DISEASES AND DISORDERS

Autonomous early detection of eye disease in childhood photographs

Micheal C. Munson[1]*, Devon L. Plewman[1]*, Katelyn M. Baumer[1], Ryan Henning[2], Collin T. Zahler[1], Alexander T. Kietzman[1], Alexandra A. Beard[1], Shizuo Mukai[3,4], Lisa Diller[4,5], Greg Hamerly[2], Bryan F. Shaw[1†]

The "red reflex test" is used to screen children for leukocoria ("white eye") in a standard pediatric examination, but is ineffective at detecting many eye disorders. Leukocoria also presents in casual photographs. The clinical utility of screening photographs for leukocoria is unreported. Here, a free smartphone application (CRADLE: ComputeR-Assisted Detector of LEukocoria) was engineered to detect photographic leukocoria and is available for download under the name "White Eye Detector." This study determined the sensitivity, specificity, and accuracy of CRADLE by retrospectively analyzing 52,982 longitudinal photographs of children, collected by parents before enrollment in this study. The cohort included 20 children with retinoblastoma, Coats' disease, cataract, amblyopia, or hyperopia and 20 control children. For 80% of children with eye disorders, the application detected leukocoria in photographs taken before diagnosis by 1.3 years (95% confidence interval, 0.4 to 2.3 years). The CRADLE application allows parents to augment clinical leukocoria screening with photography.

INTRODUCTION

The reflection of visible light by the choroidal and retinal blood vessels causes the human pupil to appear red when examined by a handheld direct ophthalmoscope (*1*) or photographed with a camera flash (*2*). Testing children for this red pupillary reflex (red reflex testing) is a required part of a regular newborn and pediatric checkup (*3*). Abnormal or asymmetrical pupillary reflexes such as leukocoria (Fig. 1), a white or yellow-orange reflex encompassing all of the pupil or semicircular or crescent-shaped, can be a symptom of common and rare childhood eye disorders (*4*–*8*). These disorders include retinoblastoma (*9*), pediatric cataract (*10*), retinopathy of prematurity (*11*), persistent fetal vasculature (*5*), Coats' disease (*12*), refractive error (*7*, *8*), amblyopia (*6*–*8*), and strabismus (*7*). Retinoblastoma, the most serious of these disorders, is a malignancy of the developing retina (*9*) that accounts for 6% of cancers in children less than 5 years old (*13*). Early diagnosis of retinoblastoma results in better outcomes for children including higher rates of survival, vision preservation, and avoidance of treatment with acute or long-term toxicity (*14*, *15*).

The effectiveness of red reflex testing during a general physical exam is limited (*3*, *16*–*20*). Abramson *et al.* (*16*) reported that signs of retinoblastoma (typically leukocoria) are first found by pediatricians in only 8% of cases, suggesting that the standard pediatric examination with a handheld direct ophthalmoscope may not result in early detection. A family member or friend detected initial signs (typically leukocoria) before pediatricians in 80% of retinoblastoma cases (*16*). The ineffectiveness of red reflex testing by primary care physicians can be due to lack of eye dilation (*18*), insensitivity to peripheral and posterior abnormalities and lesions (*3*, *18*), performing the exam too close to the face (*21*), failure to correctly identify or report abnormal reflexes (*22*), and difficulty examining noncompliant infants (*23*). Omission of this exam is common because of these limitations (*19*), and the skilled use of direct ophthalmoscopy has recently been described as a "dying art" (*1*).

Digital photography, including personal photographs collected by parents of children engaged in activities and settings typical of childhood, represents a feasible supplement for conventional leukocoria screening (*4*, *24*). Foremost, a parent will photograph their child's eyes more often than primary care physicians will perform a red reflex test. In the United States, for example, only 20 wellness checkups are recommended through adolescence (*25*). Each photograph collected by a parent—hundreds or thousands throughout childhood—has the potential to represent a frequent test for leukocoria. A parent will also photograph a child at multiple optical axes and distances, which increases the probability that light will reflect off ocular lesions regardless of location (*4*). However, photographic leukocoria can be difficult for parents to detect or recognize, especially when leukocoria appears at low intensity, low resolution, or in a small fraction of photographs (*4*). Artificial neural networks, operating autonomously on smartphones, might help parents detect leukocoria by routinely scanning personal photographs for leukocoria throughout development (as a supplement to clinical leukocoria screening) (*26*).

RESULTS

White Eye Detector/CRADLE smartphone application

A previously described convolutional neural network [network 16 in reference (*27*)] was reconfigured and retrained in the current study on a proprietary set of casual photographs of children exhibiting pathologic leukocoria and of children and adults exhibiting non-leukocoria (see Fig. 2A for representative array of training images). Individuals who were pictured in training images did not include any of the 40 children from the cohort of this current study. The newly trained and reconfigured network was embedded into a prototype application for smartphones and tablets. We refer to this application

[1]Department of Chemistry and Biochemistry, Baylor University, Waco, TX, USA. [2]Department of Computer Science, Baylor University, Waco, TX, USA. [3]Retina Service, Massachusetts Eye and Ear Infirmary, Boston, MA, USA. [4]Harvard Medical School, Boston, MA, USA. [5]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: bryan_shaw@baylor.edu

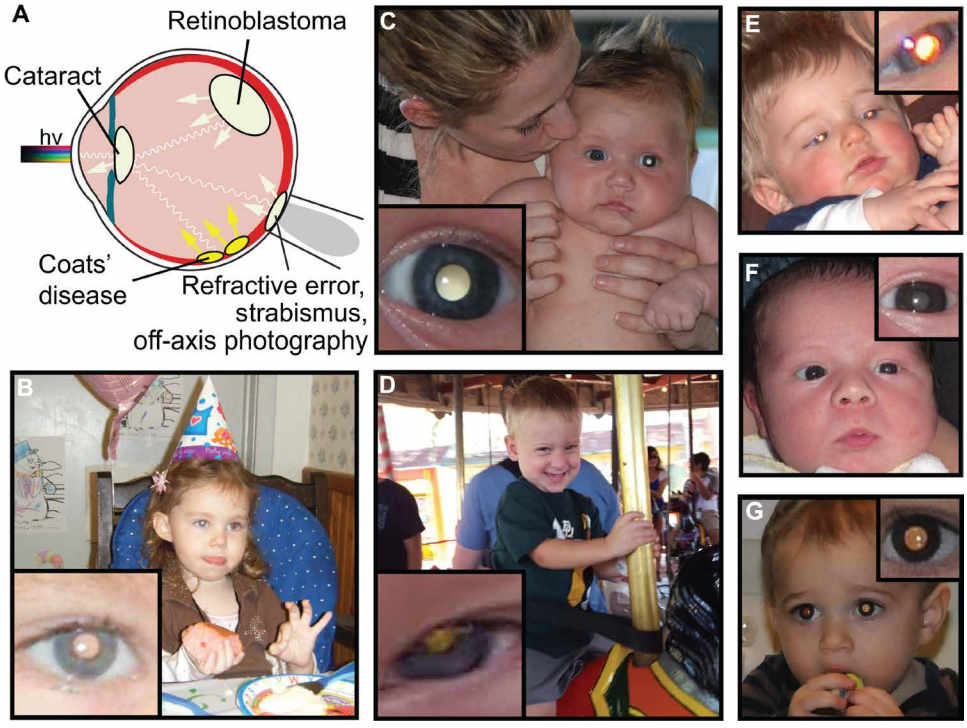

Fig. 1. Examples of pathologic and physiologic leukocoria detected in childhood photographs by the prototype CRADLE application. (**A**) Pathologic leukocoria is caused by the reflection of light off abnormal ocular surfaces including aggregated γ-crystalline (cataract), cholesterol (Coats' disease), tumor surfaces (retinoblastoma), or the abnormal reflection of light off the optic disc (refractive error, anisometropic amblyopia, or strabismus). (**B** to **F**) Examples of pathologic leukocoria in photographs of test children from this study who have been diagnosed with (B) hyperopia, (C) retinoblastoma, (D) Coats' disease, (E) anisometropic amblyopia, and (F) cataract. (**G**) Child without an eye disorder exhibiting normal red reflex (right eye) and physiologic leukocoria (left eye) caused by reflection of light off the optic disc during off-axis photography (*27*). Insets designate leukocoric pupil positively detected by the prototype CRADLE application operating on iPhone X. Photo credit: With permission from the subjects' legal guardian.

as CRADLE (Computer-Assisted Detector of Leukocoria). The CRADLE application privately analyzes digital imagery stored directly on the user's device, without the need to upload a photograph. CRADLE alerts the user when leukocoria is detected (the Android version of CRADLE operates autonomously, analyzing newly collected photographs after scanning the initial library; however, the iOS version requires the user to initiate the analysis of their photographs during each session). For humanitarian reasons, the CRADLE prototype application was immediately uploaded to the Apple App Store and Google Play (in 2014 and 2015, before the completion of this study) under the name "White Eye Detector" (fig. S1 and the Supplementary Materials). This CRADLE prototype can be downloaded at zero cost.

The current study represents the first description of the CRADLE/White Eye Detector smartphone application and the first longitudinal testing on childhood photographs collected by parents. This study is the first to assess the clinical utility of this (or any) smartphone application for detecting leukocoria in personal photographs. This study determined the longitudinal sensitivity, specificity, and accuracy of the application and, by retrospective analysis, whether the application could have detected leukocoria in a parent's personal photographs before clinical diagnosis. Prior studies on the development of leukocoria detection algorithms (*26*, *27*) did not determine sensitivity, specificity, accuracy, or clinical timing of leukocoria detection in personal photographs. The current study also more firmly establishes the natural frequency of occurrence of physiologic and pathologic leukocoria in personal photographs collected throughout childhood (by parents). Previous studies only examined the longitudinal frequency of pathologic leukocoria in personal photographs of a single child with retinoblastoma and did not examine frequency of physiologic leukocoria in healthy children (*4*).

Computer-assisted detection of pathologic and physiologic leukocoria in childhood photographs

The ability of the CRADLE prototype to detect photographic leukocoria was quantified on three different smartphones (iPhone 7, iPhone X, and Google Pixel 2XL) loaded with 52,982 facial photographs of 20 children with an eye disorder and 20 children without an eye disorder. These photographs were collected longitudinally by parents and are casual in nature (examples are shown in Figs. 1 and 2). The results described in this paper refer to the CRADLE application operating on the iPhone X smartphone. Results from other smartphones can be found in the Supplementary Materials (see tables S1 to S6).

Before the analysis of photographs by the prototype application, researchers manually analyzed each photograph and discarded images that did not contain the face of the child in question (face = one or both eyes open + part of forehead and nose present + part of either lips or chin present). The final set of photographs included 23,248 facial photographs of 20 children with an eye disorder: unilateral retinoblastoma (n = 8), bilateral retinoblastoma (n = 7), Coats' disease (n = 2), bilateral cataracts (n = 1), anisometropic amblyopia

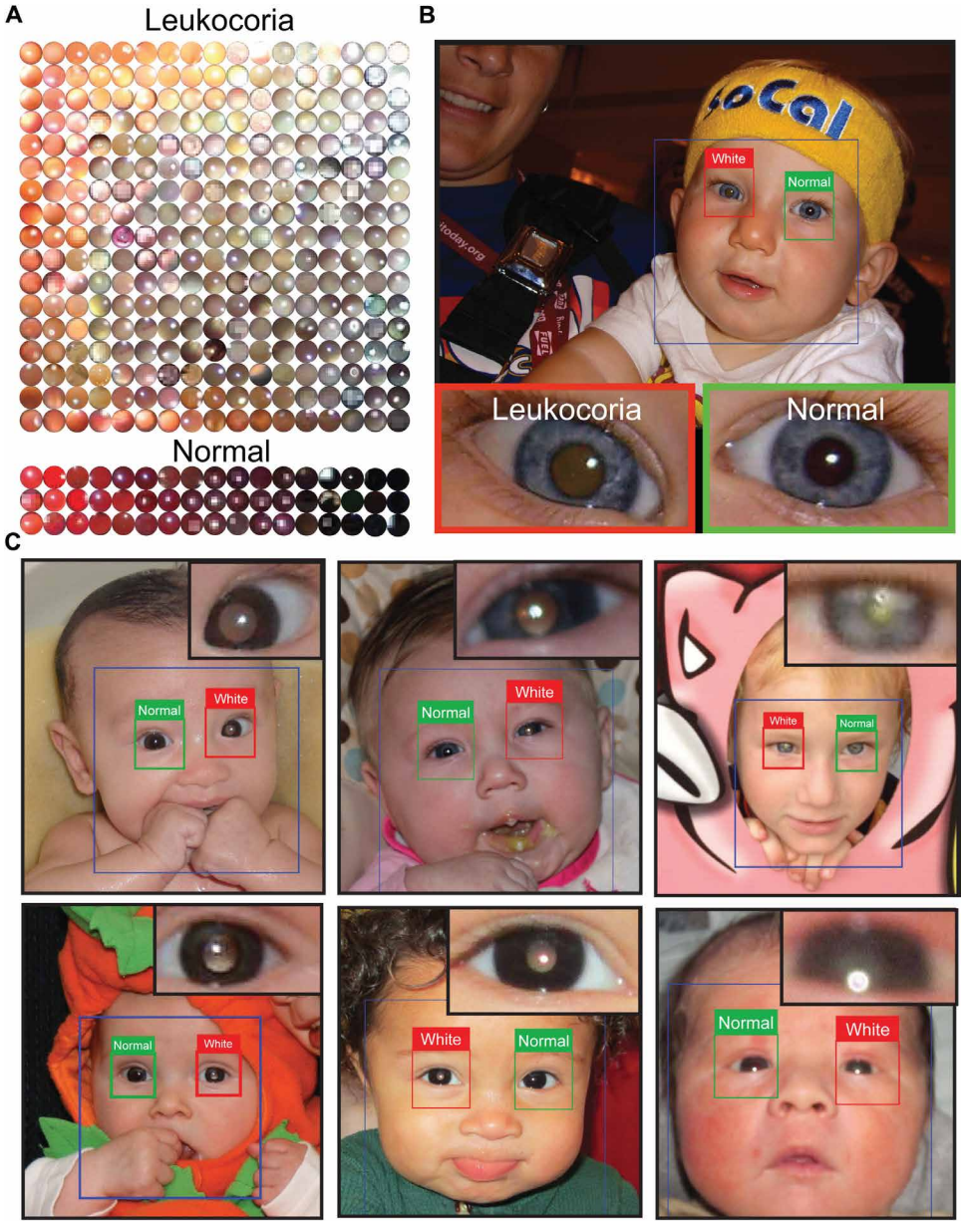

Fig. 2. Training the CRADLE prototype application to detect pathologic leukocoria in photographs of children with eye disease. (**A**) Representative arrays of leukocoric (top) and nonleukocoric (bottom) pupils used to train CRADLE to detect leukocoria. (**B**) Example of pathologic leukocoria detected by CRADLE in test child with retinoblastoma. (**C**) CRADLE detection of pathologic leukocoria caused by cataract (bottom right) and retinoblastoma (all other panels). Red box indicates positive leukocoria detection; green box indicates negative leukocoria detection. The insets in the upper right corner of each photograph shows a magnified view of the leukocoric pupil detected by CRADLE. Photo credit: With permission from the subjects' legal guardian.

(n = 1), and bilateral hyperopia (n = 1). The remaining 29,734 facial photographs were of 20 children without an eye disorder. Children with eye disorders are referred to as "test" children (children 1 to 20); children without eye disorders are referred to as "control" children (children 21 to 40). Each facial photograph was manually inspected by three researchers who classified each pupil as leukocoric or nonleukocoric by abductive reasoning (i.e., "the elephant test") (*28*). Coordinated standard criteria for defining leukocoria observed in photographs or by clinicians in a red reflex exam do not exist. However, a quantitative scale of photographic leukocoria has been previously established (*4*) and was used to arbitrate the classification of ambiguous pupils.

Examples of pathologic leukocoria detected by CRADLE can be found in multiple photographs throughout this manuscript (e.g., Figs. 1 to 3, figs. S1 and S2, and the Supplementary Materials). These photographs illustrate the intended settings and activities for CRADLE usage, for example, children eating birthday cake (Fig. 1B), riding an outdoor carousel (Fig. 1D), learning to eat solid food (Fig. 2C), receiving a bath (Fig. 2C), posing in a carnival cutout (outdoors; Fig. 2C), trick-or-treating (dressed in a pumpkin costume;

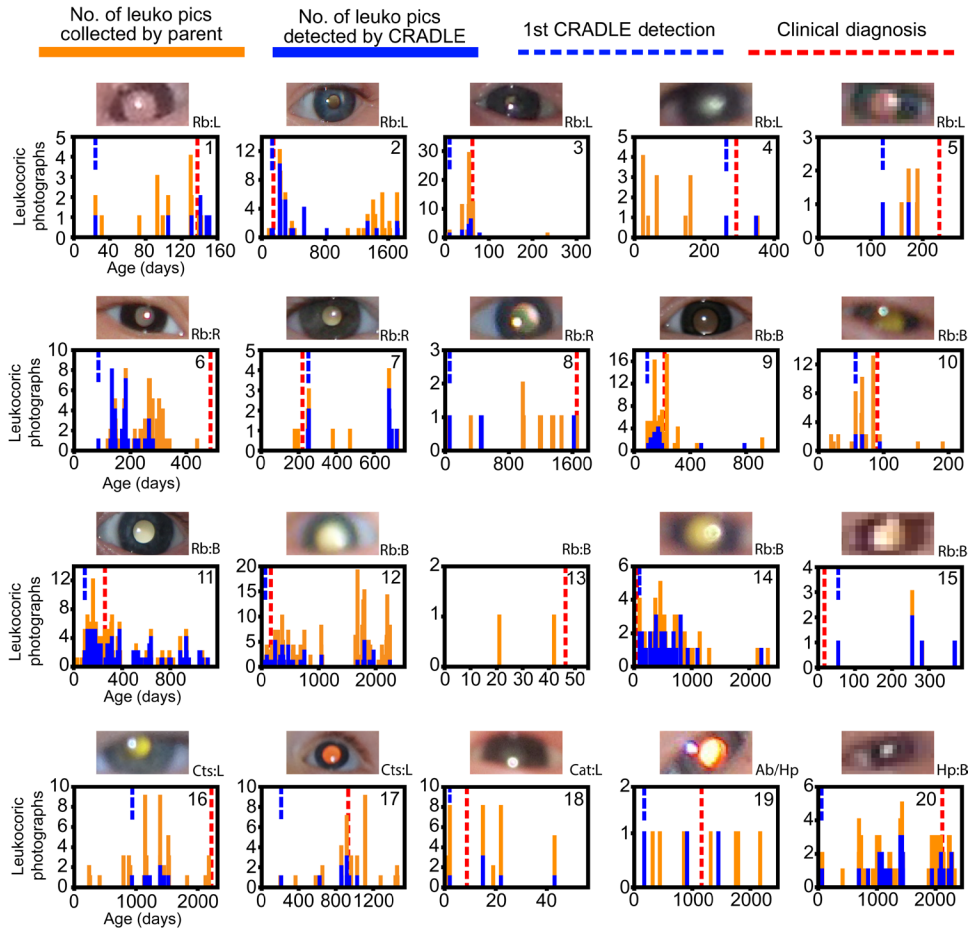

Fig. 3. Longitudinal frequency and detection of pathologic leukocoria in childhood photographs of 20 test children with an eye disorder. Plots 1 to 20 depict leukocoria occurrence among 23,248 photographs of 20 children from before to after diagnosis. Orange bars represent the total number of leukocoric photographs collected on a specific day; blue represents the fraction of those photographs detected by CRADLE. Blue dashed line denotes first detection of pathologic leukocoria by CRADLE; red dashed line denotes age of clinical diagnosis. Crop of first leukocoric pupil detected by CRADLE is shown (leukocoria was not detected by CRADLE for child 13). Rb, retinoblastoma; Cts, Coats' disease; Cat, cataract; Ab, amblyopia; Hp hyperopia. L (left), R (right), and B (bilateral) indicate which eye was affected by each disorder (child 19 exhibited hyperopia in left eye and amblyopia in right eye). The duplicate analysis of these photographic libraries by CRADLE operating on iPhone 7 and Google Pixel 2XL smartphones can be found in figs. S7 and S8. Photo credit: With permission from the subjects' legal guardian.

Fig. 2C), sitting in the jaws of a giant model shark (fig. S2 and the Supplementary Materials), and posing next to a holiday decoration (fig. S2 and the Supplementary Materials).

The CRADLE application was able to detect pathologic leukocoria presenting at low brightness and low resolution, which, otherwise, might be difficult for a parent to recognize on their smartphone display (Fig. 2B, fig. S2, and the Supplementary Materials). For example, CRADLE detected faint leukocoria that appeared as a gray pupil in an affected eye of a child with retinoblastoma (Fig. 2B). The color space coordinates of this particular leukocoric pupil were characterized by $H = 45°$, $S = 0.51$, and $V = 0.30$ (Fig. 2B). We consider this trace level of leukocoria to be at or near the current lower detection limit of CRADLE. The trace level of leukocoria in this child's right eye might be difficult for a parent to distinguish from the nonleukocoric (left) eye, which exhibited color coordinates, $H = 346°$, $S = 0.32$, and $V = 0.19$ (Fig. 2B). The CRADLE application was also able to detect pathologic leukocoria appearing at low resolution (e.g., a pupil composed of 18 pixels) in photographs with a wide field of view (fig. S2 and the Supplementary Materials).

Leukocoria can also be present in individuals without eye disorders (physiologic leukocoria) (*29*). Physiologic leukocoria (Fig. 1G) is caused by flash photography at ~15° off the optic axis, which maximizes reflection of light off the optic disc toward the camera lens (*29*, *30*). This "magic" angle was corroborated in this study (fig. S3 and the Supplementary Materials). Physiologic leukocoria can exhibit identical colorimetric properties to pathologic leukocoria (*4*). In this study, leukocoria was categorized as pathologic or physiologic based on the following criteria: (i) occurring in a child with an eye disorder in the leukocoric eye (pathologic) or (ii) occurring in a child without a known eye disorder or in a child with an eye disorder, but with leukocoria in the clinically normal eye (physiologic). We no longer use the term "pseudoleukocoria" to refer to reflection of light off the optic disc as the gross appearance and colorimetric properties can be identical to leukocoria associated with an eye disorder (*4*).

Retrospective longitudinal analysis of digital photographs by CRADLE

The date that each photograph was collected was retrieved from metadata and used to retrospectively determine how early the CRADLE prototype would have detected pathologic leukocoria in the parent's photolibrary had the application been used by parents to routinely

scan the library of photographs stored on their smartphone (Fig. 3). At least one leukocoric photograph (as classified by researchers) was collected before diagnosis for 18 of 20 test children (Fig. 3). The two test children who did not exhibit photographic leukocoria before diagnosis (children 14 and 15) had a family history of bilateral retinoblastoma and were diagnosed 2 weeks after birth (Fig. 3, table S7, and the Supplementary Materials). Of the 18 children who exhibited leukocoria before diagnosis, CRADLE detected leukocoria in photographs taken before diagnosis for 16 children, by a mean of 484 days before diagnosis [95% confidence interval (CI), 138 to 829 days; table S7 and the Supplementary Materials]. The longitudinal frequency of physiologic leukocoria in photographs of healthy control children and detection by CRADLE can be found in Supplementary Materials (figs. S4 to S6 and the Supplementary Materials).

Determining longitudinal sensitivity, specificity, and accuracy of CRADLE

Four outcomes were used to calculate the sensitivity (TPR), specificity (SPC), and accuracy (ACC) of CRADLE at detecting pathologic leukocoria (Table 1). A true-positive (TP) outcome was defined as the detection of pathologic leukocoria by CRADLE in at least one photograph (classified by researchers to exhibit pathologic leukocoria) of a child with an eye disorder. A false-negative (FN) outcome was defined as the failure of CRADLE to detect pathologic leukocoria in all photographs of a child with an eye disorder. A true-negative (TN) outcome was the absence of detection of leukocoria by CRADLE in all photographs of a child without an eye disorder, regardless of whether pictures exhibited physiologic leukocoria. A false-positive (FP) outcome was the detection of leukocoria by CRADLE in at least one photograph of the child without an eye disorder, whether the photograph exhibited physiologic leukocoria or normal pupillary reflex.

The sensitivity, specificity, and accuracy were calculated at discrete time intervals for photographs collected up to 24 months of age of test and control children (Table 1). The sensitivity of CRADLE at detecting pathologic leukocoria in facial photographs of children was 75% at age <6 months and 90% at age <2 years (specificity = 20.0% and accuracy = 55.0% at age <2 years) (Table 1).

As expected, the sensitivity of the CRADLE prototype increased, while the specificity decreased with age (Table 1). This divergence occurs for two reasons. First, most children who develop conditions such as Coats' disease (*12*) and unilateral retinoblastoma (*5*) do not have disease at birth; therefore, pathologic leukocoria will emerge in later years (bilateral retinoblastoma can be present at birth). Second, the probability of collecting and detecting at least one photograph with physiologic or pathologic leukocoria increases with time (Fig. 3, figs. S4 to S8, and the Supplementary Materials); that is, CRADLE detects physiologic leukocoria and pathological leukocoria (Table 2). For example, in the first 3 months of life of test and control children, pathologic leukocoria occurred in at least one photograph of 13 different test children, and physiologic leukocoria occurred naturally in 4 control children (Fig. 3, fig. S4, and the Supplementary Materials). Up to 2 years of age, 20 test children and 11 control children exhibited leukocoria, of which CRADLE detected leukocoria in 18 of these test children and 9 control children.

Intrinsic rates of leukocoria detection by CRADLE

The rates of true and false detection of both physiologic and pathologic leukocoria by the CRADLE prototype were also calculated and reported on a per-photograph basis (Table 2, table S8, and the Supplementary Materials). These intrinsic metrics are more accurate indicators of the gross performance of the CRADLE prototype because they are not affected by the number or timing of photographs collected by a parent. Unlike values listed in Table 1, rates of intrinsic TP (ITP), intrinsic FP (IFP), intrinsic TN (ITN), and intrinsic FN (IFN) detection express the rate that CRADLE will detect pathologic leukocoria in a test child or physiologic leukocoria in a control child. Here, a positive sample was defined as a photograph containing one or more pupils classified by researchers to exhibit pathologic leukocoria in a test child or physiologic leukocoria in a control child. A negative sample was a photograph containing no examples of photographic leukocoria. Researchers assigned one of four single classifications (ITP, IFP, ITN, and IFN) to each photograph after analysis by CRADLE based on the linear rank order: ITP > IFN > IFP > ITN. See the Supplementary Materials for further explanation (tables S9 and S10).

Table 1. Sensitivity, specificity, and accuracy of the prototype CRADLE application (iPhone X) at detecting pathologic leukocoria in childhood photographs.

Age	Sensitivity (TPR)		Specificity (SPC)		Accuracy (ACC)		No. of diagnosed†
	No./total no.*	% (95% CI)	No./total no.*	% (95% CI)	No./total no.*	% (95% CI)	
≤1 month	3/17	17.6 (0.0–35.8)	10/16	62.5 (38.8–86.2)	13/33	39.4 (22.7–56.1)	3/20
≤2 months	7/17	41.2 (17.8–64.6)	6/16	37.5 (13.8–61.2)	13/33	39.4 (22.7–56.1)	4/20
≤3 months	10/18	55.6 (32.6–78.5)	7/18	38.9 (16.4–61.4)	17/36	47.2 (30.9–63.5)	6/20
≤6 months	15/20	75.0 (56.0–94.0)	5/20	25.0 (6.0–44.0)	20/40	50.0 (34.5–65.5)	9/20
≤1 year	18/20	90.0 (76.9–100.0)	4/20	20.0 (2.5–37.5)	22/40	55.0 (39.6–70.4)	14/20
≤1.5 years	18/20	90.0 (76.9–100.0)	4/20	20.0 (2.5–37.5)	22/40	55.0 (39.6–70.4)	15/20
≤2 years	18/20	90.0 (76.9–100.0)	4/20	20.0 (2.5–37.5)	22/40	55.0 (39.6–70.4)	20/20

*Total number refers to the total number of children whose faces and open eye(s) were photographed in each age interval. The natural occurrence of physiological leukocoria in control children lowers specificity over time. Sensitivity is low in the first few months of life because certain disorders, such as unilateral retinoblastoma, Coats' disease, and refractive error, are typically not present at birth, whereas bilateral retinoblastoma can be present at birth or develop in the first few months of life. †Number of children diagnosed in each respective age period.

Table 2. Incidence and detection rate of leukocoria in photographs collected by parents of 20 test children and 20 healthy control children (iPhone X). N/A, not applicable.

	Unilateral retinoblastoma (*n* = 8)	Bilateral retinoblastoma (*n* = 7)	Coats' disease, cataract, amblyopia, hyperopia (*n* = 5)	Healthy control (*n* = 20)*
Average age at diagnosis (days)	402	114	1277	N/A
Average age of detection by CRADLE prototype (days)	117	75	276	N/A
No. of photographs collected†	9065	7129	7054	29,734
No. of leukocoric photographs collected†	444	927	304	144
No. of leukocoric photographs detected by CRADLE†	138	289	72	81
No. of leukocoric photographs collected before diagnosis†	261	293	232	N/A
No. of leukocoric photographs detected by CRADLE before diagnosis†	74	87	56	N/A
ITP rate‡	31.08%	31.18%	23.68%	56.25%
IFP rate‡	0.75%	0.61%	0.33%	0.79%

*Data listed for 20 control children without an eye disorder. Eight of 20 control children did not exhibit physiologic leukocoria during the time period of photography. Table S8 lists personalized values for each control child. †Sum of all facial photographs (with open eyes) collected from each patient from each respective category. ‡Rates of intrinsic true positive (TP) and FP leukocoria detection refer to the average true or false detection of leukocoria by CRADLE per photograph (see table S7 for personalized values for each test child).

The rate that the CRADLE prototype detected both physiologic leukocoria in photographs of control children and pathologic leukocoria in test children (ITP) was 31.9%. Thus, of the 1819 photographs that were determined by researchers to contain physiologic or pathologic leukocoria (of 52,982 photographs), leukocoria was correctly identified by CRADLE in 580 photographs (Table 2). The IFP rate—where CRADLE reports leukocoria in a photograph without physiologic or pathologic leukocoria present—was 0.7% per photograph among 51,163 nonleukocoric photographs (Table 2).

Effect of image properties on detection of leukocoria by CRADLE

The resolution of leukocoric pupils accurately detected by CRADLE (physiologic and pathologic) ranged from 9 to 9838 pixels per cropped pupil, compared with 1 to 9741 pixels for leukocoric pupils that were undetected by CRADLE (fig. S9 and the Supplementary Materials). Image resolution was not a statistically significant factor in determining the intrinsic rate of TP or FN detection of leukocoria. The average resolution (number of pixels per cropped pupil) of all leukocoric pupils that were accurately detected by CRADLE (ITP) was 845.04 ± 93.34 pixels (fig. S9 and the Supplementary Materials). The average resolution of all leukocoric pupils misidentified by CRADLE (IFN) was 818.42 ± 68.62 pixels (P = 0.6519 using two-sample t test with assumed heteroscedasticity). In particular, for photographs exhibiting pathologic leukocoria, P = 0.625 when comparing pixel counts of pupils that produced ITP and IFN results. Similarly, for photographs exhibiting physiologic leukocoria, P = 0.356 when comparing pixel counts of pupils that produced ITP and IFN results.

DISCUSSION

A leukocoria detector for parents

The White Eye Detector (CRADLE) prototype, first released to the public in 2014, is the first computer-assisted detector of leukocoria for digital imagery. The CRADLE prototype was designed for parents to augment (not replace) conventional clinical leukocoria screening. This study represents the first determination of CRADLE's sensitivity, specificity, and accuracy against longitudinal sets of personal photographs of children. Strengths of this study are that the sets of photographs were collected by parents before enrollment in this study. This status prevented bias in photography of a particular angle, distance, brightness, or setting that might promote or prevent leukocoria (*4*, *29*). The analysis of photographs by the CRADLE prototype was also blind per se. For example, each FP result that was used to calculate specificity or accuracy was an unbiased, computer-generated result.

The intended setting for using CRADLE is any setting or activity of childhood. We trained CRADLE with casual photographs of children collected by parents in diverse environments (indoors and outdoors). The utility of the CRADLE application is that it harnesses a parent's natural tendency to collect myriad pictures of their child at multiple optic axes, thus maximizing the probability that leukocoria will be present, regardless of the position of a lesion or abnormality.

We point out that this CRADLE prototype was able to detect leukocoria in children of different races. Of the three test children with retinoblastoma who were Asian, Latino, or African-American, CRADLE produced ITP rates of 25.0, 34.1, or 31.8%, respectively. These rates are similar to the average ITP rate of detection of pathologic leukocoria for all test children (29.8%; Table 2). This similarity

suggests that CRADLE can distinguish leukocoria, despite differences in eye anatomy.

Challenges of leukocoria detection in personal (casual) photographs

Two challenges of autonomously screening personal photographs for leukocoria are (i) a diverse format (e.g., photographic field of view, angle, pupil hue, saturation, value, resolution, or pose of the child) and (ii) physiologic leukocoria. Of the leukocoric photographs analyzed in this study, the photographic angle, relative to the child's optic axis, varied from 0.002° to 62.3° (horizontally); distance between the camera and eye varied 10-fold; and pupil resolution (area) varied 10,000-fold. The effect of image format on rates of leukocoria detection is discussed in greater detail in the Supplementary Materials. Despite this photographic diversity, the sensitivity of CRADLE at ≤2 years of age—when retinoblastoma typically develops and many children are incommunicable (*19*)—surpassed 80%. This 80% threshold is viewed by ophthalmologists as a general "gold" standard of sensitivity for telemedical devices (*31*) and for clinical, high-resolution machine-learning based screening devices, e.g., those that analyze retinal vasculature for diabetic retinopathy (*32*, *33*). The low sensitivity of CRADLE in the first few months of life (Table 1) is caused by the later onset of diseases such as unilateral retinoblastoma (*5*) and Coats' disease (*12*), which can develop months or years after bilateral retinoblastoma (bilateral retinoblastoma can be present at birth).

As expected, the natural occurrence of physiologic leukocoria in photographs of control children (and detection by CRADLE) lowered the specificity and accuracy of the CRADLE prototype below the 80% gold standard. However, an FP result produced by CRADLE can be resolved quickly and inexpensively with a noninvasive examination by an ophthalmologist. In future versions of CRADLE (or similar applications), this low specificity and accuracy resulting from physiologic leukocoria can be overcome by a detector of leukocoria frequency. For example, pathologic leukocoria occurred more frequently per photograph of a test child than physiologic leukocoria in a control child by a factor of 15 (7.2% of photographs of test children versus 0.48% of control children; Table 2). Introducing an optional setting that alerts the user only when a threshold rate of leukocoria (per photograph) is reached, e.g., when >0.5% of photographs exhibit leukocoria over a given period of time, can overcome this shortcoming. We withheld this feature in the current CRADLE prototype because the feature would delay leukocoria detection in children with retinoblastoma, for whom the application was originally designed.

Guidelines and recommendations for using CRADLE

We issue a few guidelines for parents (or clinicians) who wish to use the CRADLE application. First, CRADLE only analyzes facial photographs for leukocoria. CRADLE will not analyze an image for leukocoria if the entire face is not present or is not detected by the device's face detector. Photographs without a full face (e.g., "ninja mask" photographs with only one or two eyes in the field of view) will not be recognized by CRADLE regardless of whether the photograph contains leukocoria. Second, a flash is typically required to produce leukocoria. Of the leukocoric photographs analyzed in this study, metadata revealed that a flash was used to collect all but two leukocoric photographs. Red eye reduction technology (where an initial flash is delivered from the camera to constrict pupils, followed by a second flash during image recording) did not completely abolish the emergence of leukocoria. According to the metadata, ~25% of the photographs that contained leukocoria were collected with active red eye reduction technology.

Moreover, users of CRADLE should be informed that this prototype cannot discriminate pathologic and physiologic leukocoria. However, this indiscrimination does not change the current guidelines that children who present with photographic leukocoria are to be referred to an ophthalmologist (*21*). Leukocoria is already a symptom that parents detect and report via manual inspection of photographs. Will CRADLE's ability—a parent's enhanced ability—to detect physiologic leukocoria create unnecessary referrals and anxiety for parents? We believe that education about the proper context of leukocoria can help prevent and manage parent anxiety. We recommend that pediatricians inform parents of the following: (i) that pathologic leukocoria is typically associated with nonfatal, harmless conditions (refractive error or amblyopia); (ii) that physiologic leukocoria is a common artifact of off-axis photography; and (iii) that recurrent leukocoria in multiple photographs is of greater concern than a single instance of photographic leukocoria (however, all instances of photographic leukocoria should be reported and investigated).

Users should also be aware that this particular CRADLE prototype does not detect every photograph exhibiting pathologic or physiologic leukocoria. The probability that the CRADLE prototype will correctly detect leukocoria when it appears in a single facial photograph (i.e., the ITP rate) is ~30% per leukocoric photograph; that is, CRADLE will detect (on average) every third photograph that exhibits leukocoria. This modest rate is useful, considering the high frequency at which parents photograph their children and the high frequency that leukocoria appears in photographs of children with different eye disorders. For example, with an ITP rate of 31.9%, the probability (P) of accurately detecting at least one photograph with pathologic or physiologic leukocoria will exceed 90% when the library includes $n > 6$ photographs exhibiting pathologic or physiologic leukocoria, i.e., $P = 1 - (1 - \text{ITP})^n$ (Fig. 4A). The average number of photographs with pathologic leukocoria that were collected by parents in this study (before clinical diagnosis) was 39 leukocoric photographs per child (table S7 and the Supplementary Materials). The modest ITP rate of the CRADLE prototype, coupled with the high frequency of pathologic leukocoria in photographs, can explain why the CRADLE prototype was able to detect leukocoria before diagnosis for 16 test children by an average of 1.3 years.

The IFP rate of leukocoria detection is 0.7% per facial photograph; that is, CRADLE will correctly ignore 99.3% of facial photographs without physiologic or pathologic leukocoria. The probability that CRADLE falsely reports leukocoria in at least one facial photograph that does not contain physiologic or pathologic leukocoria reaches 90% after 500 nonleukocoric facial photographs are analyzed (Fig. 4B).

Therefore, the CRADLE prototype is more efficacious at detecting an instance of leukocoria within a large library of photographs that contain multiple examples of leukocoria than in a small library with only a single leukocoric photograph. We do not recommend using CRADLE to analyze a single photograph or a small set of photographs. We also do not recommend using CRADLE to analyze small numbers of clinically derived photographs depicting leukocoria (taken of fully dilated eyes in a darkened examination room). The network in CRADLE was trained on casual photographs of children in live action who were present in diverse environments, i.e., indoors, outdoors, photographed with myriad secondary light sources (sunlight, overhead lighting, table top lamps, etc.), and whose pupils were not pharmacologically dilated.

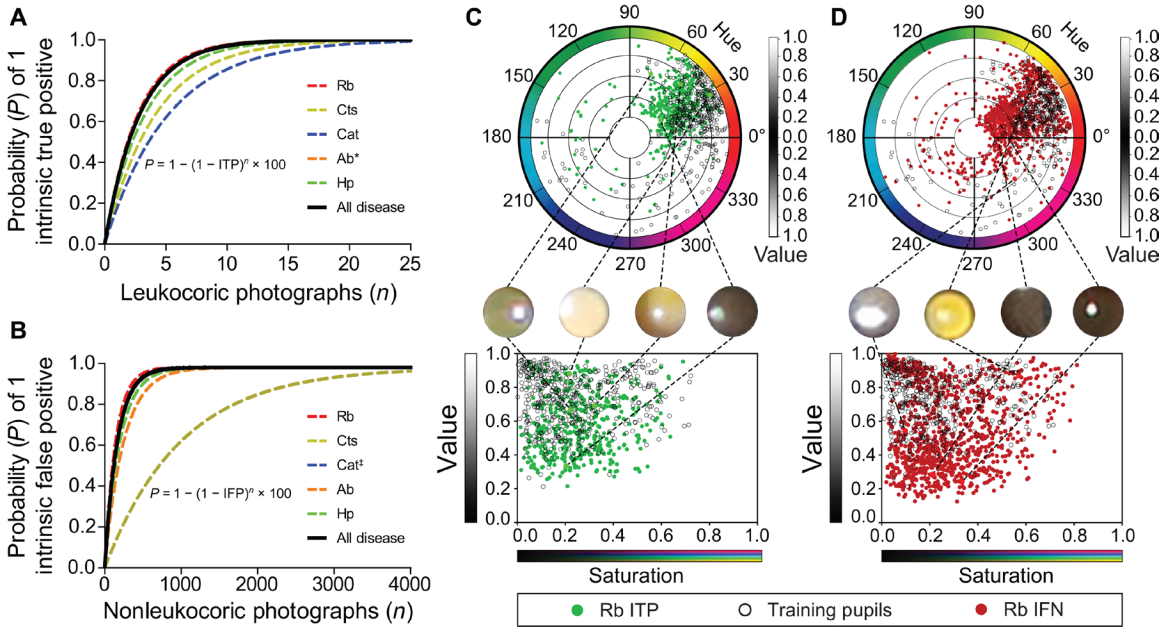

Fig. 4. Probability of CRADLE prototype to (**A**) accurately or (**B**) falsely detect a single photograph with leukocoria as a function of number of photographs analyzed (based on ITP and IFP rates from analysis by CRADLE of 23,248 photographs of 20 children with retinoblastoma, Coats' disease, cataract, anisometropic amblyopia, and hyperopia). (**C** and **D**) Comparison of average hue and value (top) and saturation and value (bottom) of cropped photographs of leukocoric pupils of children with retinoblastoma that were (C) detected and (D) undetected by CRADLE. *Ab is identical to Rb in ITP probability plot. †IFP rate = 0 due to small sample size.

Improving the prototype CRADLE application

The ITP rate and sensitivity of the CRADLE prototype will increase, and IFP rate will decrease, as additional photographs are collected to continually retrain its neural network. A colorimetric analysis of leukocoric pupils that CRADLE accurately detected (ITP; Fig. 4C) and leukocoric pupils that CRADLE failed to detect (IFN; Fig. 4D) suggests that the colorimetric coordinates of misidentified pupils (i.e., hue, >60°; value, <0.4) were not highly represented by the leukocoric photographs used to train CRADLE (i.e., hue, <60°; value, >0.4). Differences in the average colorimetric properties of accurately identified leukocoric pupils compared with misidentified leukocoric pupils were observed for other eye disorders (fig. S10 and the Supplementary Materials). These mismatches (and similar mismatches found in future studies) can be corrected by retraining CRADLE with the leukocoric photographs that it failed to detect (and nonleukocoric photographs that it falsely reports as leukocoric). This feedback is a benefit of the learning-based approach we have chosen. We request that any users who wish to donate training photographs contact the corresponding author of this study (Shaw).

Effect of early leukocoria detection on outcomes of children with retinoblastoma

Would CRADLE have helped test children if it had been used by their parents to routinely scan digital photographs? For children with retinoblastoma, CRADLE detected leukocoria 284 ± 547 days before diagnosis of unilateral retinoblastoma and 50 ± 103 days for bilateral retinoblastoma ($P = 0.272$) (table S7 and the Supplementary Materials). Immediate diagnosis and treatment of retinoblastoma are crucial because malignancies grow and metastasize rapidly (*9*, *34*–*36*). Decreasing the age of retinoblastoma diagnosis by several months can prevent vision loss and reduce the need (or dose) of interventions, such as chemotherapy and radiation therapy, that increase the risk of secondary malignancies (*14*, *37*). Considering the early dates of leukocoria detection by CRADLE, we expect that several children in this study would have experienced better clinical outcomes if diagnosed near the time at which CRADLE first detected photographic leukocoria. A discussion of individual cases is included in the Supplementary Materials.

The growing prevalence of smartphones (*38*) and shrinking physician supply (*39*) suggests that the computer-assisted screening of personal photographs for leukocoria—by parents—is a feasible strategy for parents to supplement clinical leukocoria screening. In the United States, where ~92% of adults reported owning a smartphone in 2015 (*38*), there are ~37 million parents with children ≤5 years old (*40*). This population far exceeds the ~27,000 general pediatricians (*41*) and ~123,000 family and general practitioners (*42*) in the United States who might perform a handful of red reflex exams on a child. In the nation of Kenya, the physician density is lower than the United States by an order of magnitude—only 0.2 per 1000 people (*43*)—which contributes to the late diagnosis of retinoblastoma and 3-year survival rates of 27% in certain clinical populations (*44*). However, smartphone ownership among Kenyans in 2015 was 34% and rising among persons 18 to 34 years of age (*38*).

For children with retinoblastoma, smartphone applications, such as CRADLE, might lead to better outcomes by increasing personal and public awareness and detection of photographic leukocoria. Increasing the public awareness of leukocoria, even in resource-limited environments, has been shown to improve outcomes for children with retinoblastoma by accelerating diagnosis (*45*). In Honduras, for example, a public awareness campaign on the symptoms of retinoblastoma (including leukocoria) decreased morbidity from 24 to 4% by reducing extraocular disease from 73 to 35% (leukocoria was the presenting symptom in 83% of cases) (*45*).

CONCLUSION

Since being released for Android and iOS devices in 2014–2015, the free CRADLE prototype has been downloaded onto 10^5 devices by people on every continent (excluding Antarctica). Although we cannot track and quantify the impact of these CRADLE prototypes, media outlets have reported incidents of parents who initiated early diagnoses of retinoblastoma, Coats' disease, and myelin retinal nerve fiber layer in their children by (i) using the CRADLE prototype to detect leukocoria in photographs of their children and (ii) reporting the results to their child's clinician (*46*).

This study suggests that a parent can use a smartphone application, such as CRADLE, to detect leukocoria. Current guidelines from the American Academy of Pediatrics (*47*) recommend that red reflex testing begin in neonates. In addition to this screening, healthcare providers should inform parents that personal photographs and/or a leukocoria detection application, such as CRADLE, can assist in the early diagnosis of eye disorders. Since early detection is key to better outcomes, educating parents of the importance of finding leukocoria early by screening photographs or using a leukocoria detection system such as CRADLE will greatly improve the prognosis for vision and the eye. To fully elucidate the benefits and methods of implementing CRADLE into practice, further community-based studies should be considered, particularly those in resource-limited settings, where late referral for retinoblastoma is a cause for substantial and preventable morbidity and mortality.

MATERIALS AND METHODS

Participants

Longitudinal sets of 52,982 facial photographs of 20 children previously diagnosed with an eye disorder (test children) and 20 children without an eye disorder (control children) were donated to researchers by parents in response to solicitations in mass media. To ensure that parents did not engage in the biased collection of photographs (at a particular angle, setting, or rate of collection), the researchers requested that parents donate their entire, unsorted, digital library of photographs that had been collected up to the time of enrollment. Parents of test children reported the date of diagnosis of the eye disorder, the disorder that was diagnosed, the laterality of the disorder (right, left, or bilateral), the treatment regimen, and in some cases, provided electronic copies of medical records (e.g., corrective lens prescription and fundus photographs). Test children were previously diagnosed with unilateral retinoblastoma (four males and four females), bilateral retinoblastoma (three males and four females), unilateral Coats' disease (two males), bilateral pediatric cataract (one male), bilateral hyperopia (one female), and anisometropic amblyopia (one male).

Children with retinoblastoma were clinically diverse and exhibited the full range of disease severity according to both the International Classification for Intraocular Retinoblastoma Grouping (i.e., groups A to E) and Reese-Ellsworth Staging (i.e., stages I to V) (*48*). Children received a wide range of different therapies including proton beam radiation, intra-arterial chemotherapy, subtenon chemotherapy, systemic chemotherapy, laser photoablation therapy, cryotherapy, and eye enucleation. The two children with Coats' disease were diagnosed at stages 2 and 3.

The age of clinical diagnoses for test children (table S7 and the Supplementary Materials) mirrored the typical range of dates of diagnosis for children in highly developed nations. For example, in highly developed nations, retinoblastoma is typically diagnosed in the first 2 years of life (bilateral retinoblastoma at 9 to 12 months and unilateral retinoblastoma at ~24 months) (*13*) and Coats' disease at ~5 years (*12*). The average reported age of diagnosis for pediatric cataract ranges from birth to <3 months old (*49*) [amblyopia, 1.2 years; (*50*)]; the emergence of hyperopia is uncertain and ranges from months to years (*51*). Control children did not exhibit and had not been diagnosed with any type of eye condition, including common conditions such as refractive error, amblyopia, or strabismus (according to parents). This study was determined to be exempt from review by an Institutional Review Board at the Baylor University. The parents of children whose photographs were used in this study had given written consent to use and, in some cases, publish photographs of their children. Additional details of study methodology can be found in the Supplementary Materials.

Analysis of photographs of test and control children

Researchers manually analyzed each photograph donated by parents of test and control children and removed images that did not contain the face of the test or control child. A usable photograph was defined as containing a face with at least one open eye. Criteria for defining a face were an image with part of the child's forehead and nose present, one or both eyes present, and part of either their mouth or chin present. Photographs were not analyzed or included in the study if both of the child's eyes were closed. A script was written to remove duplicate images from donated libraries by searching for files of identical name and file size.

All 52,982 facial photographs of control and test children were manually inspected for leukocoria by three researchers and classified as containing physiologic leukocoria, pathologic leukocoria (based on whether the eye was affected by disease), or nonleukocoria by abductive reasoning (i.e., the elephant test) (*28*). This manual classification of photographs as leukocoric or nonleukocoric was not blind per se because the evaluator can develop bias for a specific eye shape, skin color, or iris color that exhibits recurrent leukocoria. Final classifications were based on concordant designation of at least two researchers. A quantitative scale of photographic leukocoria (*4*) was used by researchers to arbitrate the classification of pupils with borderline classification. Leukocoria was categorized as pathologic or physiologic based on the following criteria: (i) occurring in a child with eye disease in the leukocoric eye (pathologic) or (ii) occurring in a child without known eye disease or in a child with eye disease, but leukocoria in the clinically normal eye (physiologic). The fraction of photographs that exhibited leukocoria were manually inspected a second time to ensure that no duplicate images existed in the set.

Digital photographs of the faces of test and control children (52,982 in total) were loaded, in batches, onto three different smartphones (iPhone models 7 and X, iOS 10.3.1; Android Google Pixel 2XL smartphone, Android 8.0) to simulate their storage on a parent's smartphone. Photographs were then analyzed with the CRADLE smartphone application as it was downloaded from the Apple App Store (November 2016 for iPhone 7 and April 2018 for iPhone X) and Google Play (April 2018) under the name White Eye Detector. The results of analyses were manually tabulated from the CRADLE readout, which reports the number of leukocoric photographs detected during each scan.

Metadata of each image file were used to retrieve the date when each photograph was taken, the type of camera used, the flash setting, the use of red eye reduction technology, and other parameters (focal length, resolution, etc.). According to metadata, images were

collected from a variety of different digital cameras, smartphones, and tablets, including different models of the Apple iPhone and Android Google smartphones (table S11 and the Supplementary Materials). The photographs were collected over multiple months and years, from as early as the day of birth up to 3175 days old. The range of photography of test children was from the day of birth to 3175 days old and control children from the day of birth to 2501 days old.

Quantifying colorimetric properties of cropped pupils

Adobe Photoshop (Photoshop CS5) was used to obtain the averaged red, green, blue (RGB) values of leukocoric pupils in the photographs of test and control children, as well as in all pupils used for training the neural network operating in the prototype CRADLE application. The RGB values were averaged for each pupil and then converted to hue, saturation, and value (HSV) color space coordinates, as previously described (*4*).

Calculation of sensitivity, specificity, and sensitivity of CRADLE at detecting pathologic leukocoria

The following four outcomes were used to calculate the clinical sensitivity, specificity, and accuracy of CRADLE at detecting pathologic leukocoria in photographs of children with an eye disorder. A TP outcome was defined as the detection of pathologic leukocoria by CRADLE in at least one photograph (judged to exhibit pathologic leukocoria) of a child with an eye disorder, taken during the time intervals listed in Table 1. An FN outcome was defined as the failure of CRADLE to detect pathologic leukocoria in all photographs of a child with an eye disorder in the same time interval. A TN outcome was the absence of detection of leukocoria by CRADLE in all photographs of a child without an eye disorder in the same time interval, regardless of whether pictures exhibited physiologic leukocoria. An FP outcome was the detection of leukocoria by CRADLE in at least one photograph of the child without an eye disorder in the same time interval, whether the photograph exhibited physiologic leukocoria or a normal pupillary reflex.

The sensitivity [TPR = TP/(TP + FN)], specificity [SPC = TN/(TN + FP)], and accuracy [ACC = (TP + TN)/(TP + FP + FN + TN)] of CRADLE at detecting pathologic leukocoria were calculated longitudinally for photographs collected at different time intervals up to 2 years of age of test and control children. The date of collection for each photograph was retrieved from metadata. Each test or control child was included in the calculation of TPR, SPC, and ACC at the point in time when their first photograph was collected.

A Student's *t* test was used to compare the properties of each cropped pupil (e.g., resolution, optic axis, saturation, and value) in different populations and outcomes. The Wheeler-Watson test (a nonparametric test used to compare data on polar coordinates) was used to compare the hue of photographed pupils among populations. A two-sample proportion test was used to compare the proportion of children with physiologic and pathologic leukocoria in the setting of eye disease and without eye disease. Two-tailed *P* values <0.05 (95% CI) were considered significant. All error values were reported as SEM.

Calculating intrinsic rates of pathologic and physiologic leukocoria detection by CRADLE

For calculating intrinsic rates of false or true detection of leukocoria by CRADLE, on a per-photograph basis, the number of positive samples (*P*) was defined as the number of photographs that were manually classified by researchers to contain leukocoria (either physiologic or pathologic), regardless of whether that image contained an eye with disease. The number of negative samples (*N*) was defined as the number of photographs that did not contain leukocoria, regardless of disease state. The ITP rate was calculated as ITP = TP/*P*. The ITN rate was calculated as ITN = TN/*N*. The IFP rate was calculated as IFP = FP/*N*. The IFN rate was calculated as IFN = FN/(TP + FN). A single outcome was assigned to each photograph, even in the setting of divergent finding (left eye or right eye with leukocoria) based on the following linear rank order: ITP > IFN > IFP > ITN. For example, a photograph with one leukocoric pupil classified by CRADLE as an ITP was scored as ITP detection by CRADLE, regardless of whether IFN, IFP, or ITN classifications were made by CRADLE for the other pupil.

Engineering the White Eye Detector/CRADLE smartphone application

Instructions on how to use the White Eye Detector/CRADLE prototype application on tablet and smartphone devices can be found within the application. The basic features of the application are illustrated in fig. S1. The prototype CRADLE application analyzes photographs privately; that is, CRADLE analyzes photographs that are stored directly on the user's device and does not require the image to be uploaded to a distal server for analysis. This design was chosen to permit the application to operate quickly in real time (there is no latency that would result from uploading the image and waiting for results) and to respect the privacy of users. Moreover, internet access is not required to search photographs with CRADLE or to use the video screening mode of CRADLE. The CRADLE application does not track any type of usage; however, the application provides the user with the option to upload (donate) images that are determined by CRADLE to contain leukocoria (accurately or inaccurately). These types of donated images provide feedback to our research team. The source code for the convolutional neural network that detects leukocoria within the CRADLE application has been deposited on GitHub: https://github.com/BaylorCS/baylorml/tree/master/ryan_ml.

The convolutional neural network embedded in CRADLE was a reconfigured and retrained network similar to the one described in a previous study [network #16 in (*27*)]. This new network has a more elaborate structure than the prior network. The size of the input image is 40 × 40 pixels, with three color channels, padded by 10 pixels on each side so that the input layer has a size of 60 × 60 × 3. Next come four convolutional layers that have 7, 14, 21, and 50 convolutional filters, respectively. The first three convolutional layers used a 5 × 5 receptive field for each filter, followed by 2 × 2 max pooling. The fourth layer used a 2 × 2 receptive field for each filter, followed by 3 × 3 max pooling. The intent of so many convolutional and pooling layers is to reduce the image to 50 values, achieving some invariance to input image transformations. The outputs of the last max pooling layer are fully connected to the three output nodes. The convolutional layers used a hyperbolic tangent ("tanh") activation function (before max pooling), while the output layer used a softmax to give probabilities to each class. To determine the best network structure and parameters, rather than training just one network, we trained 10 networks using 10-fold cross-validation. We partitioned the data into 10 disjoint parts ("folds") and then trained one network on 9 parts, holding out a different fold each time. Each network was trained using a specialization of the backpropagation algorithm known as RMSPROP. The result was 10 networks with different weights. To make a prediction, the results of the 10 networks were combined using a voting scheme.

This design is similar to bootstrap aggregating or "bagging" (*52*) in machine learning and was used to avoid overfitting and reduce the variance of the model. For simplicity, we referred to this ensemble of 10 networks together as a singular network.

This convolutional neural network was trained on a previously collected, proprietary set of training photographs. Training photographs did not include photographs of any of the 40 children whose photographs were used to test CRADLE; i.e., the 52,982 photographs used to test the network's ability to detect leukocoria did not include any photographs of children or adults that were used to train the network. The training photographs were casual, personal photographs collected in diverse environments (indoors and outdoors) and were inputted as cropped JPEG images of the entire eye (leukocoric eyes of children with eye disease and images depicting nonleukocoria). A previously established scale of leukocoria in HSV color space (*4*) was used as a reference to ensure that leukocoric training images were represented by all colorimetric subtypes of leukocoria, i.e., xanthocoria, rhodocoria, and cirrocoria (*4*).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/5/10/eaax6363/DC1

Supplementary Materials and Methods

Supplementary Text

Fig. S1. Free CRADLE/White Eye Detector application for iOS or Android.

Fig. S2. Examples of pathologic leukocoria detection by CRADLE in photographs with low pupil resolution.

Fig. S3. Comparison of photography angle and detection of pathologic leukocoria by CRADLE.

Fig. S4. Longitudinal frequency and detection of physiologic leukocoria among 29,734 photographs of 20 control children using the iPhone X smartphone.

Fig. S5. Longitudinal frequency and detection of physiologic leukocoria among 29,734 photographs of 20 control children using the iPhone 7 smartphone.

Fig. S6. Longitudinal frequency and detection of physiologic leukocoria among 29,734 photographs of 20 control children using the Google Pixel 2XL smartphone.

Fig. S7. Longitudinal frequency and detection of pathologic leukocoria in childhood photographs of 20 test children with eye disease using the iPhone 7 smartphone.

Fig. S8. Longitudinal frequency and detection of pathologic leukocoria in childhood photographs of 20 test children with eye disease using the Google Pixel 2XL smartphone.

Fig. S9. Effect of image resolution on classification of leukocoric eye by CRADLE using the iPhone X smartphone.

Fig. S10. Average HSV coordinates of leukocoric pupils from test and control children.

Table S1. Personalized detection rates of pathologic leukocoria by CRADLE (iPhone 7) among 23,248 photographs of 20 test children with retinoblastoma, Coats' disease, cataract, anisometropic amblyopia, and hyperopia.

Table S2. Personalized detection rates of pathologic leukocoria by CRADLE (Google Pixel 2XL) among 23,248 photographs of 20 test children with retinoblastoma, Coats' disease, cataract, anisometropic amblyopia, and hyperopia.

Table S3. Personalized detection rates of physiologic leukocoria by CRADLE (iPhone 7) among photographic libraries of 20 control children.

Table S4. Personalized detection rates of physiologic leukocoria by CRADLE (Google Pixel 2XL) among photographic libraries of 20 control children.

Table S5. Sensitivity, specificity, and accuracy of CRADLE (iPhone 7) for detecting pathologic leukocoria in childhood photographs.

Table S6. Sensitivity, specificity, and accuracy of CRADLE (Google Pixel 2XL) for detecting pathologic leukocoria in childhood photographs.

Table S7. Personalized detection rates of pathologic leukocoria by CRADLE (iPhone X) among 23,248 photographs of 20 test children with retinoblastoma, Coats' disease, cataract, anisometropic amblyopia, and hyperopia.

Table S8. Personalized detection rates of physiologic leukocoria by CRADLE (iPhone X) among photographic libraries of 20 control children.

Table S9. Example of a data sheet used to obtain intrinsic detection rates from test and control children.

Table S10. Linear ranking matrix used to assign a single outcome to a photograph containing two pupils (based on the occurrence of the highest rank of order: ITP > IFN > IFP > ITN).

Table S11. Camera make and model used to capture photographs of test and control children (extracted from metadata).

References (*53*–*55*)

REFERENCES AND NOTES

1. D. D. Mackay, P. S. Garza, B. B. Bruce, N. J. Newman, V. Biousse, The demise of direct ophthalmoscopy: A modern clinical challenge. *Neurol. Clin. Pract.* **5**, 150–157 (2015).
2. D. Murphy, H. Bishop, A. Edgar, Leukocoria and retinoblastoma–pitfalls of the digital age? *Lancet* **379**, 2465 (2012).
3. M. Sun, A. Ma, F. Li, K. Cheng, M. Zhang, H. Yang, W. Nie, B. Zhao, Sensitivity and specificity of red reflex test in newborn eye screening. *J. Pediatr.* **179**, 192–196.e4 (2016).
4. A. Abdolvahabi, B. W. Taylor, R. L. Holden, E. V. Shaw, A. Kentsis, C. Rodriguez-Galindo, S. Mukai, B. F. Shaw, Colorimetric and longitudinal analysis of leukocoria in recreational photographs of children with retinoblastoma. *PLOS ONE* **8**, e76677 (2013).
5. A. Balmer, F. Munier, Differential diagnosis of leukocoria and strabismus, first presenting signs of retinoblastoma. *Clin. Ophthalmol.* **1**, 431–439 (2007).
6. R. W. Arnold, M. D. Armitage, Performance of four new photoscreeners on pediatric patients with high risk amblyopia. *J. Pediatr. Ophthalmol. Strabismus* **51**, 46–52 (2014).
7. S. O. Bani, A. K. Amitava, R. Sharma, A. Danish, Beyond photography: Evaluation of the consumer digital camera to identify strabismus and anisometropia by analyzing the Bruckner's reflex. *Indian J. Ophthalmol.* **61**, 608–611 (2013).
8. L. W. Christian, W. R. Bobier, Case 2: A three-year-old boy with photographic leukocoria. *Paediatr. Child Health* **20**, 345–346 (2015).
9. C. L. Shields, A. M. Leahey, Detection of retinoblastoma at risk for metastasis using clinical and histopathologic features and now mRNA. *JAMA Ophthalmol.* **134**, 1380–1381 (2016).
10. S. Sheeladevi, J. G. Lawrenson, A. R. Fielder, C. M. Suttle, Global prevalence of childhood cataract: a systematic review. *Eye* **30**, 1160–1169 (2016).
11. P. K. Shah, V. Prabhu, S. S. Karandikar, R. Ranjan, V. Narendran, N. Kalpana, Retinopathy of prematurity: Past, present and future. *World J. Clin. Pediatr.* **5**, 35–46 (2016).
12. B. Morris, B. Foot, A. Mulvihill, A population-based study of Coats disease in the United Kingdom I: Epidemiology and clinical features at diagnosis. *Eye* **24**, 1797–1801 (2010).
13. E. Broaddus, A. Topham, A. D. Singh, Incidence of retinoblastoma in the USA: 1975-2004. *Br. J. Ophthalmol.* **93**, 21–23 (2009).
14. I. Al-Nawaiseh, A. Q. Ghanem, Y. A. Yousef, Familial retinoblastoma: Raised awareness improves early diagnosis and outcome. *J. Ophthalmol.* **2017**, 5053961 (2017).
15. H. Dimaras, T. W. Corson, D. Cobrinik, A. White, J. Zhao, F. L. Munier, D. H. Abramson, C. L. Shields, G. L. Chantada, F. Njuguna, B. L. Gallie, Retinoblastoma. *Nat. Rev. Dis. Primers.* **1**, 15021 (2015).
16. D. H. Abramson, K. Beaverson, P. Sangani, R. A. Vora, T. C. Lee, H. M. Hochberg, J. Kirszrot, M. Ranjithan, Screening for retinoblastoma: Presenting signs as prognosticators of patient and ocular survival. *Pediatrics* **112**, 1248–1255 (2003).
17. J. W. Simon, P. Kaw, Commonly missed diagnoses in the childhood eye examination. *Am. Fam. Physician* **64**, 623–628 (2001).
18. A. O. Khan, S. Al-Mesfer, Lack of efficacy of dilated screening for retinoblastoma. *J. Pediatr. Ophthalmol. Strabismus* **42**, 205–210 (2005).
19. C. Rodriguez-Galindo, The basics of retinoblastoma: Back to school. *Pediatr. Blood Cancer* **57**, 1093–1094 (2011).
20. R. I. Chee, R. V. P. Chan, Universal newborn eye screening: An effective strategy to improve ocular health? *Eye* **32**, 50–52 (2018).
21. American Academy of Pediatrics; Section on Ophthalmology; American Association for Pediatric Ophthalmology and Strabismus; American Academy of Ophthalmology; American Association of Certified Orthoptists, Red reflex examination in neonates, infants, and children. *Pediatrics* **122**, 1401–1404 (2008).
22. W. J. Muen, M. Hindocha, M. A. Reddy, The role of education in the promotion of red reflex assessments. *JRSM Short Rep.* **1**, 46 (2010).
23. A. C. Tongue, G. W. Cibis, Brückner test. *Ophthalmology* **88**, 1041–1044 (1981).
24. S. K. Gagan, A. K. Hutchinson, S. R. Lambert, The value of serial personal photographs in timing the onset of unilateral cataracts in children. *J. AAPOS* **13**, 459–462 (2009).
25. T. R. Coker, T. Thomas, P. J. Chung, Does well-child care have a future in pediatrics? *Pediatrics* **131**(suppl. 2), S149–S159 (2013).
26. P. Rivas-Perea, E. Baker, G. Hamerly, B. F. Shaw, Detection of leukocoria using a soft fusion of expert classifiers under non-clinical settings. *BMC Ophthalmol.* **14**, 110 (2014).
27. R. Henning, P. Rivas-Perea, B. Shaw, G. Hamerly, paper presented at the IEEE Southwest Symposium on Image Analysis and Interpretation 2014.
28. P. Thagard, C. Shelley, *Synth Library* (Springer, 1997), vol. 259, pp. 413–427.
29. J. Marshall, G. A. Gole, Unilateral leukocoria in off axis flash photographs of normal eyes. *Am J. Ophthalmol.* **135**, 709–711 (2003).
30. V. M. Asensio-Sánchez, L. Díaz-Cabanas, A. Martín-Prieto, Photoleukocoria with smartphone photographs. *Int. Med. Case. Rep. J.* **11**, 117–119 (2018).
31. D. M. Squirrell, J. F. Talbot, Screening for diabetic retinopathy. *J. R. Soc. Med.* **96**, 273–276 (2003).
32. V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, R. Kim, R. Raman, P. C. Nelson, J. L. Mega, D. R. Webster, Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **316**, 2402–2410 (2016).

33. T. N. Kim, F. Myers, C. Reber, P. J. Loury, P. Loumou, D. Webster, C. Echanique, P. Li, J. R. Davila, R. N. Maamari, N. A. Switz, J. Keenan, M. A. Woodward, Y. M. Paulus, T. Margolis, D. A. Fletcher, A smartphone-based tool for rapid, portable, and automated wide-field retinal imaging. *Transl. Vis. Sci. Technol.* **7**, 21 (2018).
34. D. H. Abramson, A. C. Schefler, K. L. Beaverson, I. S. Rollins, M. S. Ruddat, C. J. Kelly, Rapid growth of retinoblastoma in a premature twin. *Arch. Ophthalmol.* **120**, 1232–1233 (2002).
35. D. J. DerKinderen, J. W. Koten, L. K. J. Van Romunde, N. J. D. Nagelkerke, K. E. W. P. Tan, F. A. Beemer, W. Den Otter, Early diagnosis of bilateral retinoblastoma reduces death and blindness. *Int. J. Cancer* **44**, 35–39 (1989).
36. P. K. Shah, V. Narendran, N. Kalpana, In vivo growth of retinoblastoma in a newborn infant. *Indian J. Ophthalmol.* **58**, 421–423 (2010).
37. A. G. Goddard, J. E. Kingston, J. L. Hungerford, Delay in diagnosis of retinoblastoma: risk factors and treatment outcome. *Br. J. Ophthalmol.* **83**, 1320–1323 (1999).
38. J. Poushter, *Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies* (Pew Research Center, 2016).
39. N. Crisp, L. Chen, Global supply of health professionals. *N. Engl. J. Med.* **370**, 950–957 (2014).
40. J. A. Martin, B. E. Hamilton, M. J. Osterman, A. K. Driscoll, T. J. Mathews, Births: Final data for 2015. *Natl. Vital Stat. Rep.* **66**, 1 (2017).
41. 29–1065 Pediatricians, General (U.S. Bureau of Labor Statistics May 2016).
42. 29–1062 Family and General Practitioners (U.S. Bureau of Labor Statistics, May 2016).
43. Density of physicians (total number per 1000 population, latest available year), (World Health Organization, 2018).
44. E. N. Gichigo, M. M. Kariuki-Wanyoike, K. Kimani, M. M. Nentwich, Retinoblastoma in Kenya: Survival and prognostic factors. *Ophthalmologe* **112**, 255–260 (2015).
45. C. Leander, L. C. Fu, A. Peña, S. C. Howard, C. Rodriguez-Galindo, J. A. Wilimas, R. C. Ribeiro, B. Haik, Impact of an education program on late diagnosis of retinoblastoma in Honduras. *Pediatr. Blood Cancer* **49**, 817–819 (2007).
46. A. Corinthios, Devoted dad creates app that can detect eye cancer in children after his son was diagnosed at 3 months old in PEOPLE Magazine (2016); https://people.com/human-interest/father-creats-smartphone-app-to-detect-eye-cancer-after-his-son-was-diagnosed
47. S. P. Donahue, C. N. Baker; Committee on Practice and Ambulatory Medicine; Section on Ophthalmology; American Academy of Certified Orthoptists; American Association for Pediatric Opthalmology and Strabismus; American Academy of Ophthalmology; Procedures for the Evaluation of the Visual System by Pediatricians. *Pediatrics* **137**, 10.1542/peds.2015-3597 (2016).
48. A. Linn Murphree, Intraocular retinoblastoma: The case for a new group classification. *Ophthalmol. Clin. N. Am.* **18**, 41–53 (2005).
49. B. Haargaard, A. Nystrom, A. Rosensvard, K. Tornqvist, G. Magnusson, The Pediatric Cataract Register (PECARE): Analysis of age at detection of congenital cataract. *Acta Ophthalmol.* **93**, 24–26 (2015).
50. E. E. Birch, J. M. Holmes, The clinical profile of amblyopia in children younger than 3 years of age. *J. AAPOS* **14**, 494–497 (2010).
51. V. D. Castagno, A. G. Fassa, M. L. Carret, M. A. Vilela, R. D. Meucci, Hyperopia: A meta-analysis of prevalence and a review of associated factors among school-aged children. *BMC Ophthalmol.* **14**, 163 (2014).
52. L. Breiman, Bagging Predictors. *Mach. Learn.* **24**, 123–140 (1996).
53. R. S. Adam, P. J. Kertes, W. C. Lam, Observations on the management of Coats' disease: Less is more. *Br. J. Ophthalmol.* **91**, 303–306 (2007).
54. L. Sandfeld Nielsen, L. Skov, H. Jensen, Visual dysfunctions and ocular disorders in children with developmental delay. II. Aspects of refractive errors, strabismus and contrast sensitivity. *Acta Ophthalmol. Scand.* **85**, 419–426 (2007).
55. Committee on Science, Space, and Technology, Research and technology subcommittee hearing - Smarth health: Empowering the future of mobile apps (2016); www.youtube.com/watch?v=dIfMLoYiGgk

Acknowledgments: This manuscript is dedicated to the memory of M. Gottstein, who bravely fought retinoblastoma and treatment-induced leukemia. **Funding:** The NSF (CHE: 1856449; CHE: 1352122) and the Robert A. Welch Foundation (AA-1854) are acknowledged for funding research in the Shaw laboratory. **Author contributions:** B.F.S. and G.H. conceived the project and experiments and supervised the project. All authors discussed the data. B.F.S., G.H., L.D., S.M., M.C.M., C.T.Z., A.T.K., and K.M.B. wrote the manuscript. **Competing interests:** B.F.S., G.H., and R.H. are the inventors of the CRADLE/White Eye Detector Application. The authors declare that they have no other competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 9 April 2019
Accepted 10 September 2019
Published 2 October 2019
10.1126/sciadv.aax6363

Citation: M. C. Munson, D. L. Plewman, K. M. Baumer, R. Henning, C. T. Zahler, A. T. Kietzman, A. A. Beard, S. Mukai, L. Diller, G. Hamerly, B. F. Shaw, Autonomous early detection of eye disease in childhood photographs. *Sci. Adv.* **5**, eaax6363 (2019).