Original article

Telemedicine correlation in retinopathy of prematurity between experts and non-expert observers

D. Ossandón a, M. Zanolli a, J.P. López a, R. Stevenson a, R. Agurto a, C. Cartes b

a Equipo Multicéntrico Retinopatía del Prematuro, Hospital de Niños Roberto del Río, Santiago, Chile
b Fundación Oftalmológica Los Andes, Santiago, Chile

ARTICLE INFO

Article history:
Received 28 January 2014
Accepted 30 June 2014
Available online 5 March 2015

Keywords:
Telemedicine
Retinopathy of prematurity
RetCam correlation

ABSTRACT

Objective: To study the correlation between expert and non-expert observers in the reporting images for the diagnosis of retinopathy of prematurity (ROP) in a telemedicine setting.

Methods: A cross-sectional, multicenter study, consisting of 25 sets of images of patients screened for ROP. They were evaluated by two experts in ROP and 1 non-expert and classified according to telemedicine classification, zone, stage, plus disease and Ells referral criteria. The telemedicine classification was: no ROP, mild ROP, type 2 ROP, or ROP that requires treatment. Ells referral criteria are defined as the presence at least one of the following: ROP in zone I, Stage 3 in zone I or II, or plus+. For statistical analysis, SPSS 16.0 was used. For correlation, Kappa value was performed.

Results: There was a high correlation between observers for the assessment of ROP stage (0.75; 0.54–0.88) plus disease (0.85; 0.71–0.92), and Ells criteria (0.89; 0.83–1.0). However, inter-observer values were low for zone (0.41; 0.27–0.54) and telemedicine classification (0.43; 0.33–0.6).

Conclusions: When evaluating telemedicine images by examiners with different levels of expertise in ROP, the Ells criteria gave the best correlation. In addition, stage of disease and plus disease have good correlation among observers. In contrast, the correlation between observers was low for zone and telemedicine classification.

© 2014 Sociedad Española de Oftalmología. Published by Elsevier España, S.L.U. All rights reserved.

* Corresponding author.
E-mail address: mario.zanolli@me.com (M. Zanolli).
2173-5794/© 2014 Sociedad Española de Oftalmología. Published by Elsevier España, S.L.U. All rights reserved.


**Correlación en telemedicina de retinopatía del prematuro entre observadores expertos y no expertos**

**RESUMEN**

Objetivo: Estudiar la correlación de observadores expertos y no expertos en la evaluación de imágenes para el diagnóstico de retinopatía del prematuro (ROP) en un sistema de telemedicina.

**Métodos:** Estudio transversal y multicéntrico. Veinticinco sets de imágenes de pacientes con screening para ROP evaluadas por 2 expertos en ROP y uno no experto. Se clasiﬁcaron según esquema de telemedicina: zona, estadio, enfermedad plus y según criterios derivación de Ells. El esquema de telemedicina era: no ROP, ROP leve, ROP tipo 2 o ROP que requiere tratamiento. Criterios de derivación según Ells son la presencia de al menos uno de los siguientes: ROP en zona I, estadio 3 en zona I o II, o plus+. Para el análisis estadístico se utilizó SPSS® v.16.0. Para la correlación se usó el índice de Kappa.

**Resultados:** Hubo una alta correlación entre observadores para la evaluación del estadio ROP (0,75; 0,54–0,88), enfermedad plus (0,85; 0,71–0,92) y los criterios de derivación de Ells (0,89; 0,83–1,0). Sin embargo, los valores interobservadores fueron bajos para zona (0,41; 0,27–0,54) y esquema de telemedicina (0,43; 0,33–0,6).

**Conclusiones:** Al evaluar imágenes de telemedicina por examinadores con distinto grado de especialización en ROP, la clasificación de Ells tiene la mejor correlación, junto con estadio y enfermedad plus. En cambio, la correlación entre los observadores fue baja para zona de ROP y esquema de telemedicina.

© 2014 Sociedad Española de Oftalmología. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

---

**Introduction**

Retinopathy of prematurity (ROP) is a vasoproliferative disorder that affects low weight premature babies. Several multicenter studies have developed criteria for the classification and treatment of ROP. This disorder continues to be the first cause of childhood blindness in developed countries. A survey carried out by the American Ophthalmology Academy revealed that 54% of ophthalmologists are willing to carry out screening and 20% of those who do are considering discontinuing said screening, one of the reasons being the possibility of medical-legal problems. Telemedicine is a strategy to resolve the limitations of ROP screening, i.e., lack of specialist time, shortage of specialists in charge of screening and lack of clinical documentation in images. This is particularly important for the moderate to severe grade of the disorder. In addition, telemedicine has demonstrated a better cost-effectiveness ratio than ophthalmoscopy. The primary objective of this study is to analyze the correlation of expert and non-expert observers for diagnosing ROP in a system of image analysis via telemedicine.

**Methods**

A transversal and multicenter study in which an independent ophthalmologist selected 25 sets of ocular fundus images of patients screened for ROP obtained with RetCam equipment (Clarity Medical System; 2011) (Figs. 1 and 2) by trained nonmedical staff (Fig. 3). The images were assessed by 2 ROP experts and one non-expert (resident with previous training). The images were classified according to the telemedicine scheme as zone, stage, plus and according to the Ells referral system. The statistical analysis was performed with SPSS® v.16.0 (SPSS Inc; 2007. Mac version, Chicago, USA). The correlation was assessed with the Kappa index. The study was approved by the ethics committee. The reporting scheme for telemedicine was: no ROP, slight ROP, ROP without indication for treatment or type II (stage I or 2 in zone 1 without plus; stage III in zone 2 without plus), ROP which requires treatment or type I (stage III in zone 1 without plus, stage II or 3 with plus) unclassifiable. The composite referral criteria according to Ells require one of the following criteria for positive rating: presence of ROP in zone 1, stage III in zone 2.

**Fig. 1** - Image taken by neonatology nurse utilizing the RetCam system at week 37 of gestation in premature 27-week newborn with natal weight of 750 g. The image is compatible with ROP 2 due to the presence of stage II in zone 2 without the presence of plus. Note ridge hemorrhage toward temporal without visible neovessels.
zone one or 2, or presence of plus disease. In addition, the observation time was measured.

Results

The detailed values for the correlation between experts and non-expert observers are found in Table 1. A high correlation was found between observers for assessment of ROP stage (0.75; 0.54–0.88), plus disease (0.85; 0.71–0.92) and Ells referral criteria (0.89; 0.83–1.0). Utilizing the Ells criteria, the correlation between expert observers was complete (Kappa = 1.0). However, interobserver values were low for zones (0.41; 0.27–0.54) and telemedicine scheme (0.43; 0.33–0.6). The observation times for experts 1 and 2 were 25 min and 15 min, respectively, whereas for non-expert observers the observation time was 45 min.

Discussion

The correlation of ROP diagnostic between 2 ROP experts and a non-expert observer was analyzed. Previous studies yielded good correlation levels. The main findings of this study were that, when examiners with varying degrees of specialization in ROP assessed the images via telemedicine, the Ells classification exhibited the best correlation together with stage and plus disease. Particularly, the use of the Ells referral criteria featured total correlation (Kappa = 1) between expert observers. There always was a match between at least one of the following: stage III in zone one or 2, presence of plus or any stage in zone 1.

The implementation of telemedicine schemes requires strict protocols for processing undetermined cases. Each one of the criteria that were used separately in this analysis exhibited lower correlation when compared to the excellent correlation displayed by the Ells criteria, i.e., a composite criterion. This allows for raising the capacity of the examination to detect cases for referral and ocular fundus assessment, thus producing savings and optimization of specialist time.

Other correlation studies demonstrated that the interobserver agreement with the use of telemedicine is high among expert observers, the main discrepancies being in the definition of zone one and the diagnostic of plus disease. It has recently been reported that the plus disease could be imperfect in the correlation, because specialists have different interpretations of similar findings. This study exhibited high correlation for the plus disease.

Telemedicine technology facilitates reliable remote diagnostics, discriminating between cases requiring treatment, with low complication rates. In general terms, telemedicine sensitivity and specificity studies for ROP provide different results, ranging between 0.46 and 1 for sensitivity and 0.37 and one for specificity depending on factors such as observer experience and type. ROP diagnostic sensitivity

<table>
<thead>
<tr>
<th>Table 1 – Correlation between E1, E2 and NE observers.</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1 vs NE</td>
</tr>
<tr>
<td>Scheme</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Zone</td>
</tr>
<tr>
<td>Stage</td>
</tr>
<tr>
<td>Plus</td>
</tr>
<tr>
<td>Ells criterion</td>
</tr>
<tr>
<td>Pre plus</td>
</tr>
</tbody>
</table>

E: expert; NE: non-expert.
* NE did not examine pre-plus.
and specificity is variable amongst experts as well as non-experts. Williams et al.\textsuperscript{13} reported a mean of 0.95 and 0.93 respectively for experts, and 0.87 and 0.73 for non-expert residents. The same author reported that ROP cases requiring treatment had a mean sensitivity and specificity of 1 and 0.93 for experts, and 0.88 and 0.84 for non-expert residents. In addition, the time required by ophthalmology specialists was significantly lower than for ophthalmoscopists\textsuperscript{27} apart from being more cost-effective.\textsuperscript{13,14} A recent study comprising 1022 eyes for screening revealed high sensitivity and specificity in addition to a positive predictive value of 94% and a negative predictive value of 100%, demonstrating that telemedicine exhibits high diagnostic precision rates.\textsuperscript{28}

At present, discussions are focused on other areas to assess the traditional follow-up schemes vis-à-vis telemedicine in order to optimize frequency, costs and human resources which are always scarce in these scenarios. A recent study demonstrated that the use of telemedicine for ROP screening can be carried out in developing countries, enabling timely referral to ophthalmologists.\textsuperscript{26}

The limitations of this study are the absence of comparison with the gold standard (ocular fundus by independent observers), that the cases comprised variable groups of images (range 3–5), that the photographs may not be as peripheral or as clear in the periphery and that the contrast and brightness of the images were not modifiable.

Finally, it can be concluded that when examiners with varying degrees of specialization in ROP examine images via telemedicine, the Ells classification has the best correlation together with the stage and plus disease. In contrast, the correlation between observers was low for ROP zone and the telemedicine scheme. Non-expert observers with previous training exhibited good correlation with the experts.

### Conflict of interests

No conflict of interest has been declared by the authors.

### References


