Journal of Perinatology (2020) 40:1100–1108
https://doi.org/10.1038/s41372-020-0605-5

ARTICLE

Evaluation of the economic impact of modified screening criteria for retinopathy of prematurity from the Postnatal Growth and ROP (G-ROP) study

John A. F. Zupancic[1,2] · Gui-shuang Ying[3] · Alejandra de Alba Campomanes[4] · Lauren A. Tomlinson[5] · Gil Binenbaum[3,5] · on behalf of the G-ROP Study Group

Received: 16 September 2019 / Revised: 7 January 2020 / Accepted: 31 January 2020 / Published online: 28 February 2020

Abstract

Importance The Postnatal Growth and Retinopathy of Prematurity (G-ROP) Study showed that the addition of postnatal weight gain to birth weight and gestational age detects similar numbers of infants with ROP, but requires examination of fewer infants.

Objective To determine the incremental cost-effectiveness of screening with G-ROP compared with conventional screening.

Design, setting and participants We built a microsimulation model of a 1-year US birth cohort <32 weeks gestation, using data from the G-ROP study. We obtained resource utilization estimates from the G-ROP dataset and from secondary sources, and test characteristics from the G-ROP cohort.

Results Among 78,281 infants nationally, screening with G-ROP detected ~25 additional infants with Type 1 ROP. This was accomplished with 36,233 fewer examinations, in 14,073 fewer infants, with annual cost savings of approximately US $2,931,980 through hospital discharge.

Conclusions Screening with G-ROP reduced costs while increasing the detection of ROP compared with current screening guidelines.

Introduction

Retinopathy of prematurity (ROP) is one of the leading causes of severe childhood visual impairment in middle- and high-income countries. Infants who develop a significant severity of disease known as Type 1 ROP have an approximately 30% risk of adverse visual outcome (defined functionally as visual acuity four standard deviations below the mean, or structurally as posterior retinal fold or detachment involving the macula) [1]. However, with treatment, this probability can be reduced substantially [2, 3].

ROP is diagnosed almost exclusively in infants who are born at gestational age below 32 weeks or birth weight below 1500 g, but does not develop for several weeks after birth. Serial retinal examinations with indirect ophthalmoscopy can reliably detect Type 1 disease in high-risk infants in time for intervention to take place. Current consensus recommendations by professional societies recommend screening of all infants born at <1501 g, or ≤30 weeks, or those between 1500 and 2000 g with hypotension or substantial oxygen exposure [4].

Although the currently recommended guidelines have high sensitivity for detection of ROP, their specificity is low, and the great majority of screened infants never develop ROP that requires intervention. Over the last

Members of the G-ROP Study Group are listed below Funding.

Supplementary information The online version of this article (https://doi.org/10.1038/s41372-020-0605-5) contains supplementary material, which is available to authorized users.

✉ John A. F. Zupancic
jzupanci@bidmc.harvard.edu

[1] Division of Newborn Medicine, Harvard Medical School, Boston, MA, USA

[2] Department of Neonatology, Beth Israel Deaconess Medical Center, Boston, MA, USA

[3] Scheie Eye Institute, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA

[4] University of California, San Francisco, CA, USA

[5] The Children's Hospital of Philadelphia, Philadelphia, PA, USA

15 years, several predictive models have suggested that the addition of suboptimal postnatal growth to birth weight and gestational age maintained sensitivity of screening criteria while substantially improving specificity, but these studies were based on relatively small cohorts [5–10]. Recently, the Postnatal Growth and Retinopathy of Prematurity (G-ROP) Study examined a cohort of 7483 preterm infants to develop a predictive model that included such modified screening criteria, with resulting sensitivity of 100% and specificity of 32%, both superior to traditional birth weight and gestational age screening alone in the study cohort [11]. Given the substantial reduction of the number of infants requiring serial ophthalmological examinations using these criteria, we undertook a study to determine the economic implications of adopting such a screening strategy.

Methods

Study design and model specification

We developed a decision analytic microsimulation model to evaluate the proposed G-ROP revision to current screening criteria for ROP, using TreeAge Pro Suite software (TreeAge Software Inc 2017, Williamstown, MA). The model, presented in schematic form in Fig. 1, generated estimates of resource utilization and outcomes associated with ROP for a hypothetical 1-year national birth cohort under various assumptions of the criteria employed for screening.

We compared two approaches to screening. The first comparator (conventional screening, CS) was based on the current guideline of the American Academy of Ophthalmology and the American Academy of Pediatrics, which recommend screening of infants with birth weight 1500 g or less, or gestational age at birth 30 weeks or less [4]. The second comparator (G-ROP) incorporated the G-ROP criteria, in which infants are screened if they meet any one of the following criteria: gestational age at birth <28 weeks; or birth weight < 1051 g; or weight gain of <120, 180, or 170 g during ages 10–19, 20–29, or 30–39 days, respectively; or hydrocephalus [11]. Of note, a subjective third criterion is presently used clinically, in which infants with birth weight between 1500 and 2000 g may receive retinal examinations if they have had an unstable clinical course in the judgment of the treating neonatologist [4]. For a valid comparison to be made among the two screening approaches, a subjective criterion also could be added to the G-ROP criteria. However, because this criterion is subjective and nonspecific, it was not considered as part of either the current CS guidelines or the G-ROP criteria in this analysis.

The model was formulated in two versions. In the first phase ("discharge"), we sought to minimize model assumptions by limiting model efficacy inputs to those available in the G-ROP database. This version analyzed only short-term outcomes through hospital discharge, and included screening and acute treatment costs, and the number of cases of treatable ROP detected. In the second phase ("lifetime"), we employed literature sources to extend these results to the lifetime horizon, in order to include severe visual impairment and quality of life, as well as broader societal estimates of costs.

Description of primary analysis (discharge)

Framing

The planned main outcome for the primary analysis was the incremental cost (ΔC) incurred by a proposed G-ROP screening strategy relative to a CS strategy, divided by its incremental effectiveness (ΔE). The resulting incremental cost-effectiveness ratio (ΔC/ΔE) represented the cost-effectiveness of G-ROP screening, expressed in terms of the additional cost per additional infant with Type 1 ROP detected.

Secondary outcomes included the number of infants screened and the number of examinations performed annually in the United States, as well as the annual number of false positive screening results (those infants without

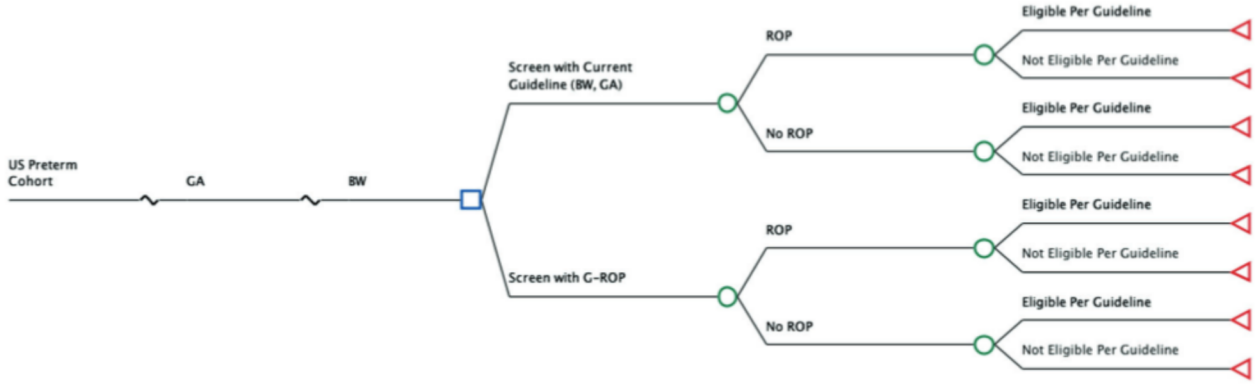

Fig. 1 Decision tree model for the base case (discharge) analysis.

Type 1 ROP who were selected for examination by each model) and the number of false negative screening results (those infants with Type 1 ROP who were missed by each model).

The primary analysis was performed from the perspective of a third-party payer, in which only direct medical costs were included, and with a time horizon to first discharge home from hospital.

Institutional review board approval was obtained and a waiver of consent was granted for the collection of deidentified data in the original study cohort at all participating hospitals.

Model inputs: effectiveness

Inputs for effectiveness are provided in Table 1. Effectiveness in our model was expressed as the proportion of cases of Type 1 ROP detected, which were defined as those diagnosed in the large, multi-center G-ROP cohort as meeting the Early Treatment of Retinopathy of Prematurity Study Type 1 ROP criteria [2]. The gestational-age-specific probability of developing Type 1 ROP was estimated directly from its incidence in the G-ROP study database [12].

In the G-ROP clinical study, the screening criteria had sensitivity 100% (95% CI, 99.2–100%) and specificity 32.3% (95% CI, 31.2–33.4%) for the detection of Type 1 ROP. In the same population, CS criteria had sensitivity 99.4% (95% CI, 98.1–99.8%) and specificity 11.8% (95% CI, 11.0–12.6%) [11].

Model inputs: resource use and costs

Estimates for resource utilization and costs are provided in Table 1. Prior to the start of the G-ROP Study, resource utilization data were collected and assessed on a sample of 100 infants undergoing ROP examinations at three hospitals in Seattle and Philadelphia. No changes in non-ROP-related resource utilization were associated with ROP examinations, including changes in respiratory support, nutrition, radiograph usage, laboratory tests, and nonophthalmological surgeries and procedures (data not shown). Therefore, we proceeded under the assumption that the only differences in resource utilization prior to discharge for ROP evaluations were those related to screening or treatment of ROP. We estimated the frequency of ophthalmological examinations per patient from G-ROP case report forms, which collected data on all eye examinations.

Because neither professional nor institutional billing reliably reflect the personnel effort, we estimated professional time input in an observational time and motion study of individuals involved in care of infants undergoing ROP diagnostic examinations [13]. Ophthalmologists, neonatal nurses, ophthalmic technicians, and ROP coordinators were timed using digital timers and standardized data collection forms for work completed at four neonatal intensive care units and two outpatient ophthalmology clinics in Philadelphia and San Francisco between February and December 2014. Primary outcomes were the overall and subtask times per infant. Seven pediatric ophthalmologists and eight ophthalmic technicians were timed performing 303 inpatient and 37 outpatient ROP exams (Table 1).

Hourly time costs of physicians, registered nurses, and administrative support staff for examinations were calculated from the Doximity physician salary survey and U.S. government data, respectively [14, 15]. A 30% fringe rate was added to these values to account for institutional overhead costs. Because infants were assumed to be resident in the NICU and therefore subject to bundled institutional reimbursement, a separate institutional technical fee for these examinations was not included. Costs for laser photocoagulation procedures and retinal procedures were derived from previously published microcosted estimates by our group [16].

Total costs were calculated as the product of the resources used and the unit prices associated with those resources. Although incurred between 2006 and 2012, all costs are expressed in this report in 2017 US dollars (USD). Where necessary, we converted costs from other dates using the Personal Consumption Expenditure health price deflator from the Center for Medicare and Medicaid Services [17, 18]. Because the time horizon for the primary analysis was <1 year, discounting was not undertaken in the primary analysis.

Model inputs: population characteristics

We used data from the U.S. National Center for Health Statistics to assign the frequency of birth at each week of gestational age (Table S1), and the distribution of birth weights within each week of gestational age (Tables S2a–k) [19].

Uncertainty and heterogeneity

In order to reflect the heterogeneity of the population to whom the results will be applied, we conducted a computer microsimulation of a 1-year United States birth cohort, in which each infant was assigned a birth weight and gestational age based on empirical data for infants between 22 and 32 weeks from the 2017 US census (the most recent year available) [19], followed by a risk of ROP based on these characteristics. Infants then proceeded through the model, one infant at a time, and were screened and accrued costs and outcomes according to their baseline characteristics and the probabilities of detection encountered in the model. At the end of the cohort, the average cost and

Table 1 Estimates of population, efficacy and resource use.

Input	Value	Type	Range or Parameter	Source
Estimates of population, efficacy and resource use				
Sensitivity				
G-ROP	100.0%	Beta distribution	$n = 4590$ $r = 4589$	[11]
CS[a]	99.4%	Beta distribution	$n = 459$ $r = 456$	[11]
Specificity				
G-ROP	32.3%	Beta distribution	$n = 7024$ $r = 2269$	[11]
CS[a]	11.8%	Beta distribution	$n = 6197$ $r = 827$	[11]
Gestational age (weeks)	22–32	Empirical distribution[b]		[19]
Birth weight by GA	400–5000	Empirical distribution[b]		[19]
ROP incidence by GA	0–0.31	Probability table		[11]
Time for examination				
MD	14 (SD 7.7)	Normal distribution	mean = 14 SD = 7.7	Observation[d]
RN	15.2 (SD 7.9)	Normal distribution	mean = 15.2 SD = 7.9	Observation[d]
Coordinator	21.9 (SD 8.0)	Normal distribution	mean = 21.9 SD = 8.0	Observation[d]
Number of examinations				
With ROP	6.1 (SD 2.9)	Normal distribution	mean = 6.1 SD = 2.9	[11]
Without ROP	2.6 (SD 1.5)	Normal distribution	mean = 2.6 SD = 1.5	[11]
Probability of laser photocoagulation				
If Type 1 detected	1.0	Point estimate	NA	Assumption
If Type 1 missed	0	Point estimate	NA	Assumption
Probability of retinal surgery				
If Type 1 detected	0.064	Point estimate	NA	[2]
If Type 1 missed	0.43	Point estimate	NA	[22]
Discount rate	0.03	Point estimate	0.015–0.045	[20]
Utility				
Good visual outcome w/o ROP	0.97	Point estimate	NA	[29]
Poor visual outcome with ROP	0.27	Point estimate	0.14–0.41	[29]
Good visual outcome with ROP	0.87	Point estimate	0.44–1.00	[29]
Probability of poor visual outcome				
Without ROP	0	Point estimate	NA	Assumption
With ROP detected and treated	0.143	Point estimate	0.143–0.198	[2]
With ROP not detected or treated	0.643	Point estimate	NA	[3]
Estimates of unit cost[c]				
Hourly personnel cost				
Ophthalmologist	254	Point estimate	127–381	[15]
RN	48	Point estimate	24–72	[14]
Coordinator	22	Point estimate	11–33	[14]
Retinal surgery	13,343	Point estimate	6672–20,015	[16]
Laser photocoagulation	3114	Point estimate	2180–4,048	[16]
Lifetime cost of poor visual outcome	791,432	Point estimate	554,002–1,028,862	[30]

[a]Conventional screening.

[b]Full details in Supplementary Tables.

[c]Values are in 2017 US dollars.

[d]Observational study ancillary to current project [13]; details in methods section above.

effectiveness across all infants, and the incremental cost-effectiveness ratio, were calculated.

We repeated this simulation with 1000 cohorts, corresponding to 1000 years. Sampling uncertainty was assessed for sensitivity, specificity, number of examinations, and professional time by entering these into the model as distributions, the details of which are provided in Table 1. At the start of each cohort, the model drew randomly from each of these distributions, and that estimate was employed for all infants in that year.

We also performed deterministic sensitivity analysis, in which a specific input to the model was changed through

the upper and lower range of plausible values, and the entire simulation re-run using that value, in order to determine the impact on the resulting costs and effectiveness. Best-case and worst-case scenarios were implemented by varying multiple inputs simultaneously.

Description of secondary analysis (lifetime)

The lifetime analysis sought to optimize generalizability of the results and the ability to compare them with other health care programs, and thus differed in several ways from the primary model. First, it employed a societal perspective, in which all costs were considered, including family out-of-pocket expenses, productivity (wage) losses to both the patient in later life and to the infant's family, and educational and other expenditures by nonmedical agents. These costs were considered over a lifetime time horizon rather than to first discharge home. Visual outcomes were converted to utilities, or the individuals' preferences for living with either normal or poor visual outcomes, and the results expressed as quality adjusted life years, defined as the utility multiplied by the average life expectancy in a given state of visual function. Costs were discounted at 3% per annum, and this rate was varied in sensitivity analysis [20]. Other aspects of the model were similar to those in the primary analysis.

Results

Each microsimulation run included a 1-year cohort of 78,281 infants [19]. Screening with G-ROP, compared with CS, detected ~25 additional infants with Type 1 ROP (Table 2). This improved detection was accomplished despite 36,233 fewer examinations being performed, in 14,073 fewer infants, and at an annual cost savings of ~USD2,931,980 in the short run and USD12,873,157 in the long run. These savings correspond to 9.5% and 2.6%, respectively, of the total annual relevant ROP-related expenditure. The lower costs and higher effectiveness of G-ROP screening indicate that it is a "dominant" option, compared with CS (Table 3).

Results for the probabilistic sensitivity analysis, reflecting sampling uncertainty for all the distributional inputs listed in Table 1, are shown in Fig. 2. Each dot on this cost-effectiveness scatterplot represents the incremental cost and effectiveness of 1 of the 1000 replications of our 1-year cohort of patients. Of these, 97.2% are in the right lower quadrant, with lower costs and higher effectiveness, indicating a 97.2% probability that G-ROP screening is a dominant option. A similar result is depicted in the Cost-Effectiveness Acceptability Curve in Fig. S1, which plots the proportion of replications that would be considered "cost-effective" for a decision maker who had the willingness to pay (WTP) for one additional case of ROP detected listed on the horizontal axis [21]. In this plot, G-ROP is a preferred option even at a WTP of 0, and rapidly approaches 100% probability of desirability at higher WTP.

Table 2 Details of screening performance through discharge in United States birth cohort.[a,b]

Parameter	Conventional screening	G-ROP screening
Total cost	$30,885,069	$27,953,089
Type 1 ROP detected	3948	3973
Type 1 ROP missed	26	1
Eligible for serial examination	68,331	54,258
Not eligible for serial examination	9950	24,023
Number of examinations	189,220	152,987
No Type 1 ROP, flagged by model	64,383	50,285
No Type 1 ROP, not flagged	9924	24,022

[a]Through 32 weeks gestational age, inclusive $n = 78,281$.

[b]Values rounded to 0 digits.

Findings from deterministic sensitivity analyses are presented in Table S3. Despite varying the cost inputs over a very broad range of 70% to 130% of the baseline value, both individually and as a group, there was minimal change in the proportion of replications in which G-ROP screening was a dominant option. Similarly, we substituted the progression to unfavorable structural outcome seen with cryotherapy in the original CRYO-ROP trial [22] for the superior efficacy in the intervention arm of the ET-ROP trial [2], as well as for a hypothetical lower progression than ET-ROP, neither of which resulted in a probability of dominance below 90%. Finally, we varied all of the above inputs simultaneously, in best and worst-case scenarios, for which the probabilities of dominance were 97.2% and 90.8%, respectively.

The lifetime model showed similar results with a lifetime time horizon and societal perspective, with G-ROP simultaneously yielding lower costs and more quality adjusted life years than the CS strategy, at a very high probability of 98.8% in probabilistic sensitivity analysis (Table S4). The direction of this result was again robust to changes in input assumptions (Table S3).

Discussion

In this microsimulation cost-effectiveness study of a US birth cohort born between 22 and 32 weeks gestational age,

Table 3 Calculation of point estimate of cost-effectiveness for primary outcome, dollars per case of severe visual impairment prevented through discharge.[a]

Comparator	Cost ($) C	Incremental Cost ΔC	Effectiveness E[c]	Inc Effectiveness ΔE	ICER[d] ΔC/ΔE
G-ROP	27,953,089	−2,931,980	3973	25	Dominant
CS[b]	30,885,069		3948		

[a]Values rounded to 0 digits.

[b]Conventional screening.

[c]Cases of Type 1 ROP detected.

[d]Incremental cost-effectiveness ratio.

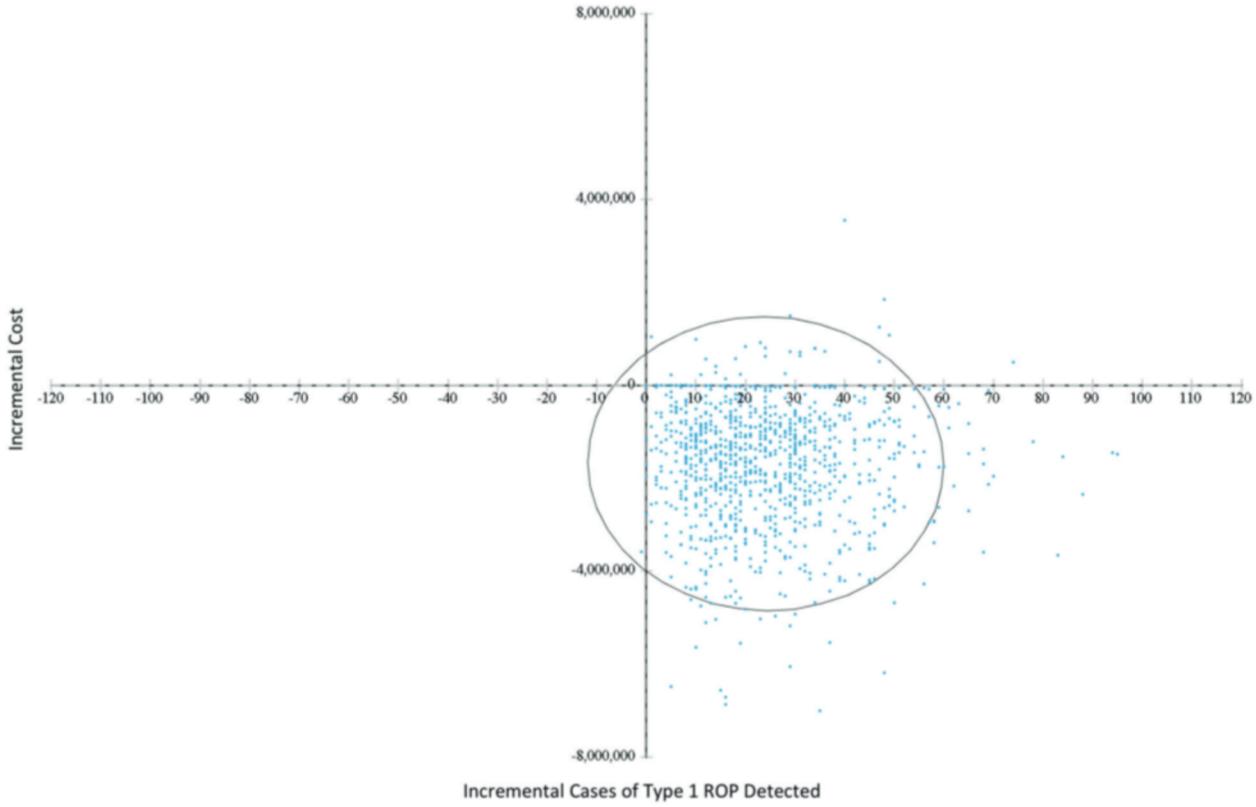

Fig. 2 Differences in mean costs and effects between conventional screening and G-ROP screening cohorts for the base case (discharge), cost per case of Type 1 ROP detected. Each point represents one run of the simulation model as described in the text. Ellipse represents 95% confidence bound for joint distribution of cost and effectiveness.

implementing screening criteria using the G-ROP model was an economically dominant strategy. Identifying infants as eligible for ophthalmological screening for ROP using the G-ROP screening criteria, which incorporate postnatal weight gain and hydrocephalus in addition to gestational age and birth weight criteria, increased the number of infants identified with Type 1 ROP while simultaneously reducing the number of examinations and decreasing costs. The results were robust to adjustments in the underlying assumptions about efficacy and costs.

Although several modeling analyses of the economic implications of ROP management have been reported, these compared either interventions (such as treatment at an earlier stage of retinal pathology), telemedicine screening devices, or combinations of screening and intervention [16, 23–27]. To our knowledge, no studies have specifically compared the cost-effectiveness of one of the newer, weight-velocity based screening algorithms with CS.

This study has several notable strengths. First, we incorporate the results of the development of a revised screening approach in a large dataset. The analysis considers both the heterogeneity of the population of eligible infants, by assigning birth weight, gestational age and gestational-age-specific ROP incidence to individual infants, as well as sampling uncertainty related to important parameters such as sensitivity and specificity, using distributions that reflect the variance in the studies that generated the estimates. Moreover, the analyses were designed

to simultaneously address the cost-effectiveness of G-ROP compared with CS, as well as the national implications in terms of numbers of infants exposed to screening, and the numbers screened who have little to no likelihood of developing serious ROP. Such explicit national estimates of the burden of screening will hopefully be of use in policy discussions of recommendations for screening strategies.

Notably, prospective validation has been completed in a similar but independent cohort of infants, and confirmed 100% sensitivity of the criteria in that population as well [28].

Certain limitations must be noted. First, although the initial G-ROP dataset included a large sample of infants in two independent samples in the US, the approach has not to date been validated in a broad variety of other populations worldwide. Generalizability to developing health care systems will remain uncertain until retested in those settings, where the characteristics of infants developing severe ROP differs considerably, and the G-ROP criteria may not be reliable. Second, in the absence of reliable national estimates, we applied the gestational age-specific incidence of ROP from the G-ROP study to the national cohort in our model. Changes in incidence will not alter the sensitivity or specificity of the model, but they could potentially alter the economic implications. Third, while in the simulation the use of the G-ROP criteria increased detection of Type 1 ROP by 25 cases per year compared with conventional BW and GA criteria, these cases are likely being captured in practice by use of the subjective third criterion of an unstable clinical course in the judgment of the neonatologist, which was not included in the model due to the infeasibility of quantification. In contrast, the G-ROP criteria had 100% sensitivity without the need for a subjective criterion. Therefore, even if such a criterion were to be added for both strategies and consequently sensitivity increased to 100% for CS, the study conclusions would not change, as the G-ROP criteria would still demonstrate "weak dominance," in which costs are lower but outcomes the same. Finally, we note that there are undoubtedly missed appointments and delayed examinations. These are included in the numbers of examinations estimated in the retrospective G-ROP dataset and should be similar in both groups, and variations are included as these were entered into the model as distributions. If the results are applied to a population that has a very different rate of compliance, however, the impact of screening might differ as well.

Conclusions

The G-ROP modified screening criteria are a dominant, economically desirable strategy when compared with the CS. If applied in clinical practice, their use would improve detection of treatment-requiring ROP while simultaneously greatly reducing the number of infants receiving retinal examinations and the number of examinations being performed, with resultant cost savings.

Disclaimer

The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the National Center for Research Resources or the National Institutes of Health. The National Institutes of Health had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Funding This study was supported by the National Institutes of Health grant R01EY021137, and the Richard Shafritz Endowed Chair in Ophthalmology Research at the Children's Hospital of Philadelphia.

G-ROP Study Group Gil Binenbaum[6], Lauren A. Tomlinson[6], Trang B. Duros[6], Anh Nguyen[6], Gui-shuang Ying[7], Maureen G. Maguire[7,41], Mary Brightwell-Arnold[7], James Shaffer[7], Yinxi Yu[7], Maria Blanco[7], Trina Brown[7], Christopher P. Helker[7], Amy Hutchinson[8], Carrie Young[8], Emily McCourt[9], Anne Lynch[9], Jennifer Cathcart[9], Ashlee Cerda[9], Levi Bonnell[9], Tamara Thevarajah[9], Gerard P. Barry[10], Marilyn Fisher[10], Maria V. Battaglia[10], Alex M. Drach[10], Kevin Hughes[10], Nachammai Chinnakaruppan[11], Andrew Meyer[11], Christina Gogal[11], Cynthia Beitler[11], Lauri Centolanza[11], Keith T. Moyer[11], Mary Sobotor[11], Pamela Donohue[12], Michael X. Repka[12], Jennifer A. Shepard[12], Megan Doherty[12], James D. Reynolds[13], Erin Connelly[13], Edward Cheeseman[14], Kinsey Shirer[14], Carol Bradham[14], Allison McAlpine[14], Sudeep Sunthankar[14], Javaneh Abbasian[15], Janet Lim[15], Michael Yang[16], Patricia Cobb[16], Elizabeth L. Alfano[16], David Rogers[17], Rachel E. Reem[17], Amanda Schreckengost[17], Rae R. Fellows[17], Kaitlyn Loh[17], Madeline A. McGregor[17], Thabit Mustafa[17], Ivy Dean[17], Rachel Miller[17], Tess Russell[17], Rebecca Stattler[17], Sara Maletic[17], Theran Jake Selph[17], David Young[18], Andrea Siu[18], Michele Kanemori[18], George Kingston[18], Megan Geloneck[19], Robert Feldman[19], Ted Baker[19], Laura Baker[19], Ephrem Melese[19], Kathryn Haider[20], Jingyun Wang[20], Elizabeth Hynes[20], Edward F. Bell[21], Alina V. Dumitrescu[21], Jonathan M. Klein[21], Gretchen A. Cress[21], Avanthi S. Ajjarapu[21], Kristine Berge[21], Eric Boeshart[21], Morgan Dorsey[21], Bethany M. Funk[21], Grace Hach[21], Claire L. Johnson[21], Kevin Kurian[21], Emily Miller[21], Angela C. Platt[21], Christine Law[22], Andrew Gissing[22], Leila Khazaeni[23], Jennifer Dunbar[23], Kelley Hawkins[23], Sharon Lee[23], Lily Sung[23], Carly Leggitt[23], Aparna Ramasubramanian[24], Rahul Bhola[24], Michelle Bottorff[24], Neviana Dimova[24], Rachel Keith[24], Laura Thomas[24], Jill Anderson[25], Raymond G. Areaux Jr.[25], Ann Marie Holleschau[25], Jordan Gross[25], Andrea Kramer[25], David Morrison[26], Sean Donahue[26], Carsyn Saige Wilkins[26], Neva Fukuda[26], Sandy Owings[26], Scott Ruark[26], R. Michael Siatkowski[27], Faizah Bhatti[27], Vanessa Bergman[27], Karen Corff[27], Kari Harkey[27], Amy Manfredo[27], Ashley Helmbrecht[27], Shrenik Talsania[27], Terri Whisenhunt[27], Donny Suh[28], Ann Anderson Berry[28], Denise Lynes[28], Kelly C. Erikson[28], Gil Binenbaum[29], Soraya Abbasi[29], Haresh Kirpalani[29], Graham E. Quinn[29], Lindsay Dawson[29], Lauren A. Tomlinson[29], Christin Sylvester[30], Kanwal Nischal[30], Lauren Runkel[30], Wendy S. Chen[31], Deidrya Jackson[31], Bradley Davitt[32], Dawn Govreau[32], Linda Breuer[32], September Noonan[32], Robert Hoffman[33], Joanna Beachy[33], Kelliann Farnsworth[33], Katie Jo Farnsworth[33],

Deborah Harrison[33], Ashlie Bernhisel[33], Bonnie Carlstrom[33], Alejandra G. de Alba Campomanes[34], Yizhuo Bastea-Forte[34], Lucia Rivera Sanchez[34], Jacquelyn Kemmer[34], Alexandra Neiman[34], Sarah Sitati-Ng'Anda[34], Kristina Tarczy-Homoch[35], Francine Baran[35], Lauren Eaton[35], Nasrin Najm-Tehrani[36], Tanya Grossi[36], Maram Isaac[36], Robin Knighton[36], Monica Ralli Khitri[37], Madeline Del Signore[37], Cynthia Dembofsky[38], Andrew Meyer[38], Karen Flaherty[38], Tracey Harris[38], Jamie Heeneke[38], Dorothy Hendricks[39], Christopher M. Fecarotta[39], Alicia Olivant Fisher[39], Mark Paullin[39], John A. F. Zupancic[40], Alejandra de Alba Campomanes[41], Edward F. Bell[41], Gil Binenbaum[41], Pamela Donohue[41], David Morrison[41], Graham E. Quinn[41], Michael X. Repka[41], David L. Rogers[41], Lauren A. Tomlinson[41], Michael Yang[41], Gui-shuang Ying[41]

[6]Office of Study Chair, The Children's Hospital of Philadelphia, Philadelphia, PA, USA; [7]Data Coordinating Center—University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA; [8]Children's Healthcare of Atlanta, Emory University School of Medicine, Atlanta, GA, USA; [9]University of Colorado Hospital, Children's Hospital Colorado, University of Colorado Denver, Denver, CO, USA; [10]Albany Medical College, Albany, NY, USA; [11]Lehigh Valley Hospital, Allentown, PA, USA; [12]Johns Hopkins Hospital, Johns Hopkins University, Baltimore, MD, USA; [13]Women & Children's Hospital of Buffalo, University at Buffalo, Buffalo, NY, USA; [14]Medical University of South Carolina, Charleston, SC, USA; [15]University of Illinois at Chicago, Chicago, IL, USA; [16]Cincinnati Children's Hospital Medical Center, Good Samaritan Hospital, and University of Cincinnati Medical Center, Cincinnati, OH, USA; [17]Nationwide Children's Hospital, Riverside Methodist Hospital, Grant Medical Center, Doctors Hospital, Nationwide Children's Hospital, Columbus, OH, USA; [18]Kapi'olani Medical Center for Women and Children, Honolulu, HI, USA; [19]Children's Memorial Hermann Hospital, University of Texas at Houston, Houston, TX, USA; [20]Riley Hospital for Children at Indiana University Health, Indiana University, Indianapolis, IN, USA; [21]University of Iowa Stead Family Children's Hospital, University of Iowa, Iowa City, IA, USA; [22]Kingston Health Sciences Center, Queen's University, Kingston, ON, Canada; [23]Loma Linda University Children's Hospital, Loma Linda University, Loma Linda, CA, USA; [24]Norton Children's Hospital, University of Louisville, Louisville, KY, USA; [25]Masonic Children's Hospital, formerly University of Minnesota—Amplatz Children's Hospital, University of Minnesota, Minneapolis, MN, USA; [26]Vanderbilt Eye Institute and Vanderbilt University Medical Center, Monroe Carell Jr. Children's Hospital at Vanderbilt, Nashville, TN, USA; [27]The Children's Hospital at OU Medical Center/The University of Oklahoma Health Sciences Center, University of Oklahoma, Oklahoma City, OK, USA; [28]Nebraska Medicine, University of Nebraska, Omaha, NE, USA; [29]The Children's Hospital of Philadelphia, Hospital of the University of Pennsylvania, Pennsylvania Hospital, The Children's Hospital of Philadelphia, Philadelphia, PA, USA; [30]Children's Hospital of Pittsburgh, Magee Women Hospital of UPMC, University of Pittsburgh, Pittsburgh, PA, USA; [31]Women and Infants Hospital of Rhode Island, Rhode Island Hospital, Providence, RI, USA; [32]Cardinal Glennon Children's Hospital, Saint Louis University, St. Louis, MO, USA; [33]Primary Children's Hospital and University of Utah Hospital, University of Utah, Salt Lake City, UT, USA; [34]UCSF Benioff Children's Hospital and Zuckerberg San Francisco General Hospital, formerly San Francisco General Hospital, University of California San Francisco, San Francisco, CA, USA; [35]Seattle Children's Hospital, and University of Washington Medical Center, Seattle, WA, USA; [36]Sick Kids, The Hospital for Sick Children, Toronto, ON, Canada; [37]Harbor-UCLA Medical Center, Los Angeles Biomedical Research Institute, Torrance, CA, USA; [38]Crozer Chester Medical Center, Delaware County Memorial Hospital, Crozer-Chester Medical Center, Upland, PA, USA; [39]Nemours/Alfred I duPont Hospital for Children, Wilmington, DE, USA; [40]Cost-Effectiveness Component, Beth Israel Deaconess Medical Center, Boston, MA, USA; [41]Executive/Editorial Committee, Boston, MA, USA

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Early Treatment for Retinopathy of Prematurity Cooperative Group. Revised indications for the treatment of retinopathy of prematurity: results of the Early Treatment for Retinopathy of Prematurity randomized trial. Arch Ophthalmol. 2003;121:1684–94.
2. Good WV, Hardy RJ, Dobson V, Palmer EA, Phelps DL, Quintos M, et al. The incidence and course of retinopathy of prematurity: findings from the Early Treatment for Retinopathy of Prematurity study. Pediatrics. 2005;116:15–23.
3. Palmer EA, Hardy RJ, Dobson V, Phelps DL, Quinn GE, Summers CG, et al. 15-year outcomes following threshold retinopathy of prematurity: final results from the Multicenter Trial of Cryotherapy for Retinopathy of Prematurity. Arch Ophthalmol. 2005;123:311–8.
4. Fierson WM, American Academy of Pediatrics Section on Ophthalmology, American Academy of Ophthalmology, American Association for Pediatric Ophthalmology and Strabismus, American Association of Certified Orthoptists. Screening examination of premature infants for retinopathy of prematurity. Pediatrics. 2018;142:e20183061.
5. Binenbaum G, Ying GS, Quinn GE, Huang J, Dreiseitl S, Antigua J, et al. The CHOP postnatal weight gain, birth weight, and gestational age retinopathy of prematurity risk model. Arch Ophthalmol. 2012;130:1560–5.
6. Binenbaum G, Ying GS, Tomlinson LA, Postnatal Growth and Retinopathy of Prematurity Study Group. Validation of the Children's Hospital of Philadelphia Retinopathy of Prematurity (CHOP ROP) model. JAMA Ophthalmol. 2017;135:871–7.
7. Cao JH, Wagner BD, Cerda A, McCourt EA, Palestine A, Enzenauer RW, et al. Colorado retinopathy of prematurity model: a multi-institutional validation study. J Am Assoc Pediatr Ophthalmol Strabismus. 2016;20:220–5.
8. Eckert GU, Fortes Filho JB, Maia M, Procianoy RS. A predictive score for retinopathy of prematurity in very low birth weight preterm infants. Eye. 2012;26:400–6.
9. Lofqvist C, Hansen-Pupp I, Andersson E, Holm K, Smith LE, Ley D, et al. Validation of a new retinopathy of prematurity screening method monitoring longitudinal postnatal weight and insulinlike growth factor I. Arch Ophthalmol. 2009;127:622–7.
10. Wu C, Lofqvist C, Smith LE, VanderVeen DK, Hellstrom A, Consortium W. Importance of early postnatal weight gain for normal retinal angiogenesis in very preterm infants: a multicenter study analyzing weight velocity deviations for the prediction of retinopathy of prematurity. Arch Ophthalmol. 2012;130:992–9.
11. Binenbaum G, Bell EF, Donohue P, Quinn G, Shaffer J, Tomlinson LA, et al. Development of modified screening criteria for retinopathy of prematurity: primary results from the Postnatal Growth and Retinopathy of Prematurity Study. JAMA Ophthalmol. 2018;136:1034–40.

12. Quinn GE, Ying GS, Bell EF, Donohue PK, Morrison D, Tomlinson LA, et al. Incidence and early course of retinopathy of prematurity: secondary analysis of the Postnatal Growth and Retinopathy of Prematurity (G-ROP) Study. JAMA Ophthalmol. 2018;136:1383–9.
13. Dawson L, De Alba Campomanes A, Zupancic J, Binenbaum G. Time and motion study for retinopathy of prematurity examinations. J Am Assoc Pediatr Ophthalmol Strabismus. 2016;20:e14.
14. Bureau of Labor Statistics US Department of Labor. Occupational employment statistics: U.S. Government Printing Office. https://www.bls.gov/oes/current/oes_nat.htm#29-0000. Accessed March 4, 2019.
15. Doximity. 2018 Physician Compensation Report. Doximity: 2018. https://s3.amazonaws.com/s3.doximity.com/careers/2018_physician_compensation_report.pdf. Accessed March 4, 2019.
16. Kamholz KL, Cole CH, Gray JE, Zupancic JA. Cost-effectiveness of early treatment for retinopathy of prematurity. Pediatrics. 2009;123:262–9.
17. Bureau of Economic Analysis U.S. Department of Commerce. Price indexes for personal consumption expenditures by function 2018. https://apps.bea.gov/iTable/. Accessed March 4, 2019.
18. Dunn A, Grosse SD, Zuvekas SH. Adjusting health expenditures for inflation: a review of measures for health services research in the United States. Health Serv Res. 2018;53:175–96.
19. United States Department of Health and Human Services (US DHHS) Centers for Disease Control and Prevention—National Center for Health Statistics (NCHS) Division of Vital Statistics. Natality Public-Use WONDER Online Database 2017. http://wonder.cdc.gov/natality-current.html. Accessed March 4, 2019.
20. Drummond M, Sculpher M, Klaxton C, Stoddart GL, Torrance G. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 2015. p. 464.
21. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ. 2001;10:779–87.
22. Multicenter trial of cryotherapy for retinopathy of prematurity. One-year outcome–structure and function. Cryotherapy for Retinopathy of Prematurity Cooperative Group. Arch Ophthalmol. 1990;108:1408–16.
23. Dunbar JA, Hsu V, Christensen M, Black B, Williams P, Beauchamp G. Cost-utility analysis of screening and laser treatment of retinopathy of prematurity. J Am Assoc Pediatr Ophthalmol Strabismus. 2009;13:186–90.
24. Jackson KM, Scott KE, Graff Zivin J, Bateman DA, Flynn JT, Keenan JD, et al. Cost-utility analysis of telemedicine and ophthalmoscopy for retinopathy of prematurity management. Arch Ophthalmol. 2008;126:493–9.
25. Rothschild MI, Russ R, Brennan KA, Williams CJ, Berrones D, Patel B, et al. The economic model of retinopathy of prematurity (EcROP) screening and treatment: Mexico and the United States. Am J Ophthalmol. 2016;168:110–21.
26. van den Akker-van Marle ME, van Sorge AJ, Schalij-Delfos NE. Cost and effects of risk factor guided screening strategies for retinopathy of prematurity for different treatment strategies. Acta Ophthalmol. 2015;93:706–12.
27. Yanovitch TL, Siatkowski RM, McCaffree M, Corff KE. Retinopathy of prematurity in infants with birth weight > or =1250 grams - incidence, severity, and screening guideline cost-analysis. J Am Assoc Pediatr Ophthalmol Strabismus. 2006;10:128–34.
28. Binenbaum G, Tomlinson LA, de Alba Campomanes AG, et al. Validation of the Postnatal Growth and Retinopathy of Prematurity Screening Criteria. JAMA Ophthalmology. 2019 [Epub ahead of print].
29. Quinn GE, Dobson V, Saigal S, Phelps DL, Hardy RJ, Tung B, et al. Health-related quality of life at age 10 years in very-low-birth-weight children with and without threshold retinopathy of prematurity. Arch Ophthalmol. 2004;122:1659–66.
30. Honeycutt A, Dunlap L, Chen H, Homsi G, Grosse S, Schendel D. Economic costs associated with mental retardation, cerebral palsy, hearing loss, and vision impairment—United States, 2003. MMWR Morb Mortal Wkly Rep. 2004;53:57–9.