

Smartphone Imaging in Ophthalmology: A Comparison with Traditional Methods on the Reproducibility and Usability for Anterior Segment Imaging

David ZY Chen, ^{1,2}MBBS, Clement WT Tan, ¹MMed (Ophth), FRCSEd, FAMS

Abstract

Introduction: This study aimed to determine the reproducibility and usability of anterior segment images taken from a smartphone stabilised on a slit-lamp with those taken from a custom-mounted slit-lamp camera. **Materials and Methods:** This was a prospective, single-blind comparative digital imaging validation study. Digital photographs of patients with cataract were taken using a smartphone camera (an iPhone 5) on a telescopic mount and a Canon EOS 10D anterior segment camera. Images were graded and compared according to the Lens Opacification Classification System III (LOCS III). **Results:** A total of 440 anterior segment images were graded independently by 2 ophthalmologists, 2 residents and 2 medical students. Intraclass correlation (ICC) between the iPhone and anterior segment camera images were fair for nuclear opalescence (NO) and nuclear colour (NC), and excellent for cortical (C) and posterior subcapsular (PSC) (NO: ICC 0.40, 95% CI, 0.16 to 0.57; NC: ICC 0.47, 95% CI, 0.16 to 0.66; C: ICC 0.76, 95% CI, 0.71 to 0.81; PSC: ICC 0.81, 95% CI, 0.76 to 0.85). There was no difference in grader impression of confidence and images usability between both cameras ($P = 0.66$ and $P = 0.58$, respectively). **Conclusion:** Anterior segment images taken from an iPhone have good reproducibility for retro-illuminated images, but fair reproducibility for NO and NC under low light settings. There were no differences in grader confidence and subjective image suitability.

Ann Acad Med Singapore 2016;45:6-11

Key words: Cataract, Clinical Ophthalmology, iPhone

Introduction

There is an increasing trend amongst medical professionals in recent years to incorporate smartphones as an informal clinical tool.^{1,2} The ubiquity and accessibility of smartphones have prompted many ophthalmologists to use them as a clinical tool, and several studies have described the techniques on how to utilise smartphone cameras during ophthalmic examinations.^{3,4}

Several studies have demonstrated qualitatively the ease of capturing ophthalmic images including fundus imaging using smartphones.^{3,5} The smartphone is also a suitable platform for reviewing images taken remotely, as demonstrated in a retinopathy of prematurity (ROP) screening study in India.⁶ However, to our knowledge, there are no studies that quantitatively compares the quality of

images taken of patients' eyes in a clinical setting using a smartphone with those taken with a custom-mounted slit-lamp camera. Ye et al has concluded recently that the image spatial resolution of custom-mounted slit-lamp cameras was better than that of smartphones (iPhone 4 and 4S), though both were inferior to direct ocular viewing through the slit-lamp.⁷ However, there are no published studies on the clinical reproducibility of smartphone images as compared to those taken with an anterior standard camera. This study describes a relatively simple method of obtaining stabilised slit-lamp images of eyes with cataract using a smartphone, and attempts to quantify its reproducibility using the validated Lens Opacification Classification System III (LOCS III), which covers important components in an anterior segment examination, namely slit examination (colour, opalescence) and retro-illumination.⁸

¹Department of Ophthalmology, National University Health System, Singapore

²Yong Loo Lin School of Medicine, National University of Singapore, Singapore

Address for Correspondence: Dr David Chen, Department of Ophthalmology, National University Health System, 1E Kent Ridge Road, Level 7, NUHS Tower Block, Singapore 119228.

Email: dzychen18@gmail.com

Materials and Methods

Study Design

This was a single centre, clinic-based, and single-blind comparative digital imaging validation study. Eligible patient subjects were recruited from the general eye clinic after their clinical consultation. Patient subjects had to be at least 21 years old, willing and able to give informed consent, and have no concurrent intraocular or lid pathologies that might obscure photo-taking. Inclusion criteria for graders were as follows: at least 21 years old, willing and able to give consent, had undergone at least 1 clinical ophthalmology posting, and familiar with grading cataract images through LOCS III. This study was conducted in accordance with the tenets of the World Medical Association's Declaration of Helsinki and had ethics approval from the National Healthcare Group Domain Specific Review Board. Informed consent was obtained from all participating subjects.

Image Capturing Protocol

As part of the routine examination, image capture was performed in the same room for 1 or both eligible eyes of the patient after dilation with 1.0% tropicamide. For images taken using the anterior segment camera, a slit-lamp BP 900® (Haag-Streit, USA) with custom-mounted 6.3-megapixel Canon® EOS 10D digital single-lens reflex camera (Canon, Japan) was used; images were taken at a resolution of 3072 x 2048 pixels. For images taken using a smartphone, an iPhone 5 (Apple, Cupertino, CA) on a telescopic mount (Orion SteadyPix Universal Camera Mount, USA) was used; images were taken at 2448 x 3264 pixels using ProCam application (ProCam, Apple Store) which allows an autofocus lock for easier focusing.

Four simultaneous images of each eligible eye were taken for each instrument: 1 diffusely illuminated image, 1 slit image of the anterior segment, and 2 retro-illuminated images focused on cortex and posterior subcortex respectively.⁸ Appropriate cropping was performed to retain only elements necessary for LOCS III grading (Fig. 1). Two anterior segment images were used to assess nuclear opalescence (NO) and nuclear colour (NC), while the remaining 2 images were used to grade cortical (C) and posterior subcapsular (PSC), respectively. Room settings and slit-lamp settings were modelled after standard conditions for LOCS III image grading, as described in greater detail elsewhere.⁹

Image Grading Protocol

Digital images were randomised separately for subject identity and for instrument used. To minimise response error, the images were shuffled through a simple random sampling (without replacement) using a computer software. A 10-minute briefing on LOCS III grading protocol was

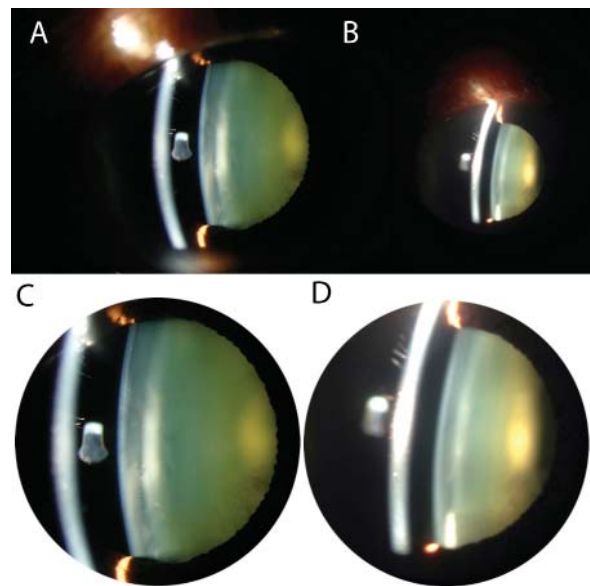


Fig. 1. Digital images taken from the anterior segment camera and the iPhone 5. Anterior segment image of the same patient taken with the anterior segment camera (A) and the iPhone 5 (B). Images were edited and cropped appropriately for both the anterior segment camera (C) and the iPhone 5 (D) to retain only necessary components for grading without giving away the source of image capture.

given prior to grading.¹⁰ Graders individually graded all images on a standardised questionnaire. Images for grading were displayed on a 13-inch MacBook Pro laptop (Apple, Cupertino, CA) with a display resolution of 2560 x 1600 and maximum screen brightness. A LOCS III grading chart was also shown on the same screen for reference. Graders were blinded to the type of instrument being used for image capturing, but were informed that 2 separate instruments were used. All gradings were completed over a minimum of 3 sessions; each session did not exceed 1 hour and involved a maximum of 40 eyes.

A standardised questionnaire and the LOCS III grading protocol were used to grade the cataract.⁸ Each image also had an option for “Can’t grade”, when image quality was too poor for grading, or when 1 region of the lens obscured another region.¹¹ In addition, 2 subjective 5-point Likert scale questions were asked for each eye and assessed holistically i.e. based on NO, NC, C and PSC images of the same eye. One question assessed grader confidence, “I am confident of the pathologies I identified on this image”, while another assessed subjective image suitability, “The quality of the image is suitable for identification of pathology”.

Outcome Measures

The primary outcome measure of this study was the reliability of LOCS III grading in assessing images taken from an iPhone and those taken from anterior segment camera.

Table 1. Agreement between 6 Graders for Images Taken with the iPhone 5 and the Anterior Segment Camera

	Inter-Rater Agreement (iPhone 5)		Inter-Rater Agreement (Anterior Segment Camera)	
	Mean (SD)	ICC Coefficient (95% CI)	Mean (SD)	ICC Coefficient (95% CI)
Nuclear opalescence images (NO)	3.93 (0.35)	0.317 (0.193, 0.468)	3.38 (0.47)	0.391 (0.252, 0.539)
Nuclear colour images (NC)	4.07 (0.24)	0.485 (0.357, 0.625)	3.44 (1.12)	0.548 (0.392, 0.687)
Cortical images (C)	2.10 (0.15)	0.741 (0.638, 0.833)	1.91 (0.16)	0.847 (0.775, 0.906)
Posterior subcapsular images (PSC)	1.64 (0.40)	0.742 (0.619, 0.970)	1.80 (0.50)	0.691 (0.547, 0.810)

CI: Confidence interval; ICC: Intraclass correlation; SD: Standard deviation

Secondary outcome measures included the number of non-gradable images taken from an iPhone and anterior segment camera, inter-rater variability in grading iPhone images as well as anterior segment camera images, and grader confidence and image suitability for LOCS III grading.

Statistical Analysis

Statistical analysis was performed using SPSS 21.0 (SPSS Inc, Chicago, Illinois). To achieve an expected agreement of kappa value (K) of at least 0.8 with a minimum power of 80%, images from 51 eyes with cataract were required. Intraclass correlation (ICC) was used to assess the between-graders and between-instrument reproducibility of the LOCS III for NO, NC, C, and PSC images. ICC interpretation was as follows: poor reproducibility if $ICC < 0.4$, fair if $0.4 \leq ICC \leq 0.75$, and excellent if $ICC > 0.75$.¹² Bland-Altman plot was also used to assess agreement between LOCS III measurements.¹³ Pearson's Chi-square test or Fisher's exact test was performed to determine associations between categorical variables and paired t-test was used to test for any difference between repeated continuous measurements. A *P* value of less than 0.05 was considered statistically significant in this study.

Results

Image Grading

A total of 440 digital images (iPhone: 220 images,

anterior segment camera: 220 images) were taken from 55 eyes of 32 patients. Images were graded by 6 separate healthcare personnel independently (2 ophthalmologists, 2 ophthalmology residents, 2 medical students), over 3 separate sessions each. All graders completed all images. Out of the 440 images, 421 images (95.7%) were gradable by 1 rater (iPhone, $n = 211$, 95.9%; anterior segment camera, $n = 210$, 95.5%, $P = 0.82$). The cumulative percentage of non-gradable images for all raters was 6.3% (166 out of 2640 images; NS: $n = 14$, 2.1%; NC: $n = 14$, 2.1%; C: $n = 61$, 9.2%; PSC: $n = 77$, 11.7%).

Reproducibility between Graders

Intraclass correlation between graders for images taken by an anterior segment camera was poor for NO, fair for NC and PSC, and excellent for C images (Table 1). Intraclass correlation between graders for images taken by iPhone was poor for NO images and fair for the other subtypes (Table 1).

Reproducibility between Instruments

The reproducibility of images taken from the iPhone, when compared to those taken by the anterior segment camera, is shown in Table 2. Compared with the anterior segment camera, images taken by the iPhone were graded with significantly higher scores for NO and NC, and significantly lower scores for PSC. However, the scores were not significantly different

Table 2. Agreement between the iPhone 5 and the Anterior Segment Camera

Test Modality	Instrument Used	Mean (SD)	<i>P</i> Value*	ICC Coefficient (95% CI)
Nuclear opalescence images (NO)	Anterior segment camera	3.41 (0.94)	<0.001	0.399 (0.162, 0.567)
	iPhone 5	4.02 (0.98)		
Nuclear colour images (NC)	Anterior segment camera	3.45 (0.95)	<0.001	0.471 (0.156, 0.659)
	iPhone 5	4.09 (0.97)		
Cortical images (C)	Anterior segment camera	1.91 (1.46)	0.37	0.760 (0.705, 0.805)
	iPhone 5	1.85 (1.54)		
Posterior subcapsular images (PSC)	Anterior segment camera	1.74 (1.71)	0.002	0.805 (0.756, 0.845)
	iPhone 5	1.54 (1.58)		

CI: Confidence interval; ICC: Intraclass correlation; SD: Standard deviation

*Paired t-test.

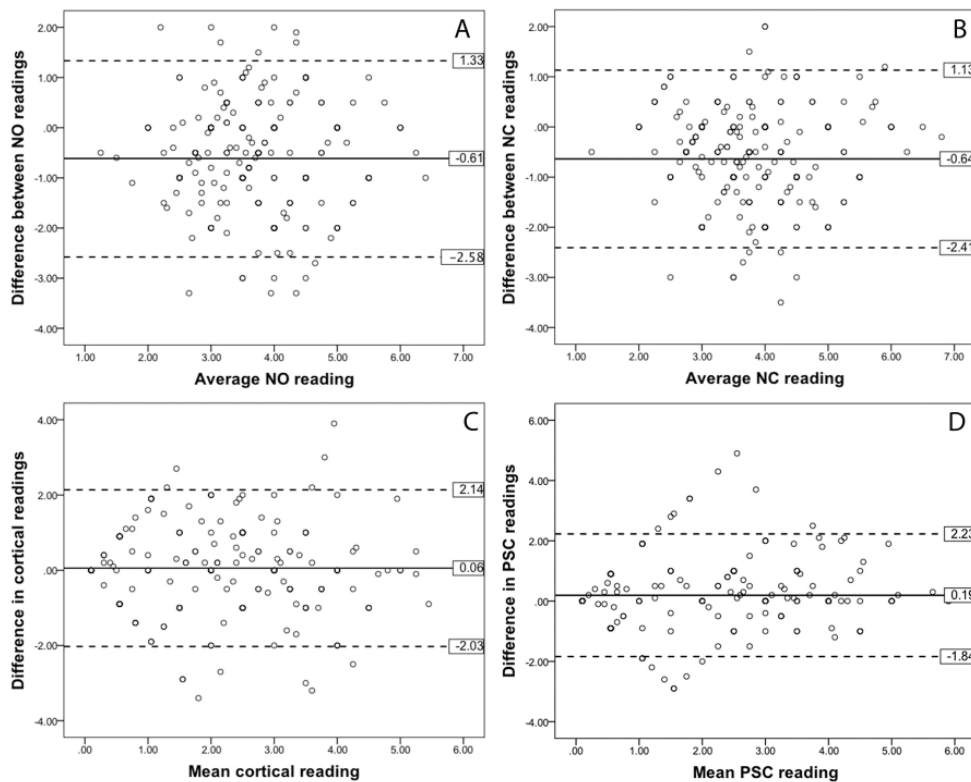


Fig. 2. Bland-Altman plot comparing images from the iPhone and from the anterior segment camera. Difference in measurement between instruments is calculated as: anterior segment camera – iPhone (where anterior segment camera represents images taken from the anterior segment camera, and iPhone represents images taken from the iPhone 5). Solid horizontal line represents the mean difference between instruments in each plot; dashed horizontal lines represent the upper and lower limits of 2 standard deviations from the mean difference. NC: Nuclear colour images; NO: Nuclear opalescence images; PSC: Posterior subcapsular images

for C images. These findings were similar and consistent for all cataract subtypes among the graders and were also reflected in the Bland-Altman plots (Fig. 2).

Overall, reproducibility for iPhone images was fair for NO and NC images, and excellent for C and PSC images. These findings were consistent with those by ophthalmologists and ophthalmology residents. Medical students found that there was poor reproducibility for NO images, and fair reproducibility for NC, C and PSC images.

Subjective Scores

Cumulative responses to the 2 subjective Likert scale questions were collected. Graders expressed their confidence

in grading images for 243 (73.6%) eyes that were taken using the anterior segment camera, and for 238 (72.1%) eyes taken using the iPhone ($P = 0.66$). On the other hand, graders felt that the images of 216 (65.5%) eyes taken by the anterior segment camera and the images of 197 (59.7%) eyes taken by the iPhone were suitable for grading ($P = 0.13$). Of the 3 grader categories, medical students responded the most positively to both questions, with more than 70% choosing “Agree” or “Strongly agree” to either question (Table 3).

Discussion

This study evaluated and compared the reproducibility of images taken using an iPhone with those captured by

Table 3. Confidence and Suitability by Proportion of Disagree (Stratified by Grader Categories)

Question	Grader Category	n (%)	
		Anterior Segment Camera	iPhone 5
“I am confident of the pathologies I identified on this image” – Agree/strongly agree	Ophthalmologist	64 (58.2)	72 (65.5)
	Resident	89 (80.9)	83 (75.5)
	Medical student	90 (81.8)	83 (75.5)
“The quality of the image is suitable for identification of pathology” – Agree/strongly agree	Ophthalmologist	60 (54.5)	63 (57.3)
	Resident	72 (65.5)	57 (51.8)
	Medical student	84 (76.4)	77 (70.0)

an anterior segment camera, using the validated LOCS III grading system on patients with cataract. The overall intraclass correlation between these 2 instruments ranged from fair to excellent, with the best reproducibility coming from the images of C and PSC cataract. Greater reproducibility was produced when graded by ophthalmologists or ophthalmology residents, while reproducibility was poorer overall when graded by medical students. Intraclass correlation between graders for anterior segment camera images ranged from poor to excellent depending on the type of image graded, and poor to fair for iPhone images. There was no significant difference in terms of grader confidence and subjective image suitability between anterior segment camera and iPhone, as measured through the 2 Likert scale questions.

This is the first study of its kind that uses ICC to grade anterior segment images through cataract assessment. Previous studies on photo grading was performed by describing the presence and absence of pathology.^{14,15} From this study, it appears that iPhone images are more reliable for estimating areas of C and PSC cataract as compared to images captured by the anterior segment camera. However, reproducibility between images taken by the iPhone and anterior segment camera for NC and NO images was only fair, with the Bland-Altman plot demonstrating a consistently higher rating for the iPhone (Fig. 2). This could possibly be due to lighting issues—the standard LOCS III setting requires a dimly lit room for image capture. As a result, the auto-ISO and auto-white balance effects of the iPhone could have been artificially increased in the absence of an external flash, resulting in oversaturation of image colour and consequently, opacification. To compensate for the phone's form factor, the iPhone 5 uses a 4.54 mm x 3.42 mm CMOS sensor while the Canon EOS 10D uses a 22.7 mm x 15.1 mm CMOS sensor, and this could explain the superior performance of the latter under lower light settings. Interestingly, this phenomenon did not appear to affect retro-illuminated images (i.e. C and PSC images), suggesting that retro-illuminated images are well taken and may benefit from auto-ISO adjustments from the iPhone even in dim lighting. In view of such lighting issues, iPhone images may not necessarily be inferior to anterior segment camera images; rather, users should be mindful of ambient lighting while taking ophthalmic images using the iPhone, and provide a sufficiently well lit environment whenever possible.

In addition, we noted that there was considerable inter-rater variability in this study. There was fair to good reproducibility for retro-illuminated images, but grader agreement for NO and NC was poor and fair, respectively. This relatively poor agreement between graders manifested in the grading for both anterior segment camera and iPhone

images (Table 1), suggesting that there is great subjective variability when grading despite using the LOCS III. Previous studies have demonstrated good inter and intrarater reliability on LOCS III; however, those studies were conducted on 35 mm film photographs instead of digital images.⁸ A study by Tan et al which compared inter-rater variability between junior and senior ophthalmologists demonstrated moderate to substantial agreement in inter-rater agreement, though the study was conducted through slit-lamp observation.¹⁰ This could imply that digital still images may not be ideal for LOCS III grading, and an alternative standardised photographic grading system such as the Wisconsin system could be used, though it may vary substantially with LOCS III measurements.¹⁶ Alternatively, the variability found in this study may be due to the small number of raters used.

The technique used to capture images from the smartphone—by securing the smartphone on a slit-lamp through a telescopic mount—significantly reduced shake and improved photo usability. At the same time, live image transmission through a high resolution screen facilitated image focus and capture, and could potentially be used for live demonstrations. However, despite the high screen resolution, the relatively small screen size of the iPhone 5 limited the eventual image focus as some images which had appeared sharp on the iPhone screen eventually turned out unfocused when enlarged on a 13-inch screen. During the image capture, we found that such focusing issues could be reduced by predetermining the camera focus settings, and manually focusing the slit-lamp, while requesting patients to hold as steady as possible.

Several authors have previously described ways to utilise the smartphone for ophthalmology in a clinical setting.³⁻⁵ The ubiquity and portability of the smartphone could prove to be a useful tool for clinical teaching. In this study, we found that medical students were equally confident when grading images taken from the iPhone and the anterior segment camera as measured through the 2 Likert scale questions, and there was fair agreement between images taken from both instruments in all categories. However, grading reproducibility between instruments ranged from fair to poor for the same group of students. This could be due to non-ideal light settings for the iPhone as explained above, though it could also be due to a poorer understanding of LOCS III grading scale in general. More experience in grading cataract images would certainly aid in enhancing reproducibility, and we believe in the potential of using the iPhone as a clinical education tool. Further studies could be conducted to evaluate the educational value of the smartphone in clinical ophthalmology.

We acknowledge the limitations of this study. Test-retest reliability was not assessed, and this could have made it

more difficult to compare intrarater reliability. Nonetheless, the purpose of this study was not so much to evaluate the best way to assess cataract, but to compare anterior segment images taken from smartphones with an anterior segment camera, using the LOCS III as a validated and standardised questionnaire. This study also confined itself to examining the graders' ability to grade photos for 1 type of pathology only, and it remains uncertain whether results could be extrapolated to other pathologies of the eye. A previous study by Kumar et al to validate digital images suggests such a possibility for smartphone images.¹⁴ Similarly, given the encouraging performances of the iPhone in capturing fundus images, as presented by Bastawrous previously using the iPhone 4,⁵ we remain very positive on the utility of the smartphone as a clinical adjunct for educational purposes or photo documentation.

Conclusion

This study has demonstrated that the iPhone has good reproducibility for retro-illuminated images, but tends to produce oversaturation of the anterior segment under low light settings. Graders reported similar levels of confidence when grading photos taken from the iPhone and the anterior segment camera. Images taken from the iPhone could be of adequate quality for teaching purposes, but further studies would be necessary to evaluate the suitability of iPhone images in demonstrating specific anterior segment pathologies, and for evaluating its utility in fundus photographs.

Acknowledgement

The authors would like to thank Ms Wong Wan Ling for her assistance in statistical analysis and Dr Dujeeva D Samarasekera for helping to review the paper.

REFERENCES

1. Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. *BMC Med Inform Decis Mak* 2012;12:67.
2. Ozdalga E, Ozdalga A, Ahuja N. The smartphone in medicine: a review of current and potential use among physicians and students. *J Med Internet Res* 2012;14:e128.
3. Lord RK, Shah VA, San Filippo AN, Krishna R. Novel uses of smartphones in ophthalmology. *Ophthalmology* 2010;117:1274-1274.e3.
4. Bastawrous A, Cheeseman RC, Kumar A. iPhones for eye surgeons. *Eye (Lond)* 2012;26:343-54.
5. Bastawrous A. Smartphone funduscopy. *Ophthalmology* 2012;119:432-3. e2; author reply 433.
6. Vinekar A, Gilbert C, Dogra M, Kurian M, Shainesh G, Shetty B, et al. The KIDROP model of combining strategies for providing retinopathy of prematurity screening in underserved areas in India using wide-field imaging, tele-medicine, non-physician graders and smart phone reporting. *Indian J Ophthalmol* 2014;62:41-9.
7. Ye Y, Jiang H, Zhang H, Karp CL, Zhong J, Tao A, et al. Resolution of slit-lamp microscopy photography using various cameras. *Eye Contact Lens* 2013;39:205-13.
8. Chylack LT Jr, Wolfe JK, Singer DM, Leske MC, Bullimore MA, Bailey IL, et al. The Lens Opacities Classification System III. The Longitudinal Study of Cataract Study Group. *Arch Ophthalmol* 1993;111:831-6.
9. Chylack LT Jr, Leske MC, McCarthy D, Khu P, Kashiwagi T, Sperduto R. Lens opacities classification system II (LOCS II). *Arch Ophthalmol* 1989;107:991-7.
10. Tan AC, Loon SC, Choi H, Thean L. Lens Opacities Classification System III: cataract grading variability between junior and senior staff at a Singapore hospital. *J Cataract Refract Surg* 2008;34:1948-52.
11. Chylack Incorporated. Frequently asked questions: grading. Available at: http://chylackinc.com/LOCS_III/LOCS_files/Grading%20FAQs.pdf. Accessed on 20 December 2015.
12. Rudnisky CJ, Tennant MT, Weis E, Ting A, Hinz BJ, Greve MD. Web-based grading of compressed stereoscopic digital photography versus standard slide film photography for the diagnosis of diabetic retinopathy. *Ophthalmology* 2007;114:1748-54.
13. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clin Rehabil* 1998;12:187-99.
14. Kumar S, Yogesan K, Constable IJ. Telemedical diagnosis of anterior segment eye diseases: validation of digital slit-lamp still images. *Eye (Lond)* 2009;23:652-60.
15. Threlkeld AB, Fahd T, Camp M, Johnson MH. Telemedical evaluation of ocular adnexa and anterior segment. *Am J Ophthalmol* 1999;127:464-6.
16. Tan AC, Wang JJ, Lamoureux EL, Wong W, Mitchell P, Li J, et al. Cataract prevalence varies substantially with assessment systems: comparison of clinical and photographic grading in a population-based study. *Ophthalmic Epidemiol* 2011;18:164-70.