AMERICAN ACADEMY™
OF OPHTHALMOLOGY

Ophthalmic Technology Assessment

Clinical Models and Algorithms for the Prediction of Retinopathy of Prematurity

A Report by the American Academy of Ophthalmology

Amy K. Hutchinson, MD,[1] Michele Melia, ScM,[2] Michael B. Yang, MD,[3] Deborah K. VanderVeen, MD,[4] Lorri B. Wilson, MD,[5] Scott R. Lambert, MD[1]

Objective: To assess the accuracy with which available retinopathy of prematurity (ROP) predictive models detect clinically significant ROP and to what extent and at what risk these models allow for the reduction of screening examinations for ROP.

Methods: A literature search of the PubMed and Cochrane Library databases was conducted last on May 1, 2015, and yielded 305 citations. After screening the abstracts of all 305 citations and reviewing the full text of 30 potentially eligible articles, the panel members determined that 22 met the inclusion criteria. One article included 2 studies, for a total of 23 studies reviewed. The panel extracted information about study design, study population, the screening algorithm tested, interventions, outcomes, and study quality. The methodologist divided the studies into 2 categories—model development and model validation—and assigned a level of evidence rating to each study. One study was rated level I evidence, 3 studies were rated level II evidence, and 19 studies were rated level III evidence.

Results: In some cohorts, some models would have allowed reductions in the number of infants screened for ROP without failing to identify infants requiring treatment. However, the small sample size and limited generalizability of the ROP predictive models included in this review preclude their widespread use to make all-or-none decisions about whether to screen individual infants for ROP. As an alternative, some studies proposed approaches to apply the models to reduce the number of examinations performed in low-risk infants.

Conclusions: Additional research is needed to optimize ROP predictive model development, validation, and application before such models can be used widely to reduce the burdensome number of ROP screening examinations. *Ophthalmology 2016;■:1–13 © 2016 by the American Academy of Ophthalmology.*

The American Academy of Ophthalmology prepares Ophthalmic Technology Assessments to evaluate new and existing procedures, drugs, and diagnostic and screening tests. The goal of an Ophthalmic Technology Assessment is to review systematically the available research for clinical efficacy and safety. After review by members of the Ophthalmic Technology Assessment Committee, relevant subspecialty societies, and legal counsel, assessments are submitted to the Academy's Board of Trustees for consideration as official Academy statements. The purpose of this assessment is to assess the ability of available retinopathy of prematurity (ROP) predictive models to detect clinically significant ROP and to what extent and at what risk these models allow for the reduction of screening examinations for ROP.

Background

Retinopathy of prematurity is a vasoproliferative retinal disorder that affects premature infants and is the leading cause of preventable childhood blindness in high- and middle-income countries.[1] Preventing blindness caused by ROP requires timely treatment, which depends on appropriate screening of infants at risk. Current United States screening guidelines recommend at a minimum examining all infants with birth weights (BWs) of 1500 g or less or estimated gestational age (GA) at birth of 30 weeks or less.[2] Although the screening criteria have high sensitivity to detect infants in need of treatment, implementation of these guidelines results in many unnecessary examinations, because only a small percentage of infants screened will meet criteria for treatment.[3,4]

A number of risk factors for ROP have been described, including BW, GA,[4] and oxygen exposure. Oxygen-dependent growth factors, such as vascular endothelial growth factor, play a primary role in the pathophysiology of ROP.[5] Deficiencies of non–oxygen-dependent growth factors, such as insulin-like growth factor-1 (IGF-1), normally passed to the developing fetus through the placenta, also play a key role in the pathophysiology of ROP.[6,7]

© 2016 by the American Academy of Ophthalmology
Published by Elsevier Inc.

http://dx.doi.org/10.1016/j.ophtha.2015.11.003
ISSN 0161-6420/15

Table 1. Levels of Evidence for Retinopathy of Prematurity Screening Predictive Model Studies

Predictive Model Development Studies

Level I: good-quality model development study including multiple (>1) cohorts
- Study cohorts are representative of population at risk of ROP*
- Study cohorts have differing risks or prevalences of disease, are from different institutions in varying geographic locations, or have differing racial or ethnic composition
- Screening was conducted using indirect ophthalmoscopy by an examiner with ROP experience, with appropriate frequency and duration
- Internal validation†
- Adequate sample size‡
- Model discrimination (sensitivity and specificity, AUC, or C index) was reported or could be calculated based on data provided§

Level II: good-quality predictive model development study including single cohort or >1 cohort but with similar ROP risk, geographic location, and racial or ethnic characteristics
- Study cohort is representative of population at risk of ROP*
- Screening was conducted using indirect ophthalmoscopy by an examiner with ROP experience, with appropriate frequency and duration
- Internal validation†
- Adequate sample size‡
- Model discrimination (sensitivity and specificity, AUC, or C index) was reported or could be calculated based on data provided§

Level III: low-quality predictive model development study
- At least 1 of the following applies:
- Study cohort(s) are not representative of population at risk of ROP*
- Screening was conducted not using indirect ophthalmoscopy, or not by an examiner with ROP experience, or not of appropriate frequency or duration
- No validation
- Inadequate sample size‡
- Model discrimination (sensitivity and specificity, AUC, or C index) was not reported or could not be calculated based on data provided

Model Validation Studies

Level I: good-quality predictive model validation study including multiple (>1) cohorts
- Study cohorts are independent of the cohort(s) used to develop the model (i.e., no overlap in infants included in the model development cohort[s] and the model validation cohort[s])
- Study cohorts were representative of population at risk of ROP*
- Study cohorts expand on the population(s) used for model development (e.g., to populations with varying risks or prevalences of disease, different institutions in varying geographic locations, or differing racial or ethnic composition, or same institutions but different periods with varying ROP risks or prevalences)
- Screening was conducted using indirect ophthalmoscopy by an examiner with ROP experience, with appropriate frequency and duration
- ROP outcome assessor was masked to model determination
- Adequate sample size‡
- Model discrimination (sensitivity and specificity, AUC, or C index) was reported or could be calculated based on data provided

Level II: good-quality predictive model validation study including single cohort or >1 cohort but with similar ROP risk, demographic, and racial or ethnic characteristics
- Study cohort is independent of model development cohort (i.e., no overlap in infants was included in the model development cohort[s] and the model validation cohort[s])
- Study cohort is representative of population at risk of ROP*
- Screening was conducted using indirect ophthalmoscopy by an examiner with ROP experience, with appropriate frequency and duration
- ROP outcome assessor was masked to model determination
- Adequate sample size†
- Model discrimination (sensitivity and specificity, AUC, or C index) was reported or could be calculated based on data provided

Level III: poor-quality predictive model validation study
- At least 1 of the following applies:
- Study cohort was not representative of population at risk of ROP
- Screening was conducted not using indirect ophthalmoscopy, or not by an examiner with ROP experience, or was not of appropriate frequency or duration
- ROP outcome assessor was not masked to model determination
- Inadequate sample size†
- Model discrimination (sensitivity and specificity, AUC, or C index) was not reported or could not be calculated based on data provided

AUC = area under the receiver operating characteristic curve; C index = concordance index; ROP=retinopathy of prematurity.

*Clinical recommendations for the population requiring screening for ROP varied by calendar time and country for the studies included in this review. A study was downgraded if the study cohort was not representative of the screening population based on screening guidelines in place at the time of the study.

†Internal validation uses the same study cohort combined with a statistical method, such as split-sample, cross-validation, or bootstrapping, to adjust for overfitting or optimism. With external validation, the predictive model is applied in a study population that is independent of and differs from the model development population by geographic location. In general, external validation is superior to internal validation and was accepted in place of internal validation if reported as part of a model development study.

‡For predictive model development, sample size was considered adequate if there were at least 10 occurrences of the outcome of interest (e.g., type 1 ROP) per predictor variable in the model. For example, for a model with 3 predictor variables, an adequate sample size included at least 30 occurrences of the outcome. For model validation, sample size was considered adequate if the width of the 95% confidence interval on sensitivity and negative predictive value was not wider than 10%.

§Ideally, model calibration also was performed, but studies were not rated based on model calibration. In model calibration, the study cohort is divided into risk groups (approximately 10 is recommended) based on predicted probability of the outcome, and the observed versus predicted probability in the risk groups are plotted and inspected for deviation from the 45° line of equality. Close adherence of all points to the 45° line indicates the model is well calibrated (i.e., it performs equally well across groups with differing risk).

Download English Version:

https://daneshyari.com/en/article/6199752

Download Persian Version:

https://daneshyari.com/article/6199752

Daneshyari.com