ORIGINAL RESEARCH

Evaluation of Streamed Hardware-to-Software Telemedicine Strabismus Consultations Utilizing Video Glasses

Joy Li[1], Angeline M Nguyen[2,3], Talia Kolin[2–4], Melinda Y Chang[2,3], Mark W Reid 🆔[2], Thomas C Lee[2,3], Sudha Nallasamy 🆔[2,3]

[1]Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; [2]The Vision Center, Children's Hospital Los Angeles, Los Angeles, CA, USA; [3]USC Roski Eye Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; [4]Veterans Affairs Los Angeles Ambulatory Care Center, Los Angeles, CA, USA

Correspondence: Sudha Nallasamy, The Vision Center, Children's Hospital Los Angeles, 4650 Sunset Boulevard, MS #88, Los Angeles, CA, 90027, USA, Tel +1 323 361 4510, Fax +1 323 361 7993, Email sudha.nallasamy@gmail.com

Background: While video glasses have been shown to be an effective tool for real-time pediatric strabismus telemedicine consultations, the high cost of the hardware-to-hardware conferencing system and bandwidth limitations may present barriers to accessibility and widespread adoption. This study evaluates the use of video glasses with a more affordable hardware-to-software video conferencing system for real-time strabismus consultations across multiple graders.

Methods: A pediatric ophthalmologist (Grader 1) wearing video glasses simultaneously performed and recorded strabismus examinations in primary gaze, with and without correction, both at distance and near. Recorded parameters included strabismus category, angle measurements, and ocular motility. Three years later, four pediatric ophthalmologists (Graders 1–4) reviewed and graded streamed video feed transmitted at 1 megabit per second (Mbps) from a hard-wired codec to software. Agreement between streamed and gold standard in-person findings was determined by weighted kappa (κ) for categorical variables, intraclass correlation coefficient (ICC) for continuous variables, and percent agreement.

Results: Eighteen patients aged 4–11 years (median, 7 years) were included. Agreement in strabismus category between in-person and streamed examinations was perfect for both horizontal and vertical deviations (κ=1.0). Almost perfect agreement was found for degree manifest (tropia vs intermittent tropia vs phoria) across graders (κ=0.91, range 0.86–0.97). Agreement for angle measurements was excellent across graders (ICC = 0.97, range 0.97–0.98). Extraocular motility agreement was 90% for all graders combined, with Grader 1 having 100% agreement between her in-person and streamed examinations.

Conclusion: Feed obtained from video glasses streamed through a hardware-to-software video conferencing system at 1 Mbps is a reliable tool for pediatric strabismus telemedicine evaluations.

Keywords: telemedicine, strabismus, ophthalmology, pediatrics

Introduction

Real-time video feed obtained with video glasses has previously been established as a reliable and effective method for evaluating strabismus category, angle, and extraocular motility in the pediatric population.[1,2] Our previous study of video glasses for strabismus examinations with two pediatric ophthalmologist graders using a hardware-to-hardware system with videos streamed at 4 megabits per second (Mbps) demonstrated that the diagnostic accuracy of telemedicine video examinations was comparable to that of in-person examinations. Perfect or almost perfect agreement was found between streamed and in-person examinations for both horizontal and vertical deviations (κ = 1.0) and degree manifest (κ = 0.94), with excellent agreement for angle measurement deviations (ICC = 0.95 and ICC = 0.91 for horizontal and vertical deviations respectively) and ocular motility (93% and 98% for inferior and superior oblique measurement respectively).[1] The high levels of agreement we demonstrated were likely in part due to the superior bandwidth used (4 Mbps), which

3927

allowed for faster and smoother transmission of larger video and image files. In contrast, prior strabismus telemedicine studies transmitting videos at lower bandwidths (128, 224, 384 Kbps) demonstrated lower levels of inter-grader agreement.[3–6] In a previous study assessing strabismus category using telemedicine videos streamed at 224 Kbps, Cheung et al demonstrated modest agreement for horizontal deviations ($\kappa = 0.66$), and poor agreement for vertical deviations ($\kappa = 0.28$).[5] Another study by Dawson et al reported satisfactory assessment of large manifest strabismus at 128 Kbps and 384 Kbps, but that latent strabismus and micro-movements were difficult to diagnose even with videos streamed at 384 Kbps.[6]

The high cost of the hardware conferencing system utilized in our prior studies[1,2] in addition to bandwidth limitations may present barriers to accessibility and widespread adoption of video glasses technology. Therefore, the current study aims to assess the validity of using a more affordable hardware-to-software system to stream videos from the video glasses at 1 Mbps for the purpose of pediatric strabismus consultations.

Materials and Methods

This prospective study was performed from October 2020 to July 2021 using telemedicine videos recorded between August 2016 and January 2017. The project was approved by the Institutional Review Board at Children's Hospital Los Angeles (CHLA) and adhered to the tenets of the Declaration of Helsinki. Written informed consent was obtained from parents or legal guardians of all participants, and assent was obtained from subjects aged seven and older.

Telemedicine System

Video glasses (Pivothead® Smart Series, Denver, CO) were used to record in-person strabismus examinations (Figure 1A). These video glasses are a wireless, wearable commercial device with a built-in high-definition camera at the nasal bridge to capture first-person point-of-view broadcasting. Touch controls on the glasses frame allow the examiner to start and stop recording with a single touch, enabling the examiner to simultaneously record and perform the examination hands-free. The video glasses connect via Wi-Fi to a local computer (Figure 1B) with the potential to record and stream videos simultaneously, allowing for voice contact between examiner and remote expert. However, for this study, we used videos recorded during a prior study. To simulate real-time telemedicine, videos were displayed in VideoLAN Client (VLC) media player (VideoLAN, Paris, France) on the local computer and securely streamed by desktop sharing through an encrypted video call from a hard-wired codec (Polycom RealPresence® Group 500 video conferencing hardware, San Jose, CA) to desktop software (Polycom RealPresence, ©2020, San Jose, CA) installed on personal computers for viewing by remote
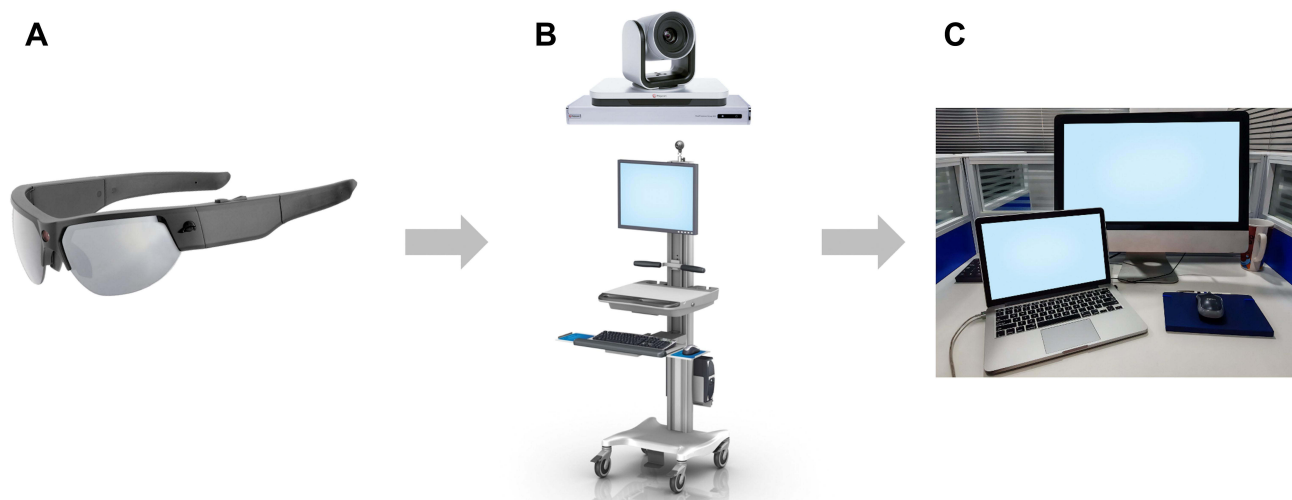

Figure 1 Feed obtained from video glasses (**A**) is streamed to the video conferencing system, including camera and hard-wired codec (**B**), and then transmitted to a laptop with conferencing software installed and larger viewing screen (**C**).

graders ([Figure 1C](#)). Retaining the hardware conference bridge allows for superior video processing. All videos were formatted at 720 pixels (p), the recommended default setting, with a frame rate of 30 frames/second. Given the software limitations, all videos were transmitted at 1 Mbps over the hospital's internal network for remote graders to view, rather than the 4 Mbps that is possible through hardware-to-hardware calls. See [Supplementary video](#) for a typical strabismus measurement using the video glasses.

Study Design

Participants were considered eligible for enrollment if they were between one month and 18 years of age with a known strabismus diagnosis. Participants had to have good enough vision in each eye to participate in prism and alternate cover testing, and had to be wearing their glasses if needed for best corrected visual acuity to be enrolled in the study. Participants were excluded if they were unable to cooperate for examination. Prior surgical procedures did not exclude participants from this study.

The study graders included four pediatric ophthalmologists, Graders 1–4 (SN, AN, TK, MYC). Grader 1 previously performed in-person strabismus examinations while wearing video glasses and recorded her results (gold standard). The videos were stored with no degradation of quality. Three years following the enrollment of the last study participant, video recordings of each examination were streamed (to simulate real-time telemedicine) via a hardware-to-software conferencing system for viewing by Graders 1–4, who were masked to the in-person clinical findings, as well as to each other's findings. All graders recorded their impressions while watching the streamed content.

Individual agreement between in-person and streamed examination results was assessed for each grader, along with inter-grader agreement across all graders combined. Intra-grader agreement was evaluated by comparing Grader 1's in-person examination results with her subsequent streamed examination results.

Strabismus measurements in primary gaze, with and without correction, both at distance and near were recorded for each participant (contributing up to 4 data points per patient for strabismus category and angle measurements). Recorded parameters include the following:[1]

- Strabismus category, which included esodeviation, exodeviation, right hyperdeviation, left hyperdeviation, and dissociated vertical deviation. Degree manifest (tropia, intermittent tropia, or phoria) were further specified as applicable.
- Angle measurements, measured during prism and alternating cover testing. Graders recorded measurements to the nearest 2 prism diopters ($^\Delta$) under 20^Δ, and to the nearest 5^Δ for measurements $\geq 20^\Delta$. If the graders recorded a range, the midpoint of that range was used for statistical analysis.
- Ocular motility, with superior and inferior oblique muscle action graded as –4 underaction to +4 overaction. For the purpose of scoring all grader responses, in-person cases were given motility scores of 0 if no deviation was noted; however, graders were not expected to rate all possible deviations. For statistical analysis, greater than 1 grade of muscle rating difference was considered as clinically significant disagreement.

Statistical Analysis

Data was entered into Excel (Microsoft Corporation, Redmond, WA) and analyzed using Stata/SE 14.2 (StataCorp LLC, College Station, TX). Agreement between streamed and in-person findings was determined by weighted kappa (κ) for strabismus categories; data from all raters were combined to assess weighted κ general agreement among all raters. The intraclass correlation coefficient (ICC) was used to assess agreement for angle measurements between each rater and the in-person results, and among all raters. Agreement on ocular muscle overaction or underaction was determined by counting the percentage of assessments where ratings were within 1 grade of each other. The 95% confidence intervals were also calculated for each measure; bootstrapping with 1000 replications was used to calculate weighted κ confidence intervals, and others were computed using the binomial exact method. Interpretation of the weighted κ statistic was based on the following guideline: <0, less than chance agreement; 0.01–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; 0.81–0.99, almost perfect agreement; 1.00, perfect agreement.[7] Interpretation of ICC was based on the following guideline: below 0.50, poor; 0.50–0.75, moderate; 0.75–0.90, good; >0.90, excellent.[8] All ICC estimates were calculated using two-way, mixed-effects models with absolute agreement for single ratings [ICC(A,1)].

Study power was calculated based on n = 18 individuals being evaluated by two graders, even though some analyses consider data from additional graders and are more powerful by design. For categorical kappa, agreement of $\kappa = 0.54$ or greater (moderate) can be detected as significantly different from change agreement ($\kappa = 0$), assuming a Type I error rate $\alpha = 0.05$ and a Type II error rate $\beta = 0.20$ (or 80% power). For angle measurement ICCs, agreement of ICC = 0.82 or greater (almost perfect) can be detected as significantly different from what would be expected if the null hypothesis was true (0.50 or lower), assuming the same Type I and II error rates. Since we expect to observe higher degrees of agreement among graders for both strabismus categories and angle measurements, the study is sufficiently powered to address its aims.

Results

In total, eighteen children were included in this study. The median age of participants was 7 years (range, 4–11 years). Eleven (61%) participants were female, and seven (39%) were male. Strabismus diagnoses included 10 esodeviations, 8 exodeviations, and 5 vertical deviations.

Strabismus Category

Agreement for strabismus categories is provided in Table 1. When comparing in-person and streamed examination results, agreement was perfect across all graders for detecting the presence of any horizontal and vertical deviations ($\kappa = 1.0$). Almost perfect agreement was found for degree manifest (tropia vs intermittent tropia vs phoria) across all graders ($\kappa = 0.86–0.97$). When comparing Grader 1's in-person examination results with her streamed examination results three years later, there was perfect agreement for detecting horizontal and vertical strabismus categories ($\kappa = 1.0$), as well as almost perfect agreement for determining degree manifest ($\kappa = 0.97$).

Angle Measurements

Angle measurement agreement is provided in Table 2. The range of horizontal measurements (n = 59) was 2^Δ-55^Δ, with a median of 18^Δ. The range of vertical measurements (n = 4) was 2^Δ-8^Δ, with a median of 5.5^Δ. All intraclass coefficients calculated were excellent (>0.90). When comparing in-person and streamed examination results, agreement for angle measurements was excellent across graders combined (ICC = 0.97), with Grader 1 having the highest agreement overall (ICC = 0.98).

Table 1 Strabismus Category Agreement for Streamed Examinations[a,b]

	Grader 1 vs In-Person	Grader 2 vs In-Person	Grader 3 vs In-Person	Grader 4 vs In-Person	All Graders
Strabismus category (horizontal)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)
Strabismus category (vertical)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)
Overall agreement in degree manifest (horizontal and vertical combined)	0.97 (0.94–1.00)	0.86 (0.80–0.93)	0.91 (0.86–0.95)	0.88 (0.84–0.92)	0.91 (0.88–0.93)

Notes: [a]Expressed as weighted kappa. [b]95% confidence interval listed in parentheses.

Table 2 Angle Measurement Agreement for Streamed Examinations[a,b]

	Grader 1 vs In-Person	Grader 2 vs In-Person	Grader 3 vs In-Person	Grader 4 vs In-Person	All Graders
Overall agreement (horizontal and vertical combined)	0.98 (0.97–0.99)	0.97 (0.95–0.98)	0.97 (0.95–0.98)	0.97 (0.95–0.98)	0.97 (0.96–0.98)

Notes: [a]Expressed as intraclass correlation. [b]95% confidence interval listed in parentheses.

Table 3 Ocular Motility Measurement Agreement for Streamed Examinations[a]

	Grader 1 vs In-Person	Grader 2 vs In-Person	Grader 3 vs In-Person	Grader 4 vs In-Person	All Graders
Percentage agreement	100% (87.7–100)	81.8% (64.5–93.0)	90.9% (78.3–97.5)	78.4% (64.7–88.7)	90.2% (85.5–93.9)

Note: [a]95% confidence interval listed in parentheses.

Ocular Motility

Ocular motility in this study focused on grading the underaction and overaction of the superior oblique (n = 1) and inferior oblique (n = 28) muscles. Due to the low number of non-zero superior oblique measurements, both superior and inferior oblique measurements were grouped together for the purpose of data analysis. Ocular motility agreement is provided in Table 3 and was excellent for all graders combined (90.2%). Graders 2 and 4 demonstrated good agreement (78.4% and 81.8%, respectively), while Graders 1 and 3 had excellent agreement (100% and 90.9%, respectively) with the in-person results, suggesting some degree of inter-grader variability. Grader 1's perfect agreement between her in-person and streamed examinations suggests that exam modality plays less of a role than inter-grader variability.

Discussion

The use of telemedicine has increased dramatically in the past decade, with ophthalmic subspecialities including pediatric ophthalmology, neuro-ophthalmology, and oculoplastics now comprising 80.5% of all reported ophthalmology telemedicine visits.[9,10] However, despite increasing adoption, one-third of providers surveyed reported feeling "not at all confident" in their ability to use telemedicine.[11] In a study by Woodward et al, ophthalmic providers expressed reservations regarding the accuracy of telemedicine, with 59% of respondents having "low confidence in remote care for providing an opinion on patient care".[12] These attitudes highlight the importance of validating the quality and accuracy of telemedicine instruments to further build confidence in their use and adoption.

In addition to dramatically improving both bandwidth capacities and video quality, advances in telemedicine technology also hold great potential for reducing long-term health system costs. However, current start-up costs including equipment and administration expenses remain high.[13] Although video glasses are relatively inexpensive (<$1000), the hardware-to-hardware conferencing system is not. Our current study with four pediatric ophthalmologist graders demonstrates that pediatric strabismus evaluations can be completed with excellent reliability using telemedicine videos transmitted at 1 Mbps through a more affordable hardware-to-software system (>$10,000 cost savings).

While written comments were occasionally noted about having excellent control or poor control, these comments were not taken into account during the data analysis process. For example, what one grader may have indicated as intermittent exotropia with excellent control, another grader may have labeled as exophoria. Although clinically near-identical, they would be recorded and treated as different categories during data analysis. Similarly, intermittent exotropia with poor control could have been labeled as exotropia by another grader. This effect is mitigated by the use of the weighted κ statistic, which treats adjacent categories such as intermittent exotropia and exotropia as more similar than distant categories such as exophoria and exotropia.

When evaluating degree manifest, subgroup analysis showed that labeling of intermittent deviations had lower agreement than labeling of tropias and phorias, even though agreement was almost perfect for all raters. Similarly, there was some variability demonstrated between graders for ocular motility measurement, with inter-grader agreement ranging from 78.4% to 100%. This is likely due to pediatric ophthalmologists having different methodologies for grading strabismus and muscle action. However, with these exceptions, the overall inter-grader agreement remained consistently high across all parameters.

In this study, intra-grader agreement was similarly excellent across all parameters. Comparing her in-person examination results with her streamed examination results three years later, Grader 1 demonstrated perfect agreement in detecting horizontal and vertical deviations (κ = 1.0), almost perfect agreement in degree manifest (κ = 0.97), and excellent agreement for angle measurements (ICC = 0.98) and ocular motility (100%). Grader 1's streamed evaluations were the most consistent with her in-person examinations compared to the other graders, suggesting that the telemedicine system introduces less variability than that already existing between different graders.

Grader 2 in this current study was also Grader 2 in our prior study.[1] In our prior study using a hardware-to-hardware system with videos streamed at 4 Mbps, Grader 2 demonstrated perfect agreement in detecting horizontal and vertical deviations (κ = 1.0), almost perfect agreement for degree manifest (κ = 0.92), excellent agreement for angle measurements (ICC = 0.98), and excellent agreement for ocular motility (93% and 98% for inferior and superior oblique measurement respectively).[1] In our current study using a hardware-to-software system with videos streamed at 1 Mbps, Grader 2 demonstrated perfect agreement in detecting horizontal and vertical deviations (κ = 1.0), almost perfect agreement for degree manifest (κ = 0.86), excellent agreement for angle measurements (ICC = 0.97), and good agreement for ocular motility (81.8%). Despite demonstrating similar or slightly lower agreement in this study compared to our previous study's results, Grader 2 nonetheless demonstrated high agreement across all parameters in both studies. This further suggests that telemedicine videos streamed at 1 Mbps using a more affordable hardware-to-software system are comparable to telemedicine videos streamed at 4–6 Mbps using a hardware-to-hardware system.

Our study has several limitations. All pediatric ophthalmologist graders were on the hospital's internal network during video transmission. While this ensured that video transmission speeds were consistent at 1 Mbps during this study, streaming over the Internet between remote sites may introduce additional bandwidth and image quality variability. In addition, strabismus category and degree were solely assessed using cover testing in the primary position. As a result, this study did not assess whether patients requiring corneal light reflex testing, such as those with sensory deviations, very young children, or patients with developmental delays, may be good telemedicine candidates. However, in our previous study assessing the accuracy of real-time telemedicine in diagnosing and managing pediatric eye conditions, surgical measurements via Krimsky and Hirschberg tests were successfully obtained via telemedicine.[2]

Despite these limitations, this study with multiple graders demonstrates that video feed from video glasses streamed through a hardware-to-software system at 1 Mbps is a highly reliable tool that may safely be used for pediatric strabismus telemedicine evaluations. Following software installation, as long as bandwidth is sufficient, providers are able to view real-time videos anywhere, including from the comfort of their homes. For remote providers, the ability to connect to and view videos on a larger screen may be helpful in evaluating the smaller and more subtle movements seen on strabismus exam. In addition, a true real-time experience using the hardware-to-software system would allow for communication between the examiner and the remote viewer (unlike in our study), further enhancing the reliability of results. Furthermore, as demonstrated in previous pediatric ophthalmology studies, telemedicine platforms also allow for trained nonphysician personnel to take part in patient care.[2,14–16] With 12–52% of pediatric ophthalmology fellowship positions going unmatched between 2000 and 2015, telemedicine could help address physician shortages by diversifying the providers delivering care and allowing ophthalmologists to co-manage remote patients with other trained, nonspecialized providers.[17] These features of telemedicine hold numerous positive implications, particularly for rural or otherwise underserved areas without immediate access to a pediatric ophthalmology specialist, or in areas where wait times are excessively long. In the future, as 5G and eventually 6G technologies will become more widespread, reducing reliance on expensive hardware and fully transitioning to a completely software-to-software system will also become increasingly feasible. Future research should evaluate the accuracy of a fully software-to-software system, videos streamed at varying bandwidths, and the role of trained nonphysician personnel to further facilitate telemedicine-based care.

Conclusion

Telemedicine feed from video glasses streamed through a hardware-to-software system at 1 Mbps enables remote strabismus and ocular motility measurements comparable to in-person clinical examinations. While the latter remains the gold standard for pediatric strabismus evaluations, advances in video and imaging technologies hold great potential in improving healthcare access and delivery.

Ethics Statement

Acknowledgments

Disclosure

References

1. Ho TC, Kolin T, Stewart C, et al. Evaluation of high-definition video smart glasses for real-time telemedicine strabismus consultations. *J Am Assoc Pediatr Ophthalmol Strabismus*. 2021;25(2):74.e1–74.e6. doi:10.1016/j.jaapos.2020.11.016
2. Stewart C, Coffey-Sandoval J, Reid MW, et al. Reliability of telemedicine for real-time paediatric ophthalmology consultations. *Br J Ophthalmol*. 2021:bjophthalmol-2020–318385. doi:10.1136/bjophthalmol-2020-318385
3. Helveston EM, Neely DE, Cherwek DH, et al. Diagnosis and management of strabismus using telemedicine. *Telemed e-Health*. 2008;14(6):531–538. doi:10.1089/tmj.2007.0086
4. Rimsza ME, Hotaling AJ, Keown ME, et al. The use of telemedicine to address access and physician workforce shortages. *Pediatrics*. 2015;136(1):202–209. doi:10.1542/peds.2015-1253
5. Cheung JC, Dick PT, Kraft SP, et al. Strabismus examination by telemedicine. *Ophthalmology*. 2000;107(11):1999–2005. doi:10.1016/S0161-6420(00)00377-8
6. Dawson E, Kennedy C, Bentley C, et al. The role of telemedicine in the assessment of Strabismus. *J Telemed Telecare*. 2002;8(1):52–55. doi:10.1258/1357633021937361
7. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005;37:360–363.
8. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155–163. doi:10.1016/j.jcm.2016.02.012
9. Patel S, Hamdan S, Donahue S. Optimising telemedicine in ophthalmology during the COVID-19 pandemic. *J Telemed Telecare*. 2020;28(7):498–501. doi:10.1177/1357633x20949796
10. Lurie N, Carr BG. The role of telehealth in the medical response to disasters. *JAMA Intern Med*. 2018;178:745–746. doi:10.1001/jamainternmed.2018.1314
11. De Lott LB, Newman-Casey PA, Lee PP, et al. Change in ophthalmic clinicians' attitudes toward telemedicine during the coronavirus 2019 pandemic. *Telemed e-Health*. 2021;27:231–235. doi:10.1089/tmj.2020.0222
12. Woodward MA, Ple-plakon P, Blachley T, et al. Eye care providers' attitudes towards tele-ophthalmology. *Telemed e-Health*. 2015;21:271–273. doi:10.1089/tmj.2014.0115
13. Snoswell CL, Taylor ML, Comans TA, et al. Determining if telehealth can reduce health system costs: scoping review. *J Med Internet Res*. 2020;22(10):e17298. doi:10.2196/17298
14. Quinn GE, Ying GS, Daniel E, et al. Validity of a telemedicine system for the evaluation of acute-phase retinopathy of prematurity. *JAMA Ophthalmol*. 2014;132:1178. doi:10.1001/jamaophthalmol.2014.1604
15. Vinekar A, Jayadev C, Mangalesh S, et al. Role of tele-medicine in retinopathy of prematurity screening in rural outreach centers in India – a report of 20,214 imaging sessions in the KIDROP program. *Semin Fetal Neonatal Med*. 2015;20(5):335–345. doi:10.1016/j.siny.2015.05.002
16. Silva RA, Murakami Y, Jain A, et al. Stanford University Network for diagnosis of retinopathy of prematurity (Sundrop): 18-month experience with telemedicine screening. *Graefe's Arch Clin Exp Ophthalmol*. 2009;247:129–136. doi:10.1007/s00417-008-0943-z
17. Dotan G, Karr DJ, Levin AV. Pediatric ophthalmology and strabismus fellowship match outcomes, 2000–2015. *J Am Assoc Pediatr Ophthalmol Strabismus*. 2017;21(3):181.e1–181.e8. doi:10.1016/j.jaapos.2017.01.004