Title: Clinician-rated quality of video otoscopy recordings and still images for the asynchronous assessment of middle-ear disease.

Authors:

Eman M. A. Alenezi [1,2]

Kathryn Jajko [2]

Allison Reid [3]

Alessandra Locatelli-Smith [3]

Courtney S. E. McMahen [1,2]

Karina F. M. Tao [2]

Julie Marsh [1,2]

Tess Bright [6]

Peter C. Richmond [1,2,3]

Robert H. Eikelboom [4,5,7]

Christopher G. Brennan-Jones [1,2,3]

1. The University of Western Australia, Nedlands, Western Australia.
2. Telethon Kids Institute, The University of Western Australia, Nedlands, Western Australia.
3. Perth Children's Hospital, Perth, Western Australia.
4. Ear Science Institute Australia, Subiaco, Western Australia.
5. Ear Sciences Centre, The University of Western Australia, Nedlands, Western Australia.
6. International Centre for Evidence in Disability, London School of Hygiene & Tropical Medicine, Keppel St, London, WC1 E7HT.
7. Department of Speech Language Pathology and Audiology, University of Pretoria, South Africa.

Contact information:
Ms Eman M. A. Alenezi
Eman.al-enezi@telethonkids.org.au
Telethon Kids Institute, The University of Western Australia
Perth Children's Hospital, 15 Hospital Avenue
Nedlands, Western Australia 6009

Abstract

Background: Video otoscopy plays an important role in improving access to ear health services via means of telemedicine. This study investigated the clinician-rated quality of video otoscopy recordings and still images, and compared their suitability to make an asynchronous diagnosis of middle-ear diseases.

Methods: Two hundred and eighty video otoscopy image-recording pairs were collected from 150 children (aged six months to 15 years) by an Ear, Nose, and Throat (ENT) specialist, audiologists and trained research assistants and independently rated by an audiologist and an ENT surgeon. On a 5-point scale, clinicians rated the cerumen amount (from Minimal to Severe), field of view (from 0% to 100%), quality, focus, light (from Excellent to Very poor) and gave an overall rating (similarly, from Excellent to Very poor) for both still images and recordings. Clinicians were also asked about the suitability of the two procedures to make accurate diagnosis of middle-ear disease.

Results: The percentage of the video otoscopy recordings rated as 'good' or 'excellent' was greater than that of the still images ratings across all domains. The mean difference between the two otoscopic procedures ratings was significant across almost all domains ($p < 0.05$), except for the 'cerumen amount' domain. The suitability to make a diagnosis significantly improved when using recordings, compared to still images, for both the audiologist ($p = 0.004$) and the ENT surgeon ($p = 0.005$). Younger age was found to have a significant, negative impact on the quality of ratings across all domains ($p < 0.03$). The role of the tester conducting video otoscopy did not have a significant impact on the quality ratings.

Conclusion: Video otoscopy recordings were found to provide clearer views of the tympanic membrane and increase the ability to make diagnoses, compared to still images, for both audiologists and ENT surgeons. Research assistants with limited practice were able to obtain video otoscopy images and recordings that were comparable to the ones obtained by audiologists and an ENT specialist.

Introduction

Video otoscopy is a procedure that involves the use of an otoscope with an built-in video camera to obtain a view of the ear canal and tympanic membrane (TM). It plays an important role in improving access to ear health services via telemedicine[1-3] by allowing images of the ear canal and TM to be sent over distances to be assessed by a clinician located elsewhere.[4] It can therefore help save time and cost of travelling for both patients and specialists.[3] Video otoscopy is also a feasible and clinically appropriate procedure that can be performed by both physicians and non-physicians.[5, 6] Evidence suggests that health care facilitators with limited training are able to obtain video otoscopy images[7] and recordings of sufficient quality to enable asynchronous diagnosis of ear disease.[7, 8] Therefore, video otoscopy can help in the process of improving access to early identification of and intervention for ear disease.[7, 9]

In rural and remote settings, the quality of the still images obtained using video otoscopy obtained by clinicians or non-clinicians has been found to be rated as good or better [1] or at least adequate or acceptable, in the majority of the cases.[2, 3] Similar findings were also found with video otoscopy recordings assessed asynchronously.[8, 10]

The quality of the video otoscopy images is important, as this is key to being able to make an accurate diagnosis and develop an appropriate treatment plan.[6] The issue of reduced recording quality has been reported as one of the main reasons behind inability to asynchronously diagnose the ear health status of children in telehealth services.[8] It has also been known to have caused some discrepancy between onsite conventional otoscopy examination findings and asynchronous findings[1], and can affect the diagnostic confidence.[3] If clinicians are unable to accurately or confidently diagnose children with ear health concerns using otoscopy, this can lead to delays in treatment and may increase the need for additional visits and interventions that are in some cases unnecessary.

Both video otoscopy still images and recordings have been found to be useful for remote assessments. However, reduced perception of depth[2, 3] and difficulties in identifying retracted TM[3] are amongst some of the limitations of still images. On the other hand, video recordings have been found to enable enhanced perception of depth by viewing the TM from different angles[2, 7] and identify retracted and perforated TM by observing changes in the reflection of the otoscopy light on the TM.[5] Video recordings may also have an added advantage of being able to be replayed and paused[8], allowing for thorough examinations. Enhanced composite images using frame selection and stitching from video otoscopy recordings have been found to increase the diagnostic ability of ENT specialists.[11] Similar to still images, identifying TM retraction was challenging using those composite images. [12]

Previous studies have compared the quality and reliability of video otoscopy still images or recordings with the results of different gold standards including conventional onsite otoscopy[1, 7, 10] or otomicroscopy.[2, 3, 8, 13, 14] Another study has compared the diagnostic accuracy and confidence using video otoscopy the enhanced composite images stitched from video otoscopy recordings collected by experienced clinicians to the original recordings.[12] However, no study has compared the quality and ability to view the TM using still images to those of video recordings. It is still unknown whether video otoscopy still images or recordings can improve clinician-rated quality of otoscopy for asynchronous telehealth assessments and enable a more accurate diagnosis. Considering this question is important as timely and appropriate management of ear disease and hearing loss depend heavily on the ability to make a definitive and accurate diagnosis. It is particularly important as non-urgent face-to-face services are increasingly being restricted amid the outbreak of COVID-19.[15] The purpose of this study is to investigate whether there are differences between the clinician-rated quality of video otoscopy still images and 10 second video otoscopy recordings. We hypothesised that short video otoscopy recordings would be rated better by clinicians than still images, and increase the ability to diagnose. If so, this will inform the video otoscopic procedure that enables higher-quality telehealth assessments.

Methods

Participants

The participants in this study were 157 children aged between six months and 15 years who attended the Telethon Weekend Expo and Cockburn Integrated Health National Aborigines and Islanders Day Observance Committee (NAIDOC) public community events in Perth, Western Australia in October 2019. On the day of the event, the parents/guardians were asked to complete an electronic informed consent before the commencement of the assessment procedure to indicate that they were willing to allow their children to participate in the study. Only the children who had informed consent provided by their parents/guardians and had no contraindications, including recent ear surgeries or experiencing discharging ears in the past six weeks, were included.

Parental questionnaire

Following obtaining informed consent and prior to performing the video otoscopy and tympanometry, the parents/guardians completed a brief electronic questionnaire administered by a research assistant. They were asked to provide information about their child's demographics (e.g. age, sex, and postal code), ear health history and any concerns regarding the child's hearing and/or speech and language development. The questionnaire also included a question about the child's Aboriginal/Torres Strait Islander (TSI) status as middle ear infections are highly prevalent in this group of children.[16] Contraindications were identified in this questionnaire.

Equipment

The video otoscopy and tympanometry devices used in this study were selected for their suitability for telehealth procedures in that they are portable, can be performed easily by research assistants or telehealth facilitators and they enable the results to be securely stored and forwarded for asynchronous assessment. A group of testers included research assistants, audiologists, and ENT specialists used a video otoscope (hearScope, HearX Group, Pretoria, South Africa) to obtain a still image, and 10 second recording of the TM and ear canal for each ear on each child. The video otoscope consisted of a hearScope attachment to Samsung Galaxy A3 smart phone (2017) with a 4.7 inch screen (resolution 1280 x 720 pixels), which managed the display and capture of still images and recordings using the hearScope app version 2.0. Tympanometry was then conducted by either a research assistant or an audiologist using a Titan Middle Ear Analyser to assess the middle-ear health and the mobility of the TM. The tympanograms obtained by research assistants were reviewed and classified immediately by a study audiologist. Data on tympanometry and case history are not presented in this paper. Research assistants assigned to conduct the video-otoscopy and tympanometry received basic training from the study audiologists, which included practical training in the procedures and accurately recording the results. Their skills were assessed by the study audiologist post-training. For children less than two years of age, video otoscopy and tympanometry were carried out by an experienced audiologist, as it is often challenging to examine the ears of younger children. All the collected data were stored in a secure server for asynchronous evaluation and diagnosis.

Review procedure

All participants were assigned a unique ID number prior to any study procedure. Data were deidentified before the images and videos were rated. An audiologist, who was also one of the testers assigned to conduct video otoscopy, and ENT surgeon independently evaluated the images and recordings at ear level asynchronously and made a diagnosis at child level based on them. The images were reviewed using 13-inch MacBook Air (1440x900-pixel resolution). The reviewers were blinded to whether the still images and video recording related to the same child. They were also blinded to the case history information and tympanometry results when they were asked to make an initial diagnosis.

The review process comprised a rating of the image Quality, Focus and Light on a scale of one to five (1 being Excellent, 2 Good, 3 Adequate, 4 Poor and 5 Very poor). The Cerumen Amount was also evaluated on a scale of one to five, with 1 being Minimal and 5 being Severe. The Field of View was rated in a range from 0% to 100%. Lastly, an Overall rating on a scale of one to five (similarly scaled, with 1 being Excellent and 5 Very poor) was made. This process was done subjectively at ear level and was followed for both the still images and the recordings separately.

The second part of the process included an evaluation of the suitability for making an initial diagnosis at child level based on the video otoscopy still images and recordings, independently, was made by asking the reviewer about their ‘ability to make an accurate diagnosis’. The reviewers then had to subjectively indicate whether the findings were normal, abnormal non-clinically significant (i.e. no referral for treatment required) or abnormal-clinically significant (i.e. referral for treatment required), and to describe all the abnormal findings.

Finally, the clinicians reviewed the video otoscopy images or recordings together with the supporting information from the participant’s case history and tympanometry findings. Clinicians were asked whether they were able to make a diagnosis of the same participant using case history information and tympanograms in addition to video otoscopy findings and if so, to make a diagnosis. Only data on clinicians' ratings of the video otoscopy still images and recordings and their suitability for making a diagnosis are presented in this paper.

Referrals

The children with clinically significant abnormal findings found on the day of testing or after the results were reviewed received a referral to a general medical practitioner (GP) for further investigation and management. GP referral letters were also provided to the parents/guardians of the children with significantly abnormal results identified on the day of the event by the study audiologist, who was also available to respond to parents' questions and concerns.

Ethics

Ethical approval was obtained from the Child and Adolescent Health Service Human Research Ethics Committee, the Western Australia Aboriginal Health Ethics Committee, and The University of Western Australia Human Research Ethics Committee.

Analysis

All statistical analysis was carried out using IBM SPSS Statistics, Version 26. Study characteristics for continuous variables were summarised using mean and standard deviation for symmetric distributions and median and range for asymmetric distributions. Categorical variables were summarised as frequencies and percentages for categorical variables. Paired Student t-tests were used to compare mean rating differences between still images and recordings for the overall rating and each of the five domains (quality, focus, light, cerumen amount and field of view) separately for each ear, assuming asymptotic properties based on the sample size. McNemar's test was used to make paired comparisons of the ability to make a diagnosis between still images and video recordings. In addition, a hierarchical logistic regression was used to identify the covariates that impact the rating of images and recordings, incorporating compound symmetry for ear side nested within the child identifier (i.e. outcomes for the left and right ears are equally correlated with each other for each child). Recorded covariates include video otoscopic procedure, ear side, age, role of tester, and role of rater. Results are reported with parameter estimates and associated 95% confidence intervals based on a significance level of 5%.

Results

Characteristics of the study population

The parents of 157 children consented to the study. Seven children had no observations recorded due to problems saving the images or recordings or poor compliance. In addition, image–recording pairs (still image and recording obtained from the same ear) that were missing or incomplete and those with incomplete ratings in at least one domain were excluded from the analyses, resulting in a total of 280 video otoscopy image–recording pairs (Right ear n =140; Left ear n =140) being collected from 150 children (mean age of 7.21 years; standard deviation (SD) 3.4; 62% female) (Figure 1) including 16 Aboriginal or Torres Strait Islander children. Each of the video otoscopy still images and recordings were reviewed and rated by a single ENT surgeon and audiologist.

Forty-two per cent of the obtained images and recordings were collected by audiologists, 30% by an ENT specialist, and 28% by research assistants. The asynchronous diagnoses made by the ENT specialist and/or audiologist, using either video otoscopy recordings or still images along with case history information and tympanometry results, show that 18% of the children were found to have some middle ear abnormality, either non-clinically significant (e.g. tympanosclerosis) or clinically significant (e.g. middle ear effusion).

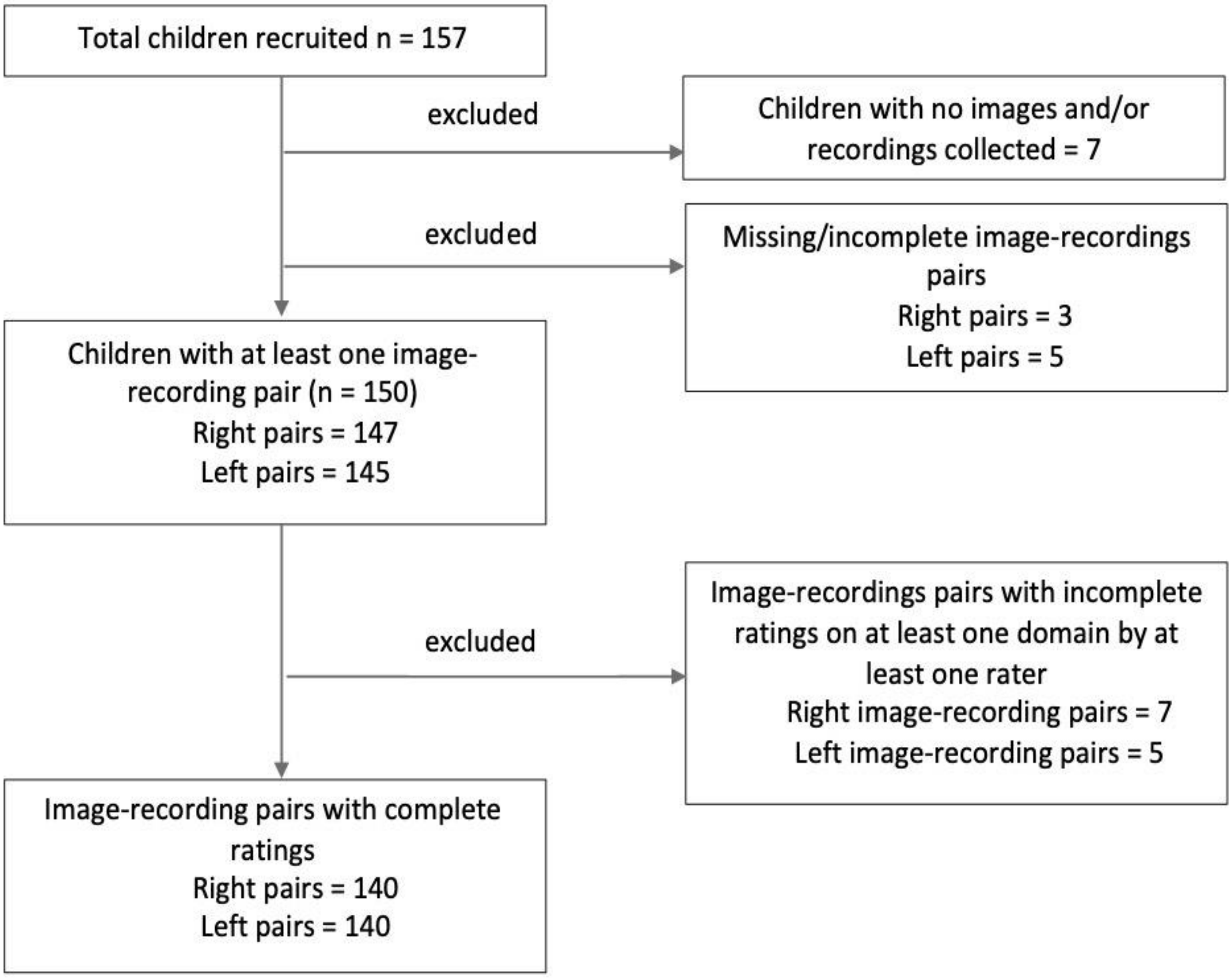

Figure 1. Flowchart illustrating the process of the inclusion/exclusion of participants and image-recording pairs.

Video otoscopy recordings consistently recorded better ratings, overall and across the five domains, compared to still images (Table 1). For the right ear, video otoscopy recordings were rated as better compared to still images for the ENT surgeon's Overall Rating (Table 2: 0.54 units, 95% confidence interval (CI) 0.36 – 0.71, $p < 0.001$). Similar findings were observed for the ENT surgeon across the domains of Quality, Focus, Light, and Field of View, in both ears. Consistent results were also observed for the audiologist except for the rating of the domain of Light on the right ear only, which did not meet statistical significance (95% CI -0.02 – 0.25, $p = 0.84$). No differences were observed in the rating of Cerumen Amount between recording and still images for either ENT specialist or audiologist in any ear ($p > 0.05$). A summary of the findings is shown in Table 2 and 3.

Table 1. Percentages of the ratings of all image-recording pairs across all the domains as rated by both clinicians (n=280)

Domain	Video otoscopic procedure	Rating Scale				
		Excellent	**Good**	**Adequate**	**Poor**	**Very poor**
Overall Rating	Image	5.1%	22.1%	20.9%	37.8%	14.0%
	Recording	18.7%	25.3%	18.2%	30.3%	7.6%
Quality	Image	5.3%	24.7%	23.8%	42.6%	3.6%
	Recording	17.6%	34.1%	22.0%	24.4%	1.9%
Focus	Image	4.8%	18.3%	24.5%	48.6%	3.8%
	Recording	17.5%	27.7%	24.6%	28.9%	1.4%
Light	Image	8.7%	29.3%	27.7%	31.2%	2.9%
	Recording	19.9%	31.8%	17.6	27.9%	2.8%
		Rating Scale				
		Minimal	**2**	**Moderate**	**4**	**Severe**
Cerumen Amount	Image	52.7%	22.7%	13.3%	6.4%	5.0%
	Recordings	56.8%	21.3%	10.7%	4.7%	5.6%
		Rating Scale				
		100%	**75%**	**50%**	**25%**	**0%**
Field of View	Image	11.0%	28.1%	20.8%	25.0%	15.1%
	Recording	27.5%	23.9%	18.7%	18.7%	11.2%

SD: standard deviation.

Table 2. Results from the paired comparison of rating domains between video otoscopy still images and recordings (right ear pairs, n=140). Lower scores indicate better ratings.

Domain	Rater	Recording Mean (SD)	Image Mean (SD)	Mean difference (SE)	95% CI of difference	*p* value
Overall Rating	ENT	2.64 (1.30)	3.18 (1.15)	0.54 (0.09)	(0.36 – .071)	<0.001
	Audiologist	2.99 (1.21)	3.41 (1.18)	0.42 (0.08)	(0.27 – 0.57)	<0.001
Quality	ENT	2.46 (1.25)	2.99 (1.18)	0.54 (0.09)	(0.35 – 0.72)	<0.001
	Audiologist	2.76 (0.88)	3.21 (0.87)	0.45 (0.07)	(0.31 – 0.0.59)	<0.001
Focus	ENT	2.60 (1.27)	3.14 (1.14)	0.54 (0.10)	(0.34 – 0.75)	<0.001
	Audiologist	2.88 (0.95)	3.33 (0.83)	0.45 (0.07)	(0.31 – 0.60)	<0.001
Light	ENT	2.51 (1.30)	2.84 (1.18)	0.32 (0.10)	(0.13 – 0.51)	0.001
	Audiologist	2.71 (0.98)	2.84 (0.92)	0.12 (0.07)	(-0.02 – 0.25)	0.084
Cerumen Amount	ENT	1.74 (1.05)	1.75 (1.03)	0.01 (0.07)	(-0.12 – 0.15)	0.837
	Audiologist	1.89 (1.23)	1.94 (1.23)	0.05 (0.08)	(-0.11 – 0.21)	0.546
Field of View	ENT	2.51 (1.28)	3.03 (1.12)	0.54 (0.09)	(0.37 – 0.72)	<0.001
	Audiologist	2.79 (1.42)	3.01 (1.45)	0.22 (0.09)	(0.05 – 0.39)	0.011

SD: standard deviation, SE: standard error, 95% CI: 95% confidence interval.

Table 3. Results from the paired comparison of rating domains between video otoscopy still images and recordings (left ear pairs, n=140). Lower scores indicate better ratings.

Domain	Rater	Recording Mean (SD)	Images Mean (SD)	Mean difference (SE)	95% CI of difference	*p* value
Overall Rating	ENT	2.59 (1.30)	3.23 (1.03)	0.64 (0.09)	(0.46 - 0.82)	<0.001
	Audiologist	2.99 (1.16)	3.48 (1.11)	0.49 (0.08)	(0.34 - 0.64)	<0.001
Quality	ENT	2.34 (1.23)	3.04 (1.08)	0.69 (0.10)	(0.50 -0 .88)	<0.001
	Audiologist	2.71 (0.89)	3.31 (0.86)	0.60 (0.07)	(0.46 - 0.74)	<0.001
Focus	ENT	2.42 (1.21)	3.21 (1.04)	0.79 (0.10)	(0.58 - 0.99)	<0.001
	Audiologist	2.79 (0.90)	3.42 (0.83)	0.64 (0.08)	(0.49 - 0.79)	<0.001
Light	ENT	2.43 (1.31)	2.91 (1.09)	0.48 (0.10)	(0.28 - 0.68)	<0.001
	Audiologist	2.74 (0.99)	2.97 (0.93)	0.24 (0.08)	(0.09 - 0.38)	0.002
Cerumen Amount	ENT	1.81 (1.23)	1.86 (1.12)	0.04 (0.07)	(-0.10 – 0.18)	0.550
	Audiologist	1.87 (1.28)	1.94 (1.27)	0.06 (0.08)	(-0.09 – 0.22)	0.415
Field of View	ENT	2.41 (1.29)	3.08 (1.11)	0.66 (0.10)	(0.46 - 0.87)	<0.001
	Audiologist	2.71 (1.41)	3.03 (1.34)	0.32 (0.08)	(0.17 - 0.48)	<0.001

SD: standard deviation, SE: standard error, 95% CI: 95% confidence interval.

Overall, the ENT surgeon reported that video otoscopy recordings was suitable to make an accurate diagnosis in 53.3% of the children compared to 42.7% when using still images (Table 4). A McNemar test showed that the two proportions were different, p = 0.005 (2-sided). Similarly, the audiologist found video otoscopy recordings were more suitable to make accurate diagnosis (Table 5, McNemar test p = 0.004).

Table 4. ENT's ability to make accurate diagnosis at child level using video otoscopy only (n=150).

Ability to make accurate diagnosis based on video otoscopy still images	Ability to make accurate diagnosis based on video otoscopy recordings		
	No	**Yes**	**Still images total**
No	63 (42%)	23 (15.3%)	86 (57.3%)
Yes	7 (4.7%)	57 (38%)	64 (42.7%)
Recordings total	70 (46.7%)	80 (53.3%)	150 (100%)

McNemar Test's exact Sig (2-sided) = 0.005.

Table 5. Audiologist's ability to make accurate diagnosis at child level using video otoscopy only (n=150)

Ability to make accurate diagnosis based on video otoscopy still images	Ability to make accurate diagnosis based on video otoscopy recordings		
	No	**Yes**	**Still images total**
No	77 (51.3%)	22 (14.7%)	99 (66%)
Yes	6 (4.0%)	45 (30.0%)	51 (34%)
Recordings total	83 (55.3%)	67 (44.7%)	150 (100%)

McNemar Test's exact Sig (2-sided) = 0.004.

Effect of co-variates

Quality and Focus

The Quality of a video otoscopy recording was more likely (odds ratio (OR)= 3.82, 95% CI 2.67 – 5.45, $p < 0.001$) to be rated as Adequate or better compared to a still image, after adjustment for age, ear side, role of tester and role of rater. A Quality rating of at least Adequate improved by, on average, 10.6% (OR= 0.89, 95% CI: 1.1% –19.2%, p= 0.029) for each one-year increase in age, after similar adjustment for covariates, including video otoscopic procedure. No differences were detected in Quality rating based on role of tester, role of rater, or ear side. Similar results and effect sizes were obtained for the Focus domain.

Light

The rating of Light was only affected by the child's age. Light rating of Adequate or better improved by 10.5%, for each one-year increase in age (OR = 0.90, 95% CI 1.2% - 18.9%, p = 0.028). No differences were detected in Light rating based on the other covariates.

Field of View and Overall Rating

The Field of View of a video otoscopy recording was more likely (OR= 1.99, 95% CI 1.45 – 2.72, $p < 0.001$) to be rated as Adequate or better compared to a still image. It was also more likely (OR= 1.71 95% CI 1.27 – 2.31, $p < 0.001$) to be rated as Adequate or better by an ENT surgeon compared to an audiologist. For each one year increase in age, Field of View rating of Adequate or better improved by 12.4% (95% CI 2% - 21.7%, p = 0.021). Similar findings were obtained for the Overall Rating.

Cerumen Amount

The amount of cerumen in the ear canal was 11.3% more likely (OR= 1.11, 95% CI 1.00 – 1.23, p = 0.043) to be rated as Moderate or more with every one-year increase in age. It was also more likely to be rated as so by an audiologist (OR = 0.70, 95% CI 0.52 – 0.93, p = 0.016) compared to an ENT surgeon. No differences were detected in Cerumen Amount rating based on other covariates.

Discussion

Overall, our findings suggest that video otoscopy recordings appear to provide a clearer view of the TM and ear canal and increase the diagnostic ability of ENT and audiology clinicians. The ratings of 10 second video otoscopy recordings exceeded the ratings of still images across all the domains. The quality of the video otoscopy images or recordings is important as reduced quality has been found to affect the diagnostic agreement between onsite and asynchronous diagnosis using video otoscopy still images.[1] It has been reported that video otoscopy recordings can be used to reduce the risk of re-examination by reducing reliance on a single image which may have reduced image quality and not be enough for asynchronous assessment.[7] It has been found that recordings have an added advantage of identifying some TM conditions, such as retraction, by observing the changes in light reflections resulting from movement in the otoscopy[10], the position of the TM being especially challenging to determine from still images.[7] The only comparison that did not show a statistically significant difference between still images and recordings was Cerumen Amount. Obtaining different views of the same ear canal at the same time point would likely show fairly similar amounts of wax and obscure the view of the TM to the same degree, regardless of the otoscopic of procedure.

We also found an increase of 10.7% in the suitability of the video otoscopy for making a diagnosis using recording compared to still images as indicated by the clinicians. In general, the ENT surgeon was more able than the audiologist to make a diagnosis using either video otoscopy still images (42.7% vs 34%) or recordings (53.3% vs 44.7%). However, in the absence of a gold standard in present study, and therefore inability to assess the accuracy of the clinicians' diagnosis, likely explanation is that the ENT was just more confident than the audiologist was in making a diagnosis using video otoscopy findings only. The ENT surgeon was also more likely to give a better rating in the Field of View as well as the Overall Rating, for both images and recordings.

Our findings show that trained research assistants were as capable as ENT specialists and audiologists in obtaining video otoscopy images and recordings for asynchronous telehealth review, with no significant difference in the rating related to the role of the tester. Others have also found that health care facilitators with limited training were able to obtain high-quality video otoscopy images[7] and recordings.[14] This supports calls to provide training for local ear health workers to facilitate ear health telehealth programs to reduce waiting times and travels costs for those living in rural and remote areas where access to ear health services are limited.[17] These results are consistent with other studies using still images[2, 3] and video recordings.[8, 13] However, our study is the first to directly compare the two video otoscopic procedures.

Age of participants was found to have a significant effect on the overall rating of images as well as across all the other domains, with younger children more likely to have lower overall and domain-specific quality ratings. Other studies have also found the quality can be significantly affected by the age of the participant.[1-3, 5, 8] Many factors might have led to such findings. For instance, it is often harder to have younger children stay still and cooperate during ear examinations. Younger children also have narrower and smaller ear canals which can make it challenging to obtain clear video otoscopic images or recordings. These findings illustrate the importance of carrying out some objective tests such as tympanometry and otoacoustic emissions (OAE), in addition to video otoscopy to increase the clinician's ability of making a diagnosis asynchronously. This additional information might contribute to more reliable telehealth services by improving diagnostic accuracy. This can also facilitate timely management, which is a main aim of telehealth services, by reducing the number of children whose video otoscopy findings are inconclusive and who, therefore, might need to be re-examined.

Limitations

Our sample of participants might not be representative of the general paediatric population, but may reflect those children and families who would be motivated to attend community hearing screening. The sample, however, might not be representative of children at higher risk of having significant middle ear complications such as Aboriginal children, or those referred to ENT clinics. Obtaining complete assessments in a busy environment was challenging. We excluded seven children and eight image-recording pairs which were missing still images and/or recordings from the analysis due to issues related to compliance during assessment and/or saving data after assessment. Some image-recording pairs were also excluded due to missing clinician ratings. It is unclear whether those ratings were missing because of technical issues, accidental omission, or because the clinicians felt unable to rate the images during the review process. Because of the exclusion criteria of the present study, our findings cannot be generalised to children with other middle ear issues such as those with discharging ears. The generalisability of our findings might be also limited by the devices used to obtain and review the images, as using different devices may result in different quality outcomes.

A further limitation of the study was the absence of a gold standard diagnosis (e.g. confirmation with otomicroscopy), which has consequently limited the ability to assess the diagnostic accuracy using each video otoscopic procedure.

As video otoscopy was only conducted by an audiologist on children less than two years of age, a future study could assess the ability of limitedly trained non-clinicians such as research assistant to obtain video otoscopy still images and recordings from children at this young age. Two clinicians from two different clinical groups (an ENT physician and a clinical audiologist) reviewed the images once. Therefore, this study could not examine intra-rater variability within the same video otoscopic procedure as rated by each clinician in two different rounds, or the potential inter-rater variability within the same clinical group.

Conclusion

The present study aimed to determine whether a video otoscopy recording can provide a clearer view of children’s TMs and ear canals for diagnosis, than a still image, for clinicians and whether it is more suitable for making an accurate asynchronous diagnosis. Our findings showed that this is indeed the case; video otoscopy recordings provided clearer otoscopic view and increased the ability of ENT specialists and audiologists to make an asynchronous diagnosis. We also found that images and recordings obtained by research assistants with limited training were not significantly different from those obtained by ENT specialists and audiologists. Incorporating objective tests should be considered, especially with younger children whose otoscopy images more commonly have reduced quality.

Acknowledgement

This work was supported by the Telethon Kids Institute. We would like to thank Ana Rakonjac, Bek Groom, Beth Arrowsmith, Danielle Darragh, Eleanor Disselkoen, Holly Richmond, Jessica Ramsay, June Doyle, Kylie Sullivan, Narelle Mullan, Natasha Morrison, Nelly Newall, Rosie Barnes, Sarah Brazier, Sharon Weeks, Sonia McAllister, Stacey Campbell and Valerie Swift for helping during the Telethon Weekend Expo and NAIDOC. We would also like to acknowledge all the children and parents who took part in this study. The authors thank the two anonymous reviewers of the manuscript for their comments and suggestions.

Reference

1. Eikelboom RH, Mbao MN, Coates HL, et al. Validation of tele-otology to diagnose ear disease in children. *Int J Pediatr Otorhinolaryngol* 2005; 69: 739-744. 2005/05/12. DOI: 10.1016/j.ijporl.2004.12.008.
2. Kokesh J, Ferguson AS, Patricoski C, et al. Digital images for postsurgical follow-up of tympanostomy tubes in remote Alaska. *Otolaryngol Head Neck Surg* 2008; 139: 87-93. 2008/07/01. DOI: 10.1016/j.otohns.2008.04.008.
3. Patricoski C, Kokesh J, Ferguson AS, et al. A Comparison of In-person Examination and Video Otoscope Imaging for Tympanostomy Tube Follow-Up. *Telemedicine journal and e-health : the official journal of the American Telemedicine Association* 2003; 9.
4. Lundberg T, Westman G, Hellstrom S, et al. Digital imaging and telemedicine as a tool for studying inflammatory conditions in the middle ear—evaluation of image quality and agreement between examiners. *International Journal of Pediatric Otorhinolaryngology* 2007; 72: 73-79. DOI: 10.1016/j.ijporl.2007.09.015.
5. Jones WS. Video otoscopy: bringing otoscopy out of the "black box". *Int J Pediatr Otorhinolaryngol* 2006; 70: 1875-1883. 2006/08/12. DOI: 10.1016/j.ijporl.2006.06.014.
6. Aronzon A, Ross AT, Kazahaya K, et al. Diagnosis of middle ear disease using tympanograms and digital imaging. *Otolaryngol Head Neck Surg* 2004; 131: 917-920. 2004/12/04. DOI: 10.1016/j.otohns.2004.07.012.
7. Biagio L, Swanepoel dW, Adeyemo A, et al. Asynchronous video-otoscopy with a telehealth facilitator. *Telemed J E Health* 2013; 19: 252-258. 2013/02/07. DOI: 10.1089/tmj.2012.0161.
8. Biagio L, Swanepoel dW, Laurent C, et al. Video-otoscopy recordings for diagnosis of childhood ear disease using telehealth at primary health care level. *J Telemed Telecare* 2014; 20: 300-306. 2014/06/25. DOI: 10.1177/1357633X14541038.
9. Swanepoel DW, Clark JL, Koekemoer D, et al. Telehealth in audiology: The need and potential to reach underserved communities. *International Journal of Audiology* 2010; 49: 195-202. DOI: 10.3109/14992020903470783.
10. Smith AC, Perry C, Agnew J, et al. Accuracy of pre-recorded video images for the assessment of rural indigenous children with ear, nose and throat conditions. *Journal of Telemedicine and Telecare* 2006; 12: 76-80. DOI: 10.1258/135763306779380138.
11. Binol H, Moberly AC, Niazi MKK, et al. SelectStitch: Automated Frame Segmentation and Stitching to Create Composite Images from Otoscope Video Clips. *Applied sciences* 2020; 10: 5894. DOI: 10.3390/app10175894.
12. Binol H, Niazi MKK, Essig G, et al. Digital Otoscopy Videos Versus Composite Images: A Reader Study to Compare the Accuracy of ENT Physicians. *The Laryngoscope* 2020. DOI: 10.1002/lary.29253.
13. Lundberg T, Biagio L, Laurent C, et al. Remote evaluation of video-otoscopy recordings in an unselected pediatric population with an otitis media scale. *Int J Pediatr Otorhinolaryngol* 2014; 78: 1489-1495. 2014/07/16. DOI: 10.1016/j.ijporl.2014.06.018.
14. Lundberg T, Biagio dJL, Swanepoel W, et al. Diagnostic accuracy of a general practitioner with video-otoscopy collected by a health care facilitator compared to traditional otoscopy. *Int J Pediatr Otorhinolaryngol* 2017; 99: 49-53. 2017/07/10. DOI: 10.1016/j.ijporl.2017.04.045.

15. Saunders GH and Roughley A. Audiology in the time of COVID-19: practices and opinions of audiologists in the UK. *International journal of audiology* 2020: 1-8. DOI: 10.1080/14992027.2020.1814432.
16. Jervis-Bardy J, Sanchez L and Carney AS. Otitis media in Indigenous Australian children: review of epidemiology and risk factors. *The Journal of Laryngology & Otology* 2014; 128: S16-S27. DOI: 10.1017/S0022215113003083.
17. Swanepoel DW and Hall JW. A Systematic Review of Telehealth Applications in Audiology. *Telemedicine journal and e-health* 2010; 16: 181-200. DOI: 10.1089/tmj.2009.0111.

Supplementary files

Appendix 1

Estimates of Fixed Effects

Target: Image **Quality** rated as either poor or less, or at least adequate.

Model Term	Odds Ratio(OR)	95% Confidence Interval for OR		p value
		Lower	Upper	
Intercept	0.499	0.192	1.298	0.154
Otoscopic procedure= Still image	3.815	2.672	5.446	<0.001*
Age	0.894	0.808	0.989	0.029*
Role of Rater= Audiologist	0.995	0.755	1.310	0.970
Role of Tester= ENT	1.254	0.527	2.981	0.609
Role of Tester= Audiologist	0.772	0.356	1.675	0.513
Ear Side: Right	1.059	0.750	1.495	0.745

*indicates a statistically significant value

Appendix 2

Estimates of Fixed Effects

Target: Image **Focus** rated as either poor or less, or at least adequate.

Model Term	Odds Ratio(OR)	95% Confidence Interval for OR		p value
		Lower	Upper	
Intercept	0.717	0.332	1.547	0.396
Otoscopic procedure= Still image	3.406	2.462	4.711	<0.001*
Age	0.920	0.848	0.997	0.042*
Role of Rater= Audiologist	1.096	0.846	1.419	0.489
Role of Tester= ENT	0.814	0.404	1.640	0.564
Role of Tester= Audiologist	0.611	0.327	1.142	0.122
Ear Side: Right	1.144	0.848	1.542	0.378

*indicates a statistically significant value

Appendix 3

Estimates of Fixed Effects

Target: Image **Light** rated as either poor or less, or at least adequate.

Model Term	Odds Ratio(OR)	95% Confidence Interval for OR		p value
		Lower	Upper	
Intercept	0.945	0.383	2.330	0.902
Otoscopic procedure= Still image	1.291	0.957	1.742	0.094
Age	0.895	0.811	0.988	0.028*
Role of Rater= Audiologist	0.947	0.722	1.242	0.694
Role of Tester= ENT	0.960	0.456	2.023	0.915
Role of Tester= Audiologist	0.604	0.287	1.273	0.185
Ear Side: Right	0.995	0.677	1.460	0.978

*indicates a statistically significant value

Appendix 4

Estimates of Fixed Effects

Target: Image **Cerumen Amount** rated as either at less than moderate, or at least moderate.

Model Term	Odds Ratio(OR)	95% Confidence Interval for OR		p value
		Lower	Upper	
Intercept	2.670	0.947	7.530	0.063
Otoscopic procedure= Still image	0.810	0.575	1.143	0.230
Age	1.113	1.003	1.233	0.043*
Role of Rater= Audiologist	0.695	0.517	0.934	0.016*
Role of Tester= ENT	1.457	0.582	3.650	0.421
Role of Tester= Audiologist	1.499	0.616	3.652	0.372
Ear Side: Right	0.936	0.624	1.405	0.751

*indicates a statistically significant value

Appendix 5

Estimates of Fixed Effects

Target: Image **Field of View** rated as either 25% or less, or at least 50%.

Model Term	Odds Ratio(OR)	95% Confidence Interval for OR		p value
		Lower	Upper	
Intercept	0.769	0.289	2.049	0.599
Otoscopic procedure= Still image	1.989	1.453	2.722	<0.001*
Age	0.876	0.783	0.980	0.021*
Role of Rater= Audiologist	1.713	1.270	2.312	<0.001*
Role of Tester= ENT	0.683	0.281	1.662	0.400
Role of Tester= Audiologist	0.507	0.211	1.218	0.129
Ear Side= Right	1.063	0.714	1.581	0.764

*indicates a statistically significant value

Appendix 6

Estimates of Fixed Effects

Target: Image **Overall Rating** rated as poor or less, or at least adequate.

Model Term	Odds Ratio(OR)	95% Confidence Interval for OR		p value
		Lower	Upper	
Intercept	1.412	0.463	4.306	0.544
Otoscopic procedure= Still image	2.750	1.955	3.870	<0.001*
Age	0.856	0.763	0.961	0.009*
Role of Rater= Audiologist	1.518	1.146	2.012	0.004*
Role of Tester= ENT	0.808	0.295	2.217	0.679
Role of Tester= Audiologist	0.572	0.220	1.489	0.252
Ear Side= Right	1.069	0.716	1.595	0.745

*indicates a statistically significant value