Teleultrasound for pre-natal diagnosis: A validation study

Nader Z. Rabie[1], **Adam T. Sandlin**[2], **Song Ounpraseuth**[3], **Wendy N. Nembhard**[4], **Curtis Lowery**[2], **Kelly San Miguel**[2] and **Everett "Pat" Magann**[2]

[1]Department of Ob-Gyn, Tripler Army Medical Center, 1 Jarrett White Road, Honolulu, Hawaii, 96859, USA
[2]Department of Obstetrics and Gynecology, University of Arkansas for Medical Sciences, 4301 West Markham Street, Little Rock, Arkansas, 72205, USA
[3]Department of Biostatistics, University of Arkansas for Medical Sciences, 4301 West Markham Street, Little Rock, Arkansas, 72205, USA
[4]Arkansas Children's Research Institute, 1 Children's Way, Little Rock, Arkansas, 72202, USA

Abstract

Introduction/Purpose: There are no large validation trials comparing teleultrasound to on-site ultrasound. We aim to compare the sensitivity and accuracy of teleultrasound and demonstrate that teleultrasound is not inferior to on-site ultrasound in the pre-natal diagnosis of fetal anomalies.

Methods: All targeted ultrasounds performed between November 2010 and December 2012 were considered. We excluded studies performed at less than 17 weeks' gestation, on multiple gestations and for reasons other than an anatomical survey. Post-natal diagnoses were obtained from a state level mandatory birth defects surveillance programme. Descriptive statistics (sensitivity, specificity, positive and negative predictive values and accuracy) were calculated for both groups. A test of non-inferiority was performed, with the non-inferiority difference set at 0.15.

Results: The teleultrasound and on-site ultrasound groups consisted of 2368 and 3145 studies, respectively. The sensitivity of teleultrasound and on-site ultrasound was 57.46% and 76.57%, and the accuracy was 95.9% and 90.97%, respectively. The observed sensitivity difference was −0.1911. The accuracy, specificity, positive and negative predictive values of teleultrasound are similar to on-site ultrasound.

Discussion: Teleultrasound is inferior to on-site ultrasound in the detection of fetal anomalies; however, it has improved accuracy, as well as higher negative and positive predictive values. A negative teleultrasound is more likely to identify a non-anomalous fetus, and a positive teleultrasound is more likely to correctly identify an anomalous fetus.

Conclusion: Teleultrasound has an important role in pre-natal diagnosis for those patients unable or unwilling to travel for an on-site ultrasound.

Keywords: pregnancy, pre-natal ultrasound, fetal anomalies, teleultrasound, telemedicine, pre-natal diagnosis.

Introduction

Telemedicine has existed for decades, but has been limited by the cost and availability of the necessary infrastructure.[1–4] Recently, telemedicine protocols have become increasingly popular as network bandwidth has increased in capacity while decreasing in cost. This is especially true in rural locations where it is more convenient and cost-effective to provide telemedicine services than transport patients to higher echelons of care.[1,5]

While interpretation of many non-obstetric radiological studies is done remotely, pre-natal ultrasound has traditionally been performed and interpreted in the same location. This is true for several reasons. Many interpreting physicians prefer to perform at least part of the study themselves, and the ability to view

Correspondence to email nader.rabie@me.com
doi: 10.1002/ajum.12175
The views expressed in this manuscript are those of the author(s) and do not reflect the official policy or position of the Department of the Army, Department of Defense, or the US Government.
This manuscript was presented at the 2018 Society for Maternal-Fetal Medicine Annual Meeting in Dallas, Texas on 01 February 2018.

remote, real-time ultrasound has been limited by technology. In addition, the ability to perform ultrasounds is dependent on both a highly skilled technician and quality equipment.[3,4] Several small studies have examined the utility of remotely interpreted pre-natal ultrasound and found a high degree of accuracy and patient satisfaction.[5–9] We recently compared the sensitivity and accuracy of 2368 teleultrasounds with the results of published literature for on-site ultrasounds and observed that they were similar.[10] To our knowledge, there are no large studies validating the accuracy of teleultrasound compared with on-site ultrasound in detecting congenital anomalies. We hypothesise that the sensitivity and accuracy of teleultrasound are not inferior to that of on-site ultrasound.

Methods and materials

Ultrasounds were excluded if they were performed prior to 17 weeks of gestation, if there was a multiple gestation or if the ultrasound was performed for any reason other than an anatomical survey (e.g. follow-up growth, Doppler evaluation, etc.). On-site ultrasound refers to ultrasounds performed in the UAMS Pre-natal Genetics Clinic, where the interpreting Maternal-Fetal Medicine (MFM) physician has immediate access to the ultrasound images, and the ability to perform hands-on scanning. Teleultrasound refers to ultrasounds performed at one of several locations remote from UAMS.

The previously published study describes the details of ultrasound performance, quality control of sonographers, generation of the study cohort and data extraction.[10] This is briefly summarised here. The same group of 13 sonographers performed ultrasounds throughout the state, and there is continuous quality control with additional training and supervision as needed. The interpreting MFM was provided with still images, cine clips and if requested, real-time viewing of the teleultrasound. The ultrasounds were viewed and interpreted using Viewpoint 5 software (Solingen, Germany).

The study cohort was defined by the overlap of data from ultrasound reports (available from November 2010) and birth registry outcome data (available through August 2012). Birth outcome data were available from the Arkansas Reproductive Health Monitoring System (ARHMS). All ultrasound reports were manually reviewed. Table S1 lists the categories and specific fetal anomalies. This list is based on the official list of defects monitored by ARHMS, which are coded using the British Pediatric Association extension of the International Classification of Diseases, Ninth Edition Clinical Modification (ICD-9-CM) coding system, as modified by the Division of Birth Defects and Developmental Disabilities of the Centers for Disease Control and Prevention[11] and by ARHMS.

All statistical analyses were performed using SAS version 9.4 (SAS Institute Inc., Cary, NC, USA). Data analysis was performed on the teleultrasound and on-site ultrasound data separately, and then, a test for non-inferiority was performed. Descriptive statistics for continuous variables are expressed as the mean $\pm$ standard deviation (SD) or median and interquartile range (IQR), where appropriate. We calculated congenital anomaly prevalence, accuracy, sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, positive predictive value (PPV) and negative predictive value (NPV), each with their respective 95% confidence intervals (CIs). Accuracy is the proportion of correctly identified cases (the sum of the true positives and true negatives). The reference standard was defined as the presence or absence of congenital anomaly(s) at birth as identified by ARHMS. The level of agreement in identifying congenital anomaly(s) through ultrasound with the presence or absence of anomaly(s) at birth was measured using Cohen's kappa statistic (κ) with levels of agreement characterised by Landis and Koch[12] as slight agreement (0–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80) and almost perfect agreement (0.81–1.00). A test of non-inferiority evaluating the differences in accuracy, sensitivity and specificity between on-site and teleultrasound was performed separately using the Farrington and Manning's likelihood score test of the difference.[13] The conjectured non-inferiority difference was set at 0.15.

Ethics approval

We conducted a retrospective observational cohort study approved by the Institutional Review Board of the University of Arkansas for Medical Sciences (UAMS) (IRB 136448). All comprehensive obstetric ultrasounds performed between November 2010 and December 2012 were considered for analysis.

Results

During the study period of November 2010 through August 2012, 3404 ultrasounds were performed on-site and 2499 ultrasounds were performed remotely. 3145 of the on-site ultrasounds met inclusion criteria, while 2368 teleultrasounds met inclusion criteria. Exclusion categories are listed in Table 1. Demographic data for both groups are presented in Table 2. Basic diagnostic statistics of the two study populations are listed in Table 3.

The results of the teleultrasound arm have been previously described and are summarised here.[10] The congenital anomaly

Table 1: Exclusion criteria.

Category	On-site ultrasound	Teleultrasound
Multiple Gestation	155	112
Gestational age < 17 weeks	41	8
Indication other than anatomic survey or follow-up	63	8
Duplicate entries	0	3

Table 2: Demographic characteristics.

Measures	On-Site Ultrasound	Teleultrasound	P-value
Age, mean ± SD	28.95 ± 6.97	29.0 ± 7.0	0.8535
Gravidity, mean ± SD	2.84 ± 1.93	2.96 ± 1.83	0.0316
Parity, mean ± SD	1.27 ± 1.38	1.39 ± 1.41	0.0028
Gestation at 1st ultrasound, mean ± SD	22.58 ± 4.54	23.36 ± 4.44	<0.0001

prevalence in the on-site arm was 11.13%, while the teleultrasound arm had a prevalence of 5.66%. The sensitivity of on-site ultrasound was 76.57%, while for teleultrasound it was 57.46%. The specificity of on-site ultrasound was 92.77%, while for teleultrasound it was 98.21%. Using a non-inferiority limit of 0.15 with a 90% CI, the observed sensitivity difference was −0.1911; therefore, the sensitivity of teleultrasound is inferior to that of on-site ultrasound. Both accuracy and specificity of teleultrasound were not inferior to that of on-site ultrasound.

Table 4 displays the number of false-positive and false-negative diagnoses by category. False positives refer to those anomalies detected on pre-natal ultrasound, but not present at birth, while false negatives refer to those anomalies present at birth but not detected pre-natally. We excluded certain anomalies from the false-negative calculation. For example, we excluded patent ductus arteriosus and patent foramen ovale because these are not pre-natal diagnoses. We also excluded atrial septal defects due to their exceedingly low detection rate.[14] Similarly, we excluded congenital hypertrophic pyloric stenosis, anal atresia and hypospadias because these are not typically diagnosed pre-natally.[15,16]

We performed a secondary manual analysis in order to determine accuracy. In the initial analysis, all cases where an anomaly was present on ultrasound and at birth were considered *true positives*. This analysis did not account for whether the anomaly was correctly identified; only whether an anomaly was present or not. Therefore, all *true positives* were manually examined and only those that were correctly identified were considered *true positives*. The remaining 'true positives' were handled as noted in Table 5.

For example, a pre-natal diagnosis of complex heart defect and a post-natal diagnosis of transposition of the great vessels were considered a correct diagnosis and therefore, a *true positive*. A pre-natal diagnosis of omphalocele and a post-natal diagnosis of a complex heart defect was considered a *false positive*. Similarly, if an omphalocele was detected pre-natally and there was both an omphalocele and a complex heart defect at birth, this was considered a *true positive* because at least one major anomaly was correctly identified. However, if mild ventriculomegaly was detected pre-natally and both mild ventriculomegaly and a complex heart defect were present at birth, this was considered a *false negative*, because the major anomaly was missed.

After the secondary analysis was complete, we were able to determine accuracy (true positives + true negatives divided by the total number of cases). The accuracy of on-site ultrasound was 90.97% while for teleultrasound it was 95.9%.

Discussion

Teleultrasound is increasing in popularity and use, particularly in Arkansas.[17] Teleultrasound is well established at our institution; therefore, we were able to compare a large number of teleultrasound and on-site ultrasound studies performed by the same sonographers, from the same population and interpreted by the same Maternal-Fetal Medicine physicians. Our aim was to demonstrate that the detection rate (sensitivity) of teleultrasound was at least equivalent to that of on-site ultrasound.

Table 3: Statistics.

Statistic	On-site ultrasound		Teleultrasound	
	Value	CI	Value	CI
Kappa	0.6030	0.5614–0.6447	0.5920	0.5183–0.6658
Sensitivity	76.57%	71.78–80.91%	57.46%	48.63–65.96%
Specificity	92.77%	71.75–93.71%	98.21%	97.57–98.72%
Positive predictive value	57.02%	53.44–60.53%	65.81%	57.98–73.00%
Negative predictive value	96.93%	96.32–97.45%	97.47%	96.93–97.91%
Disease prevalence	11.13%	10.05–12.28%	5.66%	4.76–6.67%
Accuracy	90.97%	89.91–91.95%	95.9%	

Table 4: False positives and false negatives by system.

System	False positive		False negative	
	On-Site	Teleultrasound	On-site	Teleultrasound
CNS	21	6	7	2
Cardiac	98	5	59	42
Pulmonary†	11	0	4	0
GI‡	11	1	7	0
Renal	25	16	10	8
Midline§	10	8	10	7
Extremities¶	26	8	33	10
Other††	20	2	2	0

†Includes congenital pulmonary airway malformation and congenital diaphragmatic hernia.
‡Includes gastroschisis and omphalocele.
§Includes cleft lip, cleft palate and abdominal wall defects (other than gastroschisis and omphalocele).
¶Includes musculoskeletal anomalies (e.g. skeletal dysplasia, club foot, etc.).
††Includes cystic hygroma, hydrops, amniotic band sequence and umbilical cord anomalies.

While this was not the case, we found that teleultrasound was not inferior in all other statistical measures, including specificity, positive and negative predictive value and accuracy.

The difference in the prevalence of anomalies is significant between the two groups (on-site: 11.13%, teleultrasound: 5.66%), and both of these values are greater than the general population. This is not unexpected, as both groups represent a referral population. First, the vast majority of Maternal-Fetal Medicine referrals from the entire state are seen by the UAMS Pre-natal Genetics Clinic. Second, there is probably some element of referral (selection) bias between the two groups. Patients are referred to the UAMS Pre-natal Genetics clinic primarily for one of two reasons—either an anomaly was suspected on a routine ultrasound or the patient has a significant risk factor (e.g. advanced maternal age, pregestational diabetes, etc.) for a fetal anomaly. All patients are offered a teleultrasound consultation; however, it is possible that patients from the first group, with known anomalies (especially major, surgical anomalies), are more likely to travel to a tertiary care centre for an on-site consultation, as opposed to those who need a targeted ultrasound for risk factors alone.

Teleultrasound has a higher specificity, which is reflected in a lower false-positive rate. These represent cases where the ultrasound diagnosed anomaly was not present at birth (Tables 4 and 5). This is an interesting finding, because it implies that during teleultrasound the sonographer and physician are less likely to 'overcall' a finding. This could also reflect a decreased willingness to make a subtle diagnosis without directly scanning the patient. On-site ultrasound had a lower specificity, also reflected by an increased number of false positives. In the on-site group, almost half of the false positives were presumed cardiac anomalies, and about half of those were atrial septal defects (ASDs) and ventricular septal defects (VSDs). This is not surprising, as these are difficult to detect and often resolve prior to delivery. One might argue these are true positives, as they were present in the fetus; however, the definitive diagnosis is the post-natal diagnosis.[14]

The false negatives represent the limitations of ultrasound and part of the denominator of sensitivity. This is where we aimed to demonstrate that teleultrasound was not inferior to on-site ultrasound; however, our results did not support this. In the teleultrasound group, 42.5% anomalies were missed, while only 23.4% of anomalies were missed in the on-site group. This difference is greater than our expected difference of 0.15. This may be because teleultrasound is inferior, but the previously mentioned selection bias is also a factor. Potentially more on-site anomalies would have been missed if they had not been previously detected at a routine ultrasound (which then prompted the referral for an on-site examination). Both groups were similar in the types of anomalies that were missed, with the vast majority being cardiac anomalies and half of these were ASDs and VSDs.

Interestingly, the teleultrasound group had a higher accuracy than the on-site group. This has to be interpreted with caution; however, because of how accuracy is calculated. Accuracy is the sum of true positives and true negatives divided by the total number of cases. Because the teleultrasound group had a lower prevalence of anomalies and a high true negative rate, the accuracy is higher. Teleultrasound also has a higher positive

Table 5: Discrepancies.

Pre-natal diagnosis	Post-natal diagnosis	Discrepancy	Result
Major Anomaly A	Major Anomaly B	The two anomalies are different	False positive
Major Anomaly A	Major Anomaly A and undiagnosed minor anomalies	Minor anomalies were missed, but major anomaly was correctly diagnosed	True positive
Minor Anomaly A	Minor Anomaly A and undiagnosed major anomaly	The major anomaly was missed	False negative

predictive value. Therefore, while fewer anomalies were detected, those that were detected were more likely to be present at birth.

The primary limitation of our study is the retrospective nature. The AHRMS database uses a specific list of anomalies, while the ultrasound reports were not limited to this specific list. Every ultrasound report was manually reviewed and on occasion, a clear diagnosis was not provided (e.g. a cystic structure in the fetal abdomen, not otherwise specified). While this is appropriate for clinical use, it made comparing the results difficult. A prospective study would have allowed consistent categorisation of each diagnosis. Strengths of our study include the unique nature and the large number of cases.

Conclusion

In conclusion, teleultrasound is inferior to on-site ultrasound in the detection of fetal anomalies. However, clinically, a negative teleultrasound is more likely to identify a non-anomalous fetus (specificity) and a positive teleultrasound is more likely to correctly identify an anomalous fetus (positive predictive value). Therefore, teleultrasound has an important role in pre-natal diagnosis for those patients unable or unwilling to travel for an on-site ultrasound.

Future research might include prospective studies comparing teleultrasound and on-site ultrasound. In addition, studies specifically looking at the detection of anomalies that require delivery in a tertiary care center would be especially relevant, since one of the tenets of telemedicine is providing care to remote and rural areas.

Authorship statement

The views expressed in this manuscript are those of the author (s) and do not reflect the official policy or position of the Department of the Army, Department of Defense, or the US Government.

Funding

The authors received no funding for this project and would like to acknowledge Kevin Barber, MD (University of Arkansas for Medical Sciences, no funding source, services provided complimentary) for his assistance with data collection.

Conflicts of interest

The authors have no disclosures or conflicts of interest.

References

1 Magann EF, McKelvey SS, Hitt WC, Smith MV, Azam GA, Lowery CL. The use of telemedicine in obstetrics: a review of the literature. *Obstet Gynecol Surv* 2011; 66(3): 170–8.
2 Norum J, Bergmo TS, Holdo B, Johansen MV, Vold IN, Sjaaeng EE, et al. A tele-obstetric broadband service including ultrasound, videoconferencing and cardiotocogram. A high cost and a low volume of patients. *J Telemed Telecare.* 2007; 13(4): 180–4.
3 Landwehr JB Jr, Zador IE, Wolfe HM, Dombrowski MP, Treadwell MC. Telemedicine and fetal ultrasonography: assessment of technical performance and clinical feasibility. *Am J Obstet Gynecol* 1997; 177(4): 846–8.
4 Begg L, Chan FY, Edie G, Hockey R, Wootton R. Minimum acceptable standards for digital compression of a fetal ultrasound videoclip. *J Telemed Telecare* 2001; 7(Suppl 2): 88–90.
5 Reddy ER, Bartlett PJ, Harnett JD, McManamon PJ, Snelgrove C. Telemedicine and fetal ultrasonography in a remote Newfoundland community. *CMAJ* 2000; 162(2): 206–7.
6 Chan FY, Soong B, Watson D, Whitehall J. Realtime fetal ultrasound by telemedicine in Queensland. A successful venture? *J Telemed Telecare* 2001; 7(Suppl 2): 7–11.
7 Wyatt SN, Rhoads SJ, Green AL, Ott RE, Sandlin AT, Magann EF. Maternal response to high-risk obstetric telemedicine consults when perinatal prognosis is poor. *Australian New Zealand J Obst Gynaecol* 2013; 53(5): 494–7.
8 Chan FY, Soong B, Lessing K, Watson D, Cincotta R, Baker S, et al. Clinical value of real-time tertiary fetal ultrasound consultation by telemedicine: preliminary evaluation. *Telemed J* 2000; 6(2): 237–42.
9 Hishitani T, Fujimoto Y, Saito Y, Sugamoto K, Hoshino K, Ogawa K. Accuracy of telediagnosis of fetal heart disease using ultrasound images transmitted via the internet. *Pediatr Int* 2014; 56(2): 289–91.
10 Rabie NZ, Sandlin AT, Barber KA, Ounpraseuth S, Nembhard W, Magann EF, et al. Teleultrasound: how accurate are we? *J Ultrasound Med* 2017; 36: 2329–2335.
11 (NBDPN) NBDPN. Guidelines for Conducting Birth Defects Surveillance. Atlanta, GA: National Birth Defects Prevention Network. Inc.; 2004.
12 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33(1): 159–74.
13 Farrington CP, Manning G. Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk. *Stat Med.* 1990; 9 (12): 1447–54.
14 Tegnander E, Williams W, Johansen OJ, Blaas HG, Eik-Nes SH. Prenatal detection of heart defects in a non-selected population of 30,149 fetuses–detection rates and outcome. *Ultrasound Obstet Gynecol* 2006; 27(3): 252–65.
15 Singh SJ, Trudinger B, Lam A, Zhang AL, Cass D. Antenatal prediction of hypertrophic pyloric stenosis. *Pediatr Surg Int* 2001; 17 (7): 560–2.
16 Cayan F, Cayan S. Prenatal diagnosis of penoscrotal hypospadias and review of the literature. *Turk J Urol.* 2013; 39(2): 116–8.
17 Lowery C, Bronstein J, McGhee J, Ott R, Reece EA, Mays GP. ANGELS and University of Arkansas for Medical Sciences paradigm for distant obstetrical care delivery. *Am J Obstet Gynecol* 2007; 196(6): 534.e1–534.e9.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1. List of fetal anomalies.