



REVIEW

Digital Tools for the Self-Assessment of Visual Acuity: A Systematic Review

Janneau L. J. Claessens · Judith R. Geuvers · Saskia M. Imhof ·

Robert P. L. Wisse

Received: April 27, 2021 / Accepted: June 5, 2021 / Published online: June 25, 2021
© The Author(s) 2021, corrected publication 2021

ABSTRACT

Introduction: Numerous digital tools to self-assess visual acuity have been introduced. The recent COVID-19 pandemic underlined the need for high-quality remote care. This review gives a current overview of digital tools for remotely assessing visual function and reports on their accuracy.

Methods: We searched the databases of Embase and Pubmed, and systematically reviewed the literature, conforming to PRISMA guidelines. Two preliminary papers were added from medRxiv.org. The main outcome was the agreement of the digital tools with conventional clinical charts, as expressed by mean differences and 95% limits of agreement (95% LoA).

Results: Seventeen publications included studies reported on 13 different digital tools. Most of the tools focus on distance visual acuity. The mean differences of the digital tools ranged from -0.08 to 0.10 logMAR, when compared to traditional clinical assessments. The 95% LoA differed considerably between studies: from ± 0.08 logMAR to ± 0.47 logMAR, though the variability was less pronounced for higher visual acuities.

Conclusion: The low mean differences between digital visual acuity assessments and reference charts suggest clinical equivalence, though the wide 95% LoA identify a lower precision of digital self-assessments. This effect diminishes in individuals with better visual acuities, which is a common feature of visual acuity assessments. There is great potential for the digital tools to increase access to eye care and we expect the accuracy of the current tools to improve with every iteration in technology development.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40123-021-00360-3>.

J. L. J. Claessens (✉) · S. M. Imhof · R. P. L. Wisse
Department of Ophthalmology, University Medical Center Utrecht, Heidelberglaan 100, 3508 GX
Utrecht, The Netherlands
e-mail: j.l.j.claessens@umcutrecht.nl

J. R. Geuvers
Faculty of Medicine, Utrecht University, Utrecht,
The Netherlands

Keywords: Digital health; e-Health; m-Health; Mobile health; Remote care; Self-assessment; Telemedicine; Telehealth; VA; Visual acuity

Key Summary Points

The recent COVID-19 pandemic underlined the need for high-quality remote care. This systematic review gives a current overview of digital tools for remotely assessing visual function and reports on their accuracy.

The mean differences between the included digital tools and conventional clinical charts are low, suggesting clinical equivalence.

The 95% limits of agreement (95% LoA) identify a lower precision of digital self-assessments, though this effect diminishes in individuals with better visual acuities.

There is great potential for remote self-assessments to increase access to eye care, which is acutely relevant in the current COVID-19 pandemic.

DIGITAL FEATURES

This article is published with digital features, including a summary slide, to facilitate understanding of the article. To view digital features for this article go to <https://doi.org/10.6084/m9.figshare.14731944>.

INTRODUCTION

Digital tools are of great value for enhancing access to healthcare. In ophthalmology, numerous tools for self-assessing visual function have been developed over the last decade. These tools enable laypersons to self-measure aspects of visual acuity (VA) at home, or at school, using applications on smartphones, tablets and/or computers.

VA testing is the most commonly performed examination of visual function [1]. In clinical practice, it is most commonly carried out using

a Snellen chart, which utilizes black letters or symbols (optotypes) of a range of sizes set on a white chart [2]. This chart was developed in 1862 and has been globally adopted as the standard VA test, despite some considerations regarding its design [3, 4]. The most important issues with this chart are the irregular progression of the size of the letters. Alternative charts have been introduced. LogMAR charts are the standard method in research as they are considered the most accurate [2]. The Snellen chart retains its popularity in clinical practice owing to familiarity, cost, smaller chart size and, most importantly, the short time taken to perform the test [5]. Notwithstanding, traditional VA testing with Snellen or logMAR charts requires the patient to physically attend a clinic. Figure 1 shows a visualization of the Snellen chart (left) and the alternative ETDRS chart, a well-established logMAR chart (right).

An urgent need for high-quality remote care was brought about during the COVID-19 pandemic in 2020, when most hospital care was globally suspended or postponed. Digital tools for self-assessment of VA increase access to eye care and avoid the burden for patients in quarantine, with poor mobility or without means of transportation.

A plethora of digital tools are available on the internet and in mobile app stores, which impedes choosing which tools are the most effective and reliable [6]. Before a digital self-testing tool can successfully be used in hospital care, extensive validation research and certification is needed [7]. The aim of this literature review is to provide an overview of the available scientific evidence for remote testing of visual function, and to critically appraise and report on the validity, quality and effectiveness of the available tools.

METHODS

Protocol and Registration

A review protocol was developed on the basis of the Preferred Reporting Items of Systematic Reviews and Meta-Analyses (PRISMA) statement [8]. The systematic review was registered in the

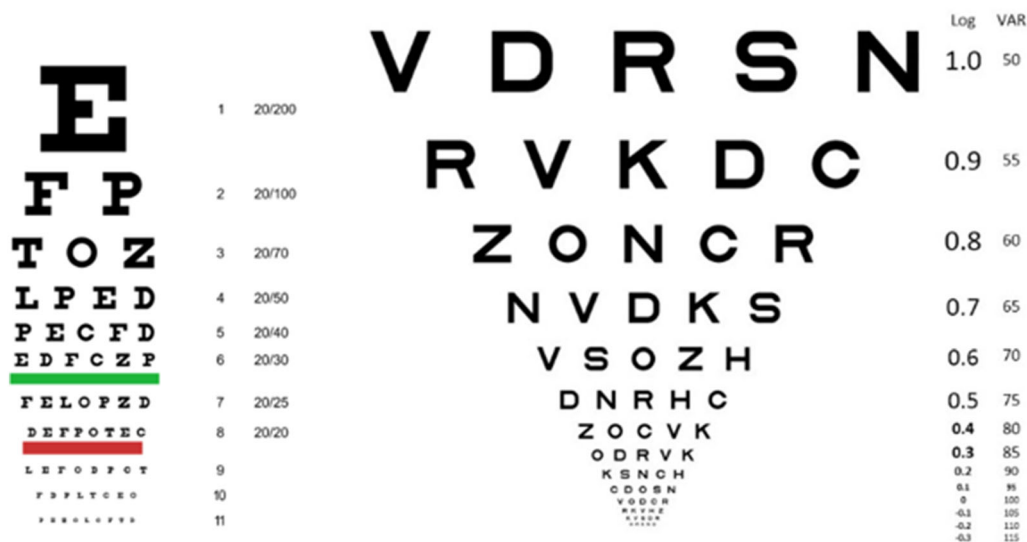


Fig. 1 Left: Snellen visual acuity chart; Right: logMAR visual acuity chart (ETDRS). Images are not-to-scale

international prospective register of systematic reviews (PROSPERO) on 28 August 2020 (ID: CRD42020201421). In accordance with Dutch law, no institutional ethical review board approval was required.

Search Strategy

A search was conducted in PubMed and Embase, including literature published up to 1 April 2021. The syntax included synonyms for the words ‘assessment’ and ‘digital’, as well as the keywords ‘refractive error’, ‘visual field’, ‘colour vision’ and ‘visual acuity’. We deliberately used more keywords than ‘visual acuity’ as we did not want to miss tools that were not primarily developed to assess VA, but had the ability to do so. The full syntax can be found in the supplementary file. Articles published before 2010 were excluded because of the rapidly changing environment of digital tools and smartphones. No further limitations or filters were applied.

Study Selection

Titles and abstracts were screened for relevance independently by two reviewers (JC and JG). The reviewers were blinded and initial title/abstract screening focused on the use of digital tools in

the field of ophthalmology in the broadest sense. Articles with a different topic were excluded. Subsequently, additional screening of titles/abstracts and full texts was performed to exclude papers about digital tools that did not include a VA assessment (i.e. different outcome) or were not self-assessments. We defined self-assessment tools as applications on a smartphone, computer or tablet that can be used without assistance of a healthcare professional. References of reviews and included studies were screened for additional citations. Articles were excluded if the full text was not available.

Our outcome of interest was the measurement accuracy of the digital tools, compared to conventional clinical charts, expressed as mean differences of VA and 95% limits of agreement (95% LoA). The mean difference can be interpreted as the systematic difference between the measurements (i.e. bias) and the 95% LoA as the range within 95% of the differences between one measurement and the other are included (i.e. random error). This methodology was first introduced by Bland and Altman and is commonly used in method comparison studies [9]. If the 95% LoA was not directly reported in the original article, it was derived from the plots or calculated manually from the mean difference and the SD using the following standard formula: $95\% \text{ LoA} = \text{mean difference} \pm 1.96 \times \text{SD}$.

In parallel to writing the current review, our group investigates a specific remote eye exam in various patient populations: in a sample of 100 keratoconus eyes and a sample of 98 uveitis eyes. To illustrate how the accuracy of one digital tool varies for different VA ranges these preliminary study data were included in this review as well. The pre-prints have been published on medRxiv.org [10, 11].

Quality Assessment

All included studies were assessed for methodological quality according to the QUADAS-2 tool by two reviewers independently [12]. Disagreements were solved in consensus. ‘Bias’ is scored in terms of patient selection, blinding of outcomes during assessment with index test or reference test and the timing of the assessments (i.e. whether both tests were done sequentially within a short interval). Applicability focuses on whether there are concerns that the patient population, index tests or reference tests match the review question.

Data Extraction and Synthesis

Two reviewers (JC and JG) independently extracted the relevant data. From each included study, we retrieved information on study and patient characteristics, type of the index test, reference test and relevant statistics including mean difference, 95% LoA and/or standard deviation (SD) of the mean difference.

If multiple comparisons were presented in one article, we only recorded the measurements that were assessed in controlled experimental settings, as these represent the agreement in the most optimal settings. Two studies compared a near vision tool with both a near VA card and a distance chart [13, 14]. Another study used both types of reference charts for evaluating a distance tool [15]. We only present the comparisons using the reference chart with the closest resemblance to the original tool, as we consider these as the best representation for the measurement accuracy. One study reported VA in ETDRS letters, which we manually converted to logMAR using standard conversion charts [13].

For the included study by our own research group, the mean difference and 95% LoA were calculated manually since we had access to the study database [16]. If the reported 95% LoA were inconsistent in different sections of the original article [17], we could not report these numerical data.

The individual studies all differed greatly with regards to the studied digital tools, reference standards and study populations. This high between-study heterogeneity precluded a meta-analysis of these outcomes, as the generalizability of a possible pooled estimate was expected to be low.

Additional Subgroup Analyses for Different Visual Acuity Ranges

Subgroup analyses illustrate how measurement accuracy of a similar tool can differ for different VA ranges. Two of the included studies reported outcomes of subgroup analyses in the original article [10, 18]. We had access to the databases of the included studies by our own research group and used these data for additional subgroup analyses [10, 16]. Subgroups were based on achieved VA of ≤ 0.5 logMAR (≥ 0.3 Snellen) and VA > 0.5 logMAR (< 0.3 Snellen). This is the cut-off value for low vision stipulated by the World Health Organization [19].

Data on Test–Retest Variability of Visual Acuity Assessments

Variation between two assessments of VA is common, partly owing to the psycho-physical nature of the test. This is demonstrated when an individual is assessed twice within a short time interval, using the same chart, and further confounds the outcomes when different charts are compared [20]. Therefore, to put in perspective the agreement between the digital tools and the clinical charts, we will also report outcomes of studies regarding repeatability (test–retest variabilities) of the conventional Snellen and ETDRS charts. A comprehensive literature search identified relevant papers [1, 21–25].

Statement of Ethics Compliance

This systematic review is based on previously conducted studies and does not contain any new studies with human participants or animals performed by any of the authors.

RESULTS

Search Strategy and Study Selection

The study flow chart can be found in Fig. 2. Our search resulted in 679 Embase and 408 PubMed

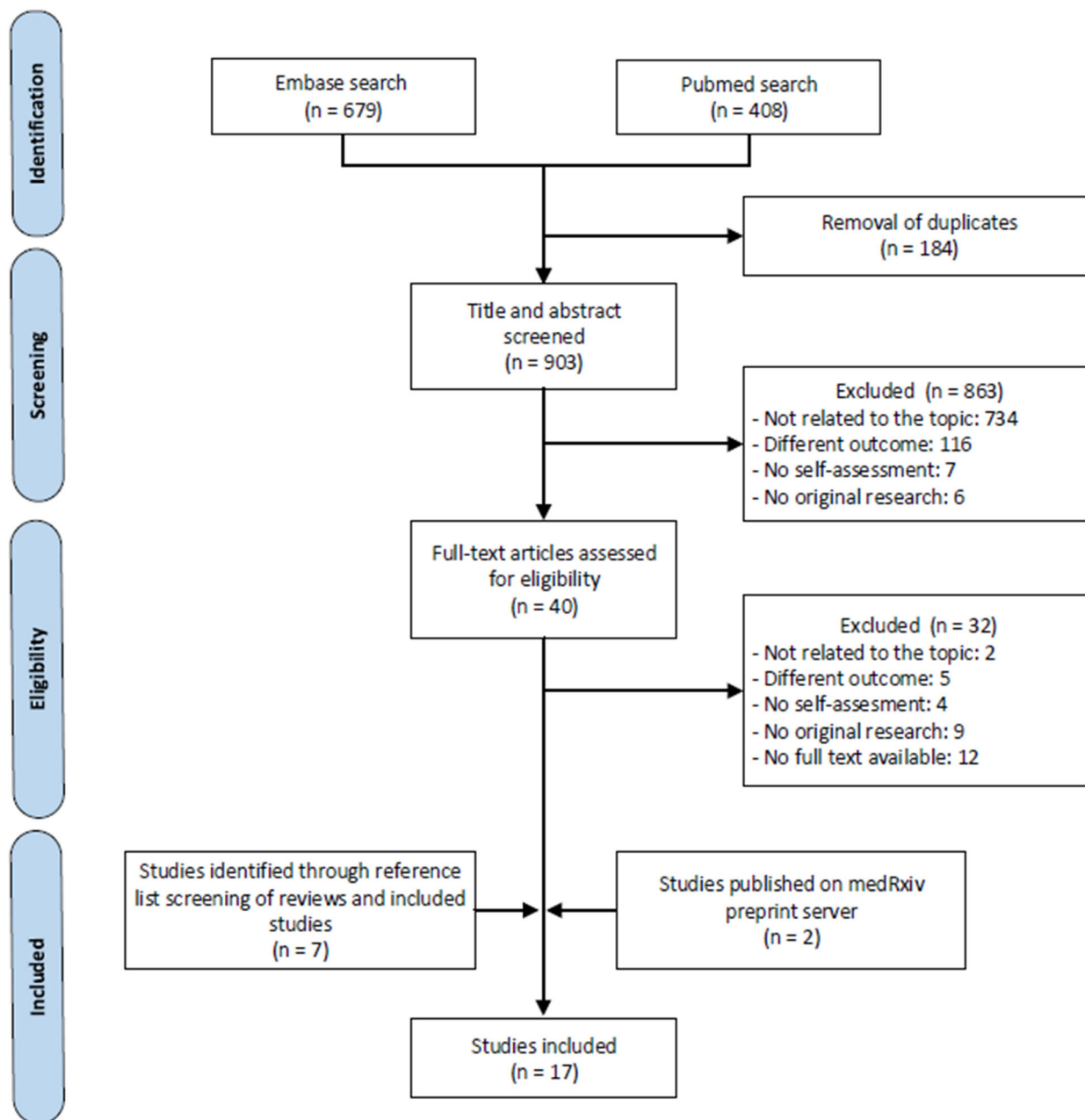


Fig. 2 Study flow diagram based on PRISMA guidelines [8]

citations. After removal of duplicates, 903 were screened for relevance by titles and abstract, followed by full-text screening of 40 potentially eligible articles. Subsequently, 32 articles were excluded on the basis of criteria depicted in the flow chart. Two articles were added from the medRxiv preprint server and seven articles were added after manual screening of references.

Quality Assessment

The evaluation of the risk of bias and applicability of the included studies is depicted in Fig. 3. In terms of risk of bias, most striking are concerns regarding the blinding of participants and/or researchers. Some studies explicitly reported on blinding, i.e. performing the index test (digital tool) without knowledge of the score of the reference test (conventional chart) and vice versa. For some studies it was unclear whether the researcher and/or subject had been blinded. Concerning patient selection, patients were mostly recruited consecutively. One study, Nik Azis et al. [26], reported a convenience

sampling method for recruitment, by excluding children who were deemed uncooperative. Overall, there were no concerns regarding applicability for any of the studies: all studies matched the review question.

Data Extraction

An overview of the study characteristics can be found in Table 1. The 17 identified studies examined 13 different tools. A total of 3591 eyes were included. There is a large heterogeneity among the included studies; participants differed in age, nationality and medical history. The majority of studies examined a population with a wide age range. Four studies focused on testing VA in children. Some studies examined only healthy participants, other studies examined patients or a mixed group of both.

Seven unique evaluated tools have been developed to test distance vision [15–18, 26–30], and four unique tools specifically focus on near vision testing [13, 14, 30–32]. One tool assesses both

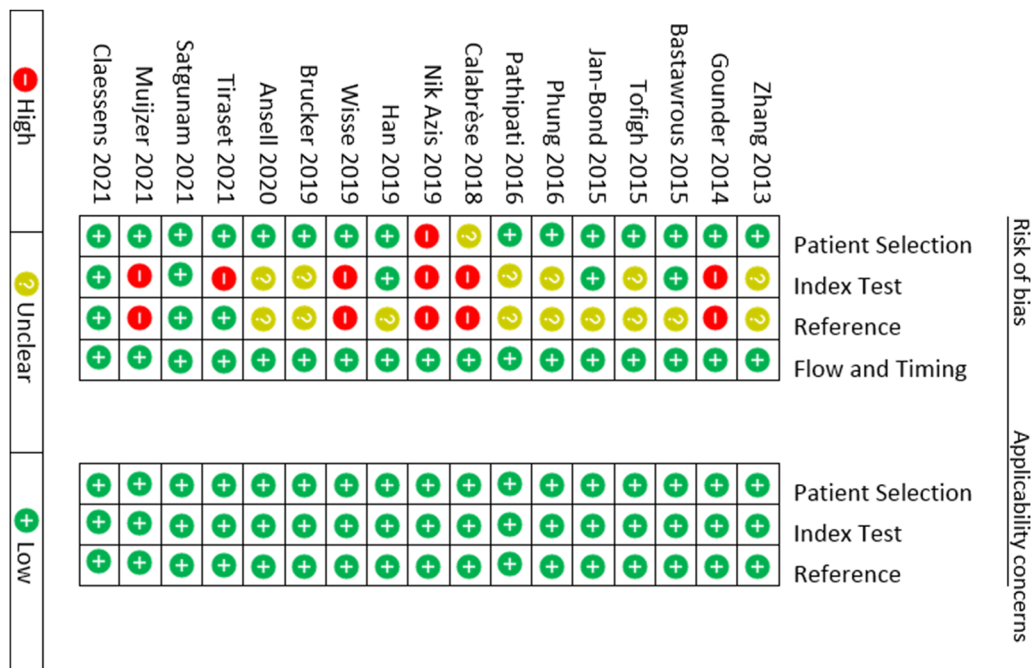


Fig. 3 Risk of bias and applicability concerns summary. Scores are based on the QUADAS-2 tool [12]

Table 1 Study characteristics of the 15 included peer-reviewed studies and two pre-print publications

Author, year, country	Name of test	No. of eyes, study population	Age (in years)	Study objectives	Platform, device	Optotypes index test	Reference standard
Zhang 2013, China [18]	Eye Chart Pro	240, patients	27–80 (mean 47)	Uncorrected VA (logMAR)	App, smartphone + tablet	Tumbling E	Standard tumbling E light box
Gounder 2014, Australia [27]	EyeSnellen app	122, patients	19–89 (mean 57)	Corrected VA (logMAR)	App, smartphone + table	Snellen chart	Snellen light box
Bastawrous 2015, Kenya [28]	Peek Acuity	544, patients	> 55	Corrected VA (logMAR)	App, Smartphone	ETDRS tumbling E	ETDRS tumbling E chart and Snellen chart
Tofigh 2015, USA [32]	EyeHandBook*	100, patients	18–89	Near VA with corrective lenses if applicable (logMAR)	App, smartphone	Tool-specific optotypes	Rosenbaum near card
Jan-Bond 2015, Malaysia [17]	REST application	202, patients and staff	5–75 (mean 37)	VA (logMAR) [unknown whether uncorrected or corrected]	App, smartphone or tablet	Tumbling E	ETDRS tumbling E chart
Phung 2016, USA [13]	Sightbook*	238, unknown population	25–93 (mean 67)	Corrected VA (approxETDRS)	App, smartphone or tablet	Tool-specific optotypes	Rosenbaum near card (near vision) and Snellen chart (distance vision)

Table 1 continued

Author, year, country	Name of test	No. of eyes, study population	Age (in years)	Study objectives	Platform, device	Optotypes index test	Reference standard
Pathipati 2016, USA [34]	Sightbook*	Phase 1: 57 Phase 2: 51, patients presenting at emergency department	Adults (mean 48.5)	Corrected VA (logMAR)	App, smartphone or tablet	Tool-specific optotypes	Phase 1: Snellen chart Phase 2: Rosenbaum near card
Calabrese 2018, USA [31]	MNREAD iPad application*	330 (normal vision) and 86 (low vision) [binocular assessments, so 165 and 43 comparisons, respectively]	Normal vision: 8–72 (mean 28)	Maximum reading speed Reading acuity Critical printing size	App, iPad	MNREAD chart	Printed MNREAD chart
Nik Azis 2019, Malaysia [26]	AAPOS Vision Screening app	390, patients and healthy participants	Low vision: 22–93 (mean 60)	Reading accessibility index	App, iPad	LEA Symbols box	LEA symbols light box
Han 2019, China/ Australia [33]	Vision @Home*	100 (adolescent Chinese), 100 (elderly Chinese) and 126 (Australian), patients and non-patients	Elderly Chinese: 50–79 Adolescent Chinese: 13–26 Australian: 8–91	Corrected VA (logMAR)	Website or app, smartphone	Tumbling E chart (distance vision) and ETDRS near chart (near vision)	ETDRS tumbling E chart (distance vision) and ETDRS near chart (near vision)

Table 1 continued

Author, year, country	Name of test	No. of eyes, study population	Age (in years)	Study objectives	Platform, device	Oprotypes index test	Reference standard
Wisse 2019, Netherlands [16]	Easee	200, healthy participants	18–40 (mean 25.4)	Refractive error, uncorrected VA (logMAR)	Website, smartphone and computer/tablet	Tumbling E and tool-specific optotypes	ETDRS chart
Brucker 2019, France [14]	Odysight*	120, patients	24–92 (mean 64.7)	Corrected VA (logMAR) Contrast sensitivity Amsler grid	App, smartphone	ETDRS Tumbling E	Sloan ETDRS chart (near vision) and ETDRS chart (distance)
Ansell 2020, UK [29]	Eye Chart	24, students	18–27 (mean 20.1)	Corrected VA (logMAR)	App, smartphone	Snellen chart	ETDRS chart
Tiraset 2021, Thailand [15]	Eye Chart	295, patients	18–85 (mean 64)	Corrected VA (logMAR)	App, smartphone	Snellen chart or Tumbling E	ETDRS chart (distance vision) and Rosenbaum near card (near vision)
Satgunam 2021, India [30]	Peek Acuity	68, healthy participants (employees of the institute)	20–60 (median 31)	Uncorrected VA (logMAR)	App, smartphone or tablet	ETDRS Tumbling E	COMPlog presenting Tumbling E
	Smart Optometry*	24, presbyopic subgroup of the population mentioned above	37–60 (median 50)	Uncorrected VA (logMAR)	App, smartphone or tablet	Tumbling E	Reduced Snellen near vision chart with tumbling E

Table 1 continued

Author, year, country	Name of test	No. of eyes, study population	Age (in years)	Study objectives	Platform, device	Optotypes index test	Reference standard
Muijzer 2021, Netherlands	Easee	100, patients with keratoconus	18–40 (mean 25.6)	Refractive error, uncorrected VA (logMAR)	Website, smartphone and computer/tablet	Tumbling E and tool-specific optotypes	ETDRS chart
Claessens 2021, Netherlands	Easee	98, patients with uveitis	Adults (mean 46.5)	Corrected VA (logMAR)	Website, smartphone and computer/tablet	Tumbling E and tool-specific optotypes	Snellen light box

VA visual acuity. Corrected = with habitual correction (if worn)

*Digital tool assesses near visual acuity (i.e. visual acuity measured at 40 cm distance)

[33]. Most of the studies assessed corrected VA, meaning that study subjects wore their habitual correction (i.e. glasses/contacts), if applicable. The digital tools involve different devices, including smartphones, tablets and computers. Most digital tools are available as an application for smartphones and tablets. Two tools are web-based and do not need installation of an app. Most of the publications were validation studies in experimental controlled environments. One study evaluated a smartphone-based near vision test to assess VA at an emergency department [34]. For one study, the digital tool was assessed unsupervised by subjects in their own home environment [11]. Different reference standards were used to assess agreement.

COMPARISONS OF DISTANCE VISUAL ACUITY ASSESSMENTS

Overall Measurement Accuracy

An overview of the comparisons of the distance VA assessments can be found in Fig. 4 [10, 11, 15, 16, 18, 26–30, 33]. Most articles reported outcomes for different subgroups or per eye, resulting in 18 comparisons. The mean differences between the digital tools and the reference standards (i.e. bias) range from –0.08 to 0.10 logMAR. Most of the digital tools provide a slightly worse VA score (i.e. higher logMAR score) than the actual VA as measured by the reference standard. The distribution of the differences between the two tests (i.e. random error), as expressed by the 95% LoA, varies greatly between the studies. It ranges from ± 0.08 logMAR (lowest variability) to ± 0.47 logMAR (highest variability) from the mean difference. Separate comparisons per eye were mostly comparable within studies.

The study by Bastawrous et al. demonstrates that the accuracy of the same digital tool (Peek Acuity) varies when compared to different reference charts. Han et al. reported on different study populations, illustrating how test accuracy slightly differs when various groups are assessed under similar conditions. The studies by our own study group (Wisse et al. [16], Muijzer et al. [10] and Claessens et al. [11]) all

focus on the same digital tool (Easee). Interestingly, the random error is much higher for *un-corrected* VA assessments in healthy individuals (of whom some have refractive errors) and patients with keratoconus than for *corrected* VA assessments in patients with uveitis.

Subgroup Analyses for Different Visual Acuity Ranges

The subgroup analyses illustrate how the measurement accuracy of a similar tool can differ for different VA ranges (Fig. 4). In all of these comparisons, the measurement accuracy appears lower in the poorer VA subgroups, illustrated by the higher mean differences and, most notably, the wider 95% LoA. In better VA ranges, these 95% LoA are smaller.

Test–Retest Variability of Clinical Wall Charts

Test–retest variabilities of logMAR and Snellen charts have been added as a reference in Fig. 4 [1, 21–25]. Test–retest variability of logMAR charts ranges from ± 0.07 to ± 0.18 logMAR (from the mean difference). Snellen charts are less consistent, with reported ranges from ± 0.18 to ± 0.34 logMAR. The line assignment in particular, often used in clinical practice, shows a great variation when measurements are repeated.

Comparisons of Near Vision Assessments

For the comparisons focusing on near VA assessments, the mean differences range from -0.03 to 0.09 logMAR [13, 30, 32, 33]. The 95% LoA range from ± 0.17 to ± 0.35 logMAR from the mean difference. One study examined the Sightbook tool in clinical practice, at an emergency department [34]. This self-administered tool provided a more accurate representation of the VA recorded by consulted ophthalmologic residents (using Rosenbaum near cards) when compared to a distance Snellen chart assessed by non-ophthalmologic personnel at the emergency department [mean difference

0.06 ± 0.40 logMAR (Sightbook app) vs. 0.21 ± 0.35 logMAR (Snellen chart)].

DISCUSSION

Many digital tools are available to self-test VA, though a clinical validation is often lacking [6]. This systematic review presents the 17 publications on 13 different tools for the self-assessment of VA currently available.

Our systematic review identified low mean differences of the digital tools when compared to reference standards for assessing distance VA, suggesting a low systematic bias. The mean differences ranged from -0.08 to 0.10 logMAR. The digital tools only slightly underestimate the VA score of the patients and we consider these low values to be negligible in clinical practice. The 95% LoA vary between studies, ranging from ± 0.08 logMAR to ± 0.47 logMAR from the mean difference. Most of the 95% LoA are rather wide, suggesting considerable variability of the VA assessments of the digital tools. As stated before, there is always a certain variation in repeated measurements in the same person [1, 20]. A study on the variability of VA assessments in a large eye clinic reported a test–retest variability of ± 0.15 logMAR, when different charts and different examiners assessed the same patient's VA within a short time interval [20]. The authors conclude that, in general, differences of less than 0.15 logMAR (i.e. 1.5 lines) are considered measurement variation and should therefore not be considered as indicative of an actual clinical change. Surprisingly, despite the different chart designs and examiners that are compared in this study, this reported variability does not substantially differ from the test–retest variability of same-chart measurements. The test–retest variability of Snellen charts is actually even wider, as depicted in Fig. 4. This figure also illustrated that the 95% LoA of most digital tools exceed the 95% LoA reflecting test–retest variability of traditional VA charts. On the basis of these findings, the digital tools appear less precise than traditional VA charts. Obviously, as with any medical technological device, quality differences of the different tools affect

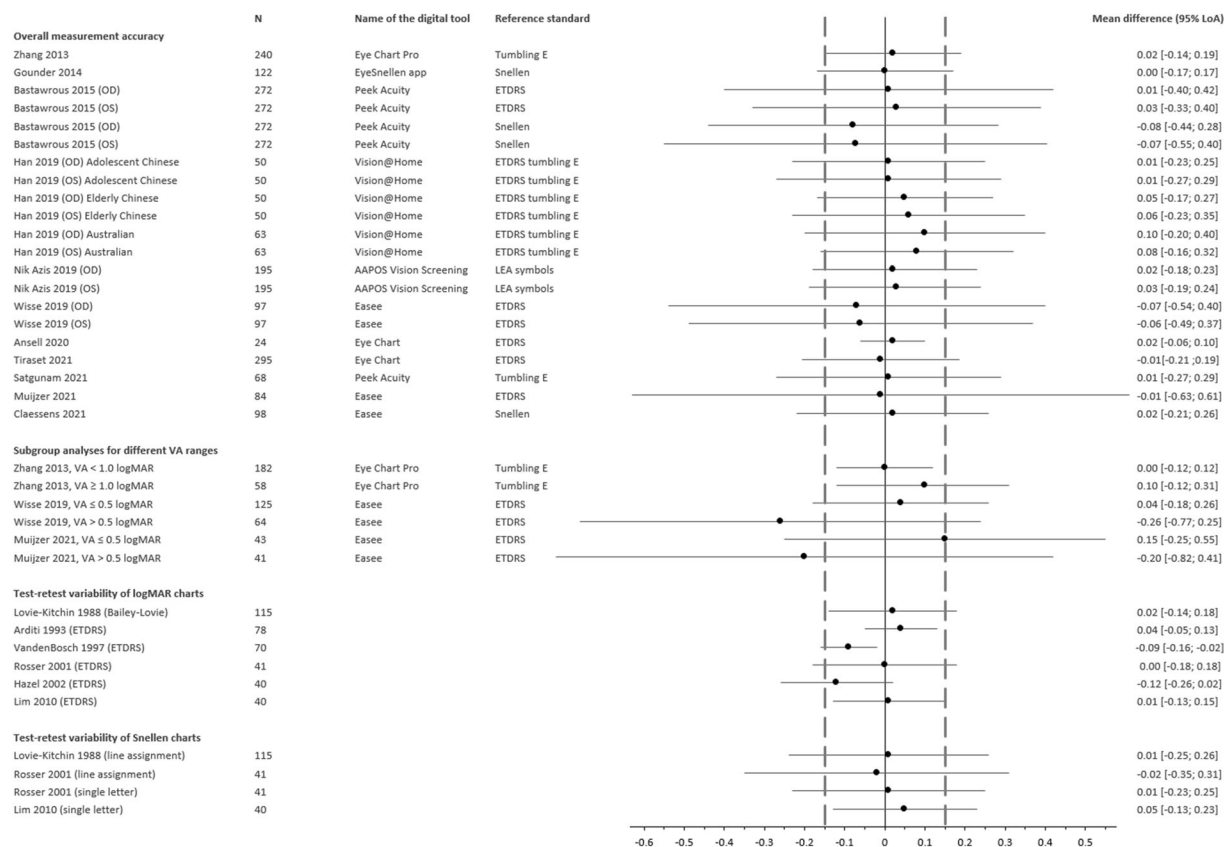


Fig. 4 Mean differences between distance visual acuity assessments (digital tool minus reference standard) and 95% limits of agreement in logMAR. Some articles reported separate comparisons per subgroup or per eye.

performance. Importantly, these technologies are developing continuously and an improvement in accuracy is attained with every design iteration.

The accuracy of the near vision tests seems better than the distance VA assessments, expressed by the smaller 95% LoA. The near vision cards and the assessments using tablets are very comparable in nature, which might explain the good agreement. Although near vision testing has a remarkably smaller role in clinical practice, these tools are very easy to use. It has been proven effective to have patients do a self-test by handing over a tablet at an emergency unit [34]. Especially in an emergency setting, convenience and time-effectiveness outweigh accuracy.

The dashed lines represent ± 0.15 logMAR, a difference that has been suggested in literature to be clinically acceptable [20]. Abbreviations: *N* number of paired observations; 95% LoA 95% limits of agreement

To the best of our knowledge, this is the first systematic review that focusses on clinically validated digital tools for self-assessing VA. We conducted a comprehensive literature search and reviewed all reference lists of included studies using PRISMA guidelines. Two reviewers independently performed the literature screening, data extraction and risk of bias assessment. The bias analysis identified concerns regarding blinding of the subject and researcher for the outcome of the tests in nine studies. We do not consider this a major problem since VA tests report a quantifiable outcome which can be recorded without interpretation. In addition, the two compared tests were different, so learning effects are not expected. One study, Nik Azis et al. [26], reported a convenience sampling method for recruitment, by excluding

children who were deemed uncooperative. This may have negatively affected generalizability of the outcomes.

When comparing VA assessments two factors are important to consider. First the type of chart used significantly affects the obtained VA outcome [4]. Consequently, this affects the observed differences between the digital tools and the reference charts. There is always variation when two different VA charts are compared, and one should be careful when using charts interchangeably [5]. This also explains the difference in agreement of the Peek Acuity tool, as studied by Bastawrous et al., when different reference charts are used (Snellen vs. ETDRS, see Fig. 4) [28].

Secondly, the precision of the assessment depends strongly on the achieved VA of the tested subject. When examining the Bland–Altman plots of individual studies, we noticed that the accuracy of the tests improves for patients with better VA scores. This was also demonstrated in the subgroup analyses in Fig. 4. For the study by Zhang et al., the variability of the difference was remarkably smaller for subjects with an achieved VA < 1.0 logMAR than for the ≥ 1.0 logMAR subgroup (± 0.12 vs. ± 0.22). From our own study data, we also learned that the measurement accuracy of one specific tool (Easee) differed for various VA ranges and study populations. For example, the additional subanalysis of the data of Wisse et al. revealed an evidently smaller 95% LoA for the healthy individuals with higher VA scores (e.g. 95% LoA of 0.00; 0.58 for visual acuities ≤ 0.5 logMAR, versus 95% LoA of -0.75 ; -0.25 for visual acuities > 0.5 logMAR) [16]. More importantly, the measurements of this specific digital tool were substantially more accurate when assessing *corrected* VA (in patients with uveitis), compared to assessing *uncorrected* VA (in healthy individuals and patients with keratoconus, with refractive errors) [10, 11, 16]. The digital self-assessments provide less accurate results for poorer VA ranges, regardless of underlying ocular comorbidities. Thus, wide 95% LoA do not necessarily imply inadequate testing conditions, or a low technical quality of the studied digital tools, but can be largely attributed to the poor vision of the tested

population. A lower accuracy in poorer ranges is a common feature of VA assessments. A recurrent clinical reason is that the person administering the conventional test might terminate the assessment too early. The Snellen chart in particular shows a poorer precision in lower VA ranges and the differences between Snellen and ETDRS increase in this range [21]. This can be explained by the chart design. When testing Snellen acuity, the tester uses a line assignment method. The poor vision lines usually contain only 1 or 2 letters. Thus, missing a letter on these lines can make a huge difference in this range [5]. In clinical practice, both time and convenience are essential, and therefore Snellen charts remain popular. We suggest that the accuracy of the digital VA self-assessments in poorer VA ranges could be improved by extending the initial assessment of individuals with poor scores for retesting (i.e. performing another assessment with different optotypes for internal validation and adjustment of the initial VA score).

Two studies show a remarkable narrow 95% LoA: Ansell et al. [29] and Zhang et al. [18] in the subgroup with VA better than 0.1 Snellen. These papers show a better agreement than well-executed test–retest studies of VA wall charts, evaluated in controlled, experimental settings.

Future research should focus on the performance of the digital tools in unsupervised conditions. Some of the included studies reported specifically that tests were performed at fixed distances, for example with the head of the participant leaning against an ophthalmologic chin piece and the smartphone or chart fixed on a desk (Brucker et al. [14]). We expect the accuracy to be different in real-world, less-controlled conditions. We encourage stratifying for different VA ranges when evaluating agreement between VA charts. We strongly recommend using logMAR charts as reference charts as they are the most accurate and consistent. For follow-up purposes, good repeatability (test–retest reliability) of the tools is important. This creates excellent opportunities for follow-up and signaling worsening of vision, as obtaining baseline measurements will allow future vision comparisons. Only two of

the included validation studies reported on repeatability [13, 28].

Digital testing with mobile technology has many important advantages that outweigh accuracy. Traditional in-hospital VA testing requires patients to physically attend a clinic and consumes substantial hospital resources. The former might be particularly difficult for patients from rural areas, the elderly and immobilized patients. With the increasing digitization and availability of mobile devices, the digital tools have the potential to identify the most important cause of visual impairment worldwide: uncorrected refractive errors. Visual impairment has a negative impact on school performance, employability and quality of life in general. Strikingly, 53% of the visually impaired people have poor vision that can be prevented or remedied by glasses or contacts [35]. For screening of visual impairment, one is mostly interested in identifying a low VA range, rather than determining an exact value. When looking at the accuracy of the Peek Acuity tool (Bastawrous et al. [28]) in the forest plot, the 95% LoA are wider than the previously discussed test–retest studies of traditional charts. Notwithstanding, the tool has been successfully implemented in school screening programmes for identifying visual impairment. This has been investigated among various study populations around the globe [36–38]. Sensitivity, specificity and predictive value varied among the studies, but included some promising results. A screening study in the USA illustrated how sensitivity differed per age category and found the highest sensitivity for detecting decreased vision in 3–5-year-olds (93–100%) [38]. In a study in Kenya, the sensitivity (77%) was reported to be similar to the conventional screening method: standard E-card assessments by school teachers [37]. Importantly, when visual impairment was detected by the smartphone application, personalized SMS reminders to attend the hospital were sent to the children’s parents or guardians, increasing adherence to hospital referral. Specificity of the digital screening was a bit lower compared to the conventional method (91% vs. 97%). The authors of the original article suggested testing strategies to reduce the false positive rate, including retesting the

children who tested positive at first and a local triage service to review all children who initially screened positive.

Our systematic review indicates that the low rate of bias makes digital tools for self-assessment of VA a promising avenue for delivering eye care remotely. The precision of most tools seems lower than traditional VA charts, though these differences diminish when assessing individuals with better VA. These self-assessments of visual function have great potential for screening purposes, particularly to increase access to eye care, which is acutely relevant in the current COVID-19 pandemic. The landscape of digital medicine has been rapidly changing, especially over the last few years: we expect the accuracy of the current tools to improve with every iteration and new tools to be introduced in the coming years.

ACKNOWLEDGEMENTS

Funding. No funding or sponsorship was received for this study or publication of this article.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Author Contributions. Conceptualization: Janneau Claessens, Robert Wisse. Literature search and critical appraisal: Janneau Claessens, Judith Geuvers. Drafting the original manuscript: Janneau Claessens, Judith Geuvers. Critical revision of manuscript: Saskia Imhof, Robert Wisse. Supervision: Robert Wisse. Final approval of manuscript: all authors.

Disclosures. Robert Wisse is a paid speaker for Thea Pharma BV, paid speaker for Carl Zeiss the Netherlands, and medical advisor/consultant/shareholder for Easee BV.

Compliance with Ethics Guidelines. This article is based on previously conducted studies and does not contain any new studies with human participants or animals performed by any of the authors.

Data Availability. Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Open Access. This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Arditi A, Cagenello R. On the statistical reliability of letter-chart visual acuity measurements. *Investig Ophthalmol Vis Sci.* 1993;34(1):120–9.
2. Salmon JF. *Kanski's clinical ophthalmology - Ninth Edition.* Edinburgh: Elsevier. 2020.
3. McGraw P, Winn B, Whitaker D. Reliability of the Snellen chart. *BMJ.* 1995;310:1481–2.
4. Colenbrander A. Measuring Vision and Vision Loss. *Duane's Clinical Ophthalmology Volume 5, Chapter 51.* Lippincott Williams & Wilkins. 2001.
5. Kaiser PK. Prospective evaluation of visual acuity assessment: a comparison of Snellen versus ETDRS charts in clinical practice (an AOS thesis). *Trans Am Ophthalmol Soc.* 2009;107:311–24.
6. Yeung WK, Dawes P, Pye A, et al. eHealth tools for the self-testing of visual acuity: a scoping review. *NPJ Digit Med.* 2019;2:82. <https://doi.org/10.1038/s41746-019-0154-5>.
7. The European Parliament and the Council of the European Union. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. 2017. <http://data.europa.eu/eli/reg/2017/745/2020-04-24>.
8. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 2009;6(7).
9. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res.* 1999;32:135–60.
10. Muijzer M, Claessens J, Cassano F, Godefrooij D, Prevooy Y, Wisse R. The evaluation of a web-based tool for measuring the uncorrected visual acuity and refractive error in keratoconus eyes: a prospective open-label method comparison study. *medRxiv.* 2021. <https://doi.org/10.1101/2021.03.01.21252664>.
11. Claessens JIJ, van Egmond J, de Boer JH, Wisse RPL. The evaluation of a novel tool to remotely assess visual acuity in chronic uveitis patients during the COVID-19 pandemic. *medRxiv.* 2021. <https://doi.org/10.1101/2021.04.14.21255457>.
12. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(4):529–36.
13. Phung L, Gregori NZ, Ortiz A, Shi W, Schiffman JC. Reproducibility and comparison of visual acuity obtained with Sightbook mobile application to near card and Snellen chart. *Retina.* 2016;36(5):1009–20.
14. Brucker J, Bhatia V, Sahel JA, Girmens JF, Mohand-Saïd S. Odysight: a mobile medical application designed for remote monitoring—a prospective study comparison with standard clinical eye tests. *Ophthalmol Ther.* 2019;8(3):461–76.
15. Tiraset N, Poonyathalang A, Padungkiatsagul T, Deeyai M, Vichitkunakorn P, Vanikiyeti K. Comparison of visual acuity measurement using three methods: standard ETDRS chart, near chart and a smartphone-based eye chart application. *Clin Ophthalmol.* 2021;15:859–69.

16. Wisse RPL, Muijzer MB, Cassano F, Godefrooij DA, Prevooy YFDM, Soeters N. Validation of an independent web-based tool for measuring visual acuity and refractive error via the MORE (Manifest Versus Online Refractive Evaluation) trial: prospective open-label noninferiority clinical trial. *J Med Internet Res*. 2019;21(11):1–13.
17. Jan-Bond C, Wee-Min T, Hong-Kee N, et al. REST – an innovative rapid eye screening test. *J Mob Technol Med*. 2015;4(3):20–5.
18. Zhang ZT, Zhang SC, Huang XG, Liang LY. A pilot trial of the iPad tablet computer as a portable device for visual acuity testing. *J Telemed Telecare*. 2013;19(1):55–9.
19. World Health Organization. Global initiative for the elimination of avoidable blindness: action plan 2006–2011. Geneva: WHO; 2007.
20. Siderov J, Tiu AL. Variability of measurements of visual acuity in a large eye clinic. *Acta Ophthalmol Scand*. 1999;77(6):673–6.
21. Lovie-Kitchin JE. Validity and reliability of visual acuity measurements. *Ophthalmic Physiol Opt*. 1989;9(4):458–458.
22. Vanden Bosch ME, Wall M. Visual acuity scored by the letter-by-letter or probit methods has lower retest variability than the line assignment method. *Eye*. 1997;11(3):411–7.
23. Rosser DA, Laidlaw DAH, Murdoch IE. The development of a “reduced logMAR” visual acuity chart for use in routine clinical practice. *Br J Ophthalmol*. 2001;85(4):432–6.
24. Hazel CA, Elliott DB. The dependency of logMAR visual acuity measurements on chart design and scoring rule. *Optom Vis Sci*. 2002;79(12):788–92.
25. Lim LA, Frost NA, Powell RJ, Hewson P. Comparison of the ETDRS logMAR, compact reduced logMAR and Snellen charts in routine clinical practice. *Eye*. 2010;24(4):673–7.
26. Nik Azis NN, Chew FLM, Rosland SF, Ramlee A, Che-Hamzah J. Parents’ performance using the AAOPOS Vision Screening App to test visual acuity in Malaysian preschoolers. *J AAOPOS*. 2019;23(5):268.e1-268.e6.
27. Gounder PA, Cole E, Colley S, Hille DM. Validation of a portable electronic visual acuity system. *J Mob Technol Med*. 2014;3(2):35–9.
28. Bastawrous A, Rono HK, Livingstone IAT, et al. Development and validation of a smartphone-based visual acuity test (Peek Acuity) for clinical practice and community-based fieldwork. *JAMA Ophthalmol*. 2015;133(8):930–7.
29. Ansell K, Maconachie G, Bjerre A. Does the EyeChart App for iPhones give comparable measurements to traditional visual acuity charts? *Br Ir Orthopt J*. 2020;16(1):19–24.
30. Satgunam P, Thakur M, Sachdeva V, Reddy S, Rani PK. Validation of visual acuity applications for teleophthalmology during COVID-19. *Indian J Ophthalmol*. 2021;69(2):385–90.
31. Calabrèse A, To L, He Y, Berkholtz E, Rafian P, Legge GE. Comparing performance on the MNREAD iPad application with the MNREAD acuity chart. *J Vis*. 2018;18(1):8.
32. Tofigh S, Shortridge E, Elkeeb A, Godley BF. Effectiveness of a smartphone application for testing near visual acuity. *Eye*. 2015;29(11):1464–8.
33. Han X, Scheetz J, Keel S, et al. Development and validation of a smartphone-based visual acuity test (Vision at Home). *Transl Vis Sci Technol*. 2019;8(4):27.
34. Pathipati AS, Wood EH, Lam CK, Sáles CS, Moshfeghi DM. Visual acuity measured with a smartphone app is more accurate than Snellen testing by emergency department providers. *Graefes Arch Clin Exp Ophthalmol*. 2016;254(6):1175–80.
35. Naidoo KS, Leasher J, Bourne RR, et al. Global vision impairment and blindness due to uncorrected refractive error, 1990–2010. *Optom Vis Sci*. 2016;93(3):227–34.
36. de Venecia B, Bradfield Y, Møller Trane R, Bareiro A, Scalomogna M. Validation of Peek Acuity application in pediatric screening programs in Paraguay. *Int J Ophthalmol*. 2018;11(8):1384–9.
37. Rono HK, Bastawrous A, Macleod D, et al. Smartphone-based screening for visual impairment in Kenyan school children: a cluster randomised controlled trial. *Lancet Glob Heal*. 2018;6(8):e924–32.
38. Zhao L, Stinnett SS, Prakalapakorn SG. Visual acuity assessment and vision screening using a novel smartphone application. *J Pediatr*. 2019;213:203-210.e1.