Ultrasound Obstet Gynecol 2022; **60**: 759–765
Published online in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.24975.

Clinical workflow of sonographers performing fetal anomaly ultrasound scans: deep-learning-based analysis

L. DRUKKER[1,2], H. SHARMA[3], J. N. KARIM[1], R. DROSTE[3], J. A. NOBLE[3] and A. T. PAPAGEORGHIOU[1]

[1]*Nuffield Department of Women's and Reproductive Health, John Radcliffe Hospital, University of Oxford, Oxford, UK;* [2]*Women's Ultrasound, Department of Obstetrics and Gynecology, Beilinson Medical Center, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel;* [3]*Institute of Biomedical Engineering, University of Oxford, Oxford, UK*

KEYWORDS: anatomy; artificial intelligence; automation; big data; clinical workflow; computer vision; data science; deep learning; image analysis; machine learning; neural network; obstetrics; pregnancy; screening; sonography; ultrasound

CONTRIBUTION

What are the novel findings of this work?
Using sonography big data and deep learning to describe video content automatically, we studied obstetric ultrasound as a data-science problem. We showed that an anomaly scan can be modeled as a non-ordered sequence of anatomical structure acquisitions.

What are the clinical implications of this work?
The lack of a universal scanning sequence supports the hypothesis that anomaly scanning is opportunistic by nature, continuously taking advantage of fetal position. Trainees may benefit from appreciating that even experts do not follow guidelines as an ordered list, but rather acquire planes and images according to visibility.

ABSTRACT

Objective *Despite decades of obstetric scanning, the field of sonographer workflow remains largely unexplored. In the second trimester, sonographers use scan guidelines to guide their acquisition of standard planes and structures; however, the scan-acquisition order is not prescribed. Using deep-learning-based video analysis, the aim of this study was to develop a deeper understanding of the clinical workflow undertaken by sonographers during second-trimester anomaly scans.*

Methods *We collected prospectively full-length video recordings of routine second-trimester anomaly scans. Important scan events in the videos were identified by detecting automatically image freeze and image/clip save. The video immediately preceding and following the important event was extracted and labeled as one of 11 commonly acquired anatomical structures. We developed and used a purposely trained and tested deep-learning annotation model to label automatically the large number of scan events. Thus, anomaly scans were partitioned as a sequence of anatomical planes or fetal structures obtained over time.*

Results *A total of 496 anomaly scans performed by 14 sonographers were available for analysis. UK guidelines specify that an image or videoclip of five different anatomical regions must be stored and these were detected in the majority of scans: head/brain was detected in 97.2% of scans, coronal face view (nose/lips) in 86.1%, abdomen in 93.1%, spine in 95.0% and femur in 92.3%. Analyzing the clinical workflow, we observed that sonographers were most likely to begin their scan by capturing the head/brain (in 24.4% of scans), spine (in 23.2%) or thorax/heart (in 22.8%). The most commonly identified two-structure transitions were: placenta/amniotic fluid to maternal anatomy, occurring in 44.5% of scans; head/brain to coronal face (nose/lips) in 42.7%; abdomen to thorax/heart in 26.1%; and three-dimensional/four-dimensional face to sagittal face (profile) in 23.7%. Transitions between three or more consecutive structures in sequence were uncommon (up to 13% of scans). None of the captured anomaly scans shared an entirely identical sequence.*

Conclusions *We present a novel evaluation of the anomaly scan acquisition process using a deep-learning-based analysis of ultrasound video. We note wide variation in the number and sequence of structures obtained during routine second-trimester anomaly scans.*

Correspondence to: Prof. A. T. Papageorghiou, Nuffield Department of Women's and Reproductive Health, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, UK (e-mail: aris.papageorghiou@wrh.ox.ac.uk)

Accepted: 10 June 2022

Overall, each anomaly scan was found to be unique in its scanning sequence, suggesting that sonographers take advantage of the fetal position and acquire the standard planes according to their visibility rather than following a strict acquisition order. © 2022 The Authors. Ultrasound in Obstetrics & Gynecology published by John Wiley & Sons Ltd on behalf of International Society of Ultrasound in Obstetrics and Gynecology.

INTRODUCTION

The continuous improvement of obstetric ultrasound is attributed to many factors, including developments in education, accreditation, guidelines, quality assurance, anatomical and physiological knowledge and imaging quality[1–4]. At the same time, innovations in ergonomic machine design, scanning protocols, automatic measurement tools and reporting have improved sonographer workflow[5–8]. However, unlike these innovations, the ultrasound acquisition process has remained relatively unchanged: routine obstetric ultrasound scans are performed by a sonographer sitting or standing next to a pregnant woman, manipulating a probe and adjusting the machine settings, following a defined protocol to acquire and interpret a series of standard imaging planes that are observed on the screen of the ultrasound machine.

To ensure uniformity of screening, sonographers are usually required to adhere to established local, national or international practice guidelines[9,10]. Such scanning guidelines comprise typically a checklist of standard planes, structures and anatomies to be surveyed, but do not dictate the order in which the checklist should be populated. Hence, each sonographer may conduct the examination in a different order: some may follow the protocol in a head-to-toe fashion, others may choose to assess structures according to personal preference, and still others may not follow a predefined order at all, taking advantage of visible structures according to the fetal lie and position. Moreover, anomaly-scan guidelines do not usually define the time allocated to the assessment of each structure or standard plane. Thus, sonographers exercise discretion in the order and proportion of time spent on each mandatory acquisition.

Artificial neural networks are particularly good at image pattern recognition and can be utilized to improve the speed and accuracy of diagnosing patient conditions in medical fields that depend heavily on images and video, as in the case of ultrasound[11,12]. In this study, we use the power of deep learning to describe automatically the clinical workflow of sonographers performing fetal anomaly ultrasound scans and to identify variations and similarities in the scanning sequences.

METHODS

This was a prospective study of routine anomaly ultrasound scans performed on women with a singleton pregnancy at 19–21 weeks of gestation. Scans were undertaken between May 2018 and February 2020 at the Maternity Ultrasound Unit, John Radcliffe Hospital, Oxford University Hospitals National Health Services Foundation Trust, Oxford, UK. Ultrasound examinations were carried out in accordance with the UK Fetal Anomaly Screening Programme (FASP) standards, as described previously[13,14]. All ultrasound scans included in this study were performed using a commercial Voluson E8 version BT18 (GE Healthcare, Zipf, Austria) ultrasound machine, equipped with standard curvilinear (C2-9-D, C1-5-D) and three-dimensional (3D)/four-dimensional (4D) (RAB6-D) probes.

This study was part of a project entitled Perception Ultrasound by Learning Sonographic Experience (PULSE)[15]. This is an innovative interdisciplinary project that is designed to apply the latest ideas from deep learning and computer vision to build computational models that describe how an expert sonographer performs a diagnostic study of a subject from multiple perceptual cues. To do this, we capture anonymously real-world full-length ultrasound scan videos in addition to other sensory data, namely probe movement using motion trackers, point-of-gaze of the sonographer on the monitor of the ultrasound scanner using eye-tracking and voice of the sonographer using directional microphones[15]. By understanding closely how experts learn and undertake diagnostic ultrasound, we aim to build considerably more powerful assistive interpretation methods than have been possible so far.

Data acquisition

Women attending for a routine second-trimester anomaly scan were offered the opportunity to take part in the study. After providing written informed consent, they underwent a standard obstetric anatomy scan. The entire scan video was recorded using the machine's high-definition multimedia interface (HDMI) output and a video-grabbing card (DVI2PCIe, Epiphan Video, Palo Alto, CA, USA), as reported previously[15]. According to the UK FASP guidelines, five mandatory structures should be captured as images or clips: head/brain, coronal face view (nose/lips), abdomen, spine and femur. In addition, sonographers should assess, but are not obliged to store an image or a clip of, the following structures: sagittal face (fetal profile), thorax/heart, kidneys, bladder, limbs and placenta/amniotic fluid[14]. Despite not being required formally, sonographers usually save these additional images or clips in our settings[15]. In addition to the UK FASP requirements, we measure maternal uterine artery Doppler pulsatility index routinely to assess the risk of obstetric complications[16,17].

Data analysis

The technique used to analyze sonographer workflow is presented in Figure 1. Initially, we analyzed the full-length anomaly scans automatically and identified sonographer actions that represent important scan events by detecting the occurrences of image freeze, image save and clip save. To identify the key sonographer actions, video processing analysis was carried out on a frame-by-frame basis[18] with a purpose-built software program implemented in Python (www.python.org, version 3.7.0) using OpenCV (www.opencv.org, version 3.4) and Tesseract (www.github.com/tesseract-ocr, version 3.05). After identifying the important scan events, we extracted 5-s clips that occurred around the event (i.e. immediately preceding and following the event).

We developed a deep-learning model to automate the labeling of scans[11], because the large number of short clips renders manual labeling impractical. First, we carried out detailed manual labeling of clips for 62 full-length scans[15], assigning each clip to one of the following: head/brain, sagittal face (profile), coronal face (nose/lips), thorax/heart, abdomen, kidneys, spine, femur, 3D/4D face, placenta/amniotic fluid, maternal anatomy, mixed (multiple structures) and other. Thereafter, we trained a deep-learning model, which allowed automatic labeling of videoclips as described previously[19]. Further details about the preparation of data for model training, model testing, automated video labeling and model validation are provided in Appendix S1. Confusion matrices were used to compare interannotator and manual *vs* automatic labeling.

Once all scan sequences were labeled, all manually and automatically labeled scans were used to create timelines in which each scan was portrayed as a time sequence of structure labels. This allowed us to compute structure metrics, including the number of scans in which a particular structure was present, the percentage of scans in which the structure was the first to be assessed, the mean number of times that the structure was acquired per scan, the order of structure acquisition and the common transitions between structures. These were studied using descriptive statistics.

Ethics approval

This study was approved by the UK Research Ethics Committee (reference 18/WS/0051), and written informed consent was given by all participating pregnant women and sonographers.

RESULTS

A total of 518 consecutive women attending a routine second-trimester anomaly scan were recruited at a mean ± SD gestational age of 20.3 ± 0.6 weeks. Of these, 22 (4.2%) were excluded because of technical problems with recordings, such as incomplete videos. Therefore, 496 second-trimester anomaly scans (one per woman) were included. Demographic characteristics of the participants are shown in Table 1.

Each scan was carried out by one of 14 operators, of whom 11 were accredited sonographers and three fetal medicine doctors, with a median of 3 years (range,

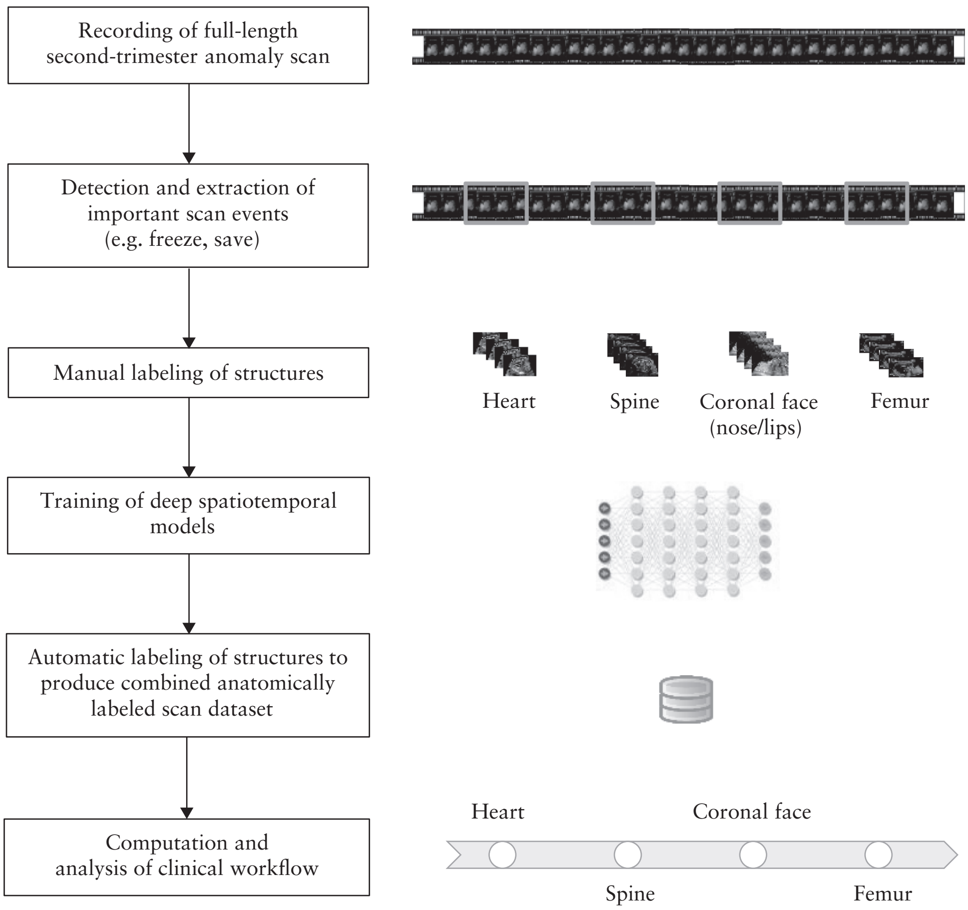

Figure 1 Outline of clinical workflow analysis pipeline.

4 months to 14 years) of clinical postaccreditation experience in sonography (Table 1). The mean ± SD anomaly scan duration was 32.1 ± 11.4 min, representing a total of approximately 265 h of scan video. The mean ± SD number of important scan events was 39 ± 13 per scan, and these were labeled according to the anatomical structure.

Table 1 Characteristics of 496 women with singleton pregnancy and 14 ultrasound operators included in study cohort

Characteristic	*Value*
Pregnant women (n = 496)	
Maternal age (years)	31.8 ± 5.5
Smoker at presentation	39/419 (9.3)
BMI at < 15 weeks (kg/m^2)	24.9 ± 4.8
Conception by IVF	5/417 (1.2)
Nulliparous	200/440 (45.5)
Pregnancy dating by CRL	409/440 (93.0)
GA at fetal anomaly scan (weeks)	20.3 ± 0.6
Operators (n = 14)	
Gender	
Female	12 (85.7)
Male	2 (14.3)
Clinical experience in scanning	
< 2 years	3 (21.4)
2–5 years	3 (21.4)
5–10 years	6 (42.9)
> 10 years	2 (14.3)
Accreditation	
Sonographer	11 (78.6)
Fetal medicine doctor	3 (21.4)

Data are given as mean ± SD, n/N (%) or n (%). BMI, body mass index; CRL, crown–rump length; GA, gestational age; IVF, *in-vitro* fertilization.

The workflow for 24 randomly selected representative anomaly scans is shown in Figure 2. Videoclip S1 shows a representative scan sequence and Table 2 presents the detection metrics of the different structures across the entire set of video scans. The structures most commonly captured during anomaly scans were thorax/heart, spine and maternal anatomy (including uterine artery), with 3254, 2430 and 1564 respective occurrences in the 496 scans. The five acquisitions that are required by the FASP guidelines to be stored as an image or clip were detected by the automated annotation technique in the majority of scans: head/brain was detected in 97.2% of scans, coronal face view (nose/lips) in 86.1%, abdomen in 93.1%, spine in 95.0% and femur in 92.3% (Table 2). For structures not included in the mandatory capture list, the prevalence of automatic detection was 78.8% for sagittal face (profile), 95.4% for thorax/heart, 50.2% for kidneys, 49.6% for 3D/4D face and 74.4% for placenta/amniotic fluid.

In 461 (92.9%) scans, at least four of five mandatory structures (head/brain, coronal face (nose/lips), abdomen, spine and femur) were detected, and in 359 (72.4%) scans, all five mandatory structures were detected. Regarding both mandatory and non-mandatory structures (head/brain, sagittal face (profile), coronal face (nose/lips), thorax/heart, abdomen, kidneys, spine, femur, placenta/amniotic fluid and maternal anatomy), at least eight structures were detected automatically in 420 (84.7%) scans and at least nine were detected in 313 (63.1%) scans.

Scans in which a structure was detected automatically usually contained multiple capture episodes, suggesting that sonographers performed an acquisition, moved to a different structure, and returned to perform an additional acquisition of the already captured structure.

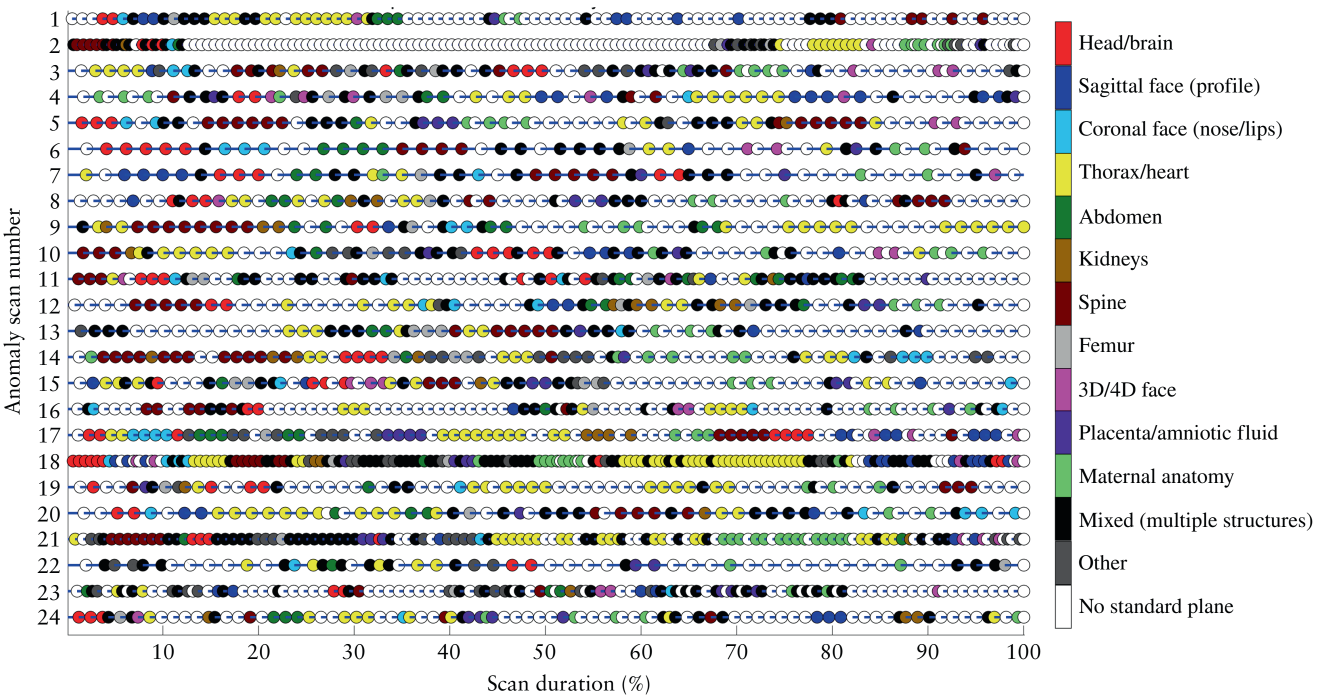

Figure 2 Sequence of anatomical planes or fetal structures obtained over time in 24 representative anomaly scans. Scan duration is normalized to percentage. 3D, three-dimensional; 4D, four-dimensional.

The structures most commonly captured multiple times were thorax/heart, spine and maternal anatomy, recorded on average seven, five and three times in each scan, respectively (Table 2).

Sonographers were most likely to begin their scan by capturing one of three structures: head/brain, spine or thorax/heart. The structures that were least likely to be captured first were kidneys, femur and coronal face (nose/lips) (Table 2). The structures most commonly evaluated last were maternal anatomy, sagittal face (profile) and spine, occurring in 111 (22.4%), 100 (20.2%) and 61 (12.3%) scans, respectively.

We also identified structure transitions, in which sonographers move from one anatomical area to another. The prevalence of two-structure transitions is presented in Table 3. Each of these transition pairs was performed by multiple sonographers, making it impossible to attribute a particular transition pair to a specific sonographer. The most commonly identified two-structure transitions were: placenta/amniotic fluid to maternal anatomy, occurring in 44.5% of scans; head/brain to coronal face (nose/lips) in 42.7%; abdomen to thorax/heart in 26.1%; and 3D/4D face to sagittal face (profile) in 23.7%. The most common three-structure transition sequences noted were: thorax/heart to abdomen to thorax/heart in 65 (13.1%) scans; head/brain to coronal face (nose/lips) to abdomen in 61 (12.3%); head/brain to coronal face to thorax/heart in 53 (10.7%); and placenta/amniotic fluid to maternal anatomy to placenta/amniotic fluid in 53 (10.7%). Four-structure transition sequences detected were: head/brain to coronal face to abdomen to thorax/heart in 21 (4.2%) scans; head/brain to coronal face to abdomen to femur in 19 (3.8%); head/brain to coronal face to thorax/heart to abdomen in 19 (3.8%); and thorax/heart to head/brain to coronal face to abdomen in 18 (3.6%). All of the most common three- and four-structure transition sequences were performed by at least three sonographers. None of the included scans shared an entirely identical pathway with another scan.

DISCUSSION

We present a novel large-scale data analysis of sonographer workflow. This was performed by capturing full-length ultrasound videos of routine clinical scans, identifying episodes of interest within the scan automatically and, using deep learning, detecting the structure present in each of these scan episodes. Therefore, we studied ultrasound anomaly scans longitudinally, as a data-science problem, and observed that each scan is a

Table 2 Structure capture, defined as screen freeze or image/videoclip save, in 496 second-trimester anomaly scans

Structure label	*Structure capture required*[14]	*Times structure was captured*	*Scans with structure capture detected*	*Mean number of captures per scan*	*Scans with structure captured as first structure*
Head/brain	Yes	1464	482 (97.2)	3.0	121 (24.4)
Sagittal face (profile)	No	1225	391 (78.8)	3.1	51 (10.3)
Coronal face (nose/lips)	Yes	793	427 (86.1)	1.9	10 (2.0)
Thorax/heart	No	3254	473 (95.4)	6.9	113 (22.8)
Abdomen	Yes	999	462 (93.1)	2.2	27 (5.4)
Kidneys	No	442	249 (50.2)	1.8	1 (0.2)
Spine	Yes	2430	471 (95.0)	5.2	115 (23.2)
Femur	Yes	645	458 (92.3)	1.4	8 (1.6)
3D/4D face	No	579	246 (49.6)	2.4	11 (2.2)
Placenta/amniotic fluid	No	845	369 (74.4)	2.3	21 (4.2)
Maternal anatomy	No	1564	475 (95.8)	3.3	18 (3.6)
Mixed (multiple structures)	—	4935	476 (96.0)	10.4	0 (0)
Other	—	1367	249 (50.2)	5.5	0 (0)

Data are given as *n* or *n* (%), unless stated otherwise. 3D, three-dimensional; 4D, four-dimensional.

Table 3 Prevalence of two-structure transition patterns in 496 second-trimester anomaly scans

	Subsequent structure										
Structure	*Head/ brain*	*Sagittal face (profile)*	*Coronal face (nose/lips)*	*Thorax/ heart*	*Abdomen*	*Kidneys*	*Spine*	*Femur*	*3D/4D face*	*Placenta/ AF*	*Maternal anatomy*
Head/brain	—	7.8	42.7	8.0	13.1	2.1	9.9	3.5	1.8	6.1	5.1
Sagittal face (profile)	12.2	—	6.3	18.3	8.4	1.0	9.9	5.0	20.8	10.3	8.0
Coronal face (nose/lips)	18.8	14.7	—	17.8	16.6	1.6	11.3	5.0	1.2	6.5	6.5
Thorax/heart	9.4	16.2	6.7	—	16.2	5.1	13.5	12.4	4.2	7.2	9.0
Abdomen	7.1	3.2	3.9	26.1	—	14.7	9.5	19.7	3.2	5.4	7.3
Kidneys	6.6	3.4	1.9	21.0	16.9	—	20.7	17.2	3.1	3.8	5.3
Spine	10.8	5.0	4.3	20.8	12.3	15.6	—	7.4	4.5	9.3	10.0
Femur	6.8	6.6	5.4	18.1	10.5	2.0	15.5	—	5.4	17.1	12.7
3D/4D face	11.9	23.7	7.9	12.9	3.6	1.4	13.7	5.0	—	3.6	16.2
Placenta/AF	8.3	12.1	3.1	8.9	6.9	0.4	7.5	6.0	2.3	—	44.5
Maternal anatomy	5.7	13.4	5.3	15.9	7.2	1.6	16.3	7.1	7.1	20.5	—

Data are given as %. 3D, three-dimensional; 4D, four-dimensional; AF, amniotic fluid.

non-ordered multistep process of anatomical structure acquisition.

Our real-world data analysis shows that sonographers usually begin their scan by acquiring one of three specific structures: head/brain, heart or spine, and rarely initiate scans by assessing structures such as the fetal femur or kidneys. This suggests that sonographers often start a scan by taking advantage of fetal position to maximize the possibility of imaging the fetal spine or heart, leaving the easier-to-acquire structures that can be visualized from various angles to later stages of the scan[20,21]. We also established that the most frequent imaging transitions occur between structures that are in proximity to each other, including placenta/amniotic fluid to maternal anatomy, head/brain to nose/lips, 3D/4D face to face profile and abdomen to thorax/heart. Three- and four-structure sequence patterns were not very common, and, overall, each of the anomaly scans included in the study followed a unique pattern.

The lack of a universal scanning sequence supports the somewhat opportunistic nature of anomaly scanning. Unlike non-obstetric ultrasound, obstetric ultrasound scans are restricted by fetal position and movement[22,23]. It is likely that while performing anomaly scans, sonographers have a preferred scanning sequence, and at the same time, take advantage of the fetal position, capturing structures that are visualized more favorably in each part of the scan. For example, while attempting to study the fetal thorax, if the fetus moves to a position that benefits fetal spine imaging, sonographers will use this opportunity before the fetus moves into a less favorable position. Trainees and infrequent users may benefit from such knowledge by appreciating that even experts do not follow guidelines as an ordered list, but rather make the best of the fetal position in each part of the scan. This understanding could reduce trainee anxiety and save valuable training time.

It is difficult to compare our findings with those of previous studies. Until recently, longitudinal analysis of ultrasound scans was hampered by the lack of an automated way to curate the data. However, large-scale datasets of full-length videos can now be processed effectively using deep-learning-based analysis methods. Previous work reported on methods for the automatic detection and localization of fetal structures or standard planes, with comparable detection rates to those presented in the current work, albeit on still images[24–29]. However, no prior work has assessed sonographer workflow as in this study.

Ultrasound departments are challenged consistently by increasing patient workload, poor recruitment and retention of qualified sonographers and repetitive stress injuries[30–32]. The ascertainment that the anomaly scan acquisition sequence is not predefined highlights the importance of workflow analysis. Automated sonographer workflow analysis holds the potential to improve our understanding of the scanning process itself, of sonographer trainee learning curves and of sonographer skills, and to measure quantitatively the competency of qualified sonographers. Currently, common methods to assess trainees and supervise certified sonographers include observing the scan in real time, measuring scan duration and manual postscan image review by a clinical expert. These methods of assessment are labor-intensive and of unknown reproducibility, yet are mandatory to provide adequate quality assurance. In the future, automated workflow analysis may aid the monitoring of trainee learning progress and identification of sonographers that require additional teaching. Ultimately, the hope is that this approach may contribute towards making ultrasound a more accessible technology to the non-expert across the world.

The main strength of this study is the novel automated analysis of ultrasound scans, which makes it possible to study anomaly scans as a data-science problem. Our analysis provides quantitative evidence of the non-ordered nature of anomaly scans. On the one hand, our analysis provides further confirmation of the importance of quality assurance and on the other, makes it possible to understand better the scanning process itself and how it may be improved. This study also has some expected limitations. We could not judge whether specific sequences perform better or result in shorter scanning times. This is because we purposefully collected prospective observational data on what happens during routine ultrasound in real-world settings with as little interference as possible. The particular sequences were carried out during normal practice and it is not possible to judge if a different sequence would have been better. Also, the study was conducted in a single maternity unit which may not represent practice at other centers. To reduce the risk of bias, we included 14 sonographers, and we noted that the findings were consistent, making external validity more likely. In addition, even though sonographers were aware that the scans were being recorded, they had not been informed of the aim of the current analysis, meaning that it is unlikely that they acted differently from their usual practice while participating in this study. Another limitation concerns the accuracy of automatic detection of structures. While most mandatory screening images defined by FASP[14] were detected automatically in the majority of scans, a relatively small proportion of these structures were not detected automatically in some scans. This may be a consequence of an incomplete scan[33], which could be due to a number of reasons including challenging fetal position, abdominal scarring, uterine fibroid, raised maternal body mass index (BMI) or simply sonographer fatigue or forgetfulness. However, it is also possible that missing structures were included in the mixed (multiple structures) label or otherwise misclassified by the algorithm. It should also be noted that the average BMI of women included in the study was 25, meaning the generalizability of our findings to women of higher BMI remains to be established[34].

In conclusion, sonographer workflow assessment allows the study of anomaly scanning as a data-science problem. The resulting analysis offers insights into increasing sonographer efficiency, improving human–computer interfaces with ultrasound machines and determining when and how automated analysis may assist manual scanning. Further research should evaluate the differences between trainees and expert sonographers and seek to establish workflow patterns that provide the best scanning results.

ACKNOWLEDGMENTS

We thank the pregnant women, sonographers, doctors and research midwives who participated in this study. Funding for this study was granted by the European Research Council (ERC-ADG-2015 694581, project PULSE).

Disclosure

J.A.N. and A.T.P. are Senior Scientific Advisors of Intelligent Ultrasound Ltd. A.T.P. is supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre, Oxford, UK. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

REFERENCES

1. Abu-Rustum RS, Abuhamad AZ. Fetal imaging: past, present, and future. A journey of marvel. *BJOG* 2018; **125**: 1568.
2. Benacerraf BR, Minton KK, Benson CB, Bromley BS, Coley BD, Doubilet PM, Lee W, Maslak SH, Pellerito JS, Perez JJ, Savitsky E, Scarborough NA, Wax J, Abuhamad AZ. Proceedings: Beyond Ultrasound First Forum on improving the quality of ultrasound imaging in obstetrics and gynecology. *Am J Obstet Gynecol* 2018; **218**: 19–28.
3. Sorrentino K. Accreditation, Credentialing, and Quality Improvement in Diagnostic Medical Sonography: A Literature Review. *J Diagn Med Sonogr* 2019; **35**: 401–411.
4. Salomon LJ, Winer N, Bernard JP, Ville Y. A score-based method for quality control of fetal images at routine second-trimester ultrasound examination. *Prenat Diagn* 2008; **28**: 822–827.
5. Espinoza J, Good S, Russell E, Lee W. Does the use of automated fetal biometry improve clinical work flow efficiency? *J Ultrasound Med* 2013; **32**: 847–850.
6. Burnett D, Campbell-Kyureghyan N. Quantification of scan-specific ergonomic risk-factors in medical sonography. *Int J Ind Ergon* 2010; **40**: 306–314.
7. Karim JN, Roberts NW, Salomon LJ, Papageorghiou AT. Systematic review of first-trimester ultrasound screening for detection of fetal structural anomalies and factors that affect screening performance. *Ultrasound Obstet Gynecol* 2017; **50**: 429–441.
8. Jone PN, Gould R, Barrett C, Younoszai AK, Fonseca B. Data-Driven Quality Improvement Project to Increase the Value of the Congenital Echocardiographic Report. *Pediatr Cardiol* 2018; **39**: 726–730.
9. Salomon LJ, Alfirevic Z, Berghella V, Bilardo CM, Chalouhi GE, Da Silva Costa F, Hernandez-Andrade E, Malinger G, Munoz H, Paladini D, Prefumo F, Sotiriadis A, Toi A, Lee W. ISUOG Practice Guidelines (updated): performance of the routine mid-trimester fetal ultrasound scan. *Ultrasound Obstet Gynecol* 2022; **59**: 840–856.
10. AIUM-ACR-ACOG-SMFM-SRU Practice Parameter for the Performance of Standard Diagnostic Obstetric Ultrasound Examinations. *J Ultrasound Med* 2018; **37**: E13–24.
11. Drukker L, Noble JA, Papageorghiou AT. Introduction to artificial intelligence in ultrasound imaging in obstetrics and gynecology. *Ultrasound Obstet Gynecol* 2020; **56**: 498–505.
12. Xie HN, Wang N, He M, Zhang LH, Cai HM, Xian JB, Lin MF, Zheng J, Yang YZ. Using deep-learning algorithms to classify fetal brain ultrasound images as normal or abnormal. *Ultrasound Obstet Gynecol* 2020; **56**: 579–587.
13. Drukker L, Droste R, Chatelain P, Noble JA, Papageorghiou AT. Expected-value bias in routine third-trimester growth scans. *Ultrasound Obstet Gynecol* 2020; **55**: 375–382.
14. Public Health England. Fetal anomaly screening programme: standards. https://www.gov.uk/government/publications/fetal-anomaly-screening-programme-standards.
15. Drukker L, Sharma H, Droste R, Alsharid M, Chatelain P, Noble JA, Papageorghiou AT. Transforming obstetric ultrasound into data science using eye tracking, voice recording, transducer motion and ultrasound video. *Sci Rep* 2021; **11**: 14109.
16. O'Gorman N, Tampakoudis G, Wright A, Wright D, Nicolaides KH. Uterine artery pulsatility index at 12, 22, 32 and 36 weeks' gestation in screening for pre-eclampsia. *Ultrasound Obstet Gynecol* 2016; **47**: 565–572.
17. Thilaganathan B, Wormald B, Zanardini C, Sheldon J, Ralph E, Papageorghiou AT. Early-pregnancy multiple serum markers and second-trimester uterine artery Doppler in predicting preeclampsia. *Obstet Gynecol* 2010; **115**: 1233–1238.
18. Drukker L, Droste R, Chatelain P, Noble JA, Papageorghiou AT. Safety Indices of Ultrasound: Adherence to Recommendations and Awareness During Routine Obstetric Ultrasound Scanning. *Ultraschall Med* 2020; **41**: 138–145.
19. Sharma H, Drukker L, Chatelain P, Droste R, Papageorghiou AT, Noble JA. Knowledge representation and learning of operator clinical workflow from full-length routine fetal ultrasound scan videos. *Med Image Anal* 2021; **69**: 101973.
20. AIUM Practice Parameter for the Performance of Fetal Echocardiography. *J Ultrasound Med* 2020; **39**: E5–16.
21. Souka AP, Pilalis A, Papastefanou I, Eleftheriadis M, Papadopoulos G. Quality assessment of the detailed anomaly ultrasound scan. *J Matern Fetal Neonatal Med* 2019; **32**: 666–670.
22. Sanyal R, Kraft B, Alexander LF, Ismail A, Lockhart ME, Robbin ML. Scanner-Based Protocol-Driven Ultrasound: An Effective Method to Improve Efficiency in an Ultrasound Department. *AJR Am J Roentgenol* 2016; **206**: 792–796.
23. Ye H, Hang J, Chen X, Di X, Chen J, Ye X, Zhang D. An intelligent platform for ultrasound diagnosis of thyroid nodules. *Sci Rep* 2020; **10**: 13223.
24. Baumgartner CF, Kamnitsas K, Matthew J, Fletcher TP, Smith S, Koch LM, Kainz B, Rueckert D. SonoNet: Real-Time Detection and Localisation of Fetal Standard Scan Planes in Freehand Ultrasound. *IEEE Trans Med Imaging* 2017; **36**: 2204–2215.
25. Maraci MA, Napolitano R, Papageorghiou A, Noble JA. Fisher vector encoding for detecting objects of interest in ultrasound videos. IEEE 12th International Symposium on Biomedical Imaging (ISBI), 2015; 651–654.
26. Yaqub M, Kelly B, Papageorghiou AT, Noble JA. Guided Random Forests for Identification of Key Fetal Anatomy and Image Categorization in Ultrasound Scans. In *Medical Image Computing and Computer-Assisted Intervention*, Navab N, Hornegger J, Wells W, Frangi A (eds). MICCAI 2015, Lecture Notes in Computer Science. Springer: Cham, 2015; 687–694.
27. Chen H, Wu L, Dou Q, Qin J, Li S, Cheng J, Ni D, Heng P. Ultrasound Standard Plane Detection Using a Composite Neural Network Framework. *IEEE Trans Cybern* 2017; **47**: 1576–1586.
28. Burgos-Artizzu XP, Coronado-Gutierrez D, Valenzuela-Alcaraz B, Bonet-Carne E, Eixarch E, Crispi F, Gratacos E. Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes. *Sci Rep* 2020; **10**: 10200.
29. Lin M, He X, Guo H, He M, Zhang L, Xian J, Lei T, Xu Q, Zheng J, Feng J, Hao C, Yang Y, Wang N, Xie H. Use of real-time artificial intelligence in detection of abnormal image patterns in standard sonographic reference planes for screening for fetal intracranial malformations. *Ultrasound Obstet Gynecol* 2022; **59**: 304–316.
30. Centre for Workforce Intelligence. Sonography workforce supply review. https://www.gov.uk/government/publications/review-of-the-sonography-workforce-supply.
31. Britton N, Miller MA, Safadi S, Siegel A, Levine AR, McCurdy MT. Tele-Ultrasound in Resource-Limited Settings: A Systematic Review. *Front Public Health* 2019; **7**: 244.
32. Singh N, Knight K, Wright C, Baird M, Akroyd D, Adams RD, Schneider ME. Occupational burnout among radiographers, sonographers and radiologists in Australia and New Zealand: Findings from a national survey. *J Med Imaging Radiat Oncol* 2017; **61**: 304–310.
33. Byrne JJ, Morgan JL, Twickler DM, McIntire DD, Dashe JS. Utility of follow-up standard sonography for fetal anomaly detection. *Am J Obstet Gynecol* 2020; **222**: 615.e1–9.
34. Yaqub M, Kelly B, Noble JA, Papageorghiou AT. The effect of maternal body mass index on fetal ultrasound image quality. *Am J Obstet Gynecol* 2021; **225**: 200–202.

SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:

Appendix S1 Training and testing of deep-learning model for automated video annotation

Videoclip S1 Important events detected in one representative full-length anomaly ultrasound scan, shown as a sequence of fetal structures over time.