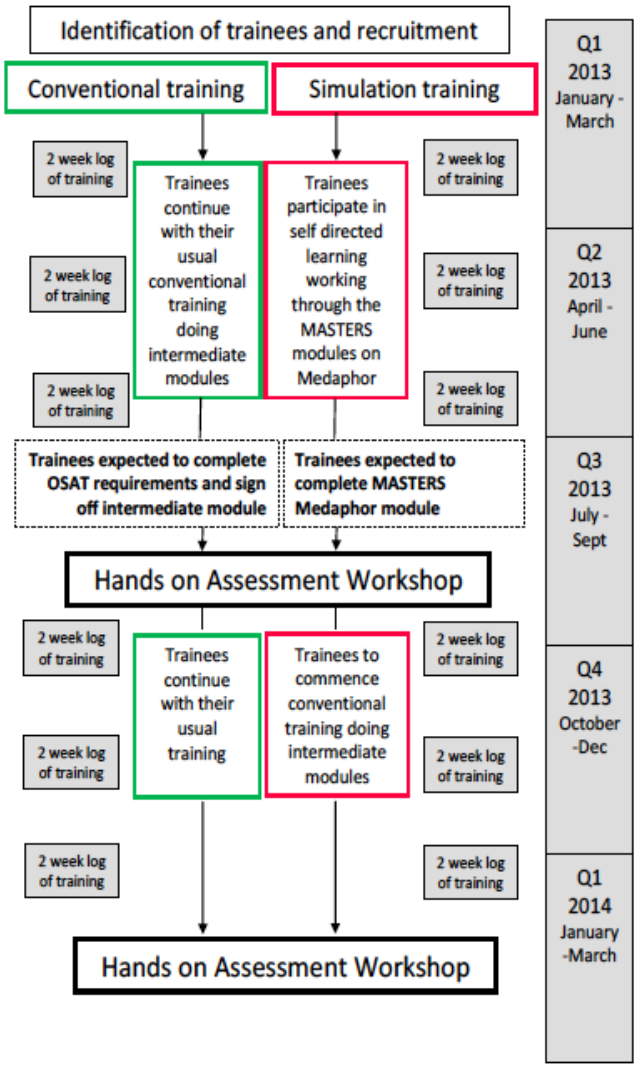
The secondary outcome was a test of skills transfer from simulator practice to real practice and a comparison of the performance of the groups by assessing the participants in a training workshop, where they would scan live models. The assessor was to be an ultrasound trainer, blinded to the group allocation, with assessment to be standardised using a summative OSAT form. These assessments were planned at six and 12 months from the beginning of the study.

To make an assessment of cost-efficiency of training methods we were looking to collect data on current clinical training. The training logs should have given data on number of scans performed with direct and indirect supervision and number of scans per hour to allow modelling based on underutilised scan appointments, assuming direct supervision of training reduces the efficiency of the scan lists. If the amount of direct supervision could be reduced by use of a simulator, then the potential cost-effectiveness could be realised by increased efficiency.

Ethics

The University of Birmingham Ethical Review Board provided ethical approval.

Figure 7 Flowchart study design



RESULTS

The study started recruiting in February 2013 and planned to run over a 12-month period. However, data collection stopped in October 2013 when it became apparent that the assessment stage was not going to be reached within the study time frame due to lack of progress in the gynaecological ultrasound training of both groups. From the start onwards this study faced several challenges in recruitment and compliance with

both the control and experimental interventions because of a variety of factors including trainee engagement, accessibility of simulation training, completion of simulated scan modules and lack of clinical ultrasound training.

In the study period 51 trainees were recruited, 17 in the conventional training group (control) and 34 in the simulation group (experimental). The average age of participants was 31 years, range 26 – 44 years. 14/51 (27.5%) were male and 37/51 (72.5%) female. The majority of the participants were in the West Midlands Deanery 26/51 (51%), East Midlands 6/51 (11.8%), Mersey 12/51 (23.5%), Severn 4/51 (7.8%), Wessex 3/51 (5.9%). Table 11 details the trainees according to speciality training year. Participants were spread across all training grades with the peak being the ST4 and ST6 cohort. Almost a quarter of trainees worked less than full time 12/51 (23.5%).

Table 11 Participants by year of speciality training

Speciality training year	Number of participants (%)
ST2	4 (7.8)
ST3	9 (17.6)
ST4	12 (23.5)
ST5	9 (17.6)
ST6	11 (21.6)
ST7	6 (11.8)

The trainees were spread across 25 different hospitals in the five regions, the number of trainees per hospital ranged from 1 – 6. There were six hospitals that had a ScanTrainer®, all deaneries had at least one ScanTrainer®, and it was generally based at a large central teaching hospital. In the West Midlands there were two hospitals

with ScanTrainers®, both were large teaching hospitals, one central hospital, the other in the north of the deanery. 18/51 (35%) participants were working in a hospital with a ScanTrainer® on site. The location of the simulator within the hospital varied and is detailed in Table 12. No simulators were within simulation centres.

Table 12 Location of the ScanTrainer® within the hospital and deanery

Deanery	Location within the hospital
East Midlands	In the doctors office
Mersey	In the ultrasound department
Severn	In an office on the gynaecology ward
Wessex	Temporary location on the neonatal unit
West Midlands	Hospital a - In the ultrasound department Hospital b - In an administrative area with consultant and secretaries offices

In the conventional training group 12 participants were undertaking both the gynaecology and early pregnancy modules, three were doing only the early pregnancy module and two only the gynaecology module. Almost one quarter (4/17; 23.5%) of the participants worked less than full time. Eleven (97.9%) of participants had done RCOG basic scan modules, which involve only transabdominal scanning, and one participant had done a prior intermediate module in gynaecology and was now doing the intermediate module in early pregnancy scanning.

In the intervention group, 11/34 (32.5%) of participants had never done a transvaginal scan on a patient, 23/34 (67.5%) had done a transvaginal scan on a patient, however

their experience was limited, 15/23 had done an estimated 0-10 scans (62.5%), 4/23 (17.4%) 10-20 scans and 4/23 (17.4%) had done between 20 up to 50 scans

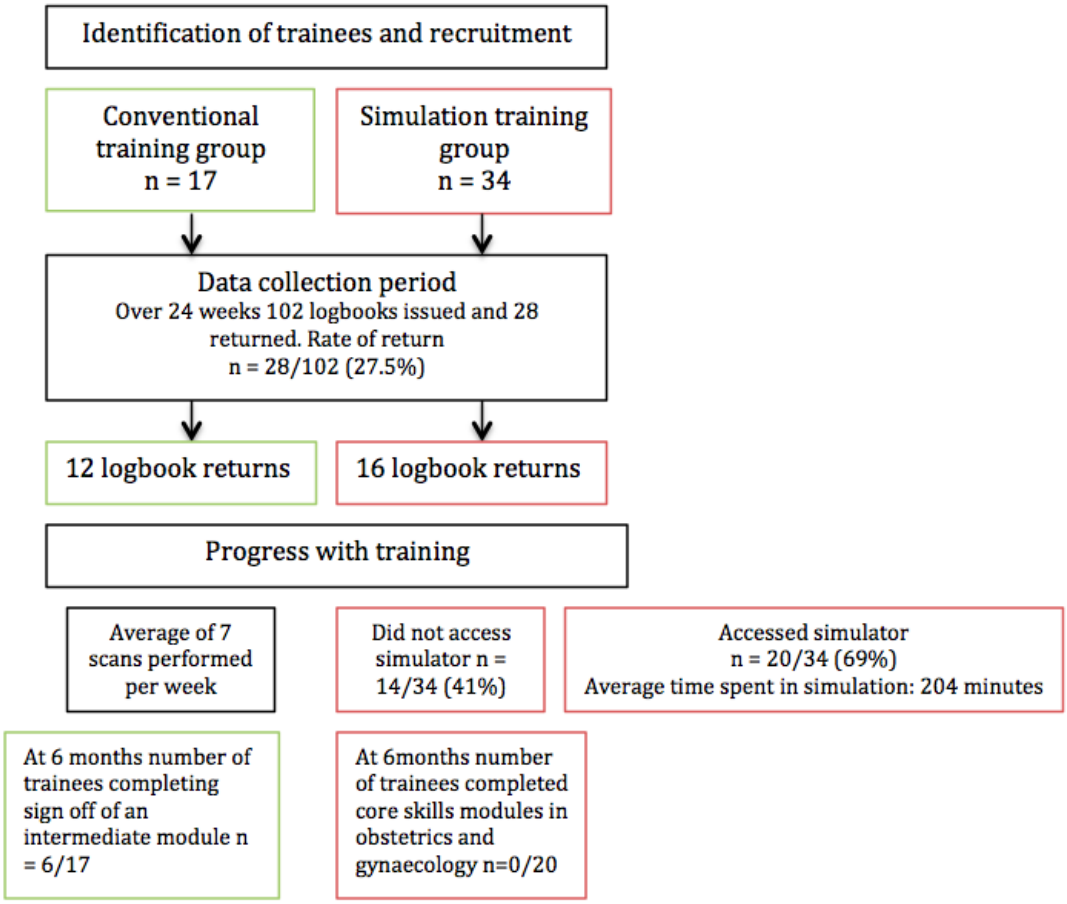
During the study period 102 training logbooks (Appendix 9) were sent out and 28/102 returned (27.5%), despite the researcher and collaborators efforts to increase the response rate with reminders and meetings. Paper logbooks were included in the welcome packs, which all participants had at enrolment, and sent on email during the two-week data collection period. Participants had the option of electronic or paper completion. Weekly reminders were sent by email from the researcher and local collaborators reminded participants face to face. After poor return rates in the first collection period, during the second data collection period text message reminders were used and the researcher called participants by telephone. There were 12 returns in the conventional group (12/34 35%) and 16 (16/68 24%) returns in the intervention group. The data from the logbooks showed the groups were getting few opportunities for scanning in clinical practice and were dissatisfied with their training. For those doing conventional training the returned logbooks covered 24 weeks of training and on average trainees doing the intermediate scan modules in early pregnancy and/or gynaecology were performing seven scans per week (range 0 – 23 scans).

In the intervention group 14/34 (41%) participants did not access the simulator. They were all contacted by email and telephone by the researcher and face-to-face where possible by the local collaborator, efforts were made to make the simulator accessible for trainees. In the West Midlands this involved getting the ScanTrainer® a permanent room in the ultrasound department, which was off the main hospital corridor not too far from the main entrance on the ground floor, as previously it was in a temporary set up in an administrators' office in the basement which was locked when the

administrator was not there. In Wessex deanery problems were encountered when the early pregnancy unit was relocated so the ScanTrainer® had to be found a temporary location in the neonatal unit. The 20 participants in the intervention group who did access the simulator spent 204 minutes (range 30 – 480 minutes) on average in simulation in this time. None of these trainees were able to pass core skills in both early pregnancy and gynaecology and no trainees attempted the advanced skills. The simulator breaks the modules down into assignments; these are individual instructed tasks, for example 'optimal demonstration of the uterus' and 'optimal assessment of the uterine cavity'. The simulator records the number of assignments attempted and the number passed to give a percentage of passed assignments, the average percentage of recorded assignments passed was 14% (range 0 – 31%). The simulator also records the number of attempts per assignment; the average number of attempts per assignment was 5.2 (range 2.2 – 18).

In the conventional group 6/17 (35%) trainees completed intermediate scan training before the study ended. Of these 4/6 (66%) were simultaneously doing module 4 (gynaecology) and module 5 (early pregnancy), so in total 10 intermediate modules were reported as signed off. It took these trainees on average 16.5 months and 148 scans to complete a module (range 47-300 scans per module).

Figure 8 Progress of participants through study



DISCUSSION

Principle findings

A prospective, multicentre, controlled observational study was set up, which was planned to run over a 12-month period. The study was curtailed after 10 months on the grounds of futility because of incomplete compliance by trainees and with gynaecological scan training precluding the collection enough data to evaluate the study outcomes.

The participants from five different training deaneries reflected the current training demographic being predominantly female with around one quarter in less than full

time training posts. All trainees had been qualified and practising as junior doctors for at least two full years with the majority of participants at Specialist Training (ST) years 4 to 6. Despite the ubiquity of transvaginal ultrasound in modern gynaecological practice, less than one in five trainees had ever performed a transvaginal ultrasound scan prior to commencing the RCOG intermediate training module and of half of those trainees who had experience of transvaginal ultrasound had limited experience performing less than 10 procedures.

The study showed that access to gynaecological scan training during clinical practice was extremely limited resulting in dissatisfaction. Most respondents in the control group, who were already enrolled on the RCOG intermediate scan modules, completing the electronic logbooks of experience during the study period, reported limited opportunities for scanning in clinical practice and were dissatisfied with their training. Only one third of these trainees completed their intermediate scan training and were signed off by their clinical supervisor during the study period.

There was also minimal use of the ultrasound simulation training resource through a combination of difficulty in accessibility and an apparent lack of motivation to use it. Less than half participants allocated access to the ScanTrainer® actually used it and of the 60% who did, the average time spent on the simulator was just over three hours. Of the pre-set simulated assignments, comprising the basic and advanced gynaecology and early pregnancy modules, on average trainees attempted them five times and overall only one in seven practice assignments attempted were passed.

This study faced several unexpected challenges and could not be completed as intended because it was not deemed feasible to obtain the pre-set outcome data. It became clear that the groups were not going to achieve the planned interventions in

the specified time frame, and that the control arm would not provide sufficient data for comparison of training outcomes and assessing the cost effectiveness of both methods of training. In the remainder of this discussion, deficiencies in the study design as well as potential factors contributing to the failure to complete the study will be explored. Practical trials in medical education are advocated give the dearth of robust, informative data on effectiveness and cost-effectiveness of educational interventions.[105] Thus, important lessons learnt from the experience of this pragmatic study and the available pilot data is presented to inform future research study designs in SBME in the field of obstetrics and gynaecology. Factors relating to individual, the task and the clinical context, will be focussed upon because the 'person-task-context' model is deemed a useful tool in designing curricula and in training assessment in postgraduate education.[106]

Study design

A pragmatic approach to study design was taken, this was deemed appropriate at the outset of the study as practical trials can inform practitioners and policymakers in making evidenced based choices.[105] At the time of designing the study the simulators had been purchased and installed in hospitals within deaneries involved in the study, but those responsible for delivering the training were not sure if it worked, who to train and how to train on a simulator. A randomised design by removing selection bias, would have been superior to test effectiveness of the simulation intervention for training than our cohort controlled design.[105] However, feasibility of being able to successfully design and run an RCT with comprehensive and relevant outcome assessments was felt to be uncertain. This was because of a lack of data pertaining to the feasibility of SBME using the ScanTrainer® at the time of inception of our study.

Moreover, these fears were realised not just because of the failure of this observational series but also given the subsequent report of an attempted RCT to assess learning curves in gynaecological ultrasound in the same UK NHS clinical population and training setting. This trial catalogued similar problems to the current study, namely with suboptimal recruitment and incomplete data precluding the strengths of clinical and educational inferences. [83]

Population and recruitment

Specialist Training registrars in obstetrics and gynaecology are attached to particular post-graduate training deaneries. They typically rotate around recognised training hospitals across the region, which often means travelling considerable distances. The current study was pragmatically designed to aid identification and importantly retention of this itinerant population by being tailored to fit within a single academic training year i.e. once trainees had been identified, agreed to participate and could be contacted within their base hospital. Moreover, the funding body dictated one-year project duration. (The Health Foundation, Shine Award 2012).

The planned recruitment of the intervention and control group was based on the West Midlands model of recruitment and enrolment to do intermediate scan modules. In this model trainees apply to do an intermediate scan module approximately six months prior to being placed in a hospital where they should then get the supervised clinical training required. The West Midlands DUC had knowledge of trainees currently undertaking an intermediate module and those who had applied. This process made it possible to identify trainees in the West Midlands and allocate to the appropriate group relatively easily. However, the same process for allocating trainees to hospitals did not exist across the country and so identification of trainees was an

initial challenge and a broader approach to recruitment had to be taken by contacting all trainees within the other four participating regions. Trainees in other deaneries trainees wanted to do scan modules so were keen to use the simulator, but they had no guarantee of a placement that would then facilitate alongside clinical training.

The study was designed to run over 12 months and some assumptions were made about progress through training being uniform and most trainees being able to sign off an intermediate scan module within a year. These assumptions were based on expert consensus from general discussion and informal consultation with those responsible for ultrasound training. Another researcher carried out a survey of expert opinion on the time taken to train by conventional training methods and reports six months was the agreed length of time needed to train by conventional methods.[83] However, this duration of training within the current study was an underestimate impacting upon the feasibility of the study. The current study found that progress through training was not uniform. Trainees' progressed at different rates, and almost one quarter of the recruited cohort was working less than full time. We found training often extended beyond 12 months in order to complete a module, and the majority of the trainees were running gynaecology and early pregnancy modules simultaneously. It took the conventional group trainees 16.5 months on average and between 47 and 300 scans to complete a module by their estimates. This is range is broadly in line with recommendations of International Society for Ultrasound in Obstetrics and Gynaecology [107] who recommend trainees spend at least 100 ultrasound examinations under direct supervision and complete a minimum of 100 ultrasound examinations before independent practice is commenced. [107].

A large international survey concerned with challenges to ultrasound learning published in 2014 after the inception of our study, surveyed 621 trainees across Denmark, Norway, Sweden and reported it took trainees of average over 24 months of clinical experience to confidently manage ultrasound examinations independently. These findings and recommendations support the importance of clinical experience to become proficient. However, acquisition of proficiency as with most craft skills will vary between individuals receiving training of a similar nature and intensity.[26] Also some simple focussed tasks may require very few supervised examinations before the operator attains a sufficient level of diagnostic accuracy.[108] Thus, there are a number of variables at play in determining the intensity of training and length of time needed to achieve basic competencies in gynaecological ultrasound. Such variables as time to achieve competency or quantitative outcomes such as the number of scans performed are therefore inappropriate outcome measure.[109]

Study conduct

This study was a pragmatic trial in a practical setting, which received a small funding award (Shine Award) making the running of a 12-month trial across multiple centres an ambitious task. The trial relied on busy clinicians, consultants and senior sonographers, in the collaborating deaneries volunteering their time to the conduct of the study. The study was multicentre across a large geographical area and not restricted to one particular training region, such that the results were expected to be generalisable across a UK NHS setting and similar health care and training systems outside of the UK. However, given the limited funding resources this dispersion, while adding methodological rigour to the research, made the day to day running of the study challenging because regular individual contact with participants was

impossible and the varied locations presented a barrier to regular trial collaborator meetings and site visits. With greater financial funding and a commensurate increase in trial staffing it may have been possible maintain better individual and group contact and to encourage participation and engagement with the study. The study was designed to be pragmatic and reflect training in the current 'real' clinical environment with the addition of access to a VR simulator. Thus, the manifest problems in the study with access to, and compliance with, intermediate ultrasound training in obstetrics and gynaecology highlights a realistic problem with practical ultrasound training in general, in a busy clinical setting where the organisation and delivery of training is reliant on busy clinicians who often are not getting ring-fenced time to participate.

Individual factors

Individual factors relate to the person/learner and include consideration of general cognitive skills, motivation and self-efficacy. The study identified more intervention group than control group participants. Recruiting eligible trainees to the simulation intervention was easy because access to a novel, additional method for practicing the fundamental skill of transvaginal ultrasound appealed to trainees. This phenomenon was presumably because the novice sonographers were potentially benefitting from participation in the study. It is not uncommon that patients given a choice between taking a 'new' intervention opt for this in preference to standard treatment or placebo alternatives and this observation presents one of the main challenges to recruiting to RCTs.[110] This same phenomenon was seen in another controlled, observational study of SBME compared with standard clinical training, more participants were recruited to the simulation intervention than the control group.[75] In a recent RCT, greater

levels of commitment to data collection was seen in the simulation intervention group compared to the control, standard clinical training group; the former group having almost double the number of patient ratings returned. [72]

There were problems with engagement of trainees who were not driven to collect data, with poor returns on the requested logs of training, Furthermore, trainees did not appear to be motivated to do the required simulation scanning. When participants were asked about why they weren't accessing the simulator, repeated issues of rota constraints, no protected time for simulator training, having other commitments and competing priorities were raised and we planned to explore these through focus group interviews. A pilot focus group was run in the West Midlands with a mixed group of trainees on general ultrasound issues but unfortunately not followed with a focus group of study participants to explore the reasons further. Alsalamah [83] followed up their study participants with an end of trial survey and report factors impacting trainees frequency of access to the simulator and potential obstacles as: 'work duties and other commitments' 'wasn't given protected training time' 'location of simulator not suitable'. The majority of the study cohort was made up of trainees ST4 seniority and above, with an average age of 31 years who were predominately female often with young families. Such postgraduate trainees have many professional and domestic constraints on their time, they generally work shift patterns with long days, nights and weekends and have requirements to pass postgraduate specialty exams and meet curriculum requirements. Ultrasound above a basic level is a desirable rather than mandatory skill. Local agreements were made to allow study leave to use the simulator but this did not make an impact on the progress of the intervention group. A UK survey of trainee obstetricians and gynaecologists in 2016 demonstrated trainees' as having favourable opinions of simulation for ultrasound training but then not

voluntarily accessing relevant simulation ultrasound training. [31] A systematic review by Gostlow et al.[111] published in 2017 concerned with voluntary participation in laparoscopic skills training found unrestricted access to simulator equipment is not effective in motivating trainees to voluntarily participate in simulation based education. They report attendance was generally better amongst junior rather than senior trainees and that by far the greatest extrinsic factor influencing voluntary training was lack of available time. The lack of available time has lead authors Burden et al.[112] to call for simulation for laparoscopy to be made mandatory prior to clinical training in order to overcome problems of voluntary engagement.

Task factors

Task factors include the instructional strategies for learning new tasks as well as the distinction of the tasks themselves. The participants were not able to pass the modules set for them on the ScanTrainer®. Participants practiced on the simulator individually and some spent considerable amounts of time on the simulator but, despite the provision of integrated system feedback, they could not pass all assessed aspects of the tasks the simulator set. In fact the average percentage of recorded assignments passed was 14% (range 0 – 31) and the average number of attempts per assignment was 5.2 (range 2.2 – 18). The reasons for this apparent poor performance were not explicitly explored but it is possible that participants continued to repeat a task and when they were unable to achieve a 'pass' became frustrated with the system and disengaged. In an attempt to mitigate for this participants were advised if an aspect of a task was not achievable after seven attempts to move on with to the next task set in the module. This approach was similar to that adopted by Madsen et al. [65] in their study on the assessment of performance measures and learning curves on the

ScanTrainer®, where they instructed participants in the learning phase to attempt a maximum of seven iterations for all the modules in order to reach expert levels and then went on to only assess participants based on metrics that had previously showed construct validity. Madsen et al.[65] established credible performance measures on the transvaginal ScanTrainer® using a group of twelve experts (eight gynaecologists using ultrasound daily and four fetal medicine consultants) and 16 novice participants. The group definitions are comparable to the novice and expert groups definitions in my study (chapter five). To assess construct validity, the group took a dichotomous pass/fail criterion for each metric measured by the simulator, and they examined 153 metrics in total across seven modules. Metrics that did not significantly discriminate between the two groups, or were passed by fewer than 50% of the obstetrics and gynaecology consultants, or were passed by more novice participants than experts were excluded from further analysis as not considered stable performance measures. Taking this approach 105 metrics were excluded, meaning almost a third of metrics demonstrated construct validity. Applying the same principle Dyre et al.[113] using the transabdominal ScanTrainer® (Medaphor, UK) for obstetric scanning also found that only around a third of the metrics demonstrated construct validity. As detailed in chapter five I conducted a post hoc analysis using the same approach as Madsen et al.[65] to establish construct validity and consistent with the aforementioned studies[65, 113] found that only 29.2% of metrics discriminated between novices and experts. Thus, reliance on manufacturers choice of metrics may not provide a valid assessment of performance. It is observed that some of the simulator metrics used are too easy to pass for all participants and hence they provide little information. The metrics that did not discriminate between groups presented in chapter five, were metrics relating to image optimisation, for example correct adjustment of image gain and the time gain

compensation function because with relatively little or no adjustment it was easy for both group to pass these elements. Furthermore, some may have failed to distinguish between participants' performances due to inherent data loss associated with the used of dichotomous items rather than more sensitive continuous scales. This problem with unsupervised, self-directed learning with the ScanTrainer® system prior to rigourous validity testing was also noted in an RCT set in the UK [83]. One in ten participants surveyed about their perception of this particular simulator reported that they *"kept failing tasks and lost interest*".

Error management in training has been studied using a transabdominal VR ultrasound simulator.[114] Medical students were randomised into two groups, either error management training (EMT) or error avoidance training (EAT). The EMT group were instructed to use the simulator freely, not to be concerned with completing the specific assignment instructions in order or sticking to instructions but just to play around. Errors were framed in a positive light as a by-product of learning. The EAT group were instructed to follow the exact simulator instructions step by step and not to try other alternatives. They were instructed to be precise and to take their time, they were told it was better to take longer and be more accurate. A transfer test was carried out 7-10 days later in clinical practice and the EMT group performed better than the EAT group, with mean performance scores of 67.7% (CI 62.4 – 72.9%) compared to 51.7% (CI 45.8 – 57.6%) P=<0.001. Thus, prescriptive approaches to simulation training focussed on perfection and avoidance of errors appears to be misplaced for two main reasons. Firstly, as the aforementioned study found, transfer of learning to the clinical setting may be compromised compared with more flexible, error recognition approaches. Secondly, training should be fun and trainees should remain motivated to strive to keep practising and the findings from our study in keeping with

others found trainees becoming demotivated when error avoidance methodologies were preferred.[114]

Contextual factors

Contextual factors include supervision, the opportunity to perform the task and support from supervisors and peers. This study encountered challenges in ensuring a permanent set up and location of the simulator systems so they were accessible to trainees across the deaneries. As detailed in Table 12 the location of the simulator system varied across deaneries but all were kept in clinical areas rather than simulation centres. There are advantages to the clinical setting as it could allow opportunistic and ad-hoc use during everyday clinical practice and potentially out of hours. However there could be frequent interruptions to use, conditions in terms of lighting and ergonomics may not be ideal and the simulator may risk getting damaged or broken if not used carefully. In our study a barrier to having simulators in clinical areas within hospitals was that only 35% of trainees were working in these hospitals with a simulator on site. The other 65% of trainees had to travel to these hospitals they weren't familiar with, they could not gain access without assistance from the local collaborator due to security and therefore access was not universal. Having the system in a simulation centre would give equal access, would offer a more controlled environment for practice without interruptions which is said to be a feature contributing to effective learning.[44] However, access out of normal working hours is unlikely to be possible in most hospital education / simulation units housing expensive equipment and therefore protected time for training would be needed. Nevertheless, greater scrutiny and control of our study, particularly aiding compliance with the intervention, troubleshooting and motivating participants may have been

possible if the simulators were housed within a specific simulation centre. Indeed, the only, high quality RCT evidence on simulation for teaching transvaginal ultrasound scanning comes from a Danish research group, study detailed in Chapter 2 [72] where the simulators are kept in a state of the art leading educational centre for simulation.[115]

One of the appeals of simulated learning, and a feature promoted in marketing ScanTrainer®, is its ability to allow learners to practice in a safe learner-centred environment [44] and reduce the need for direct supervision. Perhaps work on a simulator does not require the level of supervision needed in initial clinical practice but certainly needs some level of instruction.[92] Access to a simulator as a self-directed learning tool was not as effective as demonstrated by this study in motivating and sustaining a motivation to learn. Alsalamah [83] explored this a little deeper by surveying RCT participants and reported that a factor in enhancing the learning experience with the ScanTrainer® would be supervision. It is known that deliberate practice under expert guidance with frequent expert feedback benefits learners in the early acquisition of procedural skills. [116] In the early acquisition of skills guidance perhaps does not need to be from an expert, as there are examples of using medical students as instructors with the correct training. This is beneficial as it then means supervision is more readily accessible on evenings and weekends when trainees can find time for training.[115] Similarly experience from implementation of a laparoscopic skills curriculum found it was imperative to have dedicated supervising personnel and dedicated time for training.[117]

From the returned logbooks it was estimated trainees were doing on average around seven gynaecological scans per week. There may have been an element of over

reporting bias, trainees were asked to report on a representative 2-week period of training but when they were contacted about failure to return logbooks it was sometimes because they had done no scanning and therefore not returned the logbook. The finding that trainees were getting few opportunities to scan is in keeping with other UK surveys. At the RCOG national trainees conference in 2012 a representative sample of 170 obstetric and gynaecology trainees were asked if they would like to receive training in TVU, 89% of respondents stated yes, however only 11% reported receiving training. (Personal communication, N Woodhead) In a survey of trainees in the East Midlands region it was found that few trainees were getting dedicated time in their weekly schedule for ultrasound training, with only 7% getting at least once/weekly sessions.[31] The situation was reported similarly in Denmark when surveyed the mean satisfaction with transvaginal and transabdominal ultrasound training was low, with mean scores of 2.64 ±1.1 and 2.62 ±1.2, respectively (where 1=very dissatisfied and 5=very satisfied).[26] Trainees in the UK regarding similar types of skills learning were found not to be expected to be motivated to use a simulator if it was not going to lead to clinical practice of the learned skill.[118]

CONCLUSION

At the time of designing this study there was limited evidence on how to assess simulated performance, how much practice was needed on a simulator, what elements training should include, where simulators should be kept or what instructional designs should be used. The attempt to run this practical study examining the effectiveness of a simulation intervention for transvaginal ultrasound highlighted the complexity of interventions in postgraduate medical education. When designing the evaluation of

the intervention it had not been fully defined and developed. Ideally, any innovation should follow a course of development and evaluation reflecting a continuum with increasing evidence generation along the course leading to practical trials.[119] The sequential phases of developing complex intervention evaluations starts with searching for relevant learning theories, modelling interventions accordingly, conducting exploratory studies such as the current one to provide pilot data to then conduct definitive randomised controlled trials in controlled environments and then studying implemented interventions in real word settings.[105, 119]

In hindsight, there was much exploratory work that needed to happen before reaching a stage when a practical, feasible and rigorous comparative study using SBME alongside clinical training in gynaecological ultrasound could be carried out. The conduct of this study was extremely challenging because elements of assessment and learning using simulators had not been previously studied or considered at the time of designing the study. Nevertheless, attempting this study has provided important information that will be informative for future research in this field. The failure of this cohort controlled study led to hypothesis generation about the validity of the intervention itself culminating in the design and conduct of a validity study (see Chapter 5) and qualitative work exploring women's views on training professionals to perform transvaginal ultrasound (see Chapter 6). Furthermore, the lessons learnt from this study, in combination with data acquired from recent observational studies and RCTs [105] (see Chapter 2), can inform more optimal and feasible, larger scale studies, evaluating outcomes of relevance to trainees, trainers and patients in the future.

CHAPTER 5

Face, content and construct validity of a virtual reality simulator for transvaginal ultrasonography in gynaecology and early pregnancy

ABSTRACT

Introduction

Simulators should be validated before they can be effectively incorporated into educational programmes. The aim of this study was to establish the face, content and construct validity of a virtual reality simulation in assessing transvaginal-scanning skills in gynaecology.

Methods

Three separate groups consisting of novice, intermediate and expert level participants were recruited. All performed four modules in gynaecology and early pregnancy: two core skills and two advanced skills on a high fidelity transvaginal ultrasound simulator (ScanTrainer® Medaphor®, Cardiff, UK). The scores for the modules, time taken for the examination and accuracy of measurements were compared across the three groups to determine construct validity. The participants were then asked to complete a questionnaire with a 5-point Likert rating scale to assess the face and content validity as well as opinion on the role of simulation in ultrasound training.

Results

The three discrete groups each consisted of 15 participants. Overall the simulator was felt to be useful at a basic level, in teaching a systematic approach to ultrasound scanning and offers a user-friendly learning environment. Experts did not think the simulator resembles a realistic simulation of a real life scan. Novices felt they should first practice in a simulated setting before scanning women and simulation gave them a sense of confidence. The study demonstrated construct validity in the ability of the simulator to distinguish novices from intermediate and experts in the core and advanced skills tasks scores ($p=0.001$ and <0.001 respectively) and measurement of CRL ($p=0.01$) and a fibroid in 3-planes ($p<0.001$). Experts completed the tasks in the least amount of time, significantly shorter time than novices overall ($p=0.003$).

Conclusion

The VR transvaginal simulator, ScanTrainer® demonstrates face and content validity. However, respondents did not highly rate several aspects of the system such that users were somewhat ambivalent about the ability of the simulator to reflect the real clinical environment. Construct validity was shown for the conduct of both basic and advanced training tasks, and for precision of measurements of fibroids and speed of completion of imaging. Research needs to focus on validating comprehensive training programmes incorporating simulation

INTRODUCTION

The concept of simulation for learning is not a new one however advances in technology have led to high fidelity simulation models and simulation-based education as an emerging field.[63] Simulation allows a safe, learner-centred environment, which allows for repeated practice without patient harm or discomfort.[120] Although, learning in a simulated environment holds strong common sense appeal, having attractive programs is not enough, there has to be evidence that they work.[92] Assessment of simulators is through the process of validation; it is generally accepted that simulators need to be validated before they can be effectively incorporated into education programmes.[35, 65, 92, 121, 122] Lack of proper validation could imply that the simulator at hand does not improve skills or worse, could cause incorrect skills training.[123, 124] One of the advantages of simulated learning is that simulators can collect performance data automatically using objective metrics to build up a picture of each learner's skills base.[92] The development of such metrics is in its infancy and there is currently no uniform approach to measuring performance in surgical training or ultrasound.[65, 84, 125, 126] Simulation based learning for ultrasound has been suggested for improving basic training.[17, 101, 127] However there is limited evidence on how to assess simulated performance, knowing which metrics assessment should include or how much practice is needed.[65] The first step in answering the above questions is establishing valid and reliable performance measures.

Terminology of validation

The term validation, or validity, is increasingly seen in medical literature relating to SBME and assessment. [128] Over the last decade numerous studies exploring validation VR simulators in surgery have been published.[129] The expansion of

research in this area has been driven by the growth of minimally invasive surgical techniques requiring new types of training.[90] Validation studies for ultrasound are less numerous but are now being published.[85, 130, 131] Although relatively new in the medical arena, scientific validation is commonplace in psychology from where its origins lie.[132] Validity can be defined as:

''the property of being true, correct, and in conformity with reality''[133]

In testing validity the fundamental property of any measuring instrument, device, or test is that it ''*measures what it purports to measure*.''[90] Gallagher et al. emphasise that validity is not a simple notion rather it is comprised of a number of first principles.[90] The result is that within the literature a number of validation benchmarks have been developed. Studies exploring validity often will focus on one or two of these parameters, which include: face, content, construct, concurrent, discriminate and predictive validity. There is no uniformity in how these different types of validity are defined [90, 122, 134, 135] or guidelines on how to measure them.[129]

The following are definitions referred to in the literature:[90, 136-138]

> **Validity**: determines whether an exam or a test actually succeeds in testing the competencies that it is designed to test. A valid assessment method covering all the facets of clinical competence needs to have following attributes:
>
> **Face validity**: the extent to which the examination resembles the situation in the real world.
>
> **Content validity**: the extent to which the intended content domain is being measured by the assessment exercise.

Construct validity: the extent to which a test measures the trait that it purports to measure. One inference of construct validity is the extent to which a test discriminates between various levels of expertise.

Concurrent validity: the extent to which the results of the test correlate with the gold standard tests known to measure the same domain.

Predictive validity: the extent to which an assessment will predict future performance.

Discriminate validity: an evaluation that reflects the extent to which the scores generated by the assessment tool actually correlate with factors with which they should correlate.

In their critical review of the literature pertaining to validation of surgical simulators Schout et al.[129] distinguish between subjective and objective approaches to validation, which group many of the above principles under one umbrella term depending whether they offer an objective or subjective measure, this seems a clear and logical approach.

Subjective validity: Novices' (referents') and/or experts' opinions e.g. face, content

Objective validity: Prospective experimental studies e.g. construct, discriminative, concurrent and predictive validity

Subjective approaches to validity generally require experts (usually specialists) and novices (usually students) to perform a procedure or examination on a simulator, after which both groups are interrogated about their opinion of the experience. Objective

approaches generally involve experiments to ascertain whether a simulator can distinguish between different levels of expertise or to evaluate the effects of simulator training by examining real time performance (transfer of skill).

At the time of designing and conducting this study there were no published validity studies on the transvaginal ScanTrainer® (Medaphor, Cardiff, Wales, United Kingdom) and little published work on validity of other ultrasound simulators in obstetrics and gynaecology.[101, 102, 139] In 2012 Deo et al.[139] presented their work on developing a modular programme, incorporating virtual reality simulation, in delivering a competency based core curriculum in gynaecological ultrasound. The group, from the UK collected feedback questionnaires from 13 trainees in obstetrics and gynaecology, whom had used ScanTrainer as part of the programme for delivering training. The trainees varied in clinical experience comprising doctors from specialty trainee year one (basic training) right through to specialty year 6 (advanced training). 12/13 (92%) of trainees found the simulator a useful component of training, and they believed its use increased the speed of basic ultrasound skill acquisition prior to real time scanning. Most trainees, 8/13 (63%) felt that it reduced tutor supervision in the early phase of transvaginal scan training, that it increased the flexibility of training (82%) and that it improved their confidence prior to real time patient scanning (81%).

Study hypothesis

The hypotheses that led to the conception of this study were:

- The simulator is realistic of real life transvaginal scanning.
- The content of the skills modules provides the knowledge and psychomotor skills required for transvaginal scan training.

- The simulator can objectively differentiate between experts and novices.

The aims and objectives of the study were thus:

Aim

To validate virtual reality simulation in assessing transvaginal-scanning skills in gynaecology by establishing face, content and construct validity.

Objectives

The objectives of this study were to:

- To *subjectively* determine the realism (face validity) and training capacity (content validity) of the VR simulator (transvaginal ScanTrainer®).

- To explore opinions on the training capacity of the simulator

- To *objectively* compare the following groups performance on the VR simulator (transvaginal ScanTrainer®) in undertaking core and advanced skills modules (construct validity):

    i. Expert in gynaecological and early pregnancy scanning
    ii. Intermediate level with some experience of gynaecological and early pregnancy scanning
    iii. Novice with no experience in gynaecological and early pregnancy scanning

METHODS

In an experimental set up the validity (face, content and construct) of selected metrics on a high fidelity ultrasound simulator were assessed. In this study the types of validity being tested were defined as follows [140]:

Face validity: the extent to which the examination resembles real life situations.

Content validity: the extent to which the domain that is being measured is measured by the assessment tool.

Construct validity: the extent to which a test discriminates between various levels of expertise.

The study was conducted in the imaging department of a tertiary level women's hospital in Birmingham West Midlands, UK. All the assessments took place in a training room, an undisturbed environment with optimal lighting conditions. Each participant was assessed individually. The simulator equipment and set-up remained the same throughout the study period. The study was conducted from August – December 2013.

Participants

Purposeful sampling was used to recruit participants into three pre-defined groups for inclusion. The groups were categorised as outlined below and detailed in Table 13.

Table 13 Study group definitions of novices, intermediates and experts

Group	Definition
Novice	A person with no ultrasound experience
Intermediate	A person with some ultrasound experience, having done between 30 and 300 ultrasound scans, mostly under direct supervision.
Expert	A person with more extensive ultrasound experience, having done over 1000 scans and regularly carrying out unsupervised gynaecological scanning with a scan list at least once/week.

Medical students were recruited as the novice group. The medical curriculum at the University of Birmingham (UoB) is a five-year programme and the obstetric and gynaecology rotations are completed in the final year over a six-week period. The hospital provides clinical placements for students in groups of approximately twenty per rotation and novice participants were given the opportunity to take part in the study during their attachment. They received information about the study via email and face-to-face with a presentation during their teaching programme (see Appendix 10 and 11) all those who met the inclusion criteria were enrolled. Earlier in their medical training the students had covered pelvic anatomy but had little or no experience of ultrasound because imaging does not form part of their curriculum.

The intermediate group participants needed to have some experience but still be training to expert level. Postgraduate medical speciality training is organised by geographical deaneries and those within the West Midlands Deanery who had registered for but not yet completed an RCOG intermediate scan module in early

pregnancy or gynaecology were sent information about the study via email and asked if they would like to participate. (Appendix 10) A local University offer a certificate and diploma in medical ultrasound; practitioners registered on this course were also sent information about the study via email and face to face with a presentation given during one of their lectures. (Appendix 10 and 11)

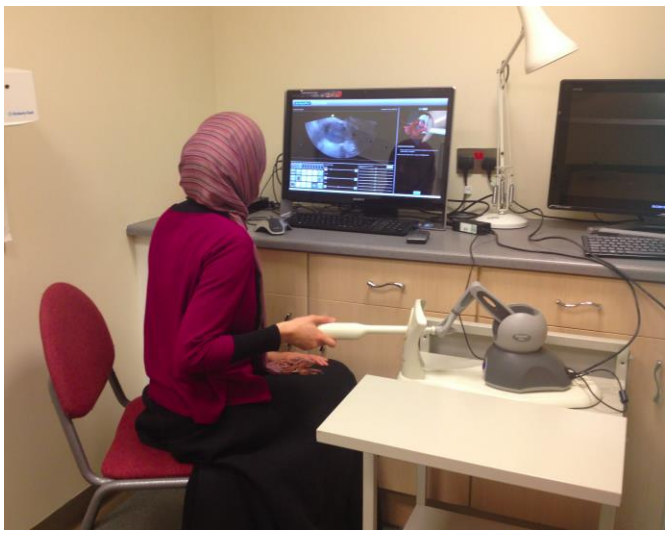
The expert group were recruited locally in the departments of gynaecology, obstetrics and radiology at the hospital and from the local university lecturers in ultrasound. All experts had a qualification in ultrasound and carried out regular ultrasound scanning, some were involved in training.

Interested potential participants were asked to complete a questionnaire detailing their ultrasound experience (Appendix 12) they were then contacted 1-2 weeks later by the researcher to arrange a convenient time to use the simulator. Participants were excluded if they had used the simulator prior to the study for duration of one hour or more.

Equipment and training exercises

A high fidelity simulator, the transvaginal ScanTrainer® was used in this study, software version 3.0. The simulator is designed for transvaginal ultrasound and consists of a monitor and a transvaginal probe docked into a haptic device that provides realistic force-feedback when moving the probe. The set up of the equipment and monitor is shown in Figure 8. The monitor provides B-mode ultrasound pictures obtained from real patients as well as a 3D animated illustration of the probe's anatomical scan position. (Appendix 8)

Figure 9 Set up of the simulator equipment and monitor



The system includes various training modules ranging from core to advanced gynaecological and early pregnancy modules. After completing a module the simulator provides feedback using dichotomous metrics in a number of task specific areas (e.g. gestational sac examined in the sagittal plane) as well as general performance aspects (e.g. image gain adjusted correctly). The study consisted of completion of two core modules, basic gynaecology and basic early pregnancy, each with their corresponding assessment, and two advanced modules measurement of an ectopic pregnancy and assessment and measurement of a fibroid. A list of the tasks and variables assessed are included in Appendix 13. These modules were chosen as they broadly aligned with the content of the RCOG ultrasound curriculum, which is detailed in Appendix 6.

All participants received a short introduction to the simulated setting. This included how to operate the simulator using the manufacturers introduction video and a short time using the unassessed practice mode. The researcher followed the same protocol with all participants (see researcher instructions Appendix 13). Participants were

asked to perform each task twice, the first run was to allow them to become familiar with the simulator and controls, and the second was the assessed task. No feedback was given between runs, the researcher remained present throughout and made sure data on each tasks was recorded on the system. The 3D animated illustration of probe orientation remained on throughout as feedback was being collected on this aspect of the simulator. Participants could have a short break between tasks but remained in the room. After completion of the assessed task the participants could view the simulator feedback on performance.

Outcomes

Construct validity

Scoring metrics from the four selected training modules were based on i) score for core tasks; ii) score for advanced tasks; iii) time taken to complete tasks; iv) accuracy of measurements.

Face and content validity

Immediately after completing the simulated tasks, all participants were asked for their opinions on the realism of the presented simulation, the training capacity of the simulator and the role of simulation in training. This allowed for assessment of face and content validity.

In the absence of a validated scale designed to specifically examine attitudes towards simulation training for ultrasound, a summated rating scale was developed following the steps set out by Robson.[141] This tool comprised of a pool of items that appeared relevant to the issue of simulation training for ultrasound. Items for inclusion were identified from validity studies relating to any technology enhanced simulation,[142-144]

a system usability scale used in industry [145] and from the small amount of work published on ultrasound simulation.[128, 139] A mix of positive and negative statements were presented in roughly equal numbers.[141] Items that were related to the same issue were allowed in order to triangulate responses and extreme statements were avoided. The resulting questionnaire contained 38 statements and responses were scored on a 5-point Likert scale from strongly agree to strongly disagree (Figure 9) (Appendix 15)

Figure 10 Rating scale used for assessment of face and content validity

Strongly agree	Agree	Undecided	Disagree	Strongly disagree

Analysis

Face, content and general opinions on the scan trainer were transformed into scores, for positive statements (5 = strongly agree, 4 = agree, 3 = undecided, 2 = disagree, 1 = strongly disagree), for negative statements (1= strongly agree, 2=agree, 3 = undecided, 4= disagree, 5= strongly disagree). Face and content validity was determined for all three groups, the averages in each group were ranked and compared using Kruskal-Wallis test. Kruskal-Wallis is a nonparametric test that compares unmatched groups. Values are ranked paying no attention to groups then the largest number gets a rank of N, where N is the total number of values in all groups. The discrepancies among the rank sums are combined to create a single value 'H statistic'; a large H statistic corresponds to a large discrepancy amongst rank sums. Construct validity was evaluated through the comparison of performance metrics between the three groups. This included examining the differences in individual variables as well as total scores and time for both basic and advanced tasks. The

group means for inaccuracy, i.e. variance from actual dimensions, were calculated for three specific advanced ultrasound measurements: a crown rump length (CRL), an ectopic pregnancy (in three planes) and a fibroid (in three planes). The mean time taken to complete the tasks between the three groups was also calculated. To establish construct validity a one-way ANOVA test was performed to see if significant differences existed between all the groups when analysed together and post-hoc Bonferroni tests to compare where differences between groups were. A P value of <0.05 was taken as statistically significant. All analyses were performed using SPSS version 21.0. (IBM Corp, 2012)

Sample size

The sample size was calculated from a small pilot study in which seven novices (medical students) completed a basic (core) skills module with nine discriminatory metrics, the average score for the novice group on this task was around 50%. The sample size calculation was based upon the assumption that an expert would pass all metrics, scoring 100%. To demonstrate this difference with a power of 0.8 and alpha set at 0.05, the group size was estimated to be 15 participants per group. The same magnitude of difference in performance was assumed for the advanced skills module between the experts and intermediate level trainees such that the same number of participants would be required assuming the same power and significance.

Ethical approval

All participants provided informed consent to participate. The study was conducted in accordance with the general terms and conditions of University of Birmingham ethical approval board: ERN_13-0185

RESULTS

Participants

Each group had 15 participants who completed all aspects of the study. No potential recruits were excluded from participating in the study on the basis of substantial prior experience with the simulator. A minority of participants in the intermediate and expert groups (2/15 (13.3%) and 3/15 (20.0%) respectively) had used the ScanTrainer® transvaginal simulator before, but in all cases it was for duration of less than 1 hour.

Table 14 shows the demographics of the groups. The groups adhered to the pre-defined definitions of novice, intermediate and expert. The novice group comprised of final year medical students. They had limited prior experience of gynaecology or early pregnancy ultrasound; 8/15 (53%) had seen a scan being performed, 7/17 (47%) recalled being shown still images in an educational setting. None had performed ultrasound or used an ultrasound simulator. The intermediate group were either trainee obstetrician and gynaecologists undertaking an intermediate scan module in early pregnancy or gynaecology 6/15 (40%) or a postgraduate professional registered with Birmingham City University enrolled on their certificate or diploma course in medical ultrasound 9/15 (60%), with a nursing, midwifery, radiography or sonography background. The expert group all had a qualification in ultrasound. For the majority this was a diploma in medical ultrasound 11/15 (73.3%), or a qualification from a medical speciality college with the relevant speciality accreditation in scanning, as well as regular ultrasound scanning as part of their job

plan, Member of the Royal College of Obstetricians and Gynaecologists 2/15 (13%) or Fellow of the Royal College of Radiologists 2/15 (13%). As expected, the median ages of each group increased according to experience. Most participants were female reflective of the current gender balance in both obstetrics and gynaecology, nursing and radiology work force and also the composition of medical undergraduates at the UoB.

Table 14 Demographics of the participants in the three groups

	Novice n=15	Intermediate n=15	Expert n=15
Average age, years [93]	23 (21 – 28)	35.3 (22 - 54)	48.1 (29 – 66)
Female	2/15 (80%)	11/15 (73%)	14/15 (93%)
Male	3/15 (20%)	4/15 (27%)	1/15 (7%)
Years of experience post qualification Average [93]	NA	NA	16.6 (1 – 37)

NA = Not applicable

Face and content validity

Mean scores and *H* values of face validity statements are detailed in Table 15. Of the nine presented statements, all scored positively by the novices and intermediates but in 4/9 statements relating to realism the experts gave negative ratings. The experts scored the ergonomic set up of the ScanTrainer® mean 2.33 H=15.95, p<0.0001 and the negatively phrased questions on resemblance of the ScanTrainer® to a real ultrasound machine mean 2.0 H=18.32 p<0.0001 and ScanTrainer® providing an unrealistic simulation of performing a real scan mean 2.73 H=9.37 p=0.009. The most positive responses with mean scores of >4 equating to agreement, were for the

statements “Images of the normal pelvic anatomy are realistic” and “Images are of a good quality”. The response to the first question was consistent across groups with mean scores ranging from 4.13 – 4.27 whereas the experts were less convinced that the images were of good quality, mean score 3.47 H=14.20 p=0.001. Responses from novices and intermediates were more positive than experts but there was no clear pattern between the former two groups with the highest mean scores being recorded in 4/9 questions by intermediates, particularly in relation to the realism of the images of the normal pelvis and fibroid, mean scores 4.27 for both statements (Table 15).

Mean scores and H values of statements relating to training capacity of the simulator (content validity) are detailed in Table 16. As with responses to realism, overall mean scores were positive for training capacity of the simulator. Whilst experts were more conservative in their mean scoring with lowest mean scores in response to 8/11 questions, this was not as marked as in this groups responses to the realism domain (Table 16) and there were no negative mean scores. The most positive responses were to the questions pertaining to the usefulness of the simulator for teaching a systematic approach to examination, mean scores across groups 4.53 – 4.07 and for providing a user-friendly learning environment, mean scores across groups 4.67 – 4.07. The experts significantly differed from the other groups in opinion on effectiveness for practicing basic skills H=10.08 p=0.006 and teaching hand-eye coordination mean H=8.55 p=0.014. The novice group found the simulator fun to use mean score 4.13 H=9.73 p=0.008 and appeared to enjoy using it more than the other groups mean 4.47 H = 11.60 p=0.004.

The experts were most sceptical of the value of the simulator feedback providing negative mean scores in three out of the four-presented statements. There were

significant differences in the opinions of the participants across the groups in feedback ratings. The novice group rated the value of the simulator feedback higher than the intermediate or expert groups. The expert group did not find the audio nor visual force feedback useful, rating 2.47 and 2.6 respectively. In contrast the novices liked this aspect of the simulator and rated its usefulness 4.53 and 4.07 respectively with the intermediates ratings falling between the mean ratings from these two groups, $H = 16.96$ $p<0.001$, $H=10.92$ $p=0.004$.

Table 15 Face validity statements

Statement	Novice n=15		Intermediate n=15		Expert n=15		*'H statistic'*	p Value
	Mean	Mean rank	Mean	Mean rank	Mean	Mean rank		
Realism of the presented simulation								
ScanTrainer is ergonomically set up (+)	4.13	31.83	3.4	23.73	2.33	13.43	15.95	<0.0001
Images are of a good quality (+)	4.67	30.00	4.4	24.80	3.47	14.20	13.52	p=0.001
Images of the normal pelvic anatomy are realistic (+)	4.13	21.83	4.27	24.23	4.13	21.37	0.801	p=0.67
The ultrasound probe has a realistic range of motion (+)	3.87	24.90	3.93	26.43	3.27	17.67	4.41	p=0.11
ScanTrainer does not resemble an ultrasound machine (-)	3.53	30.90	3.2	25.90	2.0	12.20	18.32	<0.0001
ScanTrainer provides an unrealistic simulation of a real life scan (-)	3.73	26.57	3.87	27.17	2.73	15.27	9.37	p=0.009
Images don't move as per a real machine (-)	3.60	26.43	3.43	25.87	2.73	16.70	5.67	p=0.059
Images of the fibroid are unrealistic (-)	3.87	22.83	4.27	28.17	3.33	18.00	5.64	p=0.06
The feel of the ultrasound probe is unrealistic (-)	3.67	21.80	3.86	27.20	3.40	20.00	2.92	p=0.232

Ratings on a 5 point Likert scale

Positive statements (5 = strongly agree, 4 = agree, 3 = undecided, 2 = disagree, 1 = strongly disagree)

Negative statements (1= strongly agree, 2=agree, 3 = undecided, 4= disagree, 5= strongly disagree)

Table 16 Content validity statements

Statement	Novice n=15		Intermediate n=15		Expert n=15		*'H statistic'*	p Value
	Mean	Mean rank	Mean	Mean rank	Mean	Mean rank		
Training capacity of the simulator								
ScanTrainer is useful in teaching a systematic approach to scanning (+)	4.53	27.23	4.07	20.40	4.27	21.37	3.32	p=0.189
ScanTrainer is effective for practicing basic skills (+)	4.47	26.80	4.47	26.80	3.80	15.40	10.08	p=0.006
ScanTrainer reduces the workload of those delivering scan training (+)	3.25	23.71	3.47	24.97	3.13	18.90	2.12	p=0.346
The ScanTrainer offers a user friendly learning environment (+)	4.67	27.20	4.27	23.00	4.07	18.80	4.11	p=0.128
The ScanTrainer is useful at a basic level (+)	4.07	24.03	4.29	23.60	4.07	21.37	0.48	p=0.786
The ScanTrainer is fun to use (+)	4.13	29.73	3.67	23.10	3.13	16.17	9.73	p=0.008
ScanTrainer is not useful for teaching hand eye-co-ordination (-)	4.33	28.50	3.93	23.43	3.60	17.07	8.55	p=0.014
ScanTrainer is not useful for teaching 3D orientation (-)	4.13	28.57	3.73	21.87	3.47	18.57	5.99	p=0.050
ScanTrainer would not be helpful to teach pathology recognition (-)	4.20	28.10	3.73	23.23	3.33	17.67	5.68	p=0.058
Simulated scanning makes inexperienced operators no more capable than previously (-)	3.67	22.63	3.60	22.90	3.80	23.47	0.04	p=0.982
I did not enjoy using the scan trainer (-)	4.47	29.67	4.07	24.27	3.27	15.07	11.60	p=0.004
Value of simulator feedback								
The feedback provided is helpful (+)	4.53	30.63	4.12	22.57	3.4	15.80	12.18	p=0.002
The audio force feedback is useful (+)	4.53	33.60	3.07	20.50	2.47	14.90	16.96	<0.0001
The visual force feedback gauge is not useful (-)	4.07	29.87	3.53	24.23	2.60	14.90	10.92	p=0.004
The feedback did not improve skills for subsequent scanning (-)	3.80	26.77	3.67	25.47	2.81	16.77	5.81	p=0.055

Ratings on a 5 point Likert scale

Positive statements (5 = strongly agree, 4 = agree, 3 = undecided, 2 = disagree, 1 = strongly disagree)

Negative statements (1= strongly agree, 2=agree, 3 = undecided, 4= disagree, 5= strongly disagree)

The role of simulation in ultrasound scan training was explored and means scores of statements are detailed in Table 17. In general, the novice group provided the most positive responses. The experts provided the lowest mean scores in 11/14 presented statements leading to significant differences between groups in 8/13 presented statements. The most positive responses were for the contention that simulation is useful for training, that trainees should first practice in a simulated setting before scanning women, that the simulator instils a sense of confidence and that prior simulation training is in women's best interests. However there was significant differences between groups for these statements: the usefulness of simulation for gynaecological scan training H=14.81 p=0.001, the sense of confidence it gives trainees H=9.18 p=0.010 the importance of simulated practice for trainees H=17.59 p<0.0001 and women H = 5.98 p=0.05; these were all rated less favourably by experts than other groups. Statements related to the usefulness of the ScanTrainer® for assessment, determining those with aptitude for scanning, for revalidation of skills and reducing the rate of misdiagnosis were rated negatively by experts. Indeed, the latter two assertions received the lowest overall mean scores (scores < 3). The novice group felt the simulator could accelerate a person's learning, rating 4.27, more than the intermediate group, rating 3.64 or the expert group who were undecided, rating 3.00, H=8.88 p=0.012.

Table 17 Role of simulation in ultrasound training

Statement	Novice n=15		Intermediate n=15		Expert n=15		*'H statistic'*	p Value
	Mean	Mean rank	Mean	Mean rank	Mean	Mean rank		
Role of simulation in training								
The scan trainer is useful for training in gynaecological ultrasound (+)	4.73	31.70	4.33	21.67	4.0	15.63	14.81	p=0.001
The increment of skills during training should be monitored (+)	3.8	21.23	3.87	21.53	3.65	26.23	1.77	p=0.412
Trainees should first practice in a simulated setting before scanning women (+)	4.8	34.10	3.43	16.83	3.57	18.07	17.59	<0.0001
Scan trainer could be used for assessment (+)	3.73	29.47	3.13	21.53	2.67	18.00	6.67	p=0.36
Scan trainer is suitable for evaluation during training (+)	3.93	28.17	3.4	21.63	3.00	19.20	4.48	p=0.106
Scan trainer can accelerate a person's learning (+)	4.27	29.90	3.64	21.80	3.27	17.30	8.88	p=0.012
Scan trainer gives starting operators a sense of confidence (+)	4.53	29.83	3.88	20.83	3.64	18.33	9.18	p=0.010
It is not in women's best interests for trainees to first practice in a simulated setting (-)	4.47	28.90	4.07	21.30	3.80	18.80	5.89	p=0.052
It is not important to practice scans on virtual models (-)	4.53	32.10	3.87	21.40	3.47	15.50	15.51	<0.0001
Scan trainer could not be used to determine those with aptitude for scanning (-)	3.13	26.30	3.07	23.73	2.60	18.97	2.71	p=0.258
Scan trainer could not be used for revalidation of skills (-)	3.67	32.00	3.07	24.30	2.27	12.70	18.11	<0.0001
Scan trainer will not shorten learning curves in the scan room (-)	3.33	24.03	3.33	24.33	3.0	20.63	0.81	p=0.669
Scan trainer cannot reduce the rate of misdiagnosis (-)	3.33	29.17	2.93	24.13	2.2	15.70	9.03	p=0.011
Use of Scan trainer will increase the cost of ultrasound scan training (-)	2.80	18.23	3.0	22.30	3.36	28.47	5.30	p=0.071

Ratings on a 5 point Likert scale

Positive statements (5 = strongly agree, 4 = agree, 3 = undecided, 2 = disagree, 1 = strongly disagree)

Negative statements (1= strongly agree, 2=agree, 3 = undecided, 4= disagree, 5= strongly disagree)

Construct validity

Table 18 details the task scores across the groups. A statistically significant difference between the novice and the expert groups for the core skills tasks (p= 0.001) and advanced skills tasks (p=0.004) was found. There was also a significant difference between the novice and intermediate group for these tasks, core skills (p=0.01) and advanced skills (p=0.001). No differences were observed between intermediate and expert groups (see Figure 10, 11, 12). Analysis across groups showed a statistically significant difference for both core and advanced task scores ($p = 0.001$ and $p < 0.001$ respectively).

Figure 11 Box-and-whisker plots showing median sum scores for core tasks of novices, intermediates and experts.

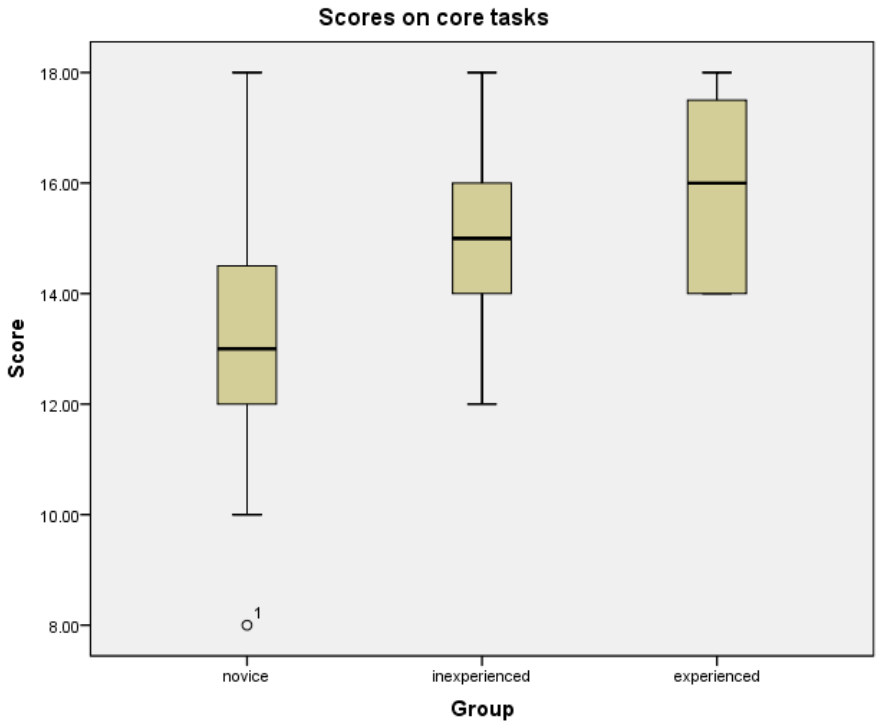


Figure 12 Box-and-whisker plots showing median sum scores for advanced tasks of novices, intermediates and experts.

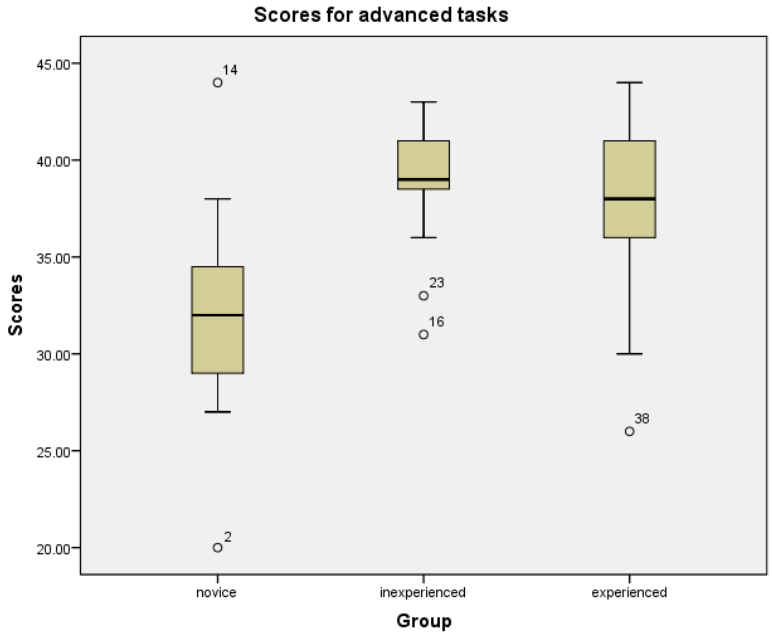


Figure 13 Box-and-whisker plots showing median sum scores for core and advanced tasks

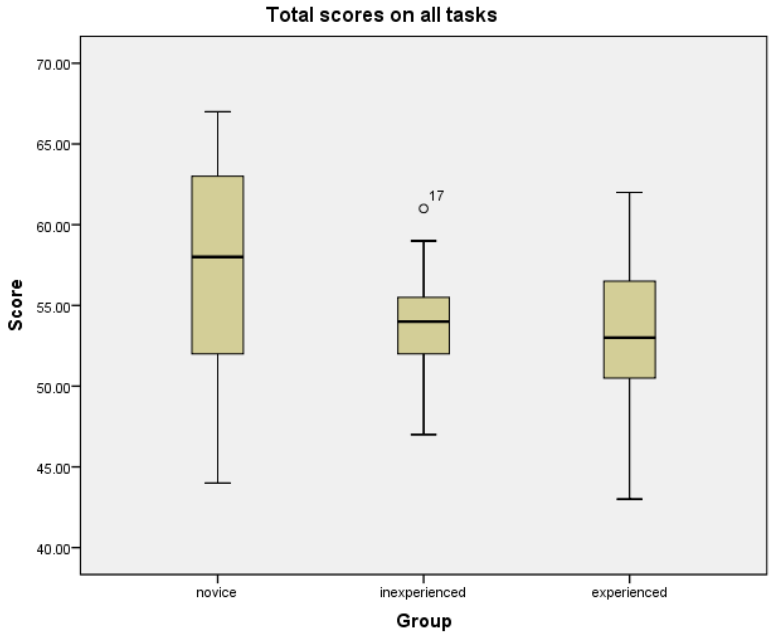
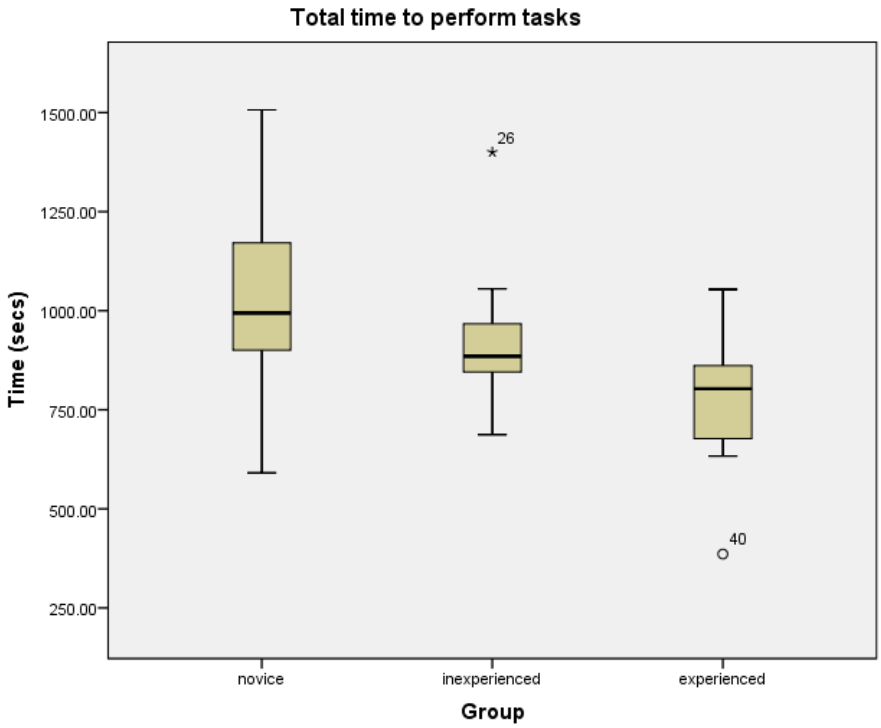


Table 18 Task scores across the three groups

	Novice (N) Mean (146)	Intermediate (I) Mean (146)	Expert (E) Mean (146)	ANOVA p value (<0.05)	Multiple comparisons	Post hoc Bonferroni p value (<0.05)
Core	13.07 (2.46)	15.27 (1.62)	16.00 (1.69)	0.001	N > I	0.01
					N > E	0.001
					I > E	0.9
Advanced	31.80 (5.47)	38.80 (3.26)	37.67 (4.92)	<0.001	N > I	0.001
					N > E	0.004
					I > E	1
Total	57.27 (7.00)	54.07 (3.43)	53.67 (5.74)	0.2	N > I	0.4
					N > E	0.3
					I > E	1

Table 19 details the mean time taken to complete the assigned tasks. Overall the experts were significantly quicker in completing the tasks than the novices (p=0.002) but no difference was observed in total completion time between expert and intermediate level groups (Figure 13).

Figure 14 Box-and-whisker plots showing total time taken to perform tasks for novices, intermediates and experts.



The intermediate group but not the novice group, took significantly longer compared to the expert group to complete the core obstetric skills task (p = 0.04). No other between group differences were observed for this particular assignment but when all the groups when analysed together a statistically significant difference was observed according to prior experience (p = 0.048). A statistically significant difference in completion time was observed between both the novices and the intermediate compared with the expert group for the core gynaecology skills task (p < 0.001 and p = 0.03 respectively) and taken together a significant difference was seen across

groups (p = <0.001). The novice group but not the intermediate group were also significantly slower to complete the advanced obstetric (ectopic pregnancy) skills tasks (p=0.008) compared to the expert group. Analysis across all groups also demonstrated a significant difference for this task (p = 0.01). No differences were observed between or across groups according to expertise in time taken to complete for the advanced gynaecology skills (fibroid measurement) task.

There were significant differences between the novice and expert groups in the accuracy of measurements of CRL (p=0.009) and the measurement of a fibroid in all three planes (p=0.001). In measuring a fibroid, the simulator was also able to discriminate between the novice and intermediate group but not the intermediate and advanced group (Table 20).

A post-hoc analysis of the metrics examined in this study was undertaken to mirror an alternative approach to estimating construct validity based upon studies [85, 113] published on validity of ultrasound simulators after completion of the current study. Scores were combined for the initial (familiarisation) and subsequent (assessment) runs on the simulator, therefore giving 30 attempts in total per group. As per the methods outlined by Madsen et al.[65] metrics were excluded if they did not show a difference between groups, where <50% of experts passed and more novices than experts passed. This approach resulted in exclusion of 49 of the 72 metrics examined because 24 showed a significant difference between groups, a further 3 of these 24 metrics were excluded as <50% of experts passed (for one of these metrics more novices than experts passed), resulting in 21 of 72 metrics (29.2%) that demonstrated construct validity by this method (Table Appendix 16).

Table 19 Mean time to complete tasks in seconds

	Novice (N) Mean (146)	Intermediate (I) Mean (146)	Expert (E) Mean (146)	ANOVA p value (<0.05)	Multiple comparisons	Post hoc Bonferroni p value (<0.05)
Core skills obs	111.9 (37.5)	132.3 (54.7)	94.8 (20.8)	0.048	N > I	0.5
					N > E	0.7
					I > E	0.04
Core skills gynae	180.07 (61.3)	153.2 (40.7)	107.8 (30.1)	<0.001	N > I	0.4
					N > E	<0.001
					I > E	0.03
Advanced skills obs ectopic	350.2 (138.1)	279.8 (64.5)	239.9 (61.3)	0.01	N > I	0.2
					N > E	0.008
					I > E	0.8
Advanced skills gynae fibroid	383.7 (92.7)	345.1 (88.7)	325.9 (76.5)	0.2	N > I	0.7
					N > E	0.2
					I > E	1
Total Time	1025.9 (241.4)	910.4 (172.6)	768.4 (159.8)	0.003	N > 2	0.3
					N > 3	0.002
					I > 3	0.2

Table 20 Mean distances from ideal measurements (mm)

	Novice (N) Mean [(146)]	Intermediate (I) Mean [(146)]	Expert (E) Mean [(146)]	ANOVA p value (<0.05)	Multiple comparisons	Post hoc Bonferroni p value (<0.05)
CRL	4.17 (3.66)	2.16 (2.58)	1.13 (0.86)	0.01	N > I	0.1
					N > E	0.009
					I > E	0.9
Ectopic A	9.74 (8.89)	5.37 (6.09)	4.73 (3.41)	0.09	N > I	0.2
					N > E	0.1
					I > E	1
Ectopic B	15.11 (11.28)	8.67 (7.59)	8.48 (5.29)	0.06	N > I	0.1
					N > E	0.1
					I > E	1
Ectopic C	8.43 (7.51)	4.99 (6.46)	5.33 (5.27)	0.3	N > I	0.5
					N > E	0.6
					I > E	1
Fibroid A	12.71 (11.16)	3.41 (3.88)	2.76 (2.09)	<0.001	N > I	0.002
					N > E	0.001
					I > E	1
Fibroid B	17.74 (13.25)	4.41 (4.24)	4.11 (2.05)	<0.001	N > I	<0.001
					N > E	<0.001
					I > E	1
Fibroid C	15.21 (14.23)	3.69 (2.32)	2.84 (1.95)	<0.001	N > I	0.002
					N > E	0.001
					I > E	1

DISCUSSION

Principal findings

This study has demonstrated that the simulator is felt to be useful at a basic level, in teaching a systematic approach to ultrasound scanning and offers a user-friendly learning environment. Experts do not think the simulator resembles an ultrasound machine, the images move as per a real machine or overall provide a realistic simulation of a real life scan. However they did recognise images were realistic and of good quality. Novices felt they should first practice in a simulated setting before scanning women and simulation gave them a sense of confidence. There was no consensus across groups of varying gynaecological ultrasound experience on how simulation should be used in a real world setting to shorten learning curves, for assessment, for evaluation, for revalidation or the impact use of simulation in training will have on cost.

The study demonstrated construct validity in the ability of the simulator to distinguish novices from intermediate and experts in the core and advanced skills tasks scores and measurement of CRL. Measurement of a fibroid in 3-planes on the simulator also distinguished the novices from the other groups, demonstrating construct validity for this task. Experts completed the tasks in the least amount of time, significantly shorter time than novices overall. In an alternative approach, where individual metrics were considered to display construct validity only if they demonstrated a difference between groups, that most experts passed and where more experts passed than novices, resulted in around a third of the 72 metrics (29.2%) retaining this psychometric property.

Strengths and limitations

The main strength of this study was the robust group definitions used. The study had three groups with strict definitions based on scanning experience. As validity research is heavily reliant on the use of experts and novices, for opinions and as the reference standard, it is important to define participants' level of expertise as if not explicitly categorised it can jeopardise the credibility of the study findings.[129] Defining three separate groups was an advantage, compared to other studies that have included only two groups,[113, 130] because it allowed inclusion of the target group for the intervention clearer discrimination between groups and inclusion of the target group for the intervention. Prior studies have used participants, often-medical students, with no experience of a procedure as the novice group.[65, 142] The advantages of using medical students is they provide a fairly uniform demographic, they will have a very similar prior knowledge and experience, they will be attached to training units where the studies are taking place and have flexibility in their timetables. Other studies have used participants with limited rather than no experience as the novice group.[130, 131, 147] Studying validity of the transvaginal ScanTrainer®, Ficquet et al.[147] included 32 gynaecologists who were divided in to two self-defined groups based upon whether scanning independently or non-independently. Alsalamah et al.[130] defined participants as experts or non-experts based on a cut-off of having more or less than 2 years scanning experience, frequency of scanning and self-identification as experts or non-experts. Al-Memar et al.[131] took a quantitative approach, defining an expert as someone who had done >500 scans, intermediate as 50-500 scans and novice as <50 scans. Similar arbitrary cut-offs based upon numbers of procedures have been used in validation studies of endoscopy and surgery.[129, 146] However, using numbers to define participants relies on participant's recall of estimates and it is likely to be

unreliable. This effect is further compounded when a continuum with an arbitrary cut off is used; the end result is likely to be non-discriminative groups with a large overlap of expertise.[129] The variation in definition of expert groups leads to questionable applicability and makes comparisons across studies difficult. In using three groups I was able to have a group with no experience, a clear definition in line with other studies.[65] The intermediate group with limited experience was particularly important to include as the target group for transvaginal simulation use in training. Whilst the current study used the number a number of scans, ranging from 30 -300, as part of the group definition based upon European guidance,[17] the more objective measure of needing supervision was also specified. The expert group carried out regular unsupervised scanning and all had a recognised qualification in ultrasound. I did include a specification of >1000 scans for the expert group which was somewhat arbitrary and subject to recall bias, but I avoided having a continuum and opted for discrete groups.

Further strengths of this study relate to its size, follow up and setting. The three groups were all of equal size and that size was based on a sample size calculation from a small pilot study. The sample size of forty-five participants across the groups represents the largest validity study done on the transvaginal ScanTrainer® to date. All participants completed all aspects of the study and there was no missing data. There were a small number of participants with prior experience of using the simulator, five in total (distributed across the intermediate (n=2) and expert group (n=3)) however the experience was limited to less than one-hour in total, for most this was in the context of having seen the system at a conference exhibition, this should not have a significant impact on the findings.

The study took place in a training room with no disruptions and optimal lighting conditions. All participants were assessed individually. This is in contrast to prior studies when the setting was a conference or course,[129, 130] which will infer an inherent degree of selection bias as conference exhibit halls by nature busy places so participants are likely to be easily distracted, and may also be observed by others, all of which could affect results.

Subjective approaches to validity have been widely critiqued.[129] The variety of scales used has been highlighted, from dichotomous yes/no questions, Likert scales containing 4 to 7 points, ordinal scales and 10-point scales. Moreover, the content of these scales is heterogeneous, with questions relating to realism, usefulness, appropriateness, acceptability and applicability, as are the thresholds used to estimate validity. A summary of studies of subjective validity on the ScanTrainer® are summarized in Appendix 17. Al-Memar et al.[131] only took the opinion of experts to determine content and face validity, Alsalmah et al.[130] only present data on content validity from experts, but face validity data for both groups. Schout et al.[129] recognise the literature offers no detailed answers in this regard and novice participants have been included in previous validity studies. However, the inclusion of novices in providing opinions on face and content validity is questionable. As found by Alsalmah et al.[130] as in the current study ratings on a 5-point Likert scale were higher from novices and intermediates than experts. This could be because the less experienced operator knows no different and is impressed by and willing to engage with a simulator, whereas the more experienced operator can more easily see the limitations. A study by Dyre et al.[113] on the validity of the transabdominal ScanTrainer® for obstetric ultrasound scanning does not explore participants' views of realism at all, stating the lack of association between realism and learning as the

reason.[86] Schout et al. take the view that it may be advisable to entrust this task to focus groups of specialists who are experts in the procedure in question and in judging simulators.[129]

A potential limitation of the study is the choice of participants to comprise the novice and intermediate groups may have introduced a degree of bias. Investigator bias may be at play because medical students may be more enthusiastic in their responses in order to please their supervisors. Furthermore, they may give overall positive feedback simply as they are getting an opportunity to do some supervised hands-on skills practice and gain feedback, rather than because the simulator was superior to alternative ultrasound training approaches. The intermediate group of participants represent the main target group of postgraduate trainees in gynaecological ultrasound. Like medical students they are also highly motivated volunteers, possibly for similar reasons to the novices, but also because they represent a body desiring accreditation in ultrasound in order to progress in their training and future careers. It is possible that training time and attempts may have been underestimated compared to a larger and less motivated group of intermediates. Whilst these factors may have overly motivated the novice and expert groups, in contrast the expert group may have been less motivated and more critical of the ScanTrainer®. This may have arisen lack of familiarity with the equipment resulting in less engagement and a higher expectation of the simulator's performance in keeping with the experience of real-life scanning. A further limitation relates to the reduced power of our study because the sample size was predicated upon an assumption that 'expert' performance on the simulator meant perfect, 100% simulator assessments scores. This was not observed in our study and indeed another study published after the design of the our project estimated expert or 'mastery' level to be 88.4% on certain validated metrics.[65]

Strengths and weaknesses in relation to other studies, discussing important differences in results

The simulator was rated highly in its usefulness at a basic level and in its utility in teaching a systematic approach to ultrasound scanning in line with other studies.[85, 101, 127] The novice group felt strongly that they should practice in a simulated environment before scanning patients, and it gave them a sense of confidence. This in keeping with the findings of another UK cohort study conducted in Bristol. In that study 88% of participants felt it was in women's best interests to practice ultrasound scanning in a simulated setting prior to undertaking imaging in real patients.[128] Others have reported the positive effect of simulation training in ultrasound on trainees confidence levels.[77]

Studies reporting the subjective assessment of validity of the Medaphor transvaginal ScanTrainer® are summarised in Appendix table 17. Prior studies [130, 131, 147] conclude the demonstration of subjective face and content validity by the simulator with broadly favourable ratings of performance from small groups of experts. However, in the current study whilst the ratings of experts for face validity were neutral or generally positive they were lower overall compared to those of intermediates and novices, and in particular in relation to set-up and realism of the scan compared to real life. This may be explained by the larger, perhaps more experienced cohort comprising the 'expert' group compared to preceding studies. However, even in those studies concluding an overall favourable opinion of simulation certain domains of realism were questioned. For example, in the study by Al-Memar et al.[131] 50% of experts were not satisfied with the realism of the instrument handling, which is in

keeping with the indeterminate ratings for the simulator's realism regarding the ultrasound probe and its range of motion.

The ScanTrainer® has audio force feedback that activates when excessive pressure is applied by the probe and a 'scream' is transmitted from the simulator. The number of times this feature was activated during the assessment was reported by Al-Memar et al.[131] who found a significant difference in the number of activations between novices compared to intermediate and expert groups. In the current study the relative number of activations was not evaluated but overall responses regarding the value of this feature were assessed and were neutral or negative amongst experts and intermediate trainees. Moreover, the amount of force applied did not show construct validity when individual metrics were examined (see supplementary analysis table, Appendix 16).

In this study, our results supported the construct that experts and intermediates and novices performed the simulated ultrasonic tasks according to their expertise. We found significant differences between the novice and expert groups in core task scores in gynaecology and early pregnancy in keeping with other studies [83, 131] Studies of objective assessment of validity on the transvaginal ScanTrainer® are summarised in Appendix table 18. The significant differences between novice and intermediate groups in completing core skills were consistent with those of Al-Memar for basic early pregnancy.[131] Previous studies did not assess the advanced skills modules but Alsalamah [83] hypothesised that differences may be revealed between intermediate and experts when more advanced skills were assessed. However in the current study we did not find any discrimination in performance of experts over intermediates for advanced skills modules (measurements of a fibroid or ectopic pregnancy), or in the

time to complete these tasks. In addition, Al-Memar did not show a difference between groups in the time taken to complete the assessments unlike my study.[131] This conflicting finding is perhaps explained by the group definitions as they followed a continuum and even the novice group had some experience of performing transvaginal ultrasound.

Examination of specific metrics used by the scan simulator allows direct comparison with other studies although the pre-specified modules and tasks varied between the few published studies using this simulator. Both this study and the evaluation by Madsen et al.[65] showed that crown rump length (CRL) measurement had construct validity. Our findings also correlate with those of Alsalamah et al.[130], both studies choosing to assess scanning simulation using the pre-set 'basic skills gynaecology case 1' package. Specifically, examination of the left ovary in two planes showed construct validity whereas measurement of the uterus in the sagittal plane and in measurement of the pouch of Douglas did not discriminate between groups.

Meaning of the study: possible explanations and implications for clinicians and policymakers

The metrics we excluded in keeping with the approach described by Madsen et al.[65] based on Messick's framework for validity.[148] This method found less than 50% of experts passed and novices outperformed experts on metrics related to image optimisation and magnification for example ensuring that the ectopic mass was correctly magnified and sufficient magnification employed to examine both the uterus and the fibroid (Table, appendix 16). Image optimisation is an area that Alsalamah et al.[130] found intermediate group participants outperformed experts in, for example in their study in examination of the yolk sac as it was not adequately magnified in the

centre of the image. These findings may be explained by the willingness of less experienced groups to engage with the simulator and its instructions. The expert operator may not take steps to get the perfect image every time in the interests of expedience and a confidence in their ability to correctly identify and interpret anatomy and pathologies. This observation is analogous experience of driving; when preparing for, or undertaking, a driving test, learners are more compliant with instructions. In contrast, once the test has been passed, more experienced and proficient drivers may not conform so explicitly to these taught rules such as ensuring a pronounced 'mirror, signal, maneuver' every time a turn in the road is executed. Whilst many metrics failed to discriminate between the relative expertise of gynaecological scanning according to experience, and so lacked utility in the assessment of trainees, this does not mean that such aspects of technique should not be taught on a simulator. These steps are integral to conducting a competent, systematic gynaecological transvaginal ultrasound examination. Thus, they highlight important aspects of the procedure but should not be used in criterion-based assessment of performance. In addition to teaching all steps considered integral to gynaecological ultrasound examination, simulated training packages should engage and motivate learners. An unduly harsh emphasis on certain arbitrarily determined pass/fail metrics may have the opposite effect if most trainees cannot pass the assessments set by the system after a reasonable time practicing in the system.

Unanswered questions and future research

Future research needs to go further in the assessment of validity beyond simple demonstration of construct validity and assess credible performance standards, that is to ensure trainees have reached well defined levels of competence before entering

clinical practice.[149] Since the design and conduct of the current study a Danish group[65] [113] have explored the learning curves of novices in 'mastery' model of learning for transvaginal scanning and transabominal scanning. The studies found that on the transvaginal simulator novices reached expert within a median of five iterations (range, 5 – 6), and on the transabdominal scan trainer novices reached the mastery learning level within a median of 4 attempts (range 3–8). Study of learning curves allows exploration of the time it takes to reach expert levels of performance on a simulator before a plateau in performance is reached, Madsen et al.[65] found this to be 3-4 hours of simulator practice, later confirmed by Tolsagaard et al.[73] Whilst such studies provide important data to help design optimal scan training programmes, the idiosyncratic nature of training junior doctors and other affiliated health professionals across countries limits the generalisability of such findings. Thus, similar studies need to be replicated in a UK based setting, with consideration of the UK curriculum and organisation of ultrasound services and levels of available supervision. Further research should also explore and compare other types and even combinations of simulator as well as how best to integrate simulation into clinical training to better address the learner's needs. Involvement of manufacturers in the development of simulation is also of key importance because computer based simulation is such a fast-moving, technology-dominated field that evaluation may trail behind innovation and as a result not be considered 'fit for purpose'.[92] Developers and evaluators working in synchrony can mitigate this risk.

CONCLUSION

The VR transvaginal simulator, ScanTrainer®, for teaching transvaginal pelvic ultrasound demonstrates face and content validity. However, respondents did not highly rate several aspects of the system such that users were somewhat ambivalent about the ability of the simulator to reflect the real clinical environment, its ability to train and its role and value in current training. Our study also showed construct validity for the conduct of both basic and advanced training tasks, and for precision of measurements of fibroids and speed of completion of imaging an ectopic pregnancy, however, the majority of individual metrics examined were unable discriminate between novices and experts. The findings are specific to ScanTrainer® and in line with other studies on the same system however are not generalisable to all simulator systems, each would need to undergo validity testing before being used. The more important bigger picture is to set clear educational goals for scan training and provide valid and reliable performance assessments for learning in the simulated environment.[120] As advocated by Schout [129] it would be better to design and evaluate a comprehensive training programme instead of validating just one part of a procedure that can be performed on a simulator. This requires an understanding of educational theories and backgrounds and a multidisciplinary approach with specialists, residents, educationalists and manufacturers.[129]

CHAPTER 6

Exploring women's views and experiences of transvaginal ultrasound with focus on the training of professionals

ABSTRACT

Introduction

The purpose of this study was to gain in-depth insight and enhance understanding of women's views and experiences of transvaginal ultrasound examination, with a particular focus on how professionals are trained in transvaginal ultrasound.

Methods

A qualitative descriptive study using semi-structured interviews was undertaken. Fifteen women who had all had a recent transvaginal ultrasound scan were interviewed face to face or by telephone. Audio-recorded individual semi-structured interviews were transcribed verbatim and analysed thematically using Nvivo 11.0.

Results

Findings suggest that transvaginal ultrasound examination is acceptable to women when it is necessary for their care. The study found the way the examiner communicated was the biggest influence on the women's experience. Women appreciated a professional yet friendly manner and were reassured when the examiner kept talking to them throughout the scan. The study-highlighted women are willing to have scans performed by trainees. Having a trainee perform the scan was an acceptable prospect to the woman if the trainee appeared calm, friendly,

communicated well and was open about their abilities and limitations Women could identify benefits and limitations to simulation for ultrasound scan training. Overall it was felt simulation should be used for the initial stages of training so trainees could practice without pressure and therefore appear calmer and more confident when scanning, but the major limitation identified was the lack of patient interaction.

Conclusion

This study has offered insight into good practice points regarding the communication during a transvaginal scan and the communication dynamic when a trainee operator is performing a scan. These practice points and the findings of this study could be taken forward into a training programme to teach communication skills for gynaecological ultrasound examinations.

INTRODUCTION

Transvaginal ultrasound (TVU) is widely used in a range of gynaecological contexts, both acute and elective. TVU allows for detailed examination and diagnosis of uterine, ovarian and adnexal pathology. [5] In pregnancy TVU can be used to measure cervical length, diagnose and date normal early pregnancy, diagnose miscarriage, ectopic and molar pregnancy. TVU has advantages over the transabdominal approach as it gives increased image resolution and does not require a full bladder. Consequently large numbers of women will undergo TVU and many healthcare practitioners need to be trained in its use, in order to meet the demand for imaging. Ultrasonography is considered safe but is highly operator dependent [17]; the images gained and their interpretation depends to a large degree on the training and experience of the operator. Training in ultrasound requires both theoretical knowledge and practical skill. [107] To develop the appropriate level of practical skill numerous scans need to be performed.[150]

There are challenges to delivering ultrasound training in the UK system. There is a lack of time and opportunity for sufficient ultrasound education due to constraints on trainers and trainees' time and the pressure on the ultrasound service. Delivering training in TVU presents additional challenges, as it is an intimate examination, a woman must undress to allow the insertion of the ultrasound probe into the vagina. As Coldicott [151] highlights: 'the teaching of intimate examinations poses ethical problems for trainees and trainers – trainees must learn but patients must be protected'.[151] Coldicott also discusses how the acceptability of practice can change over time mirroring changing expectations regarding the standard and the delivery of clinical care. Conventional teaching of intimate procedures was based upon the

principle that practising how to perform intimate examinations on patients would benefit other patients in the future once proficiency was acquired.[151] However in recent years the intimate, invasive and possible negative psychological impact of examinations such as transvaginal ultrasound has been recognised.[152] Best practice guidelines exist on intimate examinations and teaching vaginal examinations require students to obtain informed consent before undertaking any such examinations. Students are increasingly required to gain baseline competence by practising vaginal examinations, outside of the real clinical environment often utilising pelvic mannequins or 'expert patient instructors' known as gynaecological teaching associates [153] which have the added advantage of allowing communication skills to be developed. For ultrasound scan training recent advances have seen the development of simulators to teach technical ultrasound skills. Use of ultrasound simulators is a promising solution to some of the challenges faced in delivering TVU training. Empirical studies are needed to prove benefits of improved skills, training efficiency and benefit to patients through simulation for training. However an ethical argument exists for simulation use regardless, to decrease the reliance on vulnerable patients for training and has led some researchers to regard SBME as an ethical imperative.[154]

It is widely accepted that patients and the public should be engaged in healthcare and service development.[155] in order to inform and direct development of transvaginal scan training programmes. This may involve substantial costs for purchase of VR simulators. We need to ascertain the opinions of and acceptability to women of having trainee operators perform transvaginal scans on them; before potentially making investments in expensive simulated scan trainers we need to know this is in the patient and trainees interest.

Published studies, including one review,[156] have assessed women's perceptions of transvaginal ultrasound examinations using surveys. Studies have explored the acceptability of TVU to women, its predictors, pain and psychological morbidity that may arise from TVU in a range of contexts and populations.[89, 152, 157-168] The literature suggests women are accepting of TVU when it is necessary for their medical care. Most of the studies found that TVU was acceptable to women who had undergone the procedure; with for example 99% of women reporting they would have no concerns about having a future scan.[166] However the acceptability of TVU is context specific and the range of acceptance has been reported as 43 – 99% across different settings.[165] Six studies reported data on pain or physical discomfort during TVU with 36.6 - 76% of women experiencing some degree of pain or discomfort during the scan.[89, 152, 157, 158, 160, 166] The level of pain was usually mild, however 1%[158], 6%[160], 8.4%[152], 26%[89] of women rated their pain as more than mild. Four studies investigated anxiety regarding TVU and found that the majority reported some degree of anxiety.[152, 162, 163, 166] One study examined the psychological trauma following the procedure and found twelve women (1.6%) reported clinically significant levels of trauma related to the scan.[152] These are worrying findings given a trainee operator could exacerbate the anxiety and discomfort felt.

A literature search was carried out which found no published studies specifically on the experiences of women, or acceptability to women of having a TVU performed by a trainee operator. When a trainee operator carries out ultrasound examination times are longer,[30] no additional clinical information is gained and women are potentially exposing themselves to more pain, discomfort and anxiety.

Aims and objectives

Aim

The aim of this qualitative study was to explore women's views and experiences of transvaginal ultrasound examination, with a particular focus on training professionals in transvaginal ultrasound.

Objectives

The objectives of this study were:

1. To explore women's knowledge of transvaginal ultrasound.
2. To explore women's experiences of transvaginal ultrasound scans.
3. To explore the factors that influence the experience of having a transvaginal scan.
4. To gain insight into attitudes towards transvaginal ultrasound examinations performed by trainees.
5. To explore women's views on use of virtual reality simulation for training in transvaginal ultrasound.

METHODOLOGY

Data was collected using individual semi-structured interviews, a method that has been effectively applied to research in population and health.[169] Qualitative research methods such as interviews shed light on people's understanding and attitudes of issues and are deemed to provide a 'deeper' understanding of social phenomena than

would be obtained from purely quantitative methods, such as surveys.[170] Interviews were considered appropriate for this topic because detailed insights from individual participants about transvaginal scan experiences were desired and little was already known about women's attitudes to ultrasound training. Individual, rather than group, interviews were also appropriate as transvaginal scan examination could be a sensitive topic and some participants may not want to talk about such issues in a group environment.

Both telephone and face-to-face interviews were conducted. This combination has been used in prior qualitative interview studies and the interactional differences that exist between the methods has been studied.[171] Although research suggests there are some differences to be found between telephone and face-to-face interviews due to rapport and social interaction.[171] For the present study there were practical advantages to giving women the option to do a telephone interview, as they were expected to be busy women often with young families. It was felt there was possibly an advantage of telephone interviews, as women may feel less embarrassed or find it less intrusive to talk about the subject over the phone rather than face-to-face with the researcher.[172]

Sampling strategies

Pregnant and non-pregnant women were recruited from the early pregnancy assessment unit (EPAU) and acute gynaecology clinic at a tertiary level women's hospital in the Midlands, UK, where most women seen undergo a TVU scan. Both clinics run in the same department of the hospital, and the department plays a key role in training. At the time this study was done, the patient population in EPAU who had

experienced miscarriage were already being targeted for recruitment in two separate research studies therefore they were excluded.

Table 21 Details of the population from which women were sampled

	Clinical presentation
Pregnant	Early pregnancy assessment unit: usually pain and/or bleeding in early pregnancy
Non-pregnant	Acute gynaecology clinic: urgent non-emergency presentations such as pain, menstrual irregularities

Women were eligible for inclusion if they:

- had a transvaginal ultrasound scan as part of their clinical assessment
- were attending EPAU and their ultrasound showed an on-going intrauterine pregnancy
- had a reasonable command of the English language

Women were excluded if they:

- were aged under 18
- had not had a transvaginal ultrasound scan as part of their assessment
- were diagnosed with miscarriage
- needed admission to hospital
- needed on-going care or follow up by EPAU

Staff members of the clinical healthcare team in the acute gynaecology and early pregnancy clinic identified potential participants, gave them information about the study and asked if they would be happy to speak to the researcher. If so, they informed the researcher who was based onsite. The researcher then came to meet the woman following her clinical consultation to give information about the study and a study participant information sheet (Appendix 19). Consent to be recruited was taken at this stage (Appendix 20), along with details in order for the woman to be contacted again about participation in an interview. Consenting women were contacted 1-2 weeks later by their preferred method of communication to confirm if they would like to take part in an interview, either telephone or face to face, and a convenient date and time was arranged.

Interview schedule

The interview schedule (Appendix 21) was developed based on existing literature and focused on women's knowledge of TVU, their experiences of TVU, feelings about trainee operators and discussion about ultrasound simulators for training. The main focus was to explore how women felt about having a transvaginal scan, what influenced their perception of the experience, and if this altered with a trainee operator (based on lived experience or hypothetical discussion). Finally women's views on best practice for ultrasound scan training were explored. When discussing training the concept of simulation was introduced and explored. Women were provided with pictures showing the set up of a scan room with a real ultrasound machine with transvaginal probe and the set up of the simulator with transvaginal scan probe (Appendix 24) to aid discussion around training. The interviews were semi structured and interviewees were encouraged to talk freely around the questions the

researcher asked. The interviewees had all had a recent TVU so they were asked to reflect on the recent experience but drawing on previous experience was also encouraged where applicable.

The interview schedule was discussed with a patient, a public involvement group and the EPAU staff. Modifications were made based on their feedback on phrasing and acceptability of questions. The questions were also discussed in detail with the supervisor before interviews commenced. The first interview was conducted as a pilot interview. The researcher (NW) conducted the interview with supervisor present to comment on interview conduct and schedule. The participant was asked for feedback following this interview and found the questions acceptable.

Procedures and participants

Participants were recruited over a three month period (24nd September 2013 – 21st November 2013) and interviews carried out over a two month period from 22nd October 2013 to 10th December 2013. The interviews were a mixture of telephone and face-to-face interviews, depending on the participant's preference and convenience. The face to face interviews were held in the education resource centre at the hospital from which women were recruited, away from the clinical setting in which the scan took place. The researcher greeted participants and their written informed consent to participate was obtained prior to the interview (Appendix 23). The participants were asked to complete a short demographic questionnaire detailing age, ethnicity, marital status, educational qualifications, occupation, prior experience of transvaginal ultrasound examination and brief reproductive history prior to the interview. (Appendix 25) The questionnaire was used to put the interview in context. For telephone interviews the same explanations and procedures were followed to ensure

informed consent as for face-to-face interviews. The consent form, questionnaire and the pictures to be used for the interview were sent ahead of the interview in the post with a stamped and addressed envelope to be return to the researcher. Consent was recorded verbally for telephone interviews, reading the consent form together and asking participants to confirm their agreement and return the form and questionnaire.

Participants were provided with a £10 Amazon voucher to reimburse them for their time and participation.

The final sample comprised of 15 women in total, nine telephone interviews were done and six face-to-face interviews. *Interviews lasted between 20 minutes and 44 minutes (average time 31minutes).*

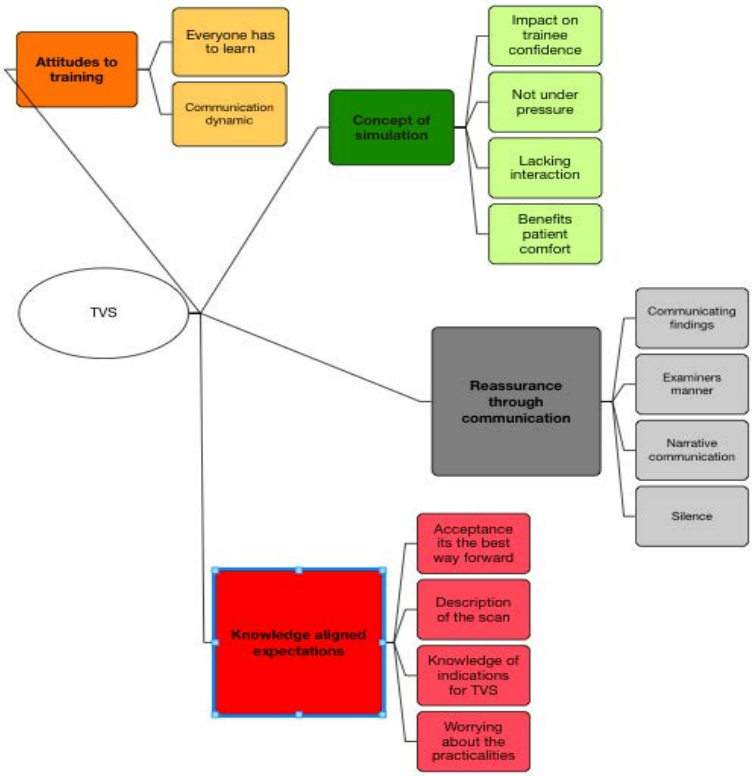
Data analysis

All interviews were audio recorded with the participants' permission and transcribed verbatim by a third party professional, then checked for quality and completeness by the researcher. Data was analysed through thematic analysis [173] following the six-steps outlined by Braun and Clarke 2006.[174] The analysis adopted an inductive approach; the codes and themes were data driven.[174] Transcripts were read and re-read to gain an overall understanding of the narrative for each participant before coding was initiated, during this process a list of initial codes was generated. Data were analysed using Nvivo 11.3.1 for Mac (Copyright © 1999-2016 QSR International Pty Ltd.) for initial coding of transcripts and organisation of subthemes and main themes.

All transcripts were coded initially through open coding creating a large number of individual codes. In a later step these codes were reviewed and discussed with a

supervisor, similar codes were categorised and subsequently merged into themes and subthemes. As a structured and systematic way to guide data analysis hierarchy charts and mind maps were generated to assist the interpretation and exploration of the data (Appendix 26). The emergent themes and subthemes are presented in Figure 14. The researcher (NW) did the coding and thematic analysis; the coding framework was discussed with a supervisor to ensure consistency. The final refinement of codes into themes resulted in four main themes: knowledge aligning expectations, reassurance through communication, attitudes to training and views on the concept of simulation.

Figure 15 Thematic mind map of emergent themes and subthemes



Positionality

The researcher (NW) has 8 years of experience as a doctor within the speciality of obstetrics and gynaecology and 3 years of experience performing gynaecological

scans. The benefits of the researcher being an O&G doctor were that she was better placed to interpret the findings with an understanding of the context within which the participants worked. However the researcher was also mindful of her 'insider' knowledge and of any possible assumptions and shared meanings.

Trustworthiness

The researcher's position was a potential source of bias and this is mitigated for by following the relevant strategies from Shenton's (2004) strategies for ensuring trustworthiness in qualitative research. The specific steps to ensure credibility of the research and how they were achieved are outlined in the table below.

Table 22 Strategies for ensuring trustworthiness

Credibility	**Trustworthiness steps within this research study**
Adoption of research methods well established	Well-recognised research methods were adopted. Semi structured interviews were used with line of questioning followed in interviews, thematic analysis using clearly defined methods was used for data analysis.
Development of early familiarity with the culture of participating organisations	Early familiarity with the culture of the participating organisation was achieved as the researcher had prior experience working with the organisation. Consulting the patient and public involvement group and EPAU staff on the study protocol and interview schedule further fostered this.
Tactics to help ensure honesty of informants	It was ensured each participant was given the opportunity to refuse participation for the research study. I only used participants that gave informed consent to participate and I maintained anonymity through transcription.
Iterative questioning	This would suggest the use of probing questions in order to obtain greater detail from the participant; this tactic was applied during interviews. Discrepancies in data will be explored and possible explanations offered.
Frequent debriefing sessions	Regular sessions with supervisor were held to discuss the study progress and test developing ideas and interpretations

Ethics

Ethical approval for the project was granted by the Proportionate Review Sub-committee of NRES, study title Exploring stakeholders perceptions of transvaginal ultrasound scan training. REC reference: 13/NW/0525. IRAS project ID: 109352.

RESULTS

The participant characteristics are outlined in Table 23

Table 23 Participant characteristics

Demographic	Number of respondents n=15 (%)
Age	
22 - 30	5 (33)
31 - 36	4 (27)
37 - 40	5 (33)
Over 40	1 (7)
Marital status	
Married	10 (66)
Unmarried	5 (33)
Ethnicity	
Asian, Pakistani/Bangladeshi	3 [141]
White, British	12 (80)
Educational level	
Below degree level	8 (53)
Degree level and above	7 (47)
Occupation	
Medical, nurse or doctor	4 (27)
Teaching	1 (7)
Public services	2 (13)
Banking and accounting	2 (13)
Management and administrative	3 [141]
Stay at home mum	2 (13)
Unemployed	1 (7)
First or repeat TV scan	
First scan	4 (27)
Repeat scan	11 (73)
Estimated total number of TV scans	
2-3	6 (40)
4 or more	5 (33)
Pregnancy	
Pregnant	10 (66)
Non-pregnant	5 (33)
No. of prior pregnancies	
Never been pregnant	3 [141]
1-3	9 (60)
3 or more	3 [141]
Previous miscarriage	
Yes	4 (27)
Attendance for routine cervical smears	
Yes	13 (87)
Below screening age	2 (13)

The sample of fifteen women were aged between 22 and 40 years old, one women did not disclose her age, she was over 40 but still of reproductive age. The majority of participants were married 10/15 (66%) and white British women 12/15 (80%). Overall the education status of the women was high, over half had a university degree 8/15 (53%); some of those who did not still had professional jobs. Most women were in employment 12/15 (80%), two women were stay at home mums and one identified as unemployed. The sample included four medical professionals, three nurses and one doctor training to be a GP. Four women had only had the recent transvaginal scan and non-before this, the majority of women 11/15 (73%) had prior experience of scans, and five women estimated they had five or more. The sample included pregnant 10/15 (66%) and non-pregnant 5/15 (33%) women; 3/5 of these had never been pregnant. Although we excluded women with recent miscarriage from recruitment in the sample there were 4/15 women who had experienced miscarriage, they had all had two or more. Where appropriate based on the lower screening age limit all women reported attending for routine cervical smears.

Four main themes were identified from the data. These are summarised as follows:

1. **Knowledge aligns expectations**
   - Acceptance it's the best way forward
   - Knowledge of and indications for TVU
   - Worrying about the practicalities

2. **Reassurance through communication**
   - Examiner's manner
   - Narrative communication
   - Dealing with silence
   - Communicating findings

3. **Attitudes to training**

- Everyone has to learn
- Communication dynamic

4. **Views on the concept of simulation**
   - Lacking interaction
   - Not under pressure

The main themes around women's knowledge and experience of TVS were about prior knowledge aligning women's expectations of the scan and how women sought reassurance through communication during and after the scan. There was a further two themes relating to training: attitudes to training and views on the concept of simulation. Findings are presented under the relevant theme and subthemes with supporting evidence from participant quotes.

Theme I: Knowledge aligning expectations

This theme captures how a woman's existing knowledge about the indications for TVU and consideration of a prior possibility of having a TVU impacted on her experience of the scan and her preparedness for undergoing the examination. Findings highlighted how women often worried about the practicalities of undergoing the examination. Within this theme information about the acceptability of transvaginal ultrasound examinations to women and women's knowledge about TVS, including descriptions of the scan, was considered. An overview of the theme and subthemes is given in Table 24.

Table 24 Theme I: Knowledge aligns expectations

Subtheme	*Illustrative quote*
Acceptance it's the best way forward	*that's the way they'll get the best results and it's probably the best way of doing it, and I've got no problem with that at all.*

	33 year old, not pregnant
Knowledge of and indications for TVU	*I guess if you're needing to assess the ovaries, or if an abdominal ultrasound doesn't give you the information that you need, I guess maybe then they'd need it, just to get a clearer picture of stuff.* *30 year old, pregnant, 1st scan*
Worrying about the practicalities	*You don't really and you don't know how to put your knees, you know, and whether to take your trousers off or whether to just take them to your ankles, or stupid stuff like that… I think I had a sanitary towel in my knickers, so I was a bit concerned, you know, my trousers were going to go on the floor at the side with this, you know, used sanitary towel on the top of my knickers, which would be on display…* 38 year old, not pregnant

Acceptance it's the best way forward

Women identified that having a transvaginal scan was not *"the most pleasant thing"* but all women were accepting of it if they were told it was clinically indicated and necessary in order to give them the information they required. One woman described it as *"a means to an end"* and another as *"mind over matter"*. It was more important for the women to get the *"bigger picture"* and find out what was happening than to worry about the scan.

> *Well, <u>fine</u> because it was just…you know, I knew I had to do it; it was something that I wanted to do, to obviously look for any fertility issues. So, you know, it wasn't the nicest, it was that I knew it had to be done so I was willing for it to happen really… I'd describe it as, sort of, like, <u>necessary</u> and*

not terribly uncomfortable… And it's good to, sort of, see what's happening. And I'd rather do that and know what's happening than not have it done.

31-year-old, pregnant woman

Women's desire to get answers by having the investigation was stronger than the preferences shown to the gender of the examiner. In general women did not mind, and would accept a male examiner but they showed a preference ideally for a female. When women voiced a desire for a female examiner they also voiced concerns that this was unfair to the male examiners and they did not want to *"kick up a fuss"* or appear silly, therefore they were more likely to go ahead with the examination on the day regardless of who the examiner was

I'd prefer…I mean, ideally I'd prefer a female to carry it out, but because I know I need the answers I'll just go ahead with it. But realistically, I would prefer a female.

37-year-old woman, not pregnant

When the examiner was a male, women showed a preference to have another person with them, a chaperone or partner. They spoke of a stronger awareness of the potential for feeling exposed or being exposed when the examiners were male.

the only thing I think that I might feel a little bit perturbed by would be having a man do it. I think that would make me feel…personally - and I know that's probably silly - but would make me feel awkward… I think in my mind maybe I'm unfair, but if I'm being absolutely honest, I think I would be uncomfortable and I think I definitely would prefer there to be another female in the room

38-year-old pregnant woman

Differentiation was made between women's first scan examination and subsequent examinations. With repeat examination women had a greater knowledge of what to expect, this made subsequent examinations easier. A positive first experience also made subsequent examinations easier, age was hinted at as a factor in acceptance, it was easier the older the woman was.

I think my early experience of internal scans were the worst, and I think probably for me because of the fact that I didn't know what was coming.

38-year-old pregnant woman

Yeah, because I was worrying about it, because obviously you don't know what's going to happen, you've never had one before. And obviously, at a young age, you don't really want anybody messing around down there really.

22-year-old woman, not pregnant

Knowledge of and indications for TVU

Women were well informed about the indications for TVU. It was recognised a TVU could give better resolution of images. This feature of transvaginal examination over transabdominal examination was often described by examiners, and therefore repeated by patients, as "allowing a closer look". Women identified that when the abdominal scan did not give all the detail a TVU may be needed. Women had some knowledge about the clinical indications for TVU such as examining the ovaries, checking for cysts, looking at early pregnancies and when undergoing fertility treatment. The knowledge came from personal lived experiences, from friends and

family and some profession knowledge, detectable through some of the medical language used such as “viability” and “menstruation.” The high level of knowledge displayed is perhaps explained by the demographic of the cohort; four women were medical professionals and many were having repeat examinations. Knowledge about the indication for the scan, particularly if they were expecting it, allowed women to accept the examination and engendered a positive experience.

> *I think the impact is having the knowledge of what is happening, so that's a positive really. I think everything else is just you've just got to take it as it comes really, haven't you, take it on the chin and just carry on really. Yeah, and just try not to let it bother you really.*
>
> *40-year-old woman, not pregnant*

Women wanted to be prepared for the possibility of TVS before the appointment. If the examination was introduced without prior knowledge it could come as a surprise to the woman and lead to worry and anxiety.

> *Just because I was expecting just to have a normal scan and wasn't really expecting to have to have an internal scan. So yeah, just having to get undressed and everything, I wasn't really ready for that [laugh] [...] I was a bit unsure about having it anyway and because I'd come for the scan on my own and I wasn't expecting it, I was a bit nervous.*
>
> *30-year-old pregnant woman, 1st scan*

Women were asked to discuss how they were told about the need for a TVS by the health professional and they were also asked to describe the scan, as if explaining to a friend. These descriptions brought out interesting descriptive language of the

ultrasound scan probe. One woman described it as *"cucumber shaped"* another as a *"wand they cover in latex and jelly"* and another as a *"camera through your vagina"*

The lay descriptions of the scan are important to inform future good practice in training through an attempt to find language that is acceptable and understandable to women when explaining the procedure of TVU examination. It was apparent from the varied descriptions and struggles to find the right words that there was no uniformity in, or ease of, description that could be easily adopted by the professional. However there were reassuring reports that the scan were not painful and this was sometimes used in the description of the scan. Some women found the scan uncomfortable, but generally less intrusive and not as bad as a smear test, which was used as a common comparator.

> *Obviously it's [pause] a vaginal probe that they're inserting vaginally, so, you know, perhaps for most women, it's not a very nice thing to have, to start off with. But otherwise, it wasn't too bad.*
>
> *35-year-old woman, not pregnant, 1st scan*

Worrying about practicalities

When discussing their experiences of TVS examinations women described how they worried about the practical issues around undressing and being exposed that impacted on their comfort. These were factors such as such as bleeding, creating mess, having a scan during menstruation and displaying worn sanitary products.

> *it was just afterwards, kind of, having my bare bum cheeks on that plastic bed wasn't the best.*

31-year-old pregnant woman

Women had practical worries about safety issues related to the insertion of the scan probe such as whether it would cause bleeding or a miscarriage. These worries were often internalised. When women spoke about their worries during the interviews they wanted the examiner to offer reassurance about these issues without necessarily being asked.

I think it would be reassuring to have the conversation, just for them to, sort of, say, you know, it's nothing they haven't seen before or that actually it makes no difference to the outcome of the scan. […] you know, just for them to reassure you, I suppose, that it's fine or whatever.

31-year-old pregnant woman

I was, like, oh, at the back of your mind, like, oh, what if it sets the bleeding off again, or something. But no one said anything about anything like that. I kept on thinking, it's probably just stupid, but I did have that thought, to be honest. […]if they'd said don't worry, this can't cause a miscarriage. Because I guess that's what you're worrying about.

30-year-old pregnant woman, 1st scan

Theme II: Reassurance through communication

Communication was a key influence on the woman's experience of the TVU. In this theme communication encompasses both verbal and non-verbal communication. The examiner's manner and verbal communication were very important and good

communication made women's experience positive. Within this theme there was commonality among all the subthemes relating to communication, the common factor was that women were seeking reassurance, this was reassurance in the examiner, the steps of the procedure and the findings. An overview of the theme and subthemes in given in Table 25.

Table 25 Theme II: reassurance through communication

Subtheme	*Illustrative quote*
Examiner's manner	*Their reassurance really, generally, just by trying to keep me comfortable and calm, and their overall, you know, professionalism that they used. It was quite good, yeah* *35 year old, not pregnant, 1st scan*
Narrative communication	*And, like, when people are talking you through it, just sort of say what they're looking for, it's reassuring* *31 year old, pregnant*
Dealing with silence	*It shows I'm looking at this, I'll let you know in a minute. So you just, kind of…like, that way, while they're quiet, because sometimes when they go quiet, you're like, what's happening. But at least this way, when they actually say, oh, I'm looking at the moment, don't worry, you know, I'll let you know.* *37 year old, not pregnant*
Communicating findings	*I think when things aren't looking so good, I think just communicating the different aspects that they're looking at, and then perhaps giving a conclusion at the end, saying what they've seen* *38 year old, pregnant*

Examiner's manner

Women spoke of a sense of the examiner's manner, this would be conveyed in part through non-verbal prompts such as facial expressions and tone of voice.

> *I think just her communication, her tone. She was very reassuring and knowledgeable and very calm, and yeah, seemed to know what she was doing [laugh].*

36-year-old pregnant woman

The perceived manner of the examiner had a strong influence on experience, positive and negative. Women appreciated the examiner to be professional yet personable in the way they communicated. Professionalism was conveyed through the knowledge the examiner displayed, their ability to be thorough and to explain things.

> *She was really professional; she took her time, explained things but was personable as well.*
>
> *38 year old, pregnant woman*

Some women had experienced unfriendly and blunt communication, this led to a negative experience. Women preferred the examiner to display a calm, caring manner, and some women referred to a *"nurturing sort of way."* This lead to a much more positive experience and greater satisfaction with the care received.

> *I can particularly remember on two occasions that the lady who - I don't know whether the term's sonographer - but she wasn't very nice. You know, she wasn't very friendly she didn't put me at ease. It seemed very much a task to her, rather than, kind of, a personal thing.*
>
> *38 year old, pregnant woman*
>
> *the way they speak as well in quite a nurturing sort of way, so you actually feel like you are cared for, rather than them just say, well, this is what's happening and you've got this, and that's not right. And it's how they tell it to you, so the way they speak.*
>
> *40 year old, pregnant woman*

Interestingly the examiner's manner, as much as gender, influenced the woman's views on the usefulness and role of a chaperone for a TVS. Many of the women identified that the chaperones role was to reassure the patient if they showed signs of discomfort or distress. Some women thought the additional person was there to check on the examiner and make sure process is followed. Others thought they were there to help the examiner, for example to input data.

> *Probably just reassurance. Not even…I don't really know why I'm saying there should be a female in the room, but yeah, I suppose just reassurance really.*
>
> *34-year-old woman, not pregnant*

When the examiner displayed the positive attributes that made women feel comfortable and reassured, they did not necessarily identify a need for a chaperone. Conversely if the examiner displayed the opposite they were more likely to identify that role of reassurance from the chaperone as it was lacking from the examiner.

> *I'm taking it that the sonographer is competent in what they're doing, so the only other bit for me is the personable bit - I think if they've got the people skills and they're, kind of, monitoring that side of it as well, I don't think there should be a need. But I think I could see where, if I looked back at my experience, it would have made a big difference had there been somebody there who was, I don't know if you want to call it the soft side or something.*
>
> *38-year-old pregnant woman*

Narrative communication

Narrative communication refers to a dialogue from the examiner describing what they are doing and seeing during the scan. Women took from this a positive reassurance that if the examiner was talking to them they had something to focus on and they felt more at ease.

> *I do think somebody's who communicating with you and just letting you know what they're doing, which I think I've had more in more recent experiences, makes a big difference […]So, you know, if you can feel that they're, kind of, stretching you at one side - it sounds a bit graphic - but, you know, say well, actually what I'm going to do now is look at your other ovary or whatever they're looking at, and just explain what they're doing*
>
> *38-year-old pregnant woman*

This descriptive, step-by-step narrative was a big feature in the interviews. Women spoke about in a positive light, it reassured them during the scan that everything was okay and they could be sure of the examiner's intentions. During an ultrasound consultation there is the added dynamic of the visualisation of the findings on the screen. Most women seemed to like to see the ultrasound screen; it was another factor in reassurance. They were aware they could not interpret the images so if the examiner explained them that further increased satisfaction in the experience. As to be expected this was a stronger factor in reassurance for the pregnant women when it allowed them to see the baby, however non-pregnant women also showed a preference for seeing the screen.

I wouldn't even know what I was looking at, I don't think. I mean, just out of curiosity, just because I'm fascinated by it, yeah, I'd love to look [laugh], but I think that's just a personal…

40-year-old woman, not pregnant

She was showing me on the screen what she was looking for, she was obviously telling me if it got painful or too uncomfortable, then tell her to stop. And she was just constantly talking, which was a bit more of reassurance

22-year-old pregnant woman

Dealing with silence

Women found verbal communication reassuring and so when the examiner became silent this transmitted non-reassuring feelings and made the women nervous, feel panic and fear the worst. This was most commonly in reference to fears about their pregnancy.

Whereas if you were lay there and they weren't saying anything, it's a horrible time, even if it's only ten seconds, it feels like forever.

34-year-old pregnant woman

Women spoke of being able to deal with silence much better if the examiner qualified the silence by continuing the narrative dialogue, saying something like *"now I am just looking at this"* or *"I will just need to concentrate on these measurements so I may go quiet for a minute"*. If the examiner explained why they were going to be quiet and prepared the woman for this then it was easier to cope with the silence that followed.

They said, we're just going to have a look now, so we'll just be quiet for a few minutes just while we look and just while we calculate or count and get what we need, and then once we've done it, we'll just show you and tell you what we've done. So as long as they say that, I think that's fine. It's just suddenly if they talk to you and then suddenly become quiet, you think, oh, hang on a second, what's wrong. But on all occasions, I have been said, you know, we'll just be quiet while I just take a look. And that's fine.

40-year-old pregnant woman

Communicating findings

Given that women were accepting of TVS as they felt it necessary to get the answers they sought; it is therefore unsurprising that women reported positive experience and most reassurance when results were given to them without delay.

It was because of the anxiety that I was having and, again, the reason why I had the scan, rather than, sort of, having to come back another day, or perhaps given the result over…you know, by letter. That was quite nice to have that information on the same day.

30 year old woman, not pregnant, 1st scan

Women appreciated the examiner being informed enough to be able to give them results, even if they could not explain the consequences and implications of those results. Basic findings were often communicated during the scan through the narrative dialogue the examiner was having with the patient. More complex results and explanations were best left until the woman was dressed so she felt comfortable and could concentrate on the information. This participant describes a bad experience she

had. The nurse conducting the examination asked a consultant to look at the images on the screen, the consultant then stood at the end of the couch talking about the implications of the scan findings:

> *I couldn't have told you ten minutes later what she'd said. That was the thing, I came away not knowing what had been said or decided because I'd just been focusing on, I feel very uncomfortable. And it was a bit, they see it every day, so you could really tell it was a bit…sometimes it feels in those scans, like, because the consultants or the nurses are doing 20 scans a day, that they've forgotten that, for you, this is a really big thing to walk in and scan. And sometimes you do want to just say, you know, I'm an individual here and I know you're doing a lot, but just remember that, for me, this is…*
>
> *34-year-old pregnant woman*

This is in contrast to another participant who described a positive experience because she had been given the opportunity to dress, then was given a clear explanation of scan findings and opportunity to discuss and ask questions.

> *So once the procedure was…the scan was performed, I was then asked to put my clothing back on and dress myself, and was asked to come back to, sort of…well, we stayed in the room but was just asked to come back and have a seat. And I was given the opportunity to discuss what the scan showed. An explanation, you know, she explained what the scan showed and gave me the opportunity to ask any questions, if I had any*
>
> *34 year old woman, not pregnant, 1st scan*

A clear explanation of findings in simple language the women could understand, using lay rather than medical terms left women satisfied. They were further satisfied if this could happen without delay, ideally by the person who had performed the scan. There were varied experiences in terms of getting results. Some women had to wait to speak to a health professional; nurse, midwife or doctor on the same day after the scan had been performed and reported. A few women had the experience of waiting to see a consultant or being told the results would go to the general practitioner, these were scans done by sonographers. When one person did the scan but then another, for example an early pregnancy nurse, explained the report and management within the same appointment it was commented by one woman that this was *"disjointed"* and would be best done the person who had performed the scan, as they would be better informed. When women were left to wait for scan reports, they felt like they were in *"limbo"* as their questions were not answered. Others described satisfaction when they had full, and timely explanations by the person who had performed the scan. When this was a nurse a couple of women described them as *"lovely"* and *"caring."* The encounter was left on a positive note if the examiner used a technique of summarising findings at the end of the scan.

> *I think maybe when you've got to the end of it. So maybe at the end, if you could say, right, you know, what I've seen here is X, Y and Z, and that would indicate that your pregnancy isn't progressing, or that it is or whatever.*
>
> *38-year-old pregnant woman*

Theme III: Attitudes to training

This theme explores women's attitudes towards professionals training in transvaginal ultrasound. When women had previously had a scan by a trainee operator this was reflected upon. When a woman had not been scanned by a trainee general opinion and attitude on acceptability was sought, with prompting if needed to consider characteristics or the trainee such as gender and the role of the supervisor. Within this theme two subthemes were identified, these are outlined in Table 26 below.

Table 26 Theme III: attitudes to training

Subtheme	*Illustrative quote*
Everyone has to learn	*No, because they've got to learn. The people that are scanning you learnt somehow, so they've all got to learn and they've all got to train* *40 year old pregnant woman*
Communication dynamic	*Yeah, I think at the same time, if they're just talking through what they can see with the trainee. The sort of thing what I found, they were actually communicating with them and I was listening as well. And then there were times when the trainer was actually communicating directly with myself as well. So yeah, I think communication's important really, they should do* *37 year old woman, not pregnant*

Everyone has to learn

The majority of women were willing to have a trainee carry out their ultrasound examination. They were of the opinion that everyone has to learn and therefore did not mind trainees *"practising"* on them. There was a recognition that *"everyone has to start somewhere"* and in order to ensure the training of a future generation they were happy for this to happen.

I realise that obviously people need to get the…you know, they need the practice and that's how they'll learn…

35-year-old woman, not pregnant

However many women would give a seemingly positive response to training and demonstrate willingness to be scanned by a trainee and then attach a caveat such as, I don't mind…. "*as long as it's not somebody early on in training, that there's obviously someone there checking, no, I wouldn't mind."*

If trainees were performing the scan the most commonly reported factor women wanted reassurance of was that adequate supervision was provided and training was being overseen. If training was being supervised and this process was open and transparent to the woman, meaning the trainee was introduced clearly as being in training and assurance offered that if the trainee reached the limitation of their ability they could ask for help, or a supervisor would step in, then that was considered adequate.

Yeah, well, I'm presuming a trainee wouldn't be in the room on their own. So to me, if somebody said, I'm just going to ask the other person just to have a look at this because it's something that I'm not familiar with, then I would be fine with that.

38-year-old pregnant woman

One women said she would prefer to know in advance if there would be a trainee scanning so she had the opportunity to say no and avoid feeling pressured on the day to allow the trainee perform the scan. This woman was the only participant to have

strong feelings that she would not allow a trainee to do her scan for fear of the examination being longer and very uncomfortable.

> *I think this is one particular thing I'd say no. I think if you were going to do that, I think a letter previously to say a student would be giving you that, so that I'd have the opportunity to say no. Only because it's so much worse when they're trying to find the image they want and they're having a good dig around, which is what they have to do. And I think if somebody is a student and they're not able to do it, it makes the whole process so much longer and so much worse, and I think I'd end up…I mean, I think I'd walk out. I think…no, I couldn't, no. I think I'd need to be told.*
>
> *40-year-old woman, not pregnant*

Some women identified the indication for the scan as a factor in their acceptance of trainees. If the scan were deemed to be for a *"routine"* indication, particularly outside of pregnancy then there was more willing than if *"it was something quite sensitive"* such as if they were pregnant and presenting with bleeding.

> *I think it depends on why you're having the scan. Because when I had this last scan, it was to see whether I was actually pregnant or not after bleeding. I think I'd be a little bit more nervous then having somebody who wasn't very experienced at all, just being nervous in the first place. But I think if it was a routine scan, I wouldn't mind as much. But yeah, so long as someone was checking it, I guess it doesn't make a difference.*
>
> *30-year-old pregnant woman*

A few women raised the trainee's prior experience as a factor in acceptance of a trainee. They were more cautious if the trainee appeared inexperienced and didn't like the idea that they could be the first patients they ever scanned. A couple of women felt it would be helpful *"maybe saying what their past experience has been. Not in too much detail."* One woman thought it would be useful to know why they wanted to get the experience, to understand what they would achieve from it. There seemed to be an automatic assumption from women that if a person was qualified to scan then that was adequate. Only one woman acknowledged that within a group of trained operators the levels of experience could be different.

> *It would be useful to know and perhaps why they were wanting to, you know, get the experience. Would it be something that they wanted to perhaps improve their skills or they wanted to go on to do further training. That might be…you know, it might be useful to know that because it's, sort of…it's better to understand, so you know it's obviously what they'll achieve from it really.*
>
> *35-year-old woman, not pregnant*

Communication dynamic

This subtheme relates to description of modification in communication, during the clinical encounter, when there was an additional person involved in a training capacity. It explored three levels of interaction: between the patient and trainee, the patient and trainer and trainee to trainer interaction. This was an interesting dynamic to explore, as it has not been previously documented in the literature. Women were specifically asked what was acceptable to them and what they would deem good practice in the scenario when a trainee was carrying out the scan.

In concordance with the theme reassurance through communication, women expected the trainee to communicate with them. It was this communication that reassured women and made the prospect of a scan by a trainee operator acceptable. A trainee making a good first impression by displaying manners, engagement in the process and confidence gave women confidence in their abilities. Conversely if the trainee appeared anxious and nervous then this feeling was conveyed to the woman and did not instil confidence.

> *Actually, yeah, saying that I would. I suppose first impressions and all that go a long way. So yeah, if that trainee was confident in telling me who they were and that they were a trainee, and they were going to do this but this person was at hand, I would actually think, well, they seem to know what they're doing, they're very confident in telling me and confident in speaking to me.*
>
> *40-year-old woman, not pregnant*
>
> *Because what sometimes happens is trainees are so…they'll be a bit anxious themselves maybe in a situation or maybe they can't find what they're looking for, so they get a bit nervous and anxious. And then that communication of I'm just trying to find this or this looks normal, if you don't really know it's normal, you know, the communication can break down a little bit.*
>
> *40-year-old woman, not pregnant*

Women ideally expected the same level of communication from the trainee as the qualified professional but they recognised this was perhaps an unfair expectation given the person was training, therefore could be nervous and focused on the technical task. In order to compensate for the short fall between the levels of

communication the woman expected and that the trainee was able to give, the role of the supervisor became very important. In this situation, the women expected the supervisor to be more reassuring and make up for a lack of communication from the trainee.

> *Yeah, because I think if somebody's training and they haven't got the experience of dealing with certain situations, then obviously that person who's got the experience can tell them this is what we're doing and why we're doing it, and this is […] I would probably expect, if they're learning, that they would maybe not be communicating with me so much, or maybe they've had somebody doing that for them […]*
>
> *31 year old, pregnant woman*

All women wanted the trainee to be introduced and their position and level of supervision made clear. Most women thought that the introduction was best coming from the trainee; this made them seem more capable and reassured women, but it could also be from the trainer and this way the trainer established a rapport with the woman and reassured them they were overseeing the training.

> *No, I think them themselves would make you feel a bit more reassured, like they were a bit more confident. Sometimes if someone else is doing all the talking, it doesn't…it makes that other person feel a bit less capable, even though…or seem a bit less capable. Whereas if they introduced themselves, I think that would be better.*
>
> *34-year-old pregnant woman*

Women did not mind the trainee having findings checked, and they absolutely expected it if the trainee was not sure. Women appreciated openness about uncertainty and the need to have supervisor input. One woman referred to it as "just a process" and they were accepting of supervisors needing to take over or to repeat scans. Women wanted to be involved in the communication dynamic between the trainee and trainer rather than for them to talk between themselves. A couple of women described increased satisfaction from having a scan by a trainee as while the trainer was teaching the student the woman felt she was learning too and left more informed than she would have been if the trainee had not been there. Another woman felt that "*knowing that both trainer and trainee were both looking at my scan at the same time would be a good thing"* assuming perhaps a more thorough examination.

Theme IV: Views on the concept of simulation

Women were shown a picture of a virtual reality transvaginal scan trainer (Medaphor® Transvaginal ScanTrainer®, see Appendix 24) and given a brief explanation of how it worked and could be used for training. The views of simulation for training in ultrasound were then explored. Women were also asked about benefits and limitations of simulation. An overview of the theme and subthemes is given in Table 27. Women identified benefits and limitations of using simulation for training but overall felt that it should be used before clinical practice.

Table 27 Theme IV: Views on the concept of simulation

Subtheme	*Illustrative quote*

Lacking interaction	*I guess the only other thing is just the reassurance side, that it doesn't…you know, you don't have to do that verbal whilst you're doing what you're doing really. And that's probably part of the training that I think is most important.* *31 year old pregnant woman*
Not under pressure	*But I do think that's quite good for just working out what's on the picture, because it must be quite nerve wracking when you've got somebody sat next to you trying to work it out and they're just being under pressure.* *30 year old pregnant woman, 1st scan*
Trainee confidence	*Yeah, I reckon it would build their confidence as well, because it is or it must be nerve wracking for them trying to do it right and, you know, get the right images, keep the patient comfortable. So if they feel confident in the machine, confident in how to use it and what they're looking for, I think the whole process would be easier for those people* *30 year old pregnant woman, 1st scan*
Benefits patient comfort	*I think, like, it, sort of, gives them the practice and that when they're pushing and moving around, if it's going to hurt and stuff, so they know, like, basically how far to go, sort of thing, when they're doing it properly* *22 year old pregnant woman*

Lacking interaction

The major limitation identified by nearly all women interviewed was the lack of patient interaction in a simulated learning environment. As the examiner's manner and how they communicated while having an ultrasound scan were identified as key contributors to the woman's experience, it followed that they felt the simulators main disadvantage was the lack of interactivity. Women felt that nothing could replace *"the personal touch"* and trainees need to learn not just how to scan but also how to interact with the woman while scanning, how to explain what they were seeing and offer appropriate reassurance. A couple of women said the feedback and response

would be better from a real patient than a simulator. Women interviewed were very insightful into how simulation may be used in the initial stages of training in order for the trainee to develop skills, then in the clinical environment they could appear more confident and focus on communication.

> *No, not really, because this is about getting the skills really more than anything else. The communication bit can come along once they actually start doing the real…but obviously, they wouldn't just go straight into, like, you know, doing it straight…they'll do it on a patient but under a trainer. So that bit of training can come from that side really when actually…*
>
> *37-year-old woman, not pregnant*
>
> *I suppose communication comes with practice and the more confident you get with something, I think the communication would get stronger. So it may be hard in the beginning because you have actually got a person in front of you rather than a machine, but if you're confident, then your communicate will…*
>
> *40-year-old pregnant woman*

Not under pressure

Women did however recognise some advantages to a simulated learning environment. The women could empathise with trainees and see that simulation would take away the pressure of learning in front of a real patient. It was recognised that patients would want to speak to the person scanning them and ask questions, if a trainee were concentrating on the skill they were learning the woman could not do that and this could lead to dissatisfaction from both parties.

Just being able to practice with being under a bit less pressure, I think. Because I know when I'm learning something new, to have somebody sat there that I'm trying to practice on would be quite nerve wracking. Whereas being able to practice and take your time, and work things out before you've actually got a real person would be a benefit.

30-year-old pregnant woman, 1st scan

Trainee confidence

Some women identified that simulation would be good for skills learning, helping trainees *"get up to speed"*. It was recognised that by practicing on a simulator trainees could *"get to know their way around"* before scanning real patients then once they started doing some real life scanning they would appear more confident. The trainee appearing more relaxed and confident in turn put the women at ease and therefore was considered advantageous to women.

I think good for the trainee and I think good for the first patient that they get. You know, the trainees themselves would come into the room and know what they're speaking about on the screen, and so, again, would put the patient at ease. So yeah, I think for both really.

40-year-old pregnant woman

Benefits patient comfort

Increased patient comfort was cited as an advantage of simulation. A few women recognised that a patient may not *"say ouch or scream or anything, thinking*

that…you know, i'll just put up with it." Other women agreed that a patient may feel discomfort and not voice it for fear of seeming impolite but feedback a complaint afterwards, for example saying *"oh yeah, that trainee did it but I just felt uncomfortable, actually I don't think she knew what she was doing."* In theory women felt simulation was advantageous in this sense but had no idea if it were actually the case that after using a simulator a scan performed by a trainee would be more comfortable.

DISCUSSION

Principle findings

The central aim of this study was to explore women's views and experiences of transvaginal ultrasound examination, with a particular focus on training professionals in transvaginal ultrasound. For most participants in this study TVU examination was not perceived to be an adverse experience. In line with other studies women were accepting of TVU when it was recommended.[166] There was a resigned acceptance that it was necessary to give them the information they wanted about their reproductive health. The factors that influenced a positive experience were centred around how the examiner communicated. An examiner who appeared calm and friendly with a professional yet personable manner evoked a positive experience. Women appreciated verbal communication. They liked the examiner to keep talking to them, to tell them what was expected, in terms of the procedure, and to keep them informed of what they were doing, by narratively describing the steps they were taking. Women could cope better with silence if the examiner first qualified this with an explanation. Women's experiences were extremely positive when they received a timely and informed discussion of results once they were comfortable and could focus

on this information. Women did not universally agree on the need for a chaperone, however they were more likely to find a chaperone reassuring if the examiner were male or they did not feel completely reassured by the examiner.

This study has highlighted that women are willing to have scans performed by trainee examiners; they are of the attitude that everyone has to learn. When a trainee was performing the scan the way the communication dynamic was handled was key to the women's experience. Trainee operators presenting themselves in a calm, friendly manner and being open about their abilities and limitations made them performing the scan an acceptable prospect to women. Some even found additional benefit from having a trainee perform the scan as they got more attention and information. Women could identify benefits and limitations in the use of a simulator to teach ultrasound training. The major limitation, given the pivotal role of communication in the woman's experience of a TVU, was the lack of interaction between the simulator and the trainee. Women identified concern that trainees would not learn how to communicate well. However they also identified potential benefits in trainees learning skills in a non-pressured simulated environment so that they could appear calmer and more confident when scanning real patients and women thought there might be benefits in increased patient comfort.

Strengths and limitations

This qualitative research study contributes important and experiential insights into a limited qualitative body of knowledge around women's experiences of TVU, particularly as regards experiences and attitudes towards trainee operators, given that no other studies were identified that examined the acceptability and experience of women having transvaginal scans by trainees. A particular strength of the study is that

the emergent understanding is grounded in the perspectives of those who have experienced scans. It emphasises what is most important to women and provides insights into what was well perceived and caninform good practice in conduction of TVU examinations and also in training healthcare professionals in TVU.

A combination of face-to-face and semi-structured interviews was used. This combination gave choice and flexibility to participants and therefore aided recruitment into the study. There were no limitations identified from using a combination of the two methods. The interviews were of comparable lengths in the two groups, there was no missing data due to quality of recording or difficulty in telephone communication and some of the more descriptive and detailed insights came from telephone interviews. The positionality of the researcher, who identified as of a similar demographic in age, ethnicity, education and profession as the participants was viewed as an advantage, as good rapport was built with the participants who were willing to share their thoughts and experiences.

Several limitations must be taken into account when interpreting the results. The study endeavoured to explore women's views and experiences of transvaginal ultrasound examination. The study sample is homogenous, all women were pre-menopausal, the majority were white, British and well educated, however similar high levels of acceptance have been reported in black populations in developing countries.[164] All were recruited from the same department within a single UK hospital therefore the findings cannot be generalised to all women who have undergone a TVU scan. The sample was further narrowed by the exclusion of women who have had a miscarriage and those who needed admission or ongoing care/follow up by EPAU, meaning ectopic pregnancies and pregnancies of uncertain viability

were excluded, leaving only women with a positive outcome and on-going pregnancy. The study therefore did not capture the views of the patient population who had a negative outcome and findings are biased towards those with a positive outcome. It can be surmised that women who experience a negative outcome following a scan would have a different experience and different needs to address than those with a positive experience. The study participants all accepted routine invitation for cervical screening and had accepted a TVU scan. However including women who refused a TVU scan or did not attend cervical screening would have broadened the sample and captured perhaps different perspectives. Some of the significant predictors of willingness of women to undergo TVU reported in the literature such as history of previous sexual abuse and previous painful vaginal examinations [175] were also not explored or captured in this study, so the findings may be bias towards women likely to accept examinations.

Interpretation and comparison with other studies

The findings of this qualitative study are largely in keeping with other studies, which report high levels of acceptance of TVU amongst women when it is necessary for their medical care.[164] Perhaps as TVU has become a more widespread practice in the last 20 years, women were expecting a TVU and were knowledgeable about the application of TVU making it more acceptable to them.[157] The study highlights what is good practice to make TVS acceptable such as due care, sensitivity and communication,[167] adequate privacy, attitude of staff toward patient dignity and providing a calm and relaxed environment.[168] One of the major factors that impacted on women's experiences was the way the examiner communicated. Prior qualitative studies on women's experiences of gynaecological examination have emphasised the

importance of good emotional contact between patients and doctors/nurses for intimate examinations.[176] Information is an important part of the examination situation and women want their doctor to talk to them about what happened during the examination. When this happened women were reassured [176, 177] and conversely when information was insufficient women worried. [178] Interestingly Thoirs et al. [179] explored the alternative perspective of the sonographers' experiences of TVS, using a critical incident technique. One of the themes from their qualitative study was related to communication and sonographers often referred to their own communication style as contributing to the patient experience.

In this study women were accepting of male examiners, but showed a preference for a female. The gender preference of the person that performs the examination has been widely reported with mixed outomes.[156, 176] Basama et al. reported gender had no influence on women's perception of the ultrasound scan[167] whereas Sharma et al. reported 84.1% of women preferred a female, this study also reported 95.2% of women surveyed did not want a chaperone,[180] it could be inferred that gender was the major influencing factor. A chaperone is an impartial observer and guidance states one should be offered to a patient when performing any intimate examination, regardless of whether or not the doctors and patient are the same gender.[181]. Previous studies have assessed women's preferences for chaperones through surveys and therefore explored quantifiable factors such as age, previous experience of examinations, parity and indication for the examination. This study due to its qualitative designed allowed a deeper understanding of what contributes to women's preferences for a chaperone and found that as well as gender the examiners manner was a factor. If the examiner was not reassuring to the woman in the way they communicated women felt a chaperone was beneficial. Perhaps a shared decision-

making approach within a consultation involving an intimate examination should be seen as more important than a rigid approach towards using a chaperone.[182] This study found women preferred to get results of examinations without delay and preferred if the implications could be explained. The professional most able to do this is a doctor and there is a move to more doctors scanning at the point of care [12] to offer this 'one-stop" model of point of care ultrasound delivery which women have been shown to prefer.[15]

Women were accepting of trainee operators, some voicing a utilitarian ethical stance of *"we all have to learn"*. However one woman did not want to have a trainee operator so acceptance was not universal. Coldicott et al.[151] describe how there has been a shift from a Utilitarian towards a more Kantian based ethical value as regards intimate examinations. Coldicott[151] explains humanity should be seen as an "end in itself, never merely as a means", so using any one person as a "means to an end" – for example using patients as teaching "aids" is unacceptable.[183] Other authors take the same view of an ethical imperative to use simulation based medical education to reduce the use of patients for training.[154] Women were accepting but a little wary of trainee operators. They wanted to ensure trainees were adequately supervised and their acceptance of a trainee examiner may depend upon indication for the examination and the experience of the trainee. Women spoke of caution if the trainee appeared inexperienced and voiced concern if there scan was the first the trainee had ever done. Therefore if guided by ethical principles it should be advocated that trainees start at least initially on simulated models for ultrasound before scanning real patients. The World Health Organisation have recommended for patient benefit institutions use simulation in the education of health professionals.[184] Women in this study were seemingly willing to allow "*practice"* and could even empathise with the

trainees' position however they did expect a level of professionalism on par with a skilled operator. Perhaps simulation at least in the initial stages of training can get a trainee to the level that they appear confident in clinical practice. Women identified that simulation may benefit patients in making the examination more comfortable for them. Tolsgaard et al. [72] have substantiated this assumption, reporting findings from an RCT that found patient reported discomfort was significantly reduced if trainees had used a simulator prior to clinical training. Patients' confidence and perceived safety in their provider were also increased, however overall satisfaction was not. This indicates that multiple factors influence patient satisfaction.

Implications for clinicians and policy makers

This study has implications for clinicians in informing good practice regarding TVU scanning. Considering women's experiences of TVU can identify strategies that contribute to high acceptance rates of TVU. Suggested advice on strategies practitioners can adopt in their own practice that have emerged from this study are:

- Offering clear explanations of the procedure in language the woman can understand.
- Initiating a communication narrative with the woman so she knows what is happening at all steps of the examination
- Offering and providing chaperones after discussion in a shared decision making process between the patient and examiner.
- Clearly explaining the indication for the examination in a way the woman understands to enable acceptance.

- Being aware of the practicalities of the scan, including the environment for the examination, ensuring comfort, privacy and availability, for example, of tissues and sanitary disposal facilities.

This study makes a unique contribution to the existing body of literature on women's experience and acceptance of TVU as it explored women's views on trainee operators. The study has identified the increased complexity of the communication dynamic that comes when a trainee operator is providing the scan and how this can be negotiated to make it an acceptable experience to women. Suggested good practice points in this regard are:

- Trainees should introduce themselves to the patient
- Trainees should be open about the their abilities and limitations, and inform the woman if they are working under supervision
- Trainers have a role to play in maintaining good communication with the woman to reassure her if the trainee is focused technical skills
- The woman should be involved in the communication dynamic if a scan is being done for training purposes.

Policy makers who decide training curriculum, requirements and guidelines need to evaluate the evidence for simulation at least in the initial stages of training. This appears to be in a woman's best interest. Its use is supported by women as demonstrated through this interview study and has shown favourable outcomes in reduced patient discomfort and need for supervision in clinical practice.[72]

Unanswered questions and future research

In terms of women's knowledge and experience of TVU examination there lacks qualitative research that gives a deeper understanding than existing survey based studies. Further research however is needed on a more diverse sample of women. It would be valuable to capture the views of the postmenopausal age group. Willingness to undergo TVU increases with age,[89] this demographic may also hold different views on chaperones,[180] experience more discomfort and have different body image attitudes than the pre-menopausal.[185] It would be important in future research to sample from a group of women having scans for varied indications as the setting and indication is known to influence acceptability.[165] Future research should also explore routine and pathological indications both inside and outside of pregnancy. Within the early pregnancy population it would be prudent to include all pregnancy outcomes, not only positive.

This study found communication was the biggest influence of women's experience of TVU. However we only considered verbal and non-verbal communication, with ultrasound there is the added level of written communication in the ultrasound report. Women will often get copies of their scan reports in their patient-held pregnancy records for example and it is advocated as good practice to offer patient copies of scans. Exploring women's perception of these written records was not done in this study but should be considered in future studies.

It was clear in this study from women's perspectives that learning on simulators risks neglecting communication training, and this is important because the way the examiner communicated was pivotal to women's experiences. Future research could address how to teach and assess best practice communication. Perhaps strategies to

teach communication in the simulated setting using patient educators or live models could be explored.

CONCLUSION

This qualitative interview study provides deeper insight into the views and experiences of women on transvaginal ultrasound than prior quantitative studies. It is the first study to qualitatively explore women's attitudes towards examinations performed by trainees and opinions on simulation training. The study found that women were accepting of TVU when it was deemed necessary and did not find TVU to be an adverse experience. Women were also willing to accept a scan by a trainee in most instances. The study highlighted that central to a woman's experience of TVU was the communication offered by the examiner. Women appreciated a professional yet friendly manner and were reassured when the examiner kept talking to them throughout the scan. The study has shed light on factors such as decision-making regarding chaperones, women were more likely to want a chaperone if they did not feel wholly reassured by the examiner. Women ideally expected the same level of communication from a trainee as they would from a trained operator. Women appreciated trainees being open about their scanning abilities and limitations and seeking supervision when required. Women could identify benefits and limitations to simulation for ultrasound scan training. Overall it was felt simulation should be used for the initial stages of training so trainees could practice without pressure and therefore appear calmer and more confident when scanning, but the major limitation identified was the lack of patient interaction. This study has offered insight into good practice points regarding the communication during a transvaginal scan and the

communication dynamic when a trainee operator is performing a scan. These practice points and the findings of this study could be taken forward into a training programme to teach communication skills for gynaecological ultrasound examinations, doing this alongside simulation training could offer trainees both the technical and non technical skills which may benefit patients in the early stages of training.

CHAPTER 7

Discussion and Conclusion

Firstly a cautionary note, throughout this thesis transvaginal ScanTrainer® (Medaphor®, Cardiff, Wales, UK) version 3.0 was studied, but there have been several updates to the system since then. Simulation technology is ever changing and interpretation of studies using simulators need to be mindful of this. The market for simulation technology is growing and the companies are large profit making multinationals.[186] The technology is becoming so advanced there now exists immersive simulation, which conveys the sense of being immersed in a task or setting as if it were real.[33] Computer based, virtual reality (VR) simulation is such a fast moving, technology dominated field that evaluation inevitably trails behind innovation.[92] New software versions are constantly being released with updates and modifications, so by the time an evaluation has been carried out, often the original product has changed beyond recognition and the product itself seems out of date.[92]

As the sophistication of simulators increases, with attention to their fidelity, that is ability to reflect reality, so does the likely cost, the simulator systems cost tens of thousands of pounds. Simulation and VR offer attractive, potential advantages to teaching craft skills and as a result are becoming widely established. However, as Kneebone et al. warn, there are dangers of embracing such new technology uncritically especially in light of the increasing demand on health care resources and budgets allocated to postgraduate medical training.[187] There is also little evidence to suggest that the outcomes afforded by sophisticated high fidelity, simulators are any

better than those of less authentic, low fidelity simulation [86] or high quality clinical training.

There is evidence that compared to nothing; SMBE improves outcomes with large effects for outcomes of knowledge, skills, behaviours and moderate effects on patient outcomes.[63, 120] In relation to ultrasound simulation in gynaecology, detailed in the systematic review in chapter two, there is not enough homogeneous evidence to aggregate data reliably by meta-analysis to estimate effect sizes with a great deal of precision. However, compared to standard clinical training there appear to be generally consistent improvements in knowledge, skills, behaviours and patient outcomes, such as discomfort with the addition of simulation.

Tolsgaard et al.,[188] reports a series of studies that led to a high quality randomised controlled trial (RCT), that has shown that the addition of 3-4 hours of practice on a simulator during the initial part of clinical training is advantageous, The study reported benefit to patients; intervention group participants had nearly 20% lower patient reported discomfort scores compared to the control group. Patients also rated intervention trainee group participants higher on perceived safety and confidence compared to the control group and the examination times were significantly reduced in the intervention group compared to the control. However this was the first, rigorous study to demonstrate an interaction between SBME, a shorter length of clinical training and a reduction in the need for both ongoing supervision and the requirement for repeat examinations by a supervisor in the first six months of clinical training. This has important implications for training, implying simulation could be used as preparation for more effective future clinical learning.[72]

Until very recently there has been little evidence to guide or substantiate the assumptions that simulation can work for gynaecological ultrasound scan training. Simulation has been heralded as a potential solution to some of the challenges in delivery of ultrasound training.[100] Investment has been made in simulation technology in UK medical training deaneries, as outlined in chapter three, but there are high trainee to simulator ratios, simulators are underutilised and not integrated into formal programmes for obstetricians and gynaecologists.

It is quite clear that the currently heavily relied upon apprenticeship model of teaching in flawed, this was highlighted in chapter four with trainees getting very little exposure to clinical ultrasound training, and has been criticised in the literature.[187, 189] It is a particularly inadequate model for ultrasound training in the UK system because the current system has barely enough sonographers to meet the growing demand for imaging in clinical practice[28] and few gynaecologists are trained and regularly perform transvaginal ultrasound to maintain skills or provide instruction.[17] The shortage of potential trainers with the time to train is one of the factors why trainees find it so hard to access ultrasound training, although they would like the opportunity to do so.[31] Historically, gynaecologists and especially fertility specialists, self-trained in obstetric and gynaecological ultrasound.[25] Indeed, until recently there was no curriculum for ultrasound and competency based assessment is a new tool. Traditionally expertise in medicine was assumed to equate to clinical experience, but research has not proven a link between length of practice and reproducibility of superior performance.[190] More recent work on surgical expertise recognises that capability is not merely acquired by experience and that there is a potential role for simulation practice.[189] In the study presented in chapter five on validity, the performance of three groups of health care professionals with varying

levels of gynaecological ultrasound experience, was compared on an ultrasound simulator. In this study the intermediate group demonstrated higher scores than the expert group for the advanced skills tasks on the simulator, which as first sight appears counter-intuitive. Using the same simulator, in a similar study designed to determine the validity of assessment of performance in a simulated setting Madsen et al.[65] also unexpectedly highlighted evidence to also support an experience-expertise inconsistency.[188] In this study the expert group consisted of 12 consultants in obstetrics and gynecology who used ultrasound on a daily basis (eight gynaecologists and four fetal medicine consultants). The experts all performed tasks on the transvaginal ScanTrainer® and it was found that the expert subgroup of fetal medicine consultants performed better than the gynaecologists with significantly higher assessment scores recorded. Although most clinicians eventually master the skills that are considered essential in their respective specialities, there is sufficient evidence to support the claim that clinicians do not display expert behaviour just because they become experienced.[188]

As assessment and learning of ultrasound skills are explored further inadequacies of the apprenticeship model are being revealed. In the randomised study by Tolsgaard et al.[73] participants were randomised to either simulation based ultrasound training and clinical training (intervention group) or clinical training only (control group). 26 participants were assessed after two months of equivalent clinical training and outcomes measured by a validated OSAUS scale. Tolsgaard et al.[73] found that only 8.3% of the control group passed a pre-established pass/fail level, compared to 87.5% of the intervention group ($p<0.001$), supporting the notion that apprenticeship training can fail to ensure acceptable clinical performance standards. The gaps between the expected levels of performance and actual performance suggest that clinical

apprenticeship learning may be insufficient when not combined with dedicated time for basic training.

A survey of 621 trainees across Denmark, Norway and Sweden exploring factors associated with trainees' confidence in performing ultrasound examinations found that the number of days spent in a specialised ultrasound unit was a predictor of trainees' confidence in performing ultrasound examinations independently.[26] It took trainees on around 24 months of clinical experience to confidently manage ultrasound examinations independently, while only 12 to 24 days in a specialised unit were needed to reach the same level of confidence.[26] There is no evidence that SBME is any better than high quality clinical training. Those opposed to the introduction of simulation perhaps would argue that the money used to buy expensive simulators would be perhaps better spent on addressing access to high quality clinical training. The problem is this is hard to address, simulators can be manufactured, mass produced and standardised in a way clinical training cannot however SMBE needs to be better aligned with subsequent clinical training where opportune clinical training still remains accepted practice. Although yet to be proven definitively in a UK setting for gynaecology ultrasound there seems to now be enough evidence to say that simulation versus standard clinical training in general is beneficial.[94] However that does not mean to say that simulation, as a teaching method is better than structured clinical training nor indeed should replace it. On the contrary Moak et al.[66] detailed in the systematic review, chapter two, found that students who practiced ultrasound skills on a pelvic mannequin compared to those who practiced on a live model had lower performance scores when then assessed on a live model. This highlights the importance of contextual similarity when assessing transfer of skills.[191]

In chapter four I present data collected from a small group of trainees who had achieved ultrasound training modules in a competency based ultrasound training curriculum and found that to complete an RCOG intermediate scan module in either gynaecology or early pregnancy it took an estimated 47 to 300 scans, a considerable range. Recommendations from international ultrasound societies on minimum standards of supervised training hours and numbers of scans do not take into account the rate at which different trainees may learn new skills. Some may reach the expected level before the recommended number; some may need to do more than the recommended number. In a competency based system a range is expected due to the variation in individual learning curves.[108, 192] It is more important to ensure that all trainees reach the same standard before independent practice. This is the concept of mastery learning, recognised to be suited to high fidelity simulation learning. [44] Mastery learning aims to produce identical outcomes for all at high performance standards. Setting the performance standard and studying the time taken to achieve mastery is the educational equation,[193] which was unknown when instructing participants to use the simulator in the observational study presented in chapter four. It has now been shown that trainees performance can be tested in a valid and reliable way using a simulator and that novices can achieve mastery levels within an average of 3-4 hours of practice on a simulator.[65] However there may be trainees who fail to reach mastery level, research has not addressed what to do about those trainees.

The study presented in chapter five aimed to validate the VR simulator in assessing transvaginal scanning skills by establishing face, content and construct validity. At the time there was no published accepted method of testing validity and reliability of performance assessment for ultrasound skills. Since 2013 this area has received considerable interest and other publications applying the educational methodology for

validity testing described by Messick[148] have come from Denmark.[60, 65, 113] In this framework subjective approaches to validity are only the preliminary stages in validity testing. However, they are a vital step as they help to ascertain the opinions of experts who will be needed to inform, engage in and direct ultrasound-training programmes. As demonstrated in the study of face and content validity reported in chapter five, there were lower ratings from the experts, particularly in some key areas such as realism of the simulation. Face and content validity testing may highlight areas of deficiency in a simulation experience. It could be that these areas can then be addressed. Some studies have used more than one training simulator to overcome issues around authenticity. In a series of studies by Tolsagaard et al.,[72-74] which have been detailed and appraised in the systematic review in chapter two, the simulation intervention groups used the transvaginal Scan Trainer®, supplemented by a 'Blue phantom™' pelvic mannequin that is designed to be used with a real ultrasound machine. The combination of the two simulators could address some of the short fallings of one type of simulator alone. This may be a reasonable approach given our findings regarding suboptimal face validity for some aspects of the ScanTrainer®. However, the studies do not attempt to establish which components of the intervention were responsible for the observed positive effects on training. In a comparison study of these two simulators Schmidt et al.[194] found 31% of their cohort preferred training on a combination of transvaginal ScanTrainer® and Blue Phantom™, and a higher proportion of experienced operators compared to medical students preferred Blue Phantom™.

Validity studies mainly focus on technical skill [129] and although important there are other factors in competence in performing ultrasound such as knowledge about equipment, differential diagnosis, communication with patients and staff and the

ability to receive and ask for supervision from experienced operators.[188] Having clear educational goals as well as providing valid and reliable performance assessments are essential to learning in a simulated environment[92]. As advocated by Schout et al.[129] it would be better to develop and evaluate a comprehensive training programme instead of validating just one part of a procedure performed on a simulator. Training programme development needs three key phases: training needs analysis, training programme design and training media specification, validation has its place in this phase.[129] This requires understanding of educational theories and a multidisciplinary approach with specialists, trainees, educationalist and manufacturers.[129]

Given the evidence of benefit of simulation in the early stages of training and the inadequacies of the current apprenticeship model SBME training should be considered for basic gynaecological ultrasound training, in the interest of trainees, trainer and patients. The benefits to patients from trainees learning in a simulated environment prior to clinical training is even more important for transvaginal scanning given the intimate nature of the examination. In chapter six I explored women's views and experiences of transvaginal ultrasound as an intimate examination. In keeping with prior studies, women were found to be generally accepting of TVU when it is necessary for their care. The qualitative interviews went on explore women's attitudes towards ultrasound examinations performed by trainees and their views on SBME, an area not previously studied. I found women are willing to have scans performed by trainee examiners; they were of the attitude everyone has to learn. The key factor in ensuring a favourable outcome when having a trainee operator perform a scan was the way in which the communication dynamic between the woman, trainee and trainer was handled. A limitation of SBME may be the lack of

patient interaction needed to develop key communication skills. The findings from my qualitative study triangulate with the findings from the study of Tolsgaard et al. who showed patient benefits in terms of reduced discomfort and increased ratings for confidence and safety in the provider when trainees had used simulation but no overall difference in patient satisfaction. The satisfaction rates could be affected by lack of attention to training in communication. Perhaps communication skills need to be taught alongside technical skills and this may involve carefully integrating simulation into clinical 'role played', case scenarios. Communication during gynaecological or early pregnancy transvaginal scans can be particularly challenging for an inexperienced operator because it represents an intimate examination, and the patient may be in pain, bleeding and anxious regarding pregnancy outcome. The trainee has to perform the scan and interpret the findings in real-time and communicate these, including sometimes breaking bad news. The insights and good practice points from chapter six could help guide the development of communication skills training to go alongside technical skills in a simulated environment [(195)] and may further benefit patient satisfaction.

Future research

There is a need for a controlled trial in a UK setting to answer the question Tolsgaard et al. asked "*What is the effect of adding initial simulation based transvaginal ultrasound training to new trainees' clinical training on quality and efficiency of care measured during the first six-month of clinical training, as compared to clinical training only?*" Conducting a study investigating a research question to which the answer is already known can be viewed as inefficient, wasting time, resources and potentially delaying the implementation of interventions with proven benefit, however medical education research is also highly context specific. [(94)] A single study is rarely

sufficient to definitively answer a question and it is generally advisable to replicate the findings.[94] Proving the benefits of simulation used to mastery level prior to clinical training in transvaginal gynaecology in a UK based controlled trial would provide powerful evidence to policymakers and experts involved in curriculum design, organisation and delivery of ultrasound training in the UK. As demonstrated in chapter three, through a national survey of deanery ultrasound coordinators (DUCs) there was a lack of consensus on the role simulation has to play in the acquisition of ultrasound scanning skills. Similarly uncertain opinions were expressed by a group of experts surveyed on the role of simulation in ultrasound training detailed in chapter five. These surveys were carried out before the publication of the RCT by Tolsgaard et al.in 2017, and perhaps repeating them now may reveal more positive opinions, although the lack of widespread, standardised implementation of simulation in the initial stages of clinical training in UK training suggests that they may not.

To evaluate the effectiveness and cost-effectiveness of SBME in gynaecological scan training requires an RCT. Replicating the RCT findings from the Danish group would allow generalisability to the UK NHS setting. Such a trial would require a sample of sufficient size to allow to reliable estimates of the magnitude and clinical relevance of the effect. The problems encountered running a feasibility controlled observational cohort described in chapter four, should make researcher considering running an RCT to use cluster-randomisation given the uneven distribution and access to simulation equipment. Furthermore, there are some important potential differences between the UK setting and the Danish setting, which would need to be considered in the design of a controlled trial, such as setting of the intervention and overcoming the fact that fewer junior gynaecologists are required or expected to perform gynaecological

ultrasound examinations, the majority are done by sonographers and specialist nurses and midwives.

As detailed in the study chapter four, the included UK deaneries did not keep the simulators set up in simulation centres. In the Danish trial the initial simulation learning was done in a simulation centre with a facilitator for assistance.[115] One of the features of high fidelity simulation leading to effective learning is learning in a controlled environment away from patient care and interruptions. [44] Most hospitals have provisions for simulation based learning however the focus has not been on ultrasound learning; perhaps there is an argument for moving ultrasound simulators currently in other areas into skills centres to allow the controlled testing. The next barrier to overcome is ensuring that trainees then complete the simulation intervention. The studies by Tolsgaard et al.[72, 73] had very few participants who did not receive the simulation intervention as planned, unlike my observational study in chapter four or the RCT by Alsalamah [83]. Successfully delivering the intervention may be helped by off site simulation centre training, which is more likely protected time. There is a difficulty in voluntarily getting trainees to participate in simulation training. [111] and so creates challenges in testing the impact of simulation interventions. Evidence is needed demonstrating benefit from SBME to inform educational policy makers if such training approaches are to become mandated, being routinely incorporated into gynaecological scan training programmes. However, without the more widespread availability of simulation, it is difficult to collect evidence, unless research bodies or industry can adequately fund future research studies so that equipment can be provided to test the intervention in a large enough population across several training centres. In the East of Denmark simulation training is now mandatory for junior level trainees in O&G for a large majority of the major

teaching hospitals.[188] Mandating training is often needed to allow sufficient time, resources and faculty to deliver it.[112]

One of the factors in motivation of trainees to use simulation is that it will then lead to clinical practice of the learned skill.[118] This is when the variances of the setting may be realised. In France it seems that GP residents during their placements in gynaecology are expected to acquire ultrasound skills quickly in order to perform emergency scans, [75] and in Denmark relatively junior gynaecologists perform scans in an acute setting.[26] Two months after starting as new residents Danish trainees were scanning out of hours and reported doing an average of 60 scans, the majority under supervision, presumably from senior doctors. In the UK we have a competency based training programme rather than logbook based and the majority of ultrasound scans are performed by sonographers and nurses, very few by gynaecologists. In the UK a relatively inexperienced ultrasound operator would not be expected, or indeed allowed, to scan unsupervised in an emergency setting until they had achieved an assessed level of competency. Therefore, how to practically carry out a trial that requires clinical outcome data [72] would need careful consideration if it is to be achievable in the UK health care setting. As demonstrated by the study detailed in chapter four, even when a group of intermediate level trainees, who were working towards achieving a recognised ultrasound module, only performed an estimated seven scans per week, most of these were within normal working hours and supervised. In a more junior cohort of obstetricians and gynaecologists, Alsalamah found trainees did on average one session every six weeks with 2-3 scans performed per supervised session. [83] Perhaps a trial would be best done in units with an established system where doctors, rather than radiologists, sonographers or specialist nurses, routinely perform the majority of ultrasound examinations such as specialised

ultrasound centres e.g Kings Hospital London. At present, it seems unlikely that a large scale, multicentre RCT would succeed across UK hospitals until there is a shift in workplace attitudes towards gynaecologists performing point of care ultrasound and a larger proportion of imaging done by gynaecologists.

The research in this thesis has highlighted the need for more, well-designed comparative effectiveness research trials to compare two active interventions, which are not without their own design pitfalls.[196] However, such studies should reflect the clinical context and so SBME should be embedded within clinical training and they should not be seen as mutually exclusive educational interventions. In addition, whilst simulation is often employed prior to clinical 'hands –on' experience, restricting its use to pre-clinical training limits the potential utility of simulation. Integrating the use of simulation during clinical training may facilitate consolidation of skills and there may be greater engagement with simulation when clinical training is commenced, as the trainees will have gained a greater insight into the technical dexterity and other skills required.

The current RCT evidence we have for simulation in gynaecology focuses largely on the transvaginal ScanTrainer® and Blue phantom™ mannequin model but other simulators are available. As well as evaluating integrated simulation and clinical training programmes in totality, future research could also look to validate and then test interventions using different types of simulator against each other. However as evaluating ever-changing technology is problematic, as research findings may become quickly outdated if they do not keep pace with the advances in technology. State of the art simulation technology is also expensive. Even if an intervention is proven effective, benefit has to be balanced against cost. It is therefore important that

economic evaluations are undertaken so that cost-utility can be examined. Indeed, some may argue, given the dearth of evidence evaluating future current clinical scan training, that research should focus on existing practices rather than focussing on testing SBME interventions, which do draw focus because they are new, innovative and look impressive. Studying existing methods for systematic clinical training and ways to improve clinical training in the same highly controlled environments afforded to trials of SBME interventions could find similar effects on quality and efficiency of care.

CONCLUSION

Ultrasound in gynaecology has many useful applications and with developments in ultrasound technology allowing more focussed point of care clinical assessment it seems logical for gynaecologists to be able to scan. There are benefits for the patient, the clinician and the service to this model of care. However, in current UK practice few gynaecologists scan regularly and so the conventional apprenticeship model of clinical training is frequently not fit for purpose. Scanning in an indiscriminate, opportunistic fashion within the clinical setting creates inadequacies of learning. Experience acquired in an ad hoc way does not allow effective feedback to be given and risks prolonging the acquisition and consolidation of skills or the loss of skills previously acquired. Training programmes need to be improved such that future generations of gynaecologists learn to scan the female pelvis in a systematic and assured way, through regular teaching and practice. Simulation based medical education may be the solution to this early systematic approach to learning because it appears to have a role both in immediate skills learning and preparation for future learning.

A systematic review of literature shows the benefits in using simulation when compared with standard approaches to learners, through increased skill, competence and confidence; to patients through reduced discomfort and increased confidence and safety in the provider and the organisation, through increased efficiency of training. Qualitative interviews reveal that women do not mind having a transvaginal scan however the use of simulation for basic level ultrasound training is supported by women, given the intimate nature of the examination and the potential for discomfort and distress there is an ethical argument for adoption of simulation, at least at a basic level.

Educational interventions in postgraduate medical education are complex and need to go through sequential phases of development. Understanding the theories behind the learning is important and educationalists are exploring and developing this area. The value of simulation based training still needs to be proven in a UK setting with an appropriately designed and executed randomised control trial incorporating an economic evaluation. A recognition of the importance of gynaecological ultrasound in contemporary practice, in light of advances in imaging technologies that allow point of care testing, should make the case to prioritise the funding of high quality research aimed at establishing optimal training strategies in 21st century obstetrics and gynaecology.

ACKNOWLEDGEMENTS

CHAPTER 2

I would like to thank the research librarians Derek Yates and Liz Askew for help and guidance with systematic database searching

CHAPTER 3

I would like to thank Geraldine Masson, RCOG ultrasound Officer 2012-2015, Alison Gale, Chair of the Simulation Advisory Network RCOG, Nandita Deo, member of the Simulation Advisory Network and Bettina Cayetano, policy coordinator at RCOG for help in designing and disseminating the survey to Deanery Ultrasound Coordinators.

CHAPTER 4

I would like to thank The Health Foundation Shine Award and Springfield Consultancy, in particular Mike Fern, for financial and quality improvement support offered through the project. I would like to thank Iain Dunbar from Medaphor UK for providing software support and assistance to run the trial collaboratively across Deaneries. I would like to thank deanery collaborators: Jo Fiquet, Judith Ten-Hof, Devender Roberts, Aileen Cope, Marianne Hamer, Judith Ten-Hof, Ghada El-Senoun, and Karen Brackley. I would also like to thank Chris Gunn at Birmingham Women's Hospital finance department.

CHAPTER 5

I would like to thank all those students and colleagues who volunteered their time to use the simulator and in particular to David Cole at Birmingham City University for his aid in recruitment of participants. I would also like to thank Paul Smith for statistical guidance.

CHAPTER 6

I would like to thank the patient and public support group and early pregnancy nursing staff at Birmingham Women's hospital for input into the interview guide and aiding recruitment of Women.

APPENDICES

Appendix 1: Systematic review study protocol

Review Q	How does use of an ultrasound simulator affect trainees learning gynaecological ultrasound
First reviewer	Natalie Woodhead
Second reviewer	Ayesha Mahmood
Supervisor	Professor Justin Clark
Project title	Ultrasound simulation for gynaecological ultrasound training: a systematic review

1. Background to review

There has been a cumulative growth in simulation literature since the 1980s. Simulation is being used because it provides a safe and supportive educational environment. It allows users of all levels from novice to expert to develop skills without the pressure and risk of the real clinical environment and encourages the acquisition of skills through experience.

Ultrasound is a technical skill that is acquired through experience. Simulation for ultrasound is appealing as some of the hours needed to train could be done away from the pressured clinical environment.

There are many commercially available simulators and their use is becoming more widespread. Views, amongst clinicians responsible for the delivery of ultrasound training, on whether simulators should be used, how they should be used and when they should be used are mixed.

Any decisions about use of simulation as part of structured ultrasound training programme should be based on best evidence.

Aim

To summarise and critically appraise the available research on the effectiveness of simulators for training in gynaecological ultrasound.

Specific objectives

1. To clarify the evidence base available for use of simulators to teach gynaecological ultrasound. Clarification will be made by a systematic review of the evidence base of journals and abstracts in this topic area, looking at all designs of study.
2. To identify factors that need to be considered in considering simulation as part of an educational programme, for example is their benefit in using a simulator? who would benefit? the timing of the intervention for maximum benefit? can it be used as a self directed learning tool? is the benefit sustained? This cannot be more specific until an examination of the evidence is done.

3. a) Criteria for including studies in the review	
i. Population or participants of interest	Health professions learner: a student, a postgraduate trainee or practitioner in a profession directly related to human health.

	This could be students and healthcare professionals from different disciplines such as medicine, nursing, midwifery, radiography and sonography.
ii. Interventions or exposures	Use of a simulator to teach gynaecological US This could be computer-based virtual reality simulators, high fidelity and static mannequins.
iii. Comparisons or controls	None, conventional training, training on one simulator compared to another or an alternative instructional method.
iv. Outcomes of interest	Any, likely to be things like changes in perceptions, attitudes, knowledge, skills, confidence etc
v. Setting	Any
vi. Study designs	Any, primary research studies of all types
3. b) Criteria for excluding studies not covered in inclusion criteria	
Simulation to teach ultrasound if it is not gynaecological (gynaecological defined as up to the end of the 1st trimester) Studies comparing one groups performance to another all on a simulator to assess validity of the system Studies including experts as the comparator group	
4. Search methods	
Electronic databases	Medline Embase CINHAL BNI PubMed Cochrane ERIC
Other methods used for identifying relevant research	Reference lists of all included papers
Journals hand searched	Simulation in healthcare Clinical simulation in nursing Advances in simulation Since inception

Appendix 2 Medline search strategy

#	Database	Search term
1	Medline	exp GYNECOLOGY/
2	Medline	exp OBSTETRICS/
3	Medline	(1 OR 2)
4	Medline	exp ULTRASONOGRAPHY/
5	Medline	(ultrasound).ti,ab
6	Medline	(transvagina*).ti,ab
7	Medline	(4 OR 5 OR 6)
8	Medline	exp "COMPUTER SIMULATION"/
9	Medline	exp "SIMULATION TRAINING"/
10	Medline	exp "HIGH FIDELITY SIMULATION TRAINING"/
11	Medline	("virtual reality").ti,ab
12	Medline	(8 OR 9 OR 10 OR 11)
13	Medline	(3 AND 7 AND 12)
14	Medline	(UltraSim OR UltraTrainer OR SonoSim3D OR SonoTrainer OR "Space Fan ST" OR SonoMom OR Scantrainer OR Schallware OR Vimedix).ti,ab

Appendix 3: Data extraction form used in systematic review

Title:

Author:

Year:

Country:

Study design and recruitment:

Population: (who, grade/year etc)

Simulator:

Study sample (total + intervention + comparator)
T=
I =
C =

Intervention (what, who, duration, amount, modules etc)

Comparator (what, who, duration, amount, modules etc)

Outcomes and assessment: (what, how, assessor, when in relation to intervention, % follow up (assessed))

Appendix 4: Tables detailing support for judgment using the Cochrane risk of bias tool

Alsalamah 2016

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated allocation ratio 1:1
Allocation concealment (selection bias)	Low risk	Central computer generated
Blinding of participants and personnel (performance bias)	High risk	Unable to blind participants Performance bias due to knowledge of the allocated interventions by participants and personnel during the study.
Blinding of outcome assessment (detection bias)	High risk	Detection bias due to knowledge of the allocated interventions by outcome assessors
Incomplete outcome data addressed (attrition bias)	High risk	Attrition bias and subgroup analysis, of the 77 recruited only 20 participants (9 control group and 11 in the intervention group) received the intervention or control as planned, subgroup analysis was performed with 6 groups and small numbers
Selective reporting (reporting bias)	High risk	One or more primary outcomes is reported using measurements, analysis methods or subsets of the data that were not pre-specified
Other		

Tolsagaard 2017

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated allocation ratio 1:1
Allocation concealment (selection bias)	Low risk	Computer generated allocation ratio 1:1, done by a different person to the one responsible for inclusion of participants
Blinding of participants and personnel (performance bias)	Low risk	Participants not blinded Data coded by blinded investigator Potential for participants to reveal allocation Primary investigator controlling integrity but not clear if blinded to intervention
Blinding of outcome assessment (detection bias)	Low risk	Assessors were blinded to the allocation (patients and nurses)
Incomplete outcome data addressed (attrition bias)	Low risk	Clearly reported in flow chart 2 lost to follow up in 1st 14days 4/26 participants did not attend final training session and use mannequin for training All 26 included in analysis
Selective reporting (reporting bias)	Low risk	All specified outcomes reported and data complete. They attempted to overcome selection bias by stratification by adjusting analysis for type of diagnosis and by hospital

Other	High risk	Patient outcomes completed 736 for intervention vs. 414 for controls

Chao 2015

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Randomisation was achieved by drawing names out of a hat"
Allocation concealment (selection bias)	High risk	Using alternate allocation to control/intervention
Blinding of participants and personnel (performance bias)	Low risk	Participants not blinded Personnel facilitating sessions not blinded Not likely to impact on results
Blinding of outcome assessment (detection bias)	Low risk	Reviewers of images 2 months after intervention were blinded to group assignment
Incomplete outcome data addressed (attrition bias)	Low risk	No missing data, all trainees assessed
Selective reporting (reporting bias)	Low risk	Primary outcome stated and reported
Other	High risk	Specific to study: all skills were assessed on virtual images generated and the intervention group had just practiced this, the control group had less practice. Some steps to overcome but likely bias in results.

Tolsgaard 2015a

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated allocation ratio 1:1
Allocation concealment (selection bias)	Low risk	Computer generated allocation ratio 1:1, done by a different person to the one responsible for inclusion of participants
Blinding of participants and personnel (performance bias)	High risk	Participants not blinded All had teaching together Primary investigator gave feedback on intervention and analysed data
Blinding of outcome assessment (detection bias)	Low risk	Blinded outcome assessors completing OSAUS
Incomplete outcome data addressed (attrition bias)	Low risk	4 lost to follow up from intervention group. 3 from control group
Selective reporting (reporting bias)	Low risk	Primary outcome stated and reported
Other		

Tolsgaard 2015 b

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation performed in blocks of 2 using a web-based randomizer
Allocation concealment (selection bias)	Low risk	Dyads paired in order of randomisation – no choice over partner

Blinding of participants and personnel (performance bias)	Low risk	Participants not blinded, not possible, not likely to have influenced the result
Blinding of outcome assessment (detection bias)	Low risk	Simulator instructor blinding not possible, used standardised instructions and feedback protocols. Blinded rater for clinical assessment, this was ensured by training either one dyad or two consecutive single participants so 2 were available for the transfer test to maintain blinding. Not clear if raters of simulator assessments had blinding maintained
Incomplete outcome data addressed (attrition bias)	Low risk	6 participants did not complete, it was generally due to lack of patients and unlikely to be a bias
Selective reporting (reporting bias)	Low risk	Primary outcome stated and reported
Other		Use of different methods of rating pre-post and transfer test performances

Moak 2014

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computed generated randomisation
Allocation concealment (selection bias)	Low risk	Computed generated randomisation 1:1
Blinding of participants and personnel (performance bias)	Low risk	Participants not blinded, Blined assessor for scanning technique and image review
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	High risk	Patient assessor not blinded Assessor of live exams blinded (but could easily be broken by patient) Blind assessor later reviewed images
Incomplete outcome data addressed (attrition bias)	Low risk	145 randomised, 134 assessed, spread across groups, one group an assessor not available.
Selective reporting (reporting bias)	Low risk	Identified
Other	High risk	Some participants had prior experience, but this was limited experience and not significantly different between groups The same patient used in the live testing as the teaching scenario Variation in assessor Not all participants had to insert the scan probe (by patient request)

Nickels 2008

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	Unclear	No details
Allocation concealment (selection bias)	Unclear	No details
Blinding of participants and personnel (performance bias)	High risk	Participants not blinded
Blinding of outcome assessment (detection bias) (patient-reported	Low risk	Single assessor blind

outcomes)		
Incomplete outcome data addressed (attrition bias)	Unclear	Insufficient detail in abstract
Selective reporting (reporting bias)	Unclear	Insufficient detail in abstract
Other	High	“The test was exactly the same as the baseline” implies same live model

Appendix 5: Summary of the RCOG basic ultrasound modules

Ultrasound Examination of Early (8- to 12-wk) Pregnancy	Fetal measurement, lie and presentation	Liquor assessment (AFI and Maximum Pool Depth)	Placental Assessment
Counsel patient about procedure Appropriate machine setup Identify bladder and right/left orientation Identify uterus Identify gestational sac and confirm its intrauterine location Measure gestational sac diameter Measure crown-rump length Confirm fetal cardiac activity Interpret ultrasound findings in the context of the clinical setting Communicate results to patients Complete a structured ultrasound report Arrange appropriate referral or follow-up Know when to refer for a transvaginal scan	Counsel patient about procedure Appropriate probe selection and machine setup/operation Confirm fetal heartbeat Establish lie Establish presentation Measure BPD transabdominally Measure FL transabdominally Measure HC transabdominally Measure AC transabdominally Communicate results and uncertainties to patients Complete a structured written ultrasound report Discuss appropriate referral if indicated Ensure images/video is recorded according to local protocol	Counsel patient about procedure Appropriate probe selection and machine setup/operation Confirm fetal heartbeat Assess fetal lie and presentation Measure amniotic fluid volume: maximum vertical pool depth Measure amniotic fluid volume: amniotic fluid index Communicate results to patients Complete a structured written ultrasound report Arrange appropriate referral follow-up	Counsel patient about procedure Appropriate probe selection and machine setup/operation Confirm fetal heartbeat Assess fetal lie and presentation Determine placental position transabdominally Communicate results to patient Communicate results to patient Complete a structured written ultrasound report Arrange appropriate referral follow-up

Appendix 6: Summary of the RCOG intermediate ultrasound modules

Intermediate ultrasound of normal fetal anatomy	Intermediate ultrasound in gynaecology	Intermediate ultrasound of early pregnancy complications
Counsel patient about procedure Safe use of US machine and correct settings Accurate measurements of BPD, HC, AC, FL, TCD and lateral atrial diameter of the cerebral Confirm normal anatomy of spine Confirm normal anatomy of abdomen Confirm normal anatomy of heart and chest Confirm normal anatomy of limbs Perform full anomaly scan Recognise common structural anomalies as stated in fetal anomaly screening programme Locate placental site & cervical canal Assess liquor volume Perform uterine artery Doppler Communicating normal results to parents Working in a multi-disciplinary team Arranging appropriate referral or follow up	Selection and set up of the US machine to ensure safe and optimal use to include appropriate probe selection and orientation Correct and appropriate transabdominal and transvaginal probe care and infection control Accurate measurement of the endometrium in the accepted sagittal plane Assessment of the adnexal regions: accurate identification of the normal ovaries, normal fallopian tube, normal pelvic fluid Accurate measurement of normal and abnormal adnexal structures: mean diameter and volume Recognise and evaluate common endometrial and myometrial abnormalities Recognise and evaluate common ovarian abnormalities Recognise and evaluate complex ovarian masses and refer on appropriately Communicating normal results to patients Communicating appropriate abnormal results to patients Producing written summary and interpretation of results Issue structured written report Arranging appropriate follow up or intervention Working in a multi-disciplinary team	Selection and set up of the US machine to ensure safe and optimal use to include appropriate probe selection and orientation Correct and appropriate transabdominal and transvaginal probe care and infection control Confirmation of the true intrauterine location of a pregnancy Confirmation of fetal viabilty Accurate measurement of MGSD, CRL Complete accurate assessment of the adnexae Accurate documentation of measurements and observations, including chart plotting Ability to identify extra-uterine pregnancy Ability to identify CS scar position Accurately record images Produce written summary and interpretation of results Issue structured written report Communicating normal results to parents Communicating abnormal results to parents Arranging appropriate follow up or intervention Arranging appropriate follow up or intervention

Appendix 7: Questions used in the electronic survey to Deanery Ultrasound Coordinators

SIMULATION IN ULTRASOUND FOR OBSTETRICS AND GYNAECOLOGY
Within your Deanery, do you have a programme for **basic obstetric** ultrasound training (basic early pregnancy and fetal size, liquor and the placenta)?
Yes/ No
If yes which of the following would it involve: (tick all the options that would apply):
Theory / Simulation / Hands on scanning
Comments: *(you may wish to comment on local delivery if regional does not apply, trainee numbers, who delivers the training and organisational issues)*

Within your Deanery, do you have a programme for **intermediate obstetric** ultrasound training (normal fetal anatomy)?
Yes/No
If yes which of the following would it involve: (tick all the options that would apply):
Theory / Simulation / Hands on scanning
Comments: *(you may wish to comment on local delivery if regional does not apply, trainee numbers, who delivers the training and organisational issues)*

Within your Deanery, do you have a programme for **intermediate early pregnancy complications** ultrasound training?
Yes/ No
If yes which of the following would it involve: (tick all the options that would apply):
Theory / Simulation / Hands on scanning
Comments: *(you may wish to comment on local delivery if regional does not apply, trainee numbers, who delivers the training and organisational issues)*

Within your Deanery, do you have a programme for **intermediate gynaecology** ultrasound training?
Yes/ No
If yes which of the following would it involve: (tick all the options that would apply):
Theory / Simulation / Hands on scanning
Comments: *(you may wish to comment on local delivery if regional does not apply, trainee numbers, who delivers the training and organisational issues)*

Which of the following **simulation equipment** for ultrasound do you have in the hospital in your Deanery,?
None
Phantoms (Blue Phantoms)
Phantoms (Ultrasim)
Haptic Simulator [197] If yes, then how many
Other, please list

The following questions are only applicable if simulation equipment is available:

How is your simulation equipment utilised? (please tick all that apply)
As part of a structured training programme

Ad hoc basis by trainees

Is the use of the simulator: (Please tick all that apply)
Supervised
Unsupervised

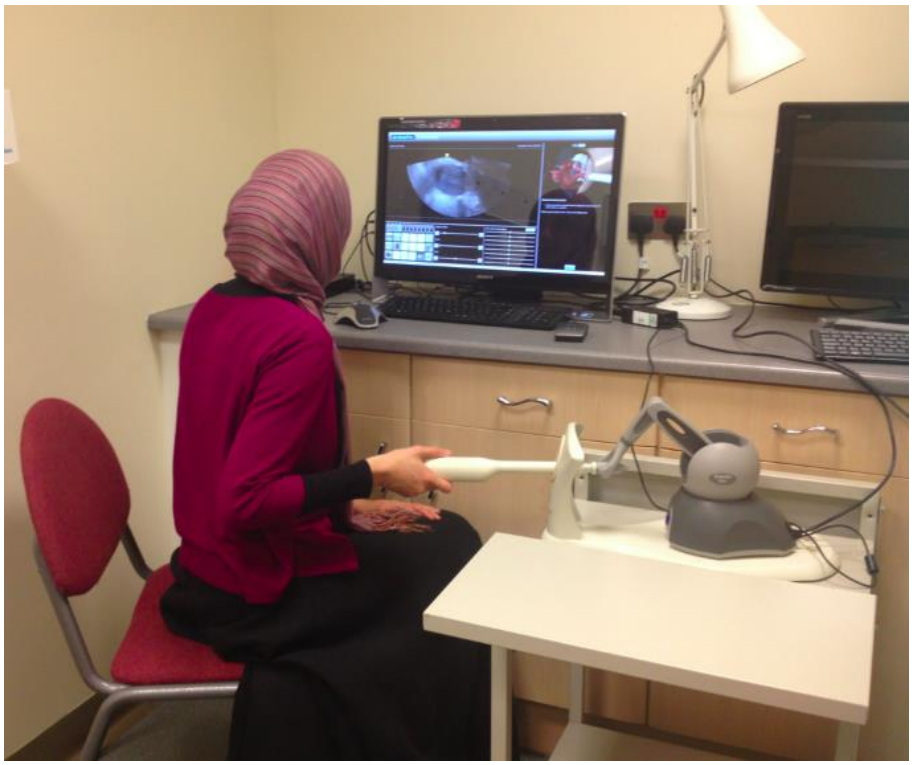
In your opinion how well is the simulation equipment utilised?
Greatly under utilised/Under utilised/Fairly well utilised/Well utilised /Very well utilised

Can you identify any benefits or constraints to delivery of ultrasound training in your Deanery?

Do you think that ultrasound **simulation** has a significant role in acquiring scanning skills in O&G trainees?
Strongly disagree/ Disagree/Neither agree nor disagree/Agree/Strongly agree

Appendix 8: Overview of the Medaphor Transvaginal ScanTrainer®

The set up of the scan trainer during the study period



The haptic device

The learning contents as they appear on the screen

The learning resources; list of tasks and assignments with computer generated pass/fail metrics and feeback

Appendix 9: Learning diary example

Participant code: EMMSC_502

Day

Date

Today was

Only complete this section if you were doing ultrasound scanning in the specified time period

	What activity were you undertaking for the most part of this hour?	Were you doing any ultrasound scanning?	Was you session scheduled or an ad-hoc opportunity?	In what setting were you scanning?	Who was your supervisor?	How many patients did you scan?	What was your level of supervison?
8-9am							
9-10am							
10-11am							
11-12am							
12-1pm							
1-2pm							
2-3pm							
3-4pm							
4-5pm							
5-6pm							
6-7pm							
7-8pm							
8-9pm							
9-10pm							
10-11pm							
11-12pm							
12-1am							
1-2am							
2-3am							
3-4am							
4-5am							
5-6am							
6-7am							
7-8am							

Did you participate in any simulated ultrasound training today?

If yes how many hours

Overall for my intermediate ultrasound training I found this day

Not at all useful — Very useful

Please indicate if this day was atypical in any way and any other comments you wish to make below

Appendix 10: Invite to participate in validity study sent via email

Establishing the face- and construct validity of virtual reality simulation of transvaginal ultrasound for gynaecology and early pregnancy

Dear colleague,

My name is Natalie Woodhead. I am a postgraduate student in the Clinical and Experimental Medicine Department at the University of Birmingham. I am conducting a research study as part of my MD, and I would like to invite you to participate.

There is a growing interest in the potential role for medical simulation in gynaecology and particularly the potential role of virtual reality. Recent advances have seen the development of haptic-feedback simulated virtual reality (VR) transvaginal scan trainers, such as the ScanTrainer® (Medaphor UK) With increasing interest in simulation programs it is recommended prior to integrating use of a simulator into an educational program, its validity first is determined. This study aims to validate virtual reality simulation in assessing transvaginal scanning skills in gynecology and early pregnancy; by establishing the extent of realism of the simulation to the actual task of scanning (face validity) and the degree to which the results of the test one uses reflects the subject tested (construct validity).

We are looking for participants with varying degrees of experience in gynecological ultrasound. If you decide to take part, depending on your prior ultrasound experience, you will be allocated to one of three groups, novice, intermediate or expert. You will then be asked to participate in a one to one session on the ScanTrainer® where you will be instructed to complete a number of tasks on the machine, such as a scan of a normal pelvis and assessment and measurement of a fibroid and your results will be recorded. Once you have completed the tasks you will be asked to complete a questionnaire about the realism and potential use of the simulator. The one to one session will take place at a mutually agreed upon time and place and should last around 1 hour.

The benefit to you of participating in this study is having the opportunity to use the scan simulator and in the future if virtual reality simulation is a proven valid educational tool it could potentially reduce the burden of training on a hard-pressed ultrasound departments. Participation is confidential. Study information will be kept in a secure location at Birmingham Women's Hospital. The results of the study may be published or presented at professional meetings, but your identity will not be revealed.

Taking part in the study is your decision. You do not have to be in this study if you do not want to. You may withdraw at anytime and participation, non-participation or withdrawal will not affect you in any way.

I will be happy to answer any questions you have about the study. You may contact me at (njw240@bham.ac.uk, 0121 371 8754) if you have study related questions or problems.

Thank you for your consideration. If you would like to participate, please complete the attached document and complete the participant questionnaire and contact details. I will call you within the next week to see whether you are willing to participate and to arrange a mutually agreeable time to meet.

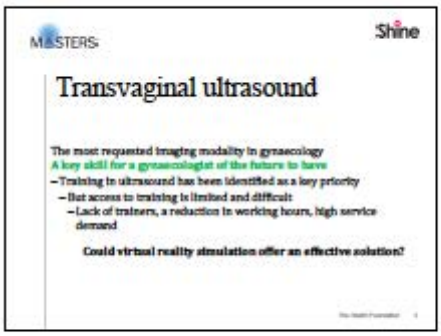
With kind regards,

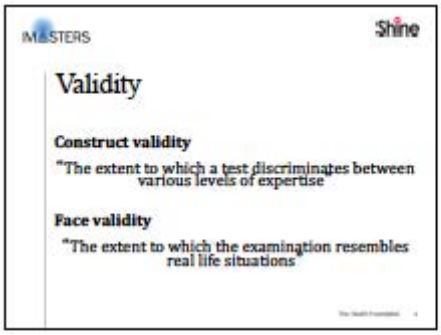
Dr Natalie Woodhead

Research fellow
Academic department, 3rd Floor
Birmingham Women's Hospital NHS Trust
Metchley Park Road
Edgbaston
Birmingham
B15 2TG

Email: nwoodhead@doctors.org.uk
Phone: [redacted]

Appendix 11: Recruitment presentation to students for validity study





Appendix 12: Questionnaire completed by potential participants detailing ultrasound experience

Establishing the face- and construct validity of virtual reality simulation of transvaginal ultrasound for gynaecology and early pregnancy

If you are interested in taking part in the above study if you could complete the following questions and return to njw240@bham.ac.uk I will then contact you within 1-2 weeks

1. Main occupation _____

If applicable:
Number of years post qualification _____

Year of study _____

2. Sex Male/Female
3. Age _____

4. Do you have any qualifications in ultrasound? Yes/No

If yes please give details _____

5. Which description best applies to you please tick one:

Novice: I have no ultrasound experience ☐

Intermediate: I have some ultrasound experience, I have done between 30 – 300 ultrasound scans, most of these have under supervision. ☐

Expert: I have more extensive ultrasound experience I have done over 1000 scans and regularly carry out unsupervised gynaecological scanning, with a regular scan list at least once per week. ☐

6. Do you have any prior experience of simulation training for USS? Yes/No

7. Have you ever before used the Medaphor Scan Trainer®

If yes for approximately how many hours? _____

8. Please give your contact details to be contacted regarding participation

Name:
Telephone:
Email:

Appendix 13: List of tasks and variables

Task 1: Core skills obs, full examination, case 2, anatomy revealed
Gestational sac examined in the sagittal plane, pass/fail (mm)
Fetal heart correctly examined in the sagittal plane, pass/fail (mm)
Fetus correctly examined in the sagittal plane, pass/fail (mm)
CRL, pass/fail, (mm)
CRL measured in the correct position, pass/fail (mm)
CRL direction is correct pass/fail (degrees)
Probe remained in the sagittal plane, pass/fail (degrees)
Examination time pass/fail (minutes and seconds)
Appropriate pressure applied, pass/fail (numerical score)
Total score (0-9)
Task 2: Core skills gynae, perform a full examination, anatomy revealed
Uterus correctly identified in the sagittal plane, pass/fail (mm)
Uterus correctly examined in the coronal plane, pass/fail (mm)
Left ovary correctly examined in the sagittal plane, pass/fail (mm)
Left ovary correctly examined in the coronal plane, pass/fail (mm)
Right ovary correctly examined in the sagittal plane, pass/fail (mm)
Right ovary correctly examined in the coronal plane, pass/fail (mm)
Pouch of Douglas correctly examined in the sagittal plane, pass/fail (mm)
Examination time, pass/fail (minutes and seconds)
Appropriate pressure applied, pass/fail (numerical score)
Total score (0-9)
Task 3: Measurement of the ectopic pregnancy
Appropriate pressure applied pass/fail (Score)
Probe correctly orientated whilst introducing the probe pass/fail (degrees)
Ectopic mass correctly positioned pass/fail (mm)
Probe correctly orientated in the sagittal plane of the ectopic mass pass/fail (degrees)
Ectopic mass correctly magnified pass/fail (number)
Image gain adjusted correctly pass/fail (score)
Image TGC adjusted correctly pass/fail
Appropriate pressure applied pass/fail (score)
Ectopic mass correctly positioned pass/fail (mm)
Probe correctly orientated in the coronal plane of the ectopic mass pass/fail (degrees)
Ectopic mass correctly magnified pass/fail (number)

Appropriate pressure applied pass/fail (score)
Maximum diameter of the ectopic mass in the anterior-posterior direction (A) pass/fail (mm)
Maximum diameter of the ectopic mass in the caudo-cephalic direction (B) pass/fail (mm)
Callipers placed at the maximum diameter of the ectopic mass pass/fail (mm)
Measurements of the ectopic taken perpendicular to each other pass/fail (degrees)
Maximum diameter of the ectopic mass from the patients right to left pass/fail (mm)
Callipers placed at the maximum diameter of the ectopic mass pass/fail (mm)
Time taken to complete
Total (0-18)
Task 4: Assessment and measurement of a fibroid
Appropriate pressure applied pass/fail (score)
Uterus correctly examined in the sagittal plane pass/fail (mm)
Probe remained in the sagittal plane whilst examining the uterus pass/fail (mm)
Uterus examined with sufficient magnification pass/fail (number)
Appropriate pressure applied pass/fail (score)
Uterus correctly examined in the coronal plane pass/fail (mm)
Probe remained in the coronal plane whilst examining the uterus pass/fail (mm)
Uterus examined with sufficient magnification pass/fail (number)
Appropriate pressure applied pass/fail (score)
Fibroid correctly examined in the sagittal plane pass/fail (mm)
Probe remained in the sagittal plane whilst examining the fibroid pass/fail (mm)
Fibroid examined with sufficient magnification pass/fail (number)
Image gain adjusted correctly pass/fail (score)
Image TGC adjusted correctly pass/fail
Fibroid frozen in the centre of the field of view pass/fail (mm)
Fibroid frozen in the sagittal plane pass/fail (mm)
Fibroid sufficiently magnified immediately prior to freezing the image pass/fail (
Fibroid frozen in its maximum diameter pass/fail
Appropriate pressure applied pass/fail (score)
Fibroid correctly examined in the coronal plane pass/fail
Probe remained in the coronal plane whilst examining the fibroid pass/fail
Fibroid examined with sufficient magnification pass/fail
Image gain adjusted correctly pass/fail (score)
Image TGC adjusted correctly pass/fail

Fibroid frozen in the centre of the field of view pass/fail
Fibroid frozen in the coronal plane pass/fail
Fibroid sufficiently magnified immediately prior to freezing the image pass/fail (number)
Fibroid frozen in its maximum diameter pass/fail (mm)
Appropriate pressure applied pass/fail (score)
Appropriate pressure applied withdrawing the probe pass/fail (score)
Diameter of the fibroid in the anterior-posterior direction (A) pass/fail (mm
Diameter of the fibroid in the caudo-cephalic direction (B) pass/fail (mm)
Callipers placed at the maximum diameter of the fibroid pass/fail (mm)
Measurements of the fibroid taken perpendicular to each other pass/fail (degrees)
Diameter of the fibroid from the patients right to left pass/fail (mm)
Callipers placed at the maximum diameter of the fibroid pass/fail (mm)
Time
Total (0-36)
Total core /18
Total advanced /54
Total combined score /72

Appendix 14: Researcher instructions

Researcher instructions

1. Introduction to study
2. Explanation of purpose
3. Explanation of tasks: 4 tasks to complete: 2 core tasks, 2 advanced tasks.

Introduction to ScanTrainer: Start with scan trainer video to allow introduction to simulator functions and controls, up to 5 minutes scanning in unassessed practice mode

2 repetitions of each task:

1st familiarisation
2nd assessed

Explanation of researcher role: I will introduce the task and be able to provide technical assistance but no instructions or feedback, my role is as an observer and I will time each task. You can have a short break between tasks but will remain in the room and I will check all the metrics are recorded.

Core skills obstetrics (TVS-O-OS-001.6)

Patient information: This is a representative scan from a 27 year old patient in her second pregnancy. She is attending a booking clinic.

Task 1: Perform a full examination (case 2, anatomy revealed)

Tutorial objectives: This tutorial consolidates your learning from the previous tutorials by asking you to carry out a full examination.

Assignment overview: In this assignment you will be asked to fully scan the patient, by:

a) Examining the gestation sac in its entirety in the sagittal plane from the right side of the patient to the left.

b) Confirming fetal cardiac activity

c) Demonstrating the fetus in its mid sagittal plane

d) Measuring the CRL of the fetus

2. Core skills gynaecology (TVS-G-CS-001.6)

Patient information: This is a representative scan from a 30 year old patient with absent periods for 3 months and a history of infrequent periods for the last 3 years. She is undergoing an ultrasound examination as part of a series of investigations.

Task 2. Perform a full examination (anatomy revealed)

Tutorial objectives: Please carry out a full examination

Examine these structures in the following order:

a) The uterus in the sagittal plane
b) The uterus in the coronal plane
c) The left ovary in the sagittal and then coronal planes
d) The right ovary in the sagittal and then coronal planes
e) The pouch of Douglas

Advanced skills obstetrics (TVS-O-AS)

Task 3: Measurement of the ectopic pregnancy (TVS-O-AS-002.2)

Tutorial objectives: In this tutorial you will learn to use ultrasound to examine an ectopic pregnancy

Assignment overview: In this assignment you learn how to measure an ectopic pregnancy

Patient information: This case contains a representative scan from a patient presenting with lower abdominal pain, vaginal bleeding and no period for 6weeks. A pregnancy test is positive. On examination, the pelvis is tender.

The participant will be taken through the following:

a) Introduce the probe
b) Optimise an image of the ectopic mass in its sagittal plane
c) Optimise an image of the ectopic mass in its coronal plane
d) Measure the ectopic mass in its maximum diameters in the sagittal plane
e) Measure the ectopic mass in its maximum diameters in the coronal plane

4. Advanced skills gynaecology (TVS-G-AS-001.2)

Task 4: Assessment and measurement of the fibroid (1.2)

Tutorial objectives: The objective of this tutorial is to familiarise you with assessing and measuring the fibroid

Assignment overview: In this assignment you will examine and measure the fibroid

Patient information: This case contains a representative scan from a patient presenting with heavy periods and pelvic pain. On examination the uterus was found to be enlarged.

The participant will be taken through the following:

a) Introduce the probe

b) Image the uterus from the patients right to left side

c) Image the uterus from the cervix to the fundus

d) Optimise the image of the fibroid in the sagittal plane

e) Optimise the image of the fibroid in the coronal plane

f) Measure the fibroid in the sagittal plane

g) Measure the fibroid in the coronal plane

Appendix 15: Questionnaire used for face and content validity

Validity of the Scan Trainer - Tell us what you think

To what extent do you agree with the following statements?	Strongly agree	Agree	Undecided	Disagree	Strongly disagree
Scan trainer does not resemble an ultrasound machine					
Images are of a good quality					
Scan trainer provides an unrealistic simulation of a real life scan					
The ultrasound probe has a realistic range of motion					
Images don't move as per a real machine					
The audio force feedback is useful (when the machine yelps or screams)					
The feel of the ultrasound probe is unrealistic					
The scan trainer is useful for training in gynaecological ultrasound					
The visual force feedback gauge is not useful					
The feedback provided is helpful					
The feedback did not improve skills for subsequent scanning					
Scan trainer is useful in teaching a systematic approach to scanning					
Scan trainer is not useful for teaching hand eye-co-ordination					
Scan trainer is effective for practicing basic skills					
Scan trainer reduces the workload of those delivering scan training					
Scan trainer is not useful for teaching 3D orientation					

1

Validity of the Scan Trainer - Tell us what you think

To what extent do you agree with the following statements?	Strongly agree	Agree	Undecided	Disagree	Strongly disagree
Scan trainer would not be helpful to teach pathology recognition					
The Scan trainer offers a user friendly learning environment					
I did not enjoy using the scan trainer					
The scan trainer is fun to use					
It is not in women's best interests for trainees to first practice in a simulated setting					
The increment of skills during training should be monitored					
It is not important to practice scans on virtual models					
Trainees should first practice in a simulated setting before scanning women					
Scan trainer could not be used to determine those with aptitude for scanning					
Scan trainer could be used for assessment					
Scan trainer could not be used for revalidation of skills					
Scan trainer is suitable for evaluation during training					
Scan trainer will not shorten learning curves in the scan room					
Scan trainer can accelerate a person's learning					
Scan trainer cannot reduce the rate of misdiagnosis					
Scan trainer gives starting operators a sense of confidence					

Validity of the Scan Trainer - Tell us what you think

To what extent do you agree with the following statements?	Strongly agree	Agree	Undecided	Disagree	Strongly disagree
Use of Scan trainer will increase the cost of ultrasound scan training					

Appendix 16: Table: Post-hoc analysis of validity of simulator test and its metrics

I have combined scores for 1st (familiarisation) and 2nd (assessment) run, therefore giving 30 attempts in total per group.
As per the methods outlined by Madsen et al. I have excluded:

- Metrics that do not show a difference between the two groups
- Metrics where <50% of experts passed
- Any metrics that more novices than experts passed

Metric	Expert Number of passes	Novice Number of passes	X^2 + Fisher exact P value
Gestational sac examined in the sagittal plane, pass/fail (mm)	30	27	0.237
Fetal heart correctly examined in the sagittal plane, pass/fail (mm)	30	30	1
Fetus correctly examined in the sagittal plane, pass/fail (mm)	30	30	1
CRL, pass/fail, (mm)	22	9	0.002
CRL measured in the correct position, pass/fail (mm)	28	23	0.145
CRL direction is correct pass/fail (degrees)	29	20	0.006
Probe remained in the sagittal plane, pass/fail (degrees)	29	25	0.195
Examination time pass/fail (minutes and seconds)	29	25	0.195
Appropriate pressure applied, pass/fail (numerical score)	30	27	0.237
Uterus correctly identified in the sagittal plane, pass/fail (mm)	23	20	0.567
Uterus correctly examined in the coronal plane, pass/fail (mm)	21	12	0.037
Left ovary correctly examined in the sagittal plane, pass/fail (mm)	24	12	0.003
Left ovary correctly examined in the coronal plane, pass/fail (mm)	25	14	0.006
Right ovary correctly examined in the sagittal plane, pass/fail (mm)	8	7	1.000

Right ovary correctly examined in the coronal plane, pass/fail (mm)	25	17	0.047
Pouch of Douglas correctly examined in the sagittal plane, pass/fail (mm)	26	29	0.353
Examination time, pass/fail (minutes and seconds)	26	15	0.454
Appropriate pressure applied, pass/fail (numerical score)	30	26	0.112
Appropriate pressure applied pass/fail (Score)	30	25	0.052
Probe correctly orientated whilst introducing the probe pass/fail (degrees)	30	30	1
Ectopic mass correctly positioned pass/fail (mm)	28	23	0.145
Probe correctly orientated in the sagittal plane of the ectopic mass pass/fail (degrees)	26	15	0.005
Ectopic mass correctly magnified pass/fail (number)	11	9	0.785
Image gain adjusted correctly pass/fail (score)	28	30	0.492
Image TGC adjusted correctly pass/fail	30	30	1
Appropriate pressure applied pass/fail (score)	25	29	0.195
Ectopic mass correctly positioned pass/fail (mm)	28	13	<0.0005
Probe correctly orientated in the coronal plane of the ectopic mass pass/fail (degrees)	27	12	<0.0005
Ectopic mass correctly magnified pass/fail (number)	12	11	1.000
Appropriate pressure applied pass/fail (score)	24	28	0.254
Maximum diameter of the ectopic mass in the anterior-posterior direction (A) pass/fail (mm)	9	11	0.785
Maximum diameter of the ectopic mass in the caudo-cephalic direction (B) pass/fail (mm)	4	2	0.671
Callipers placed at the maximum diameter of the ectopic mass pass/fail (mm)	8	1	0.026
Measurements of the ectopic taken perpendicular to each other pass/fail (degrees)	26	25	1.000
Maximum diameter of the ectopic mass from the patients right to left	15	9	0.187

pass/fail (mm)			
Callipers placed at the maximum diameter of the ectopic mass pass/fail (mm)	14	8	0.180
Appropriate pressure applied pass/fail (score)	29	25	0.195
Uterus correctly examined in the sagittal plane pass/fail (mm)	19	6	0.001
Probe remained in the sagittal plane whilst examining the uterus pass/fail (mm)	26	28	0.671
Uterus examined with sufficient magnification pass/fail (number)	9	19	0.019
Appropriate pressure applied pass/fail (score)	27	26	1.000
Uterus correctly examined in the coronal plane pass/fail (mm)	22	7	<0.0005
Probe remained in the coronal plane whilst examining the uterus pass/fail (mm)	25	14	0.006
Uterus examined with sufficient magnification pass/fail (number)	8	18	0.018
Appropriate pressure applied pass/fail (score)	28	29	1.000
Fibroid correctly examined in the sagittal plane pass/fail (mm)	16	3	0.001
Probe remained in the sagittal plane whilst examining the fibroid pass/fail (mm)	22	30	0.005
Fibroid examined with sufficient magnification pass/fail (number)	9	15	0.187
Image gain adjusted correctly pass/fail (score)	30	30	1
Image TGC adjusted correctly pass/fail	29	29	1
Fibroid frozen in the centre of the field of view pass/fail (mm)	25	13	0.182
Fibroid frozen in the sagittal plane pass/fail (mm)	27	29	0.612
Fibroid sufficiently magnified immediately prior to freezing the image pass/fail	8	15	0.110
Fibroid frozen in its maximum diameter pass/fail	23	9	0.001
Appropriate pressure applied pass/fail (score)	27	26	1.000
Fibroid correctly examined in the coronal plane pass/fail	18	5	0.001
Probe remained in the coronal plane whilst examining the fibroid pass/fail	22	17	0.279

Fibroid examined with sufficient magnification pass/fail	10	14	0.430
Image gain adjusted correctly pass/fail (score)	30	30	1
Image TGC adjusted correctly pass/fail	30	30	1
Fibroid frozen in the centre of the field of view pass/fail	19	17	0.792
Fibroid frozen in the coronal plane pass/fail	26	16	0.010
Fibroid sufficiently magnified immediately prior to freezing the image pass/fail (number)	10	15	0.295
Fibroid frozen in its maximum diameter pass/fail (mm)	22	10	0.004
Appropriate pressure applied pass/fail (score)	26	27	1.000
Appropriate pressure applied withdrawing the probe pass/fail (score)	26	27	1.000
Diameter of the fibroid in the anterior-posterior direction (A) pass/fail (mm	22	5	<0.0005
Diameter of the fibroid in the caudo-cephalic direction (B) pass/fail (mm)	14	3	0.003
Callipers placed at the maximum diameter of the fibroid pass/fail (mm)	25	8	<0.0005
Measurements of the fibroid taken perpendicular to each other pass/fail (degrees)	27	23	0.299
Diameter of the fibroid from the patients right to left pass/fail (mm)	21	6	<0.0005
Callipers placed at the maximum diameter of the fibroid pass/fail (mm)	30	14	<0.0005

Results:

Of the 72 metrics examined in total 24 showed a significant difference between groups (33.3%)
A further 3 of these 24 metrics were excluded as <50% of experts passed, for one of these metrics more novices than experts passed.

This left 21 of 72 metrics (29.2%) that demonstrated construct validity

Appendix 17: Subjective validity assessed on Transvaginal ScanTrainer (Medaphor, UK)

Author (Year)	Type of validity Definition	Simulator	Research setting	Groups, definition	Intervention	Modules	Questionnaire	Conclusion
Alsalamah (2017)	**Face** = extent of simulators realism and appropriateness when compared to the actual task **Content** = extent to which a simulator content is representative of the knowledge or skills that have to be learnt in the real environment	Medaphor TV ScanTrainer	International conference, ESGE 'simulation island' Scotland	**Total n=36** Non- expert: (n= 25) Limited experience of <2 years, occasional or limited scanning sessions/considered themselves not independent Expert: (n= 11) Subject with >2 years US experience/daily scanning/ considered themselves independent	Brief introduction 10-15 minutes exploring specific tasks within tutorials	1. Core skills gynaecology 2. Core skills early pregnancy 3. Advanced skills case studies	Completed after the intervention Given to both groups Face validity statements Content validity statements General opinions 14 statements VAS scale 0-10	Face validity: Median scores of statements ranged 7.5-9.0 and were slightly higher in experts in relation to realism of the scan and controls, slightly lower in relevance to teach scanning and simulation of movements required to perform a scan. No significant differences between the groups opinions Content validity: Median scores 8.4-9.0
Al-Memar (2017)	**Face** = How closely the task using the simulator resembles the task in reality **Content** = ability of the simulator to teach the tasks that are relevant to the trainee	Medaphor TV ScanTrainer	Department of Gynaecology, KU Leuven, Belgium	**Total n=24** Expert (n= 8) > 500 scans Intermediate (n=8) 50 – 500 scans Novice (n=8) <50 scans	Standardised instructions Complete the modules and the assessment	1. Basic early pregnancy 2. Basic gynaecology	Questionnaire of opinion of role of US simulation in training before using simulator and after would it be useful in training given to both groups Only experts assessed face and construct validity 19 statements, 5point likert scale	Face validity: Simulator rated satisfactory or very satisfactory for realism of gynae (88%) and early pregnancy (75%) module. 50% were not satisfied with realistic instrument handling. 75% satisfied with how the images looked and how the actions felt in comparison to a real patient environment overall. 63% found the simulator of use in acquiring hand-eye co-ordination skills, 88% in ambidexterity skills Content validity: 88% of experts rated the usefulness of the content in learning ambidexterity

								skills. 100% rated the usefulness of the content in learning fundamental ultrasound technical skills and 75% rated the simulator favourably for its role in ultrasound training.
Preshaw 2012 **Abstract**	**Face** no explicit definition given	Medaphor TV ScanTrainer	Department of O&G Southmead Hospital, Bristol, UK	**Total n=25** Trainees = 17 Independent ultrasound operators = 8	Orientated to the simulator and completed modules	7 modules, not specified which	Anonymous online survey to all after the intervention Statements on a 5-point likert scale	84% agreed the simulator provides a realistic simulation of an actual TVS. Functionality ranked favourably- easy to use (80%), haptic force feedback helpful (64%), fun to use (88%). Simulator felt to be useful in training (96%), assessments (68%), revalidation (64%) as well as developing ultrasonographic skills including hand-eye co-ordination (96%), 3D orientation (84%) and pathology recognition (92%). 88% felt it would be in women's best interests to practice in a simulated setting prior to scanning patient.

Appendix 18 Table: Objective validity assessed on Transvaginal ScanTrainer (Medaphor, UK)

Author (Year)	Type of validity Definition	Simulator	Method: training programme	Groups number, definition	Parameters	Results
Al-Memar (2017)	Construct= Ability of the simulator to differentiate between a novice and expert	Medaphor TV ScanTrainer	Two modules Basic gynae Basic early pregnancy Assessments on simulator	**Expert (n= 8)** > 500 scans **Intermediate (n=8)** 50 – 500 scans **Novice (n=8)** <50 scans	Overall time taken to complete the assessments Score of each section of the assessment Overall score made up of the total of scores for each section Number of times the force feedback alarmed	Significant difference between the non-expert and novice groups compared with expert groups in total module score for basic gynaecology and early pregnancy. Significant difference in score for scanning adenexa and fetus but not uterus. Significant difference in the number of screams No significant differences between groups in time taken
Madsen (2014)	Construct validity = Assess validity of simulator metrics for discriminating between different levels of competency Step 2 of 5-step approach to validity testing	Medaphor TV ScanTrainer	15 min practice then simulator test twice in sequence (with a short 10min break in between Valid metrics from basic gynaecological Ovarian cyst EP gestation sac CRL US diagnosis of ectopic Cervical length	**Experts (n=12)** Experienced practioners n=8 gynaecologists using US daily n=4 fetal med consultants **Novices (n=16)** No previous gynaecological scanning scan Final year medical students	Significant difference between the groups for each metric (153 metrics examined METRICS EXCLUDED IF <50% EXPERTS PASSED, or more novices than experts passed Valid metrics made simulator test (48metrics) Comparison of median sum scores Overall time taken to complete the first 2 simulator iterations	Significant differences between the two groups in 50 of 153 metrics, further 2 excluded as either 50% of experts didn't pass it or more novices than experts passed. 48 of 153 metrics (31.4%) showed construct validity. Significant difference in the time taken between the novice and the expert group to complete the simulator test Novices 81.5 minutes (range 51-105) compared to experts 53.0minutes (range 49-81) p=<0.001
Alsalamah (2016) PhD thesis	Construct = degree to which a test measures what it claims to measure, differention between experts and novices	Medaphor TV ScanTrainer	3 simulator assignments Gynae module 1 and 2 Early pregnancy module 7 skills in each module (21skills) Brief intro Scanning for no more than 30mins, little detail	**Experts (n=8)** independent practice >2years experience, daily scanning **Intermediate (n=10)** Competent but not yet an expert with 6months to 2 years experience **Novice (n=12)** Trainee under supervision with limited experience <6months	Differences in scores achieved by the 3 groups • scored on a simulator • scored by an observer using OSATs To set pass/fail performance standards	Median group scores Rated by simulator Gynae 1 N=2, I=3, E=6 Gynae 2 N=3.5, I=5.5, E=3.5 Early pregnancy N=3.5, I=5.5, E=3.5 Significant differences between expert and novice in all 3 modules, and between novice and intermediate but not between intermediate and expert

Birmingham Women's NHS
NHS Foundation Trust

3rd Floor, Academic
Department
Mindelsohn Way
Edgbaston
Birmingham
B15 2TG

Switchboard: 0121 472 1377
Direct Line: 0121 627 2775

Appendix 19: Participant information sheet

Participant information sheet

Research project: Exploring women's perceptions of transvaginal ultrasound scan training

Researchers: Dr Natalie Woodhead
University of Birmingham, Birmingham Women's Hospital
Dr Lorna Porcellato
Liverpool John Moores University

Transvaginal ultrasound scanning is when the ultrasound probe is inserted into the vagina to get ultrasound images of the womb and ovaries. You are being invited to take part in a research study that will explore women's experiences of having a transvaginal scan, what women think about their examination, and what they think of the examination being used as a means of training i.e. having somebody who is leaning how to scan actually doing the scan.

Before you decide whether or not you would like to take part, it is important that you understand why the research is being done and what it will involve. One of our team will go through the information sheet with you and answer any questions you have. We'd suggest this should take about 10 minutes. Please ask if anything that is not clear or if you would like more information.

What is the purpose of the study?

Transvaginal scanning is a very useful and important investigation, which many women undergo. In order for a hospital to be able to carry out the number of scans they do they need people appropriately trained in the skill of transvaginal scans. Senior doctors or ultra-sonographers through supervised scanning on patients provide currently practical training in transvaginal scanning. However transvaginal scanning is an intimate examination and no research has ever been done to determine how women feel about the experience of having a transvaginal examination particularly done by a trainee, this is what the study

3rd Floor, Academic
Department
Mindelsohn Way
Edgbaston
Birmingham
B15 2TG

Switchboard: 0121 472 1377
Direct Line: 0121 627 2775

aims to explore. The study will form the basis for future work on alternative and innovative means of training such as using simulated models rather than patients.

Do I have to take part?

No. It is up to you to decide whether or not to take part in this research. After reading this information and discussing it with the researcher if you think you would like to take part we will go through a consent form with you to allow us to contact you again to invite you to an interview. You are still free to withdraw at any time and without giving a reason. A decision to withdraw or not to take part will not affect you in any way. In particular, it will not affect the way that you are treated by any of the people involved in your current or future care in this hospital.

What will happen if I take part?

If you agree to take part you will be asked to participate in an individual interview. You will be asked to give your opinion on various issues about transvaginal ultrasound experience and training. There are no right or wrong answers; we are interested in hearing everyone's views and opinions. The interview will last up to 1 hour and be over the telephone or held at the Birmingham Women's Hospital. The interview will be audio recorded and later transcribed into a written document, which will be anonymous so you will not be identifiable from it.

Expenses and payments

You will be offered a £10 Amazon voucher as a token to reimburse you for your time.

Are there any risks or benefits from taking part?

There are no risks involved. The research will be used to help inform future ultrasound training programmes.

Will my taking part be kept confidential?

Birmingham Women's NHS
NHS Foundation Trust

3rd Floor, Academic
Department
Mindelsohn Way
Edgbaston
Birmingham
B15 2TG

Switchboard: 0121 472 1377
Direct Line: 0121 627 2775

Yes. Any personal information collected during the research will be anonymised and remain confidential. It is expected that the results of this study will be published but your individual details will not be mentioned.

What will happen to the results of the study?

It is anticipated that the results of this study will be published in scientific journals and presented at meetings and conferences. If interested, a summary of the study will be sent to you upon completion of the research.

Who is organising and funding the research?

This project has been funded by The Health Foundation, Shine award. It is being organised by the project team at Birmingham Womens Hospital.

Contact details of the researcher

If you have any questions about the research or would like further details please feel free to contact me

Dr Natalie Woodhead

Tel: 0121 371 8754 / [redacted]

Email: [redacted]

PALS (Patient advice and liaison service)

Tel: 0121 627 2747

An answerphone will take out of hours' messages. All calls should be answered within two working days. If your call is urgent, please contact the mobile phone 07720 703 511 The Patient Advice and Information Centre is sited in the front of the house area of the hospital. Or you can leave a message in one of our suggestion boxes, sited in all main patient areas. An online feedback form is also available.

Thank you for giving consideration to participating in this research projec

Birmingham Women's NHS
NHS Foundation Trust

3rd Floor, Academic
Department
Mindelsohn Way
Edgbaston
Birmingham
B15 2TG

Switchboard: 0121 4721377
Direct Line: 0121 627 2775

Appendix 20: Consent to be recruite

Recruitment consent form

Title of project: Exploring women's perceptions of transvaginal ultrasound scan training

Names of researchers: Dr Natalie Woodhead
Dr Lorna Porcellato

Please initial

1. I confirm that I have read and understand the information provided for the above research study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving a reason and that this will not affect me in any way ☐
3. I understand that any personal information collected during the research study will be anonymised and remain confidential ☐
4. I give permission for the researcher to contact me after discharge from the clinic to take part in an interview at a time convenient to me ☐
5. I agree to take part in the above research study should I be contacted for an interview ☐

Name of Participant	Date	Signature
Name of person taking consent	Date	Signature

When completed 1 copy for participant, 1 copy for researcher, 1 copy for medical records

Birmingham Women's NHS
NHS Foundation Trust

3rd Floor, Academic
Department
Mindelsohn Way
Edgbaston
Birmingham
B15 2TG

Switchboard: 0121 4721377
Direct Line: 0121 627 2775

Appendix 21 : Letter of invite to participate

Dear Name,

Exploring women's perceptions of transvaginal ultrasound scan training Research Study

Thank you for expressing an interest in this research study. As you are aware, the study will explore women's experiences of having a transvaginal scan, what women think about their examination, and what they think of the examination being used as a means of training i.e. having somebody who is leaning how to scan actually doing the scan. The study will form the basis for future work on alternative and innovative means of training such as using simulated models rather than patients.

I would like to invite you to take part in an interview to talk broadly about your experience and perceptions of having a transvaginal scan. The discussion is expected to last approximately 1 hour. The purpose of the discussion is to determine how women feel about the experience of having a transvaginal examination particularly done by a trainee, this is what the focus group will aim to explore, we will not explore in any detail individual participants clinical details.

Enclosed with this invitation letter is copy of the **Participant Information Sheet** you were given in the hospital**.** Please read it again to help you decide if you would like to be part of the research. You will be contacted in one weeks' time and asked if you would like to be interviewed for the study. If you do decide to participate, a date and details of the venue will be arranged. Anyone who agrees to be interviewed will be asked to sign a **Consent Form** prior to the interview.

We understand that giving feedback to people that have given up their time to talk to us is important, and so we commit to provide all participants, who wish to receive one, a summary of the key research findings.

If you have any queries about the research, please contact me on: 0121 371 8754 or email me on: [redacted]. Your interest in this research project is much appreciated.

Regards,

Dr Natalie Woodhead
Clinical Research Fellow

Appendix 22: Interview guide

Introductions/overview of the process

Introduction of interviewer by name
Thank you for agreeing to take part in this interview today.
We are interested in speaking to you because you have had a transvaginal (internal) scan and we are interested in hearing your views and experience, it will help to help to improve training for future practitioners.
The interview should last about 1 hour.

We will be digitally recording it, if that is ok and I want to reassure you that what you say will be confidential and all identifiable data will be removed so you will all remain anonymous.

Before interview begins, ask participants to complete demographic questionnaire and ensure consent signed (and if telephone posted)

Any questions before we start?

Knowledge of TVS

- **As we said before, you have had a transvaginal scan (TVS). Lets begin by talking about what a TVS is...**

    - When you were told you would need a TVS at your appointment, did you know what it was? Any previous knowledge or experience? Expectations that you might need one when you went to your appointment? How did you feel about that?
    - How might you describe what a TVS is to someone who has never had one before?
    - Any ideas on when a TVS may be needed?
    - How did you feel about needing to have one done?

- **Experience of TVS**

    Let's talk about the TVS itself – the process you went through...
    (Pictures of the scan room set up and the ultrasound machine will be available as aide memoires photo A)

    - Can you remember who explained the need to have a TVS to you? How did they explain/describe it? Prompt: language used, did you understand, did you have any concerns?
    - Describe the process of having the scan – what happened? Prompt: getting undressed, privacy, did you have anyone with you?
    - Was there another person in the room acting as a chaperone? How do you feel about the use of a chaperone? Prompt: does it matter if

they are male or female? What do you see as the role of a chaperone?
- How did you feel during the scan?
- What happened after the scan?

Think about the person scanning you:

- Who was doing the scan? (consultant? Trainee? Sonographer?)
- Would you mind who did the scan? Prompt: gender, age, ethnicity?
- Would there be any circumstances when you would say no to having a TVS?
- What sorts of things did they do to make you feel comfortable? (Probe: talk to you whilst scanning? What would you want them to talk about, small talk or talking about the findings or not at al? show you the screen? Would you want to see the screen?)

Reflection

- How would you describe your transvaginal scan experience?
- Any impact (positive and/or negative) on you in anyway (ask for examples)
- Once you'd had your appointment if a friend asked you how it was was what would you say to them?
- Would you be willing to undergo the examination again?
- What things helped make the whole experience easier?
- What things could have made the experience better or would you like to see done differently next time?

In order to do TVS, clinicians have to be trained. They develop their skills by practicing on female patients under the supervision of a trainer who is a qualified scanner

- Did you knowingly have a trainee scan you? If yes,
  - How was this person introduced? What happened? What was your experience of them doing the scan?

- For those of you who have not been scanned by a trainee, would you mind having your scan done by a trainee being supervised? (why? / why not?)
- Under what conditions would you be happy to be scanned by a trainee? (Prompt: does gender of the trainee matter, the age, their experience)
- When would you say no to having a trainee scan you?

Way forward

As you know, the aim of this research is to form the basis for future work on alternative and innovative means of training such as using simulated models rather than patients. Introduce the Medaphor Scan Trainer (a photograph will be available as an aide memoire – photo B) Briefly explain what it does and how it is used and what benefit it gives the trainee. The scan trainer is a virtual reality scanner – some people would liken it to a computer game or simulator. The trainee uses the scan trainer away from a clinical setting, they move the probe around and see images on the screen, they are set a series or tasks and learn to scan as they use the machine. The machine tries to replicate real life – the images are real images taken from women and the machine if you push to hard will give a ouch or a scream. The idea is trainees can practice and upskill themselves on this before scanning on real women.

- What are your thoughts on models like this?
- Any awareness that these were available to teach people?
- Knowing they are available, would you expect trainees to develop initial TVS skills first and then practice on real patients or not? Should it be compulsory or mandatory? Can it replace hands on training?
- What do think might be the benefits / limitations for 1) the trainees 2) the female patients?
- Given that they are expensive, should the use of such models be supported and why?

Any thing else you would like to add

Birmingham Women's NHS
NHS Foundation Trust

3rd Floor, Academic
Department
Mindelsohn Way
Edgbaston
Birmingham
B15 2TG

Switchboard: 0121 472 1377
Direct Line: 0121 627 2775

Appendix 23 Consent form

Participant interview consent form

Research project: Exploring women's perceptions of transvaginal ultrasound scan training

Researchers: Dr Natalie Woodhead
University of Birmingham, Birmingham Women's Hospital
Dr Lorna Porcellato
Liverpool John Moores University

1. I confirm that I have received the participant information sheet; version 2.0 received *(17/10/13),* and I understand the information provided. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving a reason and that this will not affect me in any way ☐

3. I understand that any personal information collected during the study will be anonymised and remain confidential ☐

4. I give permission for the use of audio taping during the interview ☐

5. I give permission for anonymised verbatim quotes from my interview to be used in presentations and publications about the research ☐

6. I give permission for anonymised audio clips of my interview to be used in presentations about the research ☐

7. I agree to take part in the above study ☐

Birmingham Women's NHS
NHS Foundation Trust

3rd Floor, Academic
Department
Mindelsohn Way
Edgbaston
Birmingham
B15 2TG

Switchboard: 0121 472 1377
Direct Line: 0121 627 2775

Name of Participant	Date	Signature
Name of Researcher	Date	Signature

Appendix 24 Photographs to reference during interviews

Photo A

Photo B

Appendix 25: Demographic questionnaire for interview participant

Participant code:

Demographic questionnaire for interview participants

If you could take the time to answer the questionnaire whist having refreshments, please feel free to ask the researchers to clarify any questions you are unsure of

What is your date of birth?

Date/Month/year

____/_____/_____

What is your marital status?

Never married and never registered a same-sex civil partnership	
Married	
Separated, but still legally married	
Divorced	
Widowed	
In a registered same-sex civil partnership	
Separated, but still legally in a same-sex civil partnership	
Formerly in a same-sex civil partnership which is now legally dissolved	
Surviving partner from a same-sex civil partnership	

What is your ethnic group?

A. White	
English / Welsh / Scottish / Northern Irish / British	
Irish	
Gypsy or Irish Traveller	
Any other White background, write in	

B. Mixed / multiple ethnic groups	
White and Black Caribbean	
White and Black African	
White and Asian	
Any other Mixed / multiple ethnic background, write in	

C. Asian / Asian British	
Indian	
Pakistani	
Bangladeshi	
Chinese	
Any other Asian background, write in	

D. Black / African / Caribbean / Black British	
African	
Caribbean	
Any other Black / African / Caribbean background, write in	

E. Other ethnic group	
Arab	
Any other ethnic group, write in	

Which of these qualifications do you have?

1 - 4 O levels / CSEs / GCSEs (any grades), Entry Level, Foundation diploma	
NVQ Level 1, Foundation GNVQ, Basic Skills	
5+ O levels (passes) / CSEs (grade 1) / GCSEs (grades A*- C), School Certificate, 1 A level / 2 - 3 AS levels / VCEs, Higher Diploma	
NVQ Level 2, Intermediate GNVQ, City and Guilds Craft, BTEC First / General Diploma, RSA Diploma	
Apprenticeship	
2+ A levels / VCEs, 4+ AS levels, Higher School Certificate, Progression / Advanced Diploma	
NVQ Level 3, Advanced GNVQ, City and Guilds Advanced Craft, ONC,OND, BTEC National, RSA Advanced Diploma	
Degree (for example BA, BSc), Higher degree (for example MA, PhD, PGCE)	
NVQ Level 4 - 5, HNC, HND, RSA	

Higher Diploma, BTEC Higher Level	
Professional qualifications (for example teaching, nursing, accountancy	
Other vocational / work-related qualifications	
Foreign qualifications	
No qualifications	

Occupation:

_____

Have you had a transvaginal scan prior to the one you had in clinic recently?

Yes	
No	

How many scans would you estimate you have had?

1	
2	
3	
4	
5	
6 or more	

Are you currently pregnant?

Yes	
No	

Prior pregnancies:

Number of total pregnancies	
Number of pregnancies carried beyond 24 weeks	
Number of miscarriages	
Number of tubal pregnancies (ectopic pregnancies)	
Number of voluntary terminations	

Number of living children	

Cervical screening (smear test) history:

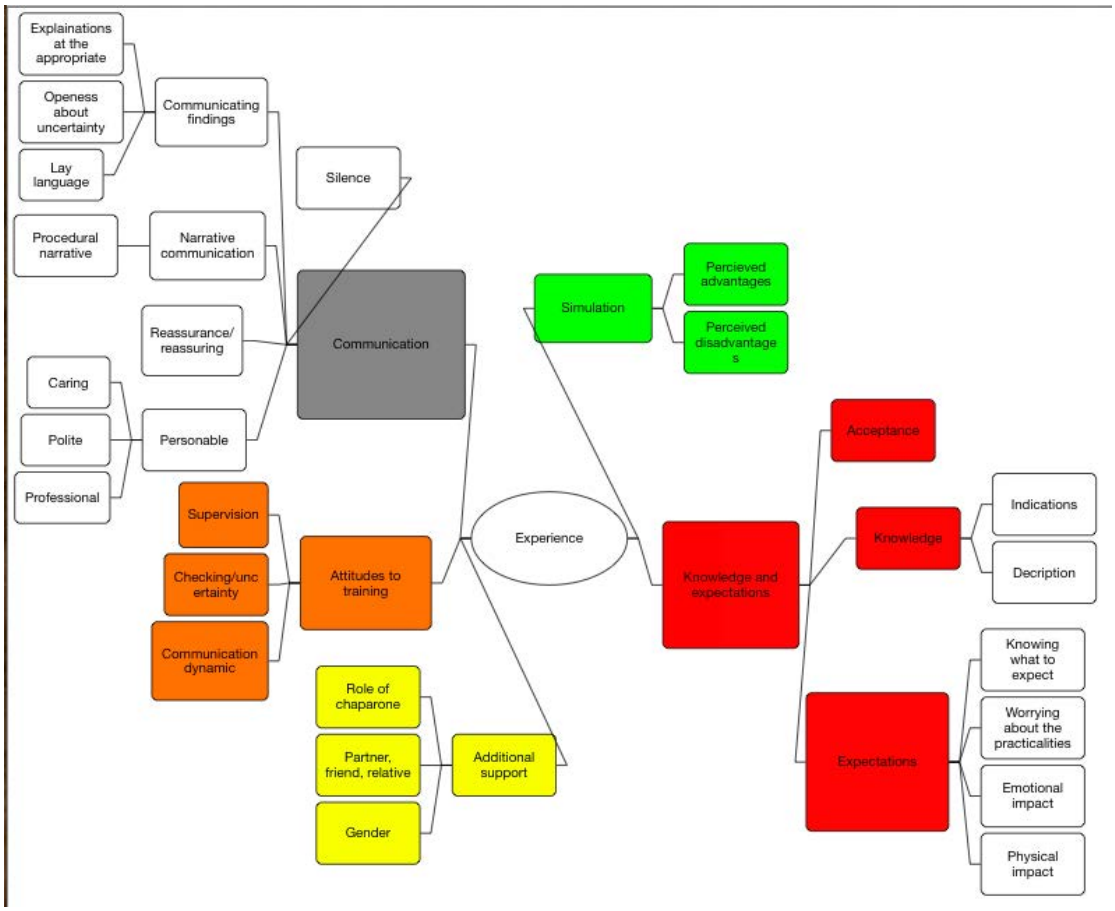
Are you aged between 25-60?

Yes	
No	

Have you attended for routine cervical smears in the past?

Yes	
No	

Appendix 26: Thematic mind map of emerging themes



REFERENCES

1. Stewart RA, Thistlethwaite J. Routine pelvic examination for asymptomatic women. Exploring the evidence. Australian family physician. 2006;35(11):873-7.

2. Campbell S. A Short History of Sonography in Obstetrics and Gynaecology. Facts, Views & Vision in ObGyn. 2013;5(3):213-29.

3. Bourne T, Hamberger L, Hahlin M, Granberg S. Ultrasound in gynecology: endometrium. International Journal of Gynaecology and Obstetrics. 1997(56):115-27.

4. Brenner DJ, Hall EJ. Computed tomography— an increasing source of radiation exposure. N Engl J Med 2007(357):2277-84.

5. Amso NN, Griffiths A. The role and applications of ultrasound in ambulatory gynaecology. **Best** Pract Res Clin Obstert Gynaecol 2005;19(5):693 - 711.

6. Royal College of Obstetricians and Gynaecologists. Ultrasound from conception to 10+0 weeks of gestation. Scientific Impact Paper No. 49. London 2015 [Accessed 2018 Feb 27]. Available from: https://www.rcog.org.uk/en/guidelines-research-services/guidelines/sip49/.

7. Kirk E, Papageorghiou AT, Condous G, Tan L, Bora S, Bourne T. The diagnostic effectiveness of an initial transvaginal scan in detecting ectopic pregnancy. Hum Reprod 2007;22(11):2824-8.

8. Knight M, Bunch K, Tuffnell D, Jayakody H, Shakespeare J, Kotnis R, et al. Saving Lives, Improving Mothers' Care - Lessons learned to inform maternity care from the UK and Ireland Confidential Enquiries into Maternal Deaths and Morbidity 2014-16. Oxford: National Perinatal Epidemiology Unit, University of Oxford 2018.

9. Jurkovic D, Wilkinson H. Diagnosis and management of ectopic pregnancy. BMJ 2011;342.

10. Smith-Bindman R, Kerlikowske K, Feldstein V. Endovaginal ultrasound to exclude endometrial cancer and other endometrial abnormalities. JAMA 1998;280(17):1510-7.

11. Van Gorp T, Veldman J, Van Calster B, Cadron I, Leunen K, Amant F, et al. Subjective assessment by ultrasound is superior to the risk of malignancy index

(RMI) or the risk of ovarian malignancy algorithm (ROMA) in discriminating benign from malignant adnexal masses. Eur J Cancer 2012;48(11):1649-56.

12. Moore CL, Copel JA. Point-of-Care Ultrasonography. N Engl Journal of Med 2011;364(8):749-57.

13. Filly RA. Is it time for the sonoscope? If so, then let's do it right! J Ultrasound Med. 2003;22(4):323-5.

14. Wittenberg M. Will ultrasound scanners replace the stethoscope? BMJ. 2014;348.

15. Abu JI, Habiba MA, Baker R, Halligan AWF, Naftalin NJ, Hsu R, et al. Quantitative and qualitative assessment of women's experience of a one-stop menstrual clinic in comparison with traditional gynaecology clinics. BJOG 2001;108(9):993-9.

16. Bolondi L. Message from the president. Eur J Ultrasound. 1997;5:63-4.

17. Salvesen KA, Lees C, Tutschek B. Basic European ultrasound training in obstetrics and gynecology: where are we and where do we go from here? Ultrasound Obstet Gynecol. 2010;36(5):525-9.

18. Preisler J, Kopeika J, Ismail L, Vathanan V, Farren J, Abdallah Y, et al. Defining safe criteria to diagnose miscarriage: prospective observational multicentre study. BMJ 2015;351: h4579.

19. Pexsters A, Luts J, Van Schoubroeck D, Bottomley C, Van Calster B, Van Huffel S, et al. Clinical implications of intra-and interobserver reproducibility of transvaginal sonographic measurement of gestational sac and crown-rump length at 6-9 weeks' gestation. . Ultrasound Obstet Gynecol. 2011;38:510-5.

20. Bobdiwala S, Al-Memar M, Farren J, Bourne T. Factors to consider in pregnancy of unknown location. Women's **Health** 2017;13(2):27-33.

21. Yazbek JJ, Raju SKSK, Ben-Nagi JJ, Holland TKTK, Hillaby KK, Jurkovic DD. Effect of quality of gynaecological ultrasonography on management of patients with suspected ovarian cancer: a randomised controlled trial. Lancet Oncol. 2008;9(2):124-31.

22. Konge L, Albrecht-Beste E, Nielsen MB. Virtual-reality simulation-based training in ultrasound. Ultraschall in der Medizin. 2014/4;95(2).

23. Leerentveld RA, Gilberts EC, Arnold MJ, Wladimiroff JW. Accuracy and safety of transvaginal sonographic placental localization. Obstet Gynecol. 1990;76(5 Pt 1):759-62.

24. Royal College of Obstetricians and Gynaecologists. [London]Ultrasound training: introduction [Accessed 2018 Feb 28]. Available from: https://www.rcog.org.uk/en/careers-training/specialty-training-curriculum/ultrasound-training/introduction/.

25. Roberts M, Hughes T. Who should perform the ultrasound examinations in gynaecology? The Obstetrician & Gynaecologist. 2012;14(4):237-42.

26. Tolsgaard MG, Rasmussen MB, Tappert C, Sundler M, Sorensen JL, Ottesen B, et al. Which factors are associated with trainees' confidence in performing obstetric and gynecological ultrasound examinations? Ultrasound Obstet Gynecol. 2014;43(4):444-51.

27. Royal College of Obstericians and Gynaecologists. [London][2010] Joint Statement on ultrasound training [Accessed 2018 Feb 18] [Available from: https://www.rcog.org.uk/en/careers-training/specialty-training-curriculum/ultrasound-training/introduction/joint-statement-on-ultrasound-training/.

28. Parker PC, Harrison G. Educating the future sonographic workforce: membership survey report from the British Medical Ultrasound Society. Ultrasound. 2015;23(4):231-41.

29. Royal College of Radiologists. Investing in the Clinical Radiology Workforce – The Quality and Efficiency Case. [London] 2012 [Accessed 2017 Sept 13]. Available from: https://www.rcr.ac.uk/sites/default/files/RCR_CRWorkforce_June2012.pdf.

30. The Society and College of Radiographers. London[2012] Ultrasound examinations lengths survey analysis [cited 2018 Feb 14]. Available from: https://www.sor.org/learning/document-library/ultrasound-examination-lengths-survey-analysis/executive-summary.

31. Patel H, Chandrasekaran D, Myriokefalitaki E, Gebeh A, Jones K, Jeve YB. The Role of Ultrasound Simulation in Obstetrics and Gynecology Training: A UK Trainees' Perspective. Simul Healthc 2016;11(5):340-4.

32. Anderson A. Ten years of maternity claims: an analysis of the NHS Litigation Authority data – key findings. Clinical **Risk** 2013;19(1):24-31.

33. Gaba DM. The future vision of simulation in healthcare. Simul Healthc. 2007;2(2):126-35.

34. The Royal Aeronautical Society. [2009] The Impact of Flight Simulation in Aerospace [Accessed 2017 Oct 13]. Available from: https://www.aerosociety.com/Assets/Docs/Publications/DiscussionPapers/The_impact_of_flight_simulation_in_aerospace.pdf.

35. Bradley P. The history of simulation in medical education and possible future directions. Med Educ. 2006;40(3):254-62.

36. Society for Simulation in Healthcare. [Date unknown] About simulation [Accessed 2017 Oct 10]. Available from: http://www.ssih.org/About-SSH/About-Simulation.

37. Maran NJ, Glavin RJ. Low- to high-fidelity simulation – a continuum of medical education? Med Educ 2003;37:22-8.

38. Tjomsland N, Baskett P. Asmund S. Laerdal. **Resuscitation** 2002;53(2):115-9.

39. Abrahamson S, Denson JS, Wolf RM. Effectiveness of a simulator in training anesthesiology residents. J Med Educ 1969;44(6):515-9.

40. Chalouhi GE, Bernardi V, Ville Y. Ultrasound simulators in obstetrics and gynecology: state of the art. **Ultrasound** Obstet Gynecol 2015;46(3):255-63.

41. Jang TB, Jack Casey R, Dyne P, Kaji A. The Learning Curve of Resident Physicians Using Emergency Ultrasonography for Obstructive Uropathy. Acad Emerg Med 2010;17(9):1024-7.

42. Gaba DM, DeAnda A. A comprehensive anesthesia simulation environment: re-creating the operating room for research and training. Anesthesiology. 1988;69(3):387-94.

43. Gaba DM, Howard SK, Fish KJ, Smith BE, Sowb YA. Simulation-Based Training in Anesthesia Crisis Resource Management (ACRM): A Decade of Experience. **Simul Gaming** 2001;32(2):175-93.

44. Issenberg BS, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach 2005;27(1):10-28.

45. Stunt JJ, Wulms PH, Kerkhoffs GM, Dankelman J, van Dijk CN, Tuijthof GJM. How valid are commercially available medical simulators? Adv Med Educ 2014;5:385-95.

46. Gordon JA, Wilkerson WM, Shaffer DW, Armstrong EG. "Practicing" medicine without risk": students' and educators' responses to high-fidelity patient simulation. Acad Med 2001;76(5):469-72.

47. Klob D. Experiential learning: experience as a source of learning and development. New Jersey: Prentice Hall; 1984 [Accessed 2017 Sept 12]. Available from: http://www.learningfromexperience.com/images/uploads/process-of-experiential-learning.pdf

48. Okuda Y, Bryson EO, DeMaria S, Jr., Jacobson L, Quinones J, Shen B, et al. The utility of simulation in medical education: what is the evidence? Mt **Sinai** J Med New York 2009;76(4):330-43.

49. Zendejas B, Wang AT, Brydges R, Hamstra SJ, Cook DA. Cost: the missing outcome in simulation-based medical education research: a systematic review. Surgery. 2013;153(2):160-76.

50. Nagendran M, Gurusamy KS, Aggarwal R, Loizidou M, Davidson BR. Virtual reality training for surgical trainees in laparoscopic surgery. Cochrane Database Systematic Reviews [Internet]. 2013 Aug 27 [Accessed 2018 Oct 17]:[Art. No.: CD006575 p.]. Available from: 10.1002/14651858.CD006575.pub3.

51. Blum T, Rieger A, Navab N, Friess H, Martignoni M. A review of computer-based simulators for ultrasound training. Simul Healthc. 2013;8(2):98-108.

52. Hocking G, Hebard S, Mitchell CH. A Review of the Benefits and Pitfalls of Phantoms in Ultrasound-Guided Regional Anesthesia. Anaesth Pain Med. 2011;36(2):162-70.

53. Kyoto Kagaku CO LTD; [2012] Fetus Ultrasound Examination Phantom "Spacefan-ST" [Accessed 2018 Jan 2]. Available from: https://www.kyotokagaku.com/products/detail03/us-7.html.

54. CAE Healthcare; [2015] By Speciality - Obstetrics & Gynaecology [cited 2018 Feb 14]. Available from: http://www.bluephantom.com/category/By-Specialty_Obstetrics-AND-Gynecology.aspx.

55. Clau-Terre F, Sharma V, Cholley B, Gonzalez-Alujas T, Galinanes M, Evangelista A, et al. Can simulation help to answer the demand for echocardiography education? Anesthesiology 2014;120(1):32-41.

56. Terkamp C, Kirchner G, Wedemeyer J, Dettmer A, Kielstein J, Reindell H, et al. Simulation of Abdomen Sonography. Evaluation of a New Ultrasound Simulator. Ultraschall in Med. 2003;24(4):239-44.

57. Richards JR, McGahan JP. Focused Assessment with Sonography in Trauma (FAST) in 2017: What Radiologists Can Learn. Radiology. 2017;283(1):30-48.

58. Knudson MM, Sisley AC. Training residents using simulation technology: experience with ultrasound for trauma. **J Trauma Acute Care** Surgery 2000;48(4):659-65.

59. Chaudery M, Clark J, Dafydd Da, Dunn J, Bew D, Wilson MH, et al. The Face, Content, and Construct Validity Assessment of a Focused Assessment in Sonography for Trauma Simulator. J Surg Educ 2015;72(5):1032-8.

60. Jensen JK, Dyre L, Jørgensen ME, Andreasen LA, Tolsgaard MG. Collecting Validity Evidence for Simulation-Based Assessment of Point-of-Care Ultrasound Skills. J Ultrasound Med 2017;36(12):2475-83.

61. Wang Y, Helminen E, Jiang J. Building a virtual simulation platform for quasistatic breast ultrasound elastography using open source software: A preliminary investigation. Med Phys 2015;42(9):5453-66.

62. Sidhu HS, Olubaniyi BO, Bhatnagar G, Shuen V, Dubbins P. Role of Simulation-Based Education in Ultrasound Practice Training. J Ultrasound Med 2012(31):785-91.

63. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. **JAMA** 2011;306(9):978-88.

64. Gurusamy K, Aggarwal R, Palanivelu L, Davidson BR. Systematic review of randomized controlled trials on the effectiveness of virtual reality training for laparoscopic surgery. Br J Surg 2008;95(9):1088 - 97.

65. Madsen ME, Konge L, Nørgaard LN, Tabor A, Ringsted C, Klemmensen ÅK, et al. Assessment of performance measures and learning curves for use of a virtual-reality ultrasound simulator in transvaginal ultrasound examination. Ultrasound Obstet Gynecol 2014;44(6):693-9.

66. Moak JH, Larese SR, Riordan JP, Sudhir A, Yan G. Training in Transvaginal Sonography Using Pelvic Ultrasound Simulators Versus Live Models: A Randomized Controlled Trial. Academic Medicine. 2014;89(7):1063-8.

67. Khan K, Kunz R, Kleijnen J, Antes G. Systematic reviews to support evidenced based medicine. How to review and apply findings of healthcare research. Second edition ed. London: Hodder Arnold; 2011.

68. Kirkpatrick DI. Evaluating Training Programs: The Four Levels. 2nd Edition ed. San Francisco Berrett-Koehler Publishers; 1998.

69. Chapter 8: Assessing risk of bias in included studies. 2011. In: Cochrane Handbook for Systematic Reviews of Interventions Version 510 [updated March 2011] [Internet]. The Cochrane Collaboration. Available from: http://handbook-5-1.cochrane.org.

70. Wells G, Shea B, O'Connell D, Peterson J, Welch V. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analysis.: The Ottawa Hospital Research Institute; 2014 [Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

71. Chao C, Chalouhi GE, Bouhanna P, Ville Y, Dommergues M. Randomized Clinical Trial of Virtual Reality Simulation Training for Transvaginal Gynecologic Ultrasound Skills. J Ultrasound Med 2015;34(9):1663-7.

72. Tolsgaard MG, Ringsted C, Rosthoj S, Norgaard L, Moller L, Freiesleben NC, et al. The Effects of Simulation-based Transvaginal Ultrasound Training on Quality and Efficiency of Care: A Multicenter Single-blind Randomized Trial. Ann Surg 2017;265(3):630-7.

73. Tolsgaard M, Ringsted C, Dreisler E, Nørgaard LN, Petersen JH, Madsen ME, et al. Sustained effect of simulation-based ultrasound training on clinical performance: a randomized trial. Ultrasound Obstet Gynecol. 2015;46.

74. Tolsgaard MG, Madsen ME, Ringsted C, Oxlund BS, Oldenburg A, Sorensen JL, et al. The effect of dyad versus individual simulation-based ultrasound training on skills transfer. Med Educ 2015;49(3):286-95.

75. Le Lous M, De Chanaud N, Bourret A, Senat MV, Colmant C, Jaury P, et al. Improving the quality of transvaginal ultrasound scan by simulation training for general practice residents. Adv in Simu (London) 2017;2(1):24.

76. Maul H, Scharf A, Baier P, Wüstemann M, Günter HH, Gebauer G, et al. Ultrasound simulators: experience with the SonoTrainer and comparative review of other training systems. Ultrasound Obstet Gynecol. 2004;24.

77. Williams CJ, Edie JC, Mulloy B, Flinton DM, Harrison G. Transvaginal ultrasound simulation and its effect on trainee confidence levels: a replacement for initial clinical training? **Ultrasound** 2013;21.

78. Arya S, Mulla ZD, Kupesic Plavsic S. Role of pelvic ultrasound simulation. Clin Teach 2017:1-5.

79. Bahl R, Hinks J, Halliwell M, Phillips R, Jenkins J. Place of an Ultrasonography Simulator in Ultrasonography Training for Fertility Nurses. **Ultrasound** 2006;14(4):213-5.

80. Madsen ME, Nørgaard LN, Tabor A, Konge L, Ringsted C, Tolsgaard MG. The Predictive Value of Ultrasound Learning Curves Across Simulated and Clinical Settings. J Ultrasound Med 2017;36(1):201-8.

81. Vallabh-Patel V, Mendez M, Kupesic Plavsic S. The Importance of Multimodality Pelvic Ultrasound Simulation in Teaching of Obstetrics and Gynecology Residents. Donald School J Ultrasound Obstet Gynecol 2014;8(1):1-5.

82. Nickels L, Leech S, Cassidy D, Briscoe J, Papa L, Sawyer S, et al. A Randomized Controlled Trial of an Ultrasound Mannequin versus a Live Model in Teaching Endovaginal Ultrasound to Emergency Physicians. Acad Emerg Med 2008;15(Supplement 2008 SAEM Annual Meeting Abstracts):S69-70.

83. Alsalamah A. Validation and determination of the influence of a virtual simulator on the acquisition of ultrasound skills and comparison of learning curves of those using simulation- supported training with a conventional training approach [Thesis (PhD)]: Cardiff University; 2016.

84. Tolsgaard MG, Todsen T, Sorensen JL, Ringsted C, Lorentzen T, Ottesen B, et al. International multispecialty consensus on how to evaluate ultrasound competence: a Delphi consensus survey. PLoS One 2013;8(2):e57687.

85. Tolsgaard MG, Ringsted C, Dreisler E, Klemmensen A, Loft A, Sorensen JL, et al. Reliable and valid assessment of ultrasound operator competence in obstetrics and gynecology. Ultrasound Obstet Gynecol 2014;43(4):437-43.

86. Norman G, Dore K, Grierson L. The minimal relationship between simulation fidelity and transfer of learning. Med Educ. 2012;46:636 - 47.

87. Salomon LJ, Nassar M, Bernard JP, Ville Y, Fauconnier A. A score-based method to improve the quality of emergency gynaecological ultrasound examination. Eur J of Obstet and Gynecol Reprod Biol 143(2):116-20.

88. Yardley S, Dornan T. Kirkpatrick's levels and education 'evidence'. Med Educ. 2012;46(1):97-106.

89. Bennett CC, Richards DS. Patient acceptance of endovaginal ultrasound. Ultrasound Obstet and Gynecol [Internet]. 2000; [Accessed 2018 Jan 3]15(1):[52-5 pp.]. Available from: http://dx.doi.org/10.1046/j.1469-0705.2000.00010.x.

90. Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. Surg Endosc 2003;17(10):1525-9.

91. Savran MM, Sorensen SM, Konge L, Tolsgaard MG, Bjerrum F. Training and Assessment of Hysteroscopic Skills: A Systematic Review. J Surg Educ. 2016;73(5):906-18.

92. Kneebone R. Simulation in surgical training: educational issues and practical implications. Med Educ 2003;37(3):267-77.

93. Osborne B, Parange N, Thoirs K. The effectiveness of the use of high fidelity simulators in obstetric ultrasound training: A systematic review. Australas J Ultrasound Med [Internet]. 2015; [Accessed 2018 Jan 3]18(3):[107-11 pp.]. Available from: http://dx.doi.org/10.1002/j.2205-0140.2015.tb00209.x.

94. Cook DA. How much evidence does it take? A cumulative meta-analysis of outcomes of simulation-based education. Med Educ 2014;48(8):750-60.

95. Royal College of Obstetricians and Gynaecologists. Obstertrics and Gynaecology Workforce Report London2017 [Accessed 2018 Jan 12]. Available from: https://www.rcog.org.uk/globalassets/documents/careers-and-training/workplace-and-workforce-issues/rcog-og-workforce-report-2017.pdf.

96. Green J, Kahan M, Wong S. Obstetric and Gynecologic Resident Ultrasound Education Project. J **Ultrasound** Med 2015;34(9):1583-9.

97. Eisen LA, Leung S, Gallagher AE, Kvetan V. Barriers to ultrasound training in critical care medicine fellowships: a survey of program directors. Crit Care Med 2010;38(10):1978-83.

98. American Institiute of Ultrasound in Medicine [2015] Training guidelines for physicians who evaluate and interpretate diagnostic ultrasound examinations of the female pelvis [Accessed 2018 Jan 30][Available from: http://www.aium.org/officialStatements/58.

99. Temple J. Time for training: a review of the impact of the european working time directive on the quality of training. 2010 [cited 2017 Mar 23]. In: Time for Training: A Review of the Impact of the European Working Time Directive on the Quality of Training [Internet]. London, [Accessed 2017 Mar 23]. Available from: https://www.hee.nhs.uk/sites/default/files/documents/Time for training report_0.pdf.

100. Messenger S. Ultrasound training boosted by 'revolutionary' simulator: BBC News; 2014 [Accessed 2018 Feb 14]. Available from: http://www.bbc.co.uk/news/uk-wales-south-east-wales-27102993.

101. Burden C, Preshaw J, White P, Draycott TJ, Grant S, Fox R. Validation of virtual reality simulation for obstetric ultrasonography: a prospective cross-sectional study. Simul Healthc 2012;7(5):269-73.

102. Heer IM, Middendorf K, Müller-Egloff S, Dugas M, Strauss A. Ultrasound training: the virtual patient. Ultrasound Obstet Gynecol 2004;24(4):440 - 4
.

103. Royal College of Obstetricians and Gynaecologists. OSATS London2018 [Accessed 2018 Feb 21]. Available from: https://www.rcog.org.uk/en/careers-training/about-specialty-training-in-og/assessment-and-progression-through-training/workplace-based-assessments/osats/.

104. Royal College of Obstericians and Gynaecologists. OSATS, intermediate ultrasound in gynaecology London2018 [Accessed 2018 Jan 21]. Available from: https://www.rcog.org.uk/en/careers-training/specialty-training-curriculum/ultrasound-training/ultrasound-module-intermediate-ultrasound-in-gynaecology/osats - pelvis.

105. Tolsgaard MG, Kulasegaram KM, Ringsted C. Practical trials in medical education: linking theory, practice and decision making. Med Educ 2017;51(1):22-30.

106. Ringsted C, Skaarup AM, Henriksen AH, Davis D. Person–task–context: a model for designing curriculum and in-training assessment in postgraduate education. Med Teach 2006;28(1):70-6.

107. ISUOG Education committee. ISUOG Education Committee recommendations for basic training in obstetric and gynecological ultrasound. Ultrasound Obstet Gynecol. 2014;43(1):113-6.

108. Jang TB, Ruggeri W, Dyne P, Kaji AH. The Learning Curve of Resident Physicians Using Emergency Ultrasonography for Cholelithiasis and Cholecystitis. Acad Emerg Med 2010;17(11):1247-52.

109. Jensen JK, Dyre L, Jorgensen ME, Andreasen LA, Tolsgaard MG. Simulation-based point-of-care ultrasound training: a matter of competency rather than volume. Acta Anaesthesiol Scand. 2018; 62(6): 811-819

110. Petticrew M, McKee M, Lock K, Green J, Phillips G. In search of social equipoise. BMJ 2013;347.

111. Gostlow H, Marlow N, Babidge W, Maddern G. Systematic Review of Voluntary Participation in Simulation-Based Laparoscopic Skills Training: Motivators and Barriers for Surgical Trainee Attendance. J Surg Educ74(2):306-18.

112. Burden C, Fox R, Hinshaw K, Draycott TJ, James M. Laparoscopic simulation training in gynaecology: Current provision and staff attitudes - a cross-sectional survey. J Obstet Gynaecol 2016;36(2):234-40.

113. Dyre L, Nørgaard LN, Tabor A, Madsen ME, Sørensen JL, Ringsted C, et al. Collecting Validity Evidence for the Assessment of Mastery Learning in Simulation-Based Ultrasound Training. Ultraschall in Med. 2016;37(04):386 - 92.

114. Dyre L, Tabor A, Ringsted C, Tolsgaard MG. Imperfect practice makes perfect: error management training improves transfer of learning. Med Educ. 2017;51(2):196-206.

115. Konge L, Ringsted C, Bjerrum F, Tolsgaard MG, Bitsch M, Sørensen JL, et al. The Simulation Centre at Rigshospitalet, Copenhagen, Denmark. J Surg Educ 72(2):362-5.

116. Bosse HM, Mohr J, Buss B, Krautter M, Weyrich P, Herzog W, et al. The benefit of repetitive skills training and frequency of expert feedback in the early acquisition of procedural skills. BMC Medical Educ 2015;15:22.

117. Stefanidis D, Acker CE, Swiderski D, Heniford BT, Greene FL. Challenges during the implementation of a laparoscopic skills curriculum in a busy general surgery residency program. J Surg Educ 2008;65(1):4-7.

118. Burden C, Appleyard T-L, Angouri J, Draycott TJ, McDermott L, Fox R. Implementation of laparoscopic virtual-reality simulation training in gynaecology: a mixed-methods design. Eur J Obstet Gynecol Reprod Biol 170(2):474-9.

119. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. BMJ 2000;321(7262):694-6.

120. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003-2009. Med Educ 2010;44(1):50-63.

121. Bradley P, Bligh J. Clinical skills centres: where are we going? Medical Education. 2005;39(7):649-50.

122. McDougall EM. Validation of surgical simulators. J Endourol. 2007;21(3):244-7.

123. Schijven MP, Jakimowicz JJ. Validation of virtual reality simulators: Key to the successful integration of a novel teaching technology into minimal access surgery. Minim Invasive Ther Allied Technol 2005;14(4):244-6.

124. Chmarra MK, Klein S, de Winter JC, Jansen FW, Dankelman J. Objective classification of residents based on their psychomotor laparoscopic skills. Surg Endosc 2010;24(5):1031-9.

125. Paisley AM, Baldwin PJ, Paterson-Brown S. Validity of surgical simulation for the assessment of operative skill. Br J Surg 2001;88(11):1525-32.

126. Gallagher AG, Richie K, McClure N, McGuigan J. Objective psychomotor skills assessment of experienced, junior, and novice laparoscopists with virtual reality. **World** J Surg 2001;25(11):1478-83.

127. Nitsche JF, Brost BC. Obstetric ultrasound simulation. Seminars in perinatology. 2013;37(3):199-204.

128. Preshaw J, Ficquet J, Burden C, Overton C, Grant S. O556 Would it be best practice for trainees to learn ultrasound scanning in a simulated setting prior to a clinical setting? . BJOG 2012;10(01):S457.

129. Schout BMA, Hendrikx AJM, Scheele F, Bemelmans BLH, Scherpbier AJJA. Validation and implementation of surgical simulators: a critical review of present, past, and future. Surg Endosc 2010;24(3):536-46.

130. Alsalamah A, Campo R, Tanos V, Grimbizis G, Van Belle Y, Hood K, et al. Face and content validity of the virtual reality simulator 'ScanTrainer®'. Gynecol Surg 2017;14(1):18.

131. Al-Memar M, Saso S, Bobdiwala S, Ameye L, Guha S, Joash K, et al. Validation of a virtual reality simulator for the use of transvaginal ultrasonography in gynaecology and early pregnancy. Australas J Ultrasound Med 2017;20 (3):97-105
.

132. American Psychological Association. Standards for educational and psychological tests. Washington, DC.1974.

133. Reber AS, Allen R, Rebr ES. The Penguin Dictionary of Psychology: Fourth Edition: Penguin; 2009.

134. Bajka M, Tuchschmid S, Streich M, Fink D, Szekely G, Harders M. Evaluation of a new virtual-reality training simulator for hysteroscopy. Surg Endosc. 2009;23(9):2026-33.

135. Carter FJ, Schijven MP, Aggarwal R, Grantcharov T, Francis NK, Hanna GB, et al. Consensus guidelines for validation of virtual reality surgical simulators. Surg Endosc. 2005;19.

136. Ahmed K, Miskovic D, Darzi A, Athanasiou T, Hanna GB. Observational tools for assessment of procedural skills: a systematic review. Am J Surg. 2011;202(4):469-80.e6.

137. Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. **Lancet** 2001;357(9260):945-9.

138. Ahmed K, Ashrafian H, Hanna GB, Darzi A, Athanasiou T. Assessment of specialists in cardiovascular practice. Nat Rev Cardiol 2009;6(10):659-67.

139. Deo ND, Mirza M, Lawin O'Brien A, Shahid A, Visvanathan SD. OP36.05: A modular programme incorporating virtual reality simulation in delivery of competency based core curriculum in gynaecology ultrasound. Ultrasound in Obstet Gynecol 2012;40(S1):165

140. Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. BMJ. 2003;327(7422):1032-7.

141. Robson C. A resource for users of social research methods in applied settings. 2011. In: Real world research [Internet]. Cornwall, UK: TJ International Ltd; [302].

142. Schreuder HW, Dongen KW, Roeleveld SJ, Schijven MP, Broeders IA. Face and construct validity of virtual reality simulation of laparoscopic gynecologic surgery. Obstet Gynecol 2009;200(5):540e1-e8

143. Schreuder H, van den Berg C, Hazebroek E, Verheijen R, Schijven M. Laparoscopic skills training using inexpensive box trainers: which exercises to choose when constructing a validated training course. BJOG 2011;118(13):1576 - 84.

144. Ayodeji ID, Schijven M, Jakimowicz J, Greve JW. Face validation of the Simbionix LAP Mentor virtual reality training module and its applicability in the surgical curriculum. Surg Endosc. 2007;21(9):1641-9.

145. Brooke J. SUS - A quick and dirty usability scale 1986 [cited 2018 Jan 28]. Available from: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

146. Verdaasdonk EG, Stassen LP, Monteny LJ, Dankelman J. Validation of a new basic virtual reality simulator for training of basic endoscopic skills: the SIMENDO. Surg Endosc 2006;20(3):511-8.

147. Ficquet J, Preshaw J, Day F, Burden C, Lenguerrand E, Grant S, et al. OP36.06: What is the construct validity for Medaphor ScanTrainer's gynaecological and early pregnancy modules? Ultrasound Obstet Gynecol 2012;40(S1):165 - 6.

148. Messick S. Validity. ETS Research Report Series. 1987;1987(2):i-208.

149. Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G, et al. Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. Ann Surg 2005;241(2):364-72.

150. European Federation of Societies for Ultrasound in Medicine and Biology. Minimum Training Requirements for the Practice of Medical Ultrasound in Europe. Ultraschall Med. 2010;31 (04):426-7.

151. Coldicott Y, Nesheim B-I, MacDougall J, Pope C, Roberts C. The ethics of intimate examinations—teaching tomorrow's doctors. BMJ. 2003;326(7380):97.

152. Clement S, Candy B, Heath V, To M, Nicolaides KH. Transvaginal ultrasound in pregnancy: its acceptability to women and maternal psychological morbidity. Ultrasound Obstet Gynecol 2003;22(5):508-14.

153. Pickard S, Baraitser P, Rymer J, Piper J. Can gynaecology teaching associates provide high quality effective training for medical students in the United Kingdom? BMJ 2003;327(7428):1389-92.

154. Ziv A, Wolpe PR, Small SD, Glick S. Simulation-Based Medical Education: An Ethical Imperative. Acad Med 2003;78(8):783-8.

155. Department of Health. Putting people at the heart of care London2009 [Accessed 2013 Jan 12]. Available from: http://webarchive.nationalarchives.gov.uk/20130123200554/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_106038.

156. Deed K, Childs J, Thoirs K. What are the perceptions of women towards transvaginal sonographic examinations? Sonography. 2014;1(2):33-8.

157. Braithwaite JM, Economides DL. Acceptability by patients of transvaginal sonography in the elective assessment of the first-trimester fetus. Ultrasound Obstet Gynecol 1997;9(2):91-3.

158. Onwude JL, Johnson N, Bark M, Bates J. The acceptability of transvaginal and transabdominal ultrasound in women with first trimester bleeding. J Obstet Gynaecol 1993(13):350–1.

159. Krubel R, Freedman D, Bursch B. Effects of transvaginal sonography on pregnant patients' perceptions of care in the ED. Am J Emerg Med 1996;14(2):232-3.

160. Heath VCF, Southall TR, Souka AP, Novakov A, Nicolaides KH. Cervical length at 23 weeks of gestation: relation to demographic characteristics and previous obstetric history. Ultrasound Obstet Gynecol 1998;12(5):304-11.

161. O'Shea RT, Forbes KL, Scopacasa L, Jones WR. Comparison of transabdominal and transvaginal pelvic ultrasonography for ovarian follicle assessment in in vitro fertilisation. Gynecol Obstet Invest 1988;26(1):52-5.

162. Kipps S. Transvaginal screening: what do women feel about ovarian and endometrial cancer screening? Nurs **Times** 1994;90:45 - 8.

163. Wolfe CD, Raju KS. The attitudes of women and feasibility of screening for ovarian and endometrial cancers in inner city practices. Eur J Obstet Gynecol Reprod Biol 1994;56(2):117-20.

164. Atalabi OM, Morhason-Bello IO, Adekanmi AJ, Marinho AO, Adedokun BO, Kalejaiye AO, et al. Transvaginal ultrasonography: a survey of the acceptability and its predictors among a native African women population. Int J Womens Health 2012;4:1-6.

165. Rosati P, Guariglia L. Acceptability of early transvaginal or abdominal sonography in the first half of pregnancy. Arch Gynecol Obstet. 2000;264(2):80-3.

166. Dutta RL, Economides DL. Patient acceptance of transvaginal sonography in the early pregnancy unit setting. Ultrasound Obstet Gynecol 2003;22(5):503-7.

167. Basama FMS, Crosfill F, Price A. Women's Perception of transvaginal sonography in the first trimester; in an early pregnancy assessment unit. Arch Gynecol Obs. 2004;269.

168. Gentry-Maharaj A, Sharma A, Burnell M, Ryan A, Amso NN, Seif MW, et al. Acceptance of transvaginal sonography by postmenopausal women participating in the United Kingdom Collaborative Trial of Ovarian Cancer Screening. Ultrasound Obstet Gynecol 2013;41(1):73-9.

169. Britten N. Qualitative Interviews in Healthcare. In: Pope C, Mays N, editors. Qualitative Research in Health Care. 2nd ed. London: BMJ Books; 1999. p. 11-9.

170. Silverman D. Doing qualitative research. London: Sage publications; 2000.

171. Irvine A, Drew P, Sainsbury R. Mode effects in qualitative interviews: a comparison of semi-structured face-to-face and telephone interviews using conversation analysis. York: Social Policy Research Unit, University of York; 2010.

172. Irvine A. Realities Tollkit 14: Using telephone interviews. Manchester, UK: Economic and Social Research Council (ESRC); 2010.

173. Green J, Thorogood N. Qualitative Methods for Health Research. 3rd ed: Sage Publications Ltd; 2014.

174. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006;3(2):77-101.

175. Weitlauf JC, Finney JW, Ruzek JI, Lee T-T, Thrailkill A, Jones S, et al. Distress and Pain During Pelvic Examinations: Effect of Sexual Violence. Obstetrics and gynecology. 2008;112(6):1343-50.

176. Yanikkerem E, Özdemir M, Bingol H, Tatar A, Karadeniz G. Women's attitudes and expectations regarding gynaecological examination. Midwifery. 2009;25(5):500-8.

177. Larsen M, Oldeide CC, Malterud K. Not so bad after all....Women's experiences of pelvic examinations. Fam Pract 1997;14(2):148-52.

178. Wendt E, Bengt F, Lidell E. Trust and confirmation in a gynecological examination situation: a critical incident technique analysis. Acta Obstetricia et Gynecologica Scandinavica 2004;83:1208-15.

179. Thoirs K, Deed K, Childs J. Transvaginal sonography: Sonographer reflections on patient experience using a critical incident technique. Sonography. 2017;4(2):55-62.

180. Sharma A, Beveridge HA, Fallowfield LJ, Jacobs IJ, Menon U. Postmenopausal women undergoing transvaginal ultrasound screening prefer not to have chaperones. BJOG 2006;113(8):954-957

181. General Medical Council. Intimate examinations and chaperones 2013 [Accessed 2018 Jan 31]. Available from: https://www.gmc-uk.org/guidance/ethical_guidance/21168.asp.

182. Whitford DL, Karim M, Thompson G. Attitudes of patients towards the use of chaperones in primary care. Br J Gen Pract 2001;51(466):381-3.

183. Korsgaard CM. Introduction to the groundwork to the metaphysics of morals (I Kant). Cambridge: Cambridge University Press; 1997.

184. World Health Organization (WHO). Transforming and scaling up health professionals' education and training: World Health Organization guidelines 2013 [Accessed 2018 Jan 26]. Available from: http://apps.who.int/iris/bitstream/10665/93635/1/9789241506502_eng.pdf?ua=1.

185. Utian WH, Maamari R. Attitudes and approaches to vaginal atrophy in postmenopausal women: a focus group qualitative study. **Climacteric** 2014;17(1):29-36.

186. Medaphor® Plc. The ScanTrainer 2016 [Accessed 2016 Jun 30]. Available from: http://www.medaphor.com/scantrainer.

187. Kneebone RL, Scott W, Darzi A, Horrocks M. Simulation and clinical practice: strengthening the relationship. Med Educ 2004;38(10):1095-102.

188. Tolsgaard MG. Assessment and learning of ultrasound skills in Obstetrics & Gynecology. Dan Med J 2018;65(2).

189. Kolozsvari NO, Feldman LS, Vassiliou MC, Demyttenaere S, Hoover ML. Sim one, do one, teach one: considerations in designing training curricula for surgical simulation. J Surg Educ 2011;68(5):421-7.

190. Ericsson KA. An expert-performance perspective of research on medical expertise: the study of clinical performance. Med Educ. 2007;41(12):1124-30.

191. Gentner D, Rattermann MJ, Forbus KD. The roles of similarity in transfer: separating retrievability from inferential soundness. Cognitive psychology. 1993;25(4):524-75.

192. Saul T, Lewiss RE, Rivera MDR. Accuracy of emergency physician performed bedside ultrasound in determining gestational age in first trimester pregnancy. Crit **Ultrasound J** 2012;4(1):22.

193. McGaghie WC, Issenberg SB, Cohen ER, Barsuk JH, Wayne DB. Medical education featuring mastery learning with deliberate practice can lead to better health for individuals and populations. Acad Med 2011;86(11):e8-9.

194. Schmidt J, Liu R, Glodflam K, Moore C. Comparison of Haptic Technology Versus Traditional Mannequin Simulation for Resident Training in Pelvic Ultrasound. Acad Em Med 2014;21(S1):S121-2.

195. Sullivan AM, Rock LK, Gadmer NM, Norwich DE, Schwartzstein RM. The Impact of Resident Training on Communication with Families in the Intensive Care Unit. Resident and Family Outcomes. Ann Am Thorac Soc 2016;13(4):512-21.

196. Cook DA. If you teach them, they will learn: why medical education needs comparative effectiveness research. Advances in Health Sciences Education. 2012;17(3):305-10.

197. Medaphor plc MG. Annual report for the year ended 31 December 2016 [Available from: http://investors.medaphor.com/wp-

content/uploads/Medaphor-Annual-Report-2016-and-Notice-of-Annual-General-Meeting-2017_2.pdf.

PUBLICATIONS

1. Woodhead N., Mahmoud A., Clark T.J., Ultrasound simulation for gynaecological ultrasound training: a systematic review. *Ultrasound Obstet Gynecol*. 2018 Submitted

2. Woodhead N., Smith P., Clark T.J., Face, content and construct validity of a virtual reality simulator for transvaginal ultrasonography in gynaecology and early pregnancy. *Acta Obstet Gynecol Scand*. 2018. Submitted

3. Woodhead N., Porcellato P., Clark T.J., Women's views and experiences of transvaginal ultrasound: a qualitative interview study. *Midwifery*. 2018.Submitted

4. Woodhead N., Porcellato P., Clark T.J., Everyone has to learn: womens attitudes towards transvaginal ultrasound training. *Medical Education*. 2018. Submiited

PUBLISHED ABSTRACTS

Simulation training – Trainees want it but don't use it: A study by Midlands Research Collaborative in Obstetrics & Gynaecology. Chandrasekaran; Patel, H; Myriokefalitaki, E; Woodhead, N; Jones, K; Gebeh, A; Jeve, D. (2016), Quality Improvement, Education, Simulation and Patient Safety (EP12). BJOG: Int J Obstet Gy, 123: 198–220. doi:10.1111/1471-0528.14110

Deo N.D, Woodhead N, Gale, A, Masson, G. UK-wide survey on the role of simulation in obstetrical and gynaecological ultrasound training. Supplement: Abstracts of the 24th World Congress on Ultrasound in Obstetrics and

Gynecology, 14–17 September 2014, Barcelona, Spain. September 2014. Volume 44, Issue S1. Pages 1–369

LETTERS

Woodhead, N. and Masson, G. (2013), Re: Who should perform the ultrasound examinations in gynaecology?. The Obstetrician & Gynaecologist, 15: 279–280. doi: 10.1111/tog.12057_3

NATIONAL ORAL PRESENTATIONS

Woodhead, N. Use of Transvaginal Ultrasound Simulation in Speciality Training. British Association of Gynaecological Imaging (BSGI) Annual Scientific Meeting RCOG, 16th April 2015

Woodhead, N. Medaphor Assessment in Specialist Training and its Economic Recompsense study. National ultrasound co-ordinators meeting. Royal College of Obstetricians and Gynaecologists. June 2014

Woodhead, N. Use of virtual reality simulation for transvaginal scan training. An overview of the MASTERS study. The Health Foundation Shine Award Showcase Event. April 2014

POSTER PRESENTATIONS

Simulation training – Trainees want it but don't use it: A study by Midlands Research Collaborative in Obstetrics & Gynaecology. Chandrasekaran; Patel, H;

Myriokefalitaki, E; **Woodhead, N**; Jones, K; Gebeh, A; Jeve, D. RCOG World Congress, Birmingham 2016

Evaluation of a virtual-reality training simulator for transvaginal ultrasound. **Woodhead N**, Masson G, Ismail K. Birmingham and Midland Obstetrical and Gynaecological Society. Autumn meeting. October 20