

Use of photography for support of trachoma grading

Progress report

Emma M. Harding-Esch¹

Fahd Naufal²

Martha Saboya³

Cristina Jimenez⁴

Sheila K West²

1- London School of Hygiene & Tropical Medicine

2- Dana Center for Preventive Ophthalmology, Johns Hopkins University

3- Pan American Health Organization

4- SightSavers

TABLE OF CONTENTS

TABLE OF CONTENTS	1
ACRONYMS	2
1. BACKGROUND	3
2. OBJECTIVES	4
3. REVIEW OF PAST AND ONGOING WORK	4
4. OPPORTUNITIES AND CHALLENGES FOR USING PHOTOGRAPHY TO SUPPORT TRAINING AND SUPERVISION OF TRACHOMA GRADERS	6
A. Opportunities	6
B. Challenges	8
5. OPPORTUNITIES AND CHALLENGES OF SUBSTITUTING FIELD GRADING WITH IMAGE CAPTURE AND GRADING IN PREVALENCE SURVEYS	9
A. Opportunities	10
B. Challenges	11
6. PRIORITIZE KEY OPERATIONAL RESEARCH QUESTIONS THAT WILL NEED TO BE ADDRESSED TO ALLOW SUBSTITUTING IMAGE CAPTURE AND GRADING FOR FIELD GRADING	14
7. ACTIONS AND NEXT STEPS TO ADDRESS CHALLENGES AND MAXIMIZE OPPORTUNITIES FOR USING PHOTOGRAPHY FOR TRAINING AND SUPERVISION OF FIELD GRADERS AND FOR REPLACING FIELD GRADERS IN TRACHOMA PREVALENCE SURVEYS	16
8. CONCLUSIONS	20
9. REFERENCES	21
APPENDICES	22
Appendix 1: Participants in one or more of the workshops and subsequent meetings, and Agendas for the workshops	23
Appendix 2: Timeline of workshops and meetings	26
Appendix 3: Summaries of research study presentations from the workshops	27

ACRONYMS

AFRO	(World Health Organisation) Regional Office for Africa
AI	Artificial intelligence
ICAPS	Image Capture and Processing System
JHU	Johns Hopkins University
M&E	Monitoring and evaluation
NTD	Neglected tropical diseases
PAHO	Pan American Health Organisation
R&D	Research and development
SD	Secure digital
SLR	Single-lens reflex
TD	Tropical Data
TF	Trachomatous inflammation—follicular
TI	Trachomatous inflammation—intense
TS	Trachomatous scarring
TT	Trachomatous Trichiasis
UCSF	University of California San Francisco
UNC	University of North Carolina
VA	Visual Acuity
WHO	World Health Organization

1. BACKGROUND

Population-based surveys are the World Health Organization (WHO)-recommended method of estimating prevalence and documenting trachoma reduction and elimination.¹ The prevalence surveys currently require the standardization of field graders using live participants in the field to ensure a high degree of accuracy of detecting trachomatous inflammation—follicular (TF).² As the prevalence of trachoma declines in districts and in countries, it is becoming increasingly difficult to train new field graders and to recertify existing graders with live participants. It is harder to find sufficient numbers of TF cases, forcing trainees to travel further in-country, or even to other, more endemic, countries.^{3,4} This makes grader training and certification, in the short-term, more difficult, expensive, and time-consuming. In the long-term, as an increasing number of countries near and achieve the elimination goal, it will no longer be possible to depend on live TF cases to train and certify graders. In addition, regardless of TF prevalence, there are other situations where the use of trained graders may not be ideal or sustainable: these include areas with limited numbers of graders and high attrition, such as the Americas; conflict zones that need surveys where graders may not be allowed in; and hard to reach areas.

It has been proposed that the use of photography and image grading could potentially overcome these challenges in the future. Photographs could be used more extensively within trainings and as part of supervision during surveys.⁵ Furthermore, the use of photography and image grading as a replacement for TF grading in the field has been proposed, as has been done in numerous research projects and clinical trials.^{6–8} However, there is a lack of standardization and validated protocols outside of research projects for the use of photographs to train graders, to assess quality of field grades, and to serve as an adjunct to, or substitute for, field grading of TF.

To explore the potential use of photography to address these issues and understand more clearly what has and is currently being done, two workshops (July and November 2020) were held as well as follow-up working group meetings. The list of participants and agendas of the workshops are included in [Appendix 1](#), and the timelines of all meetings are presented in [Appendix 2](#). This report presents: the objectives of the workshops; a summary of the review of existing data and experiences with using photography and image grading for trachoma; the outline of opportunities and challenges identified through workshop discussions; key operational research topics proposed; and a series of next steps.

2. OBJECTIVES

The workshops' objectives were:

- 1) To review existing data and experiences with using photography and image grading for trachoma and for other eye health problems.
- 2) To assess the options for use of photographic images and image grading to overcome the challenges of training graders and for supervision and quality assurance.
- 3) To assess the opportunities and challenges of replacing field grading during surveys with photography.
- 4) To identify the key operational research questions to be addressed to allow photographic images to be used for training and replacing field grading.

3. REVIEW OF PAST AND ONGOING WORK

This section summarizes reviews presented in each workshop; presentation details are provided in [Appendix 3](#).

A literature review on photography to support both training and as a substitute for field grading found that when image quality is good, the levels of agreement between image grading and field grading are good for TF and trachomatous inflammation—follicular (TFI). Agreement between trained field graders on the presence or absence of signs in the same children and between trained image graders on the presence or absence of signs in the same good quality images were also reportedly good.

Work from Ethiopia and Tanzania demonstrated that the use of photography and image assessment was feasible within routine monitoring, but judicious selection and adequate training of photographers and image graders was key. First, from an ophthalmic photography perspective, although there are formal training programs for photographers available, most success appeared to be with on-the-job training, with applicants who had “people skills” and some familiarity with

photography. Trainees were taught basic anatomy of the eye, specifics of the camera system, optimal lighting, and settings, and then practiced with one another. Certification of each trainee was based on written test and practicum. Experience specifically in taking trachoma images using the Image Capture and Processing System (ICAPS) system suggested the length of training should be at least two days, with half of a day focused on gaining familiarity with the equipment and procedures. The rest should be focused on training in the field, practicing with children and especially those aged 1-4 years where most of the poor-quality images occur and where high-quality images are most important, since this is the age group that is most likely to have TF. Second, ideally a large cadre of applicants would be screened (experience using a camera system essential) and trained to identify the best candidates, rather than training only a small group of individuals, as skills needed for trachoma grading or data recording may not translate to being a good photographer.

High quality images that allow TF to be graded are essential to the successful use of photography and image grading. Work on defining and quantifying quality used the following metrics: lighting/shadow, glare/tear film, focus, eyelid eversion, blanching, finger obstruction, debris in grading zone. The importance of standardizing the assessment of image quality is necessary for evaluating camera systems and for training using image examples. Preliminary work on specific camera systems that could take high quality images found advantages and disadvantages for SLR (single-lens reflex) camera systems and smartphones. Either system will likely require additional personnel to flip lids (unless a hands-free system is used, or the trained photographer is also the data recorder).

In preliminary work, the community and survey team members seemed to find taking images acceptable. There is ongoing work to investigate the use of crowd sourcing and Artificial Intelligence (AI) to grade images, in addition to a more traditional approach of using reading centers with graders.

4. OPPORTUNITIES AND CHALLENGES FOR USING PHOTOGRAPHY TO SUPPORT TRAINING AND SUPERVISION OF TRACHOMA GRADERS

Key points relating to opportunities and challenges of using photography to support training and supervision from the workshop and subsequent meeting discussions are presented below.

A. Opportunities

Opportunities were identified for using photography to support training and supervision, including the comparative advantages of photos over the need for live cases with TF, existing experience and resources, and different options for photo grading, which is considered under replacing field grading with photography (Table 1).

Table 1. Opportunities to use photography to support training and supervision of trachoma grading.

Main aspect	Opportunities
Training: Statement of need	- Lack of live TF cases is an increasingly recognised problem, there is collaboration and engagement by multiple stakeholders in the trachoma community (as evidenced by these workshops).
Relative ease of implementation into training methods	- Photographs already form part of Tropical Data (TD) prevalence survey training processes for TF; and the trachomatous trichiasis (TT)-only survey training is reliant entirely on photographs for teaching and assessing TT diagnosis. - As this is for training rather than in-service diagnosis purposes, there are no concerns in terms of changing WHO recommendations and guidelines.
Standardization and training	- Comparative advantages of using images (supports standardization, guards against grader drift, ability to train in places regardless of trachoma prevalence).

	- Several groups have experience using photos for the training of graders. Lessons learned can be used to develop standardized methodologies for training. - A large set of graded photos, including clear and borderline cases from different regions and countries that can be shared would be an important support for training.
Share information	- Share information on how standard sets have been tested (and means to update these as time passes), and training procedures, including certification. - Sharing of methods and lessons learnt in developing and implementing the process, to support others and obtain feedback to evaluate and continually improve our methods.
Supervision and auditing: Statement of need	- Photos for supervision are used by some programmes and more formal guidance on appropriate use is needed. (Note that many issues around taking images for supervision purposes are also the same, on a different scale, as those for replacing field graders with photography and are also presented in that section).
Auditing	- Allows auditability if images are taken during the survey, monitoring and evaluation (M&E). - Opportunity for decision-making post-survey, to support supervision and M&E.
Standardization of supervision	- Several groups have experience using photos for supervision. Lessons learned can be used to develop standardized methodologies for supervision.
Share information	- Share list of specifications and minimum requirements for systems to take good quality images, store photos and define optimal photo characteristics. - Sharing methods and lessons learned in developing and implementing processes of supervision to support others and obtain feedback to improve methods.

B. Challenges

The first main challenge identified was the development of a library of appropriate images that can be used for training and standardization; current consensus sets tend to include higher quality photos with clearer diagnosis (TF/not TF). Several groups have existing stores of images, but are not collated in a way that permits easy use. In some cases, ethical clearances may be needed to share images. Other challenges include how to properly assess the suitability of training using images only, and how to implement a supervision component that uses images (Table 2).

Table 2. Challenges to use photography to support training and supervision of trachoma grades

Main aspect	Challenges
Sets of photographic images for standardization and training	- Lack of sets or libraries of photos for standardization and training with both clear and borderline cases, and from multiple countries/regions. Parameters on quality of the images, metadata needed, and ethical clearance are some of the aspects to review to develop such a library.
Photographic image grading consensus	- Need to develop standardized methodologies for assessment of agreement of photographic images, especially for borderline cases.
Standardization and training using photographic images	- There is a need to develop, test, and implement the methodology for standardization and training of graders using photographic images. Currently (2021), there are no comparative studies that evaluate trachoma grading training systems using images versus training using live participants, where there is some “gold standard” to try and determine ground truth.
Acquiring and grading images for supervision of field graders	- Several challenges were identified that were common to acquiring and grading images both for supervision and for using photography in place of field graders for prevalence surveys. These are presented more fully in the next section. - Unique to the task of supervision are the following:

	- Determining the number of images and the selection of children to be photographed (All age groups or just 1–9-year-olds? TF field cases and a sample of non-TF to avoid methodological bias by only verifying positive cases? A random sample? Borderline cases?) - Timely transmission of images to provide feedback, and assuring a supervisor is qualified to grade the images and provide feedback, etc.
Acceptability	- What further research is needed to support training graders to diagnose TF with photos alone (removing the need for cases in the field; as we have done for TT)?

In general, there was consensus that photographic images can help to overcome increasing challenges for training and supervision of trachoma graders. Several projects have used photographic images for trachoma and other eye health problems, some of which were presented in the workshop presentations. However, it was noted that some aspects remain to be addressed, such as: the development of a photo library; establishing a consensus grading method of photographic images; standardization of training methods using photos; and organization, transfer, and upload of photos in different settings for supervision (e.g., internet connection); and identifying options for assessment of photos (grader centers, crowdsourcing, artificial intelligence (AI), etc.) for the photo library. Development of a photo set or library and standardization of consensus grading of photographic images for training were identified as the most urgent tasks.

5. OPPORTUNITIES AND CHALLENGES OF SUBSTITUTING FIELD GRADING WITH IMAGE CAPTURE AND GRADING IN PREVALENCE SURVEYS

Key points relating to opportunities and challenges of using photography to replace field grading from the workshop presentations and discussions are presented below.

A. Opportunities

Several opportunities were identified if photography and image grading were used to replace field grading. These are outlined in Table 3. Many opportunities and issues were common to issues around the use of photography for supervision, but there are unique ones as well. In general, the use of photography is not new and has been used to overcome biases in several research projects, and that experience can be mined to develop acceptable, feasible and high-quality methods for trachoma prevalence surveys in the future.

Table 3. Opportunities with use of image capture and grading in prevalence surveys

Main aspect	Opportunities
Statement of need	- Potentially larger and more diverse pool of personnel that can be trained to take images than are available as field graders. - As prevalence of trachoma declines, in hard-to-reach areas, areas that do not have trachoma graders, photography can be a useful substitute. - Addresses concerns for grader biases especially when disease prevalence is low.
Experience using images	- Several groups have experience using photos for grading trachoma in the field, with tools available, and lessons learned which can be used to help ensure quality of photos being taken. - Metrics of quality image, such as: lighting/shadow, glare/tear film, focus, eyelid eversion, blanching, finger obstruction, debris in grading zone have been developed and can be used to train photographers and assess quality.
Consistent field collection methods for all key trachoma indicators	- Strong correlation between TT photo and field grading. Could include both TF and TT grading in one system (and expand if wished to include TI and trachomatous scarring (TS)).
Assessment of photos	- Several groups have reading center experience or have standardized grading centres. - Opportunities to use new methods to assess images: - Latent class analyses- in assessment of images - Expand crowdsourcing for grading potentially increases the pool of trained graders and reduces the need for

Main aspect	Opportunities
	personnel to be diverted from their usual duties, as well as enhanced public engagement - Long term opportunity for use of AI and automated assessment of images, which might remove the requirement for photographs to be sent to a center.
Opportunities for operational research projects	- Several operational research questions were identified and summarized in the next section, that could be carried out with country partnerships.

B. Challenges

There were several challenges identified with readiness to substitute field grading with photography for program (rather than research) use. These include standards on how to train skilled photographers, standardizing metrics for camera systems, and creating systems for assessment of images. Several of these challenges are also common for the use of photography to supervise trachoma graders. Costs of taking to scale this approach to field surveys must also be considered and balanced against costs of field training and re-standardization. Research to prove equivalency of surveys conducted using image grading and field grading are needed before endorsement of this approach for programs. These challenges are summarized in Table 4.

Table 4. Challenges of substituting image capture and grading for field grading in prevalence surveys

Main aspect	Challenges
Training of photographers	- Methods to choose the right personnel to be trained: - Pre-test of potential skill? Visual Acuity check? Etc. - No standard protocol for training photographers (but groups have experience). - Criteria for determining good quality photos have been developed and tested, but need to be implemented for training.
Survey team	- Composition of team with a photographer is unclear. Can the photographer be the data recorder as well, then just need a lid

Main aspect	Challenges
	flipper? Who does the TT assessment, or can photos be taken for this too?
Standards for camera equipment	- Need minimal standards for camera/attachments to ensure quality images. - Camera systems, especially SLR systems, are expensive (balanced against costs of training, especially remotely). - Many types of phone cameras that could be tested and technology always changes. - Avoiding contamination of equipment with SARS-CoV-2 or other pathogens. - Unclear what is ideal lighting - natural lighting, ring flash, etc.
Assessment of photos	- Need standards for determining gradable and ungradable images. - What is an acceptable rate of ungradable images? - Development of protocols for grading and adjudication. - Determine acceptable delay in getting data if using remote grading. - Image storage in field and upload to a center – need to analyse options in different settings (e.g., access to internet and Wi-Fi in remote areas). - Image compression issues and impact on quality. - Linking field data (like Tropical Data identification barcodes) to images. - Need a method to arrive at consensus grade (on the same photos, as well as if multiple images are taken from the same eye), acknowledging that field graders have variation as well.
Acceptability	- Acceptability of using photographic equipment by communities (some technologies (smartphones) preferred over others (SLR cameras)). - Ethical implications of taking photos as opposed to field grading, for example whether consent procedures need modifying, how images are transferred and stored, how images are linked to underlying demographic data, how long images are stored for,

Main aspect	Challenges
	and who manages individual trachoma cases that are diagnosed remotely (for example, TT referral and tetracycline administration).
Costs and feasibility	- Studies needed on cost, time, and feasibility of instituting photography into surveys. - What is added cost, times for surveys, considering it is balanced against training graders, higher per diems for trained graders?
WHO & international guidelines	- Delineation of data needed to support transition to photo acquisition and grading for surveys. - Does the camera system become a “medical device”? - Differing international norms that need to be considered. - Questions on confidentiality need to be addressed.
Further evidence needed	- What further research is needed to enable replacing graders in the field with image capture and grading (as is done in research settings)? - Determination of correlation with field grading and acceptable limits of uncertainty (what is the gold standard?).
Resources needed and integration	- Need to identify resources needed (time, personnel, supervision, training, cost, scalability). - Work out integration (technological) of photos, data platforms, multiple platforms- including storage and upload capacity; many different systems currently in place. - Being able to adapt and integrate long term with crowd sourcing, Artificial Intelligence etc.

In general, there was consensus that there are significant opportunities for using photos/images for TF grading to replace field graders (with enthusiasm to also do this for TT grading). There is agreement on elements of quality for an image, especially the emphasis on focus, that feeds into training and standards for camera equipment. We need to compile more evidence and reach a consensus on quality parameters and number of images that need to be taken per eye. There are advantages and disadvantages for SLR cameras versus smart phones, with agreement that more research on the newer models is needed. A key point to emphasize is that

there is variation among field graders and among image graders, as well as between field and photo graders, with no clear “gold standard”. More work is needed to determine the relative strengths and weaknesses of each grading approach for use in routine surveys: comparison to only field graders as a gold standard ignores error in measurement in that group as well.

6. PRIORITIZE KEY OPERATIONAL RESEARCH QUESTIONS THAT WILL NEED TO BE ADDRESSED TO ALLOW SUBSTITUTING IMAGE CAPTURE AND GRADING FOR FIELD GRADING

Based on discussions during the workshops and ongoing work, a series of critical operational research questions have emerged whose results are critical for moving the work forward. Further questions will emerge as the ongoing work proceeds, so the list in Table 5 should not be considered all inclusive.

Table 5. Key operational research questions to be addressed prior to any transition to image capture and grading for prevalence surveys

Main aspect	Key operational research
Specifications for camera systems (SLR, Smartphones) to be used for optimal images	- What is comparable in terms of quality images? - What are the minimum set of specifications needed to result in good quality images? -Test equipment using expert photographers to tease out differences between camera system and expertise of photographer. -Definition of quality images and percentage of ungradable images with different systems. - Need data on newer smartphone models, including costs, replacements, repairs against performance. - Would 3D photos be better than 2D photos for TF grading, as they are for TT grading?
Standardized training for photographers	- What are the minimum selection criteria for trainees?

Main aspect	Key operational research
	- What is the optimal training program that yields high quality photographers and minimizes % ungradable images? - How should photographers be certified? - Frequency of re-certification or is monitoring image quality enough? - Costs of training.
Systems for assuring photo quality	- Can photo quality be determined as the image is captured, so the photographer retakes immediately if quality is inadequate?
Image grading	- Evaluation of alternative systems for grading images. - How to determine a final grade for an image, determination of a consensus grading system? - Could a follicle size guide be built into an image grading interface?
Feasibility and cost-effectiveness analyses (consider integrating in all OR studies)	Feasibility and cost, within routine prevalence surveys, of: - field vs image grading - taking 1 vs 2 or more photos. - storage (SD cards). - transferring photos, if not a simple upload. - grading photos, especially hybrid systems. - integration of systems (TD plus photos; upgrade of phones). - system to guarantee good quality images. What is the relative cost (short term (initial outlay), as well as long-term (maintenance/updating)) of substituting technology, training photographers, and grading images?
System scalability	- Imaging and grading system should be scalable. - Photographer training programs and certification.
Comparison of field grading and image grading at scale	- Formal studies comparing photos and field grading. Percentage agreement in determining prevalence of TF in different TF settings using methodology approved by WHO that provides support for transition to image-based system.

Main aspect	Key operational research
Utility of photos in post-elimination surveillance	- Value of images for identifying/confirming recrudescence for post-elimination surveillance purposes. - Opportunities to combine images with other alternative indicators for trachoma (serology and <i>C. trachomatis</i> testing).

7. ACTIONS AND NEXT STEPS TO ADDRESS CHALLENGES AND MAXIMIZE OPPORTUNITES FOR USING PHOTOGRAPHY FOR TRAINING AND SUPERVISION OF FIELD GRADERS AND FOR REPLACING FIELD GRADERS IN TRACHOMA PREVALENCE SURVEYS

Several actions and next steps were identified. There has been progress in addressing the action items from the workshops, and these are shown as of October 2021. The COVID-19 pandemic has slowed the ability of teams to test alternatives in the field, but it is expected that progress will accelerate, depending on funding, as opportunities for surveys move forward and some progress has been made since the workshops. Table 6 outlines the action items identified, contributors, progress, and status.

Table 6. Actions and next steps for using photography for training and supervision of graders and for trachoma prevalence surveys

Action item	Leads and contributors	Progress and status
Building an open-access set of consensus photos that can be used for training purposes, including borderline cases	Tropical Data (lead), Scott Nash (The Carter Center), Chris Brady (University of Vermont), Jeremy Keenan (University of California	- Discussed feasibility and drafted a plan for photo repository. - Completed the first draft of terms and conditions (ongoing). - Developer identified and funding confirmed to compile the database.

	San Francisco (UCSF)), Thomas Lietman (UCSF), Martha Saboya (Pan American Health Organisation (PAHO))	- Created a shared spreadsheet for project repository and photo metadata (completed). - Reviewed the list of metadata associated with photos (ongoing). - Initiated the compilation of information on research projects and countries where photos can come from (ongoing). - Funding applied for: Funds obtained to build the photo repository database, with additional funding opportunities being applied for (ongoing). - Future actions: Set up database; source photos; upload photos; consensus grade photos (linked to action 2 in table); create new IGA sets; write protocol manuscript.
Discuss the best way of determining consensus grades for training purposes	Tropical Data, Scott Nash (The Carter Center), Chris Brady (University of Vermont), Jeremy Keenan (UCSF), Thomas Lietman (UCSF), Martha Saboya (PAHO), Anthony Solomon (WHO)	- Discussed grading scale and analysis for graders certification purposes; short and long-term actions identified (ongoing).
Develop photo-based training system. Discuss the training of graders using photos, to agree on what further evidence/research is	Tropical Data, including Principal Trainers, Anthony Solomon (WHO), Amir Kello (AFRO), Scott Nash	- Discussed how training system would be adapted to be photo-based only; short and long-term actions identified (ongoing).

required, and how to move forward. (To include relevant follow-up with WHO on how to adapt global guidelines as a result.)	(The Carter Center), Jeremy Keenan (UCSF), John Neseman (UCSF), Thomas Lietman (UCSF), Paul Courtright (Sightsavers)	
Compile experience on training photographers for surveys and have metrics to test in operational research. - Criteria for photographers (comfortable with smart phone, prior photography experience, screening tool, visual acuity (VA)) - Certification procedures - Characteristics of the training program	Sheila West, Fahd Naufal, Meraf Wolle (JHU), Chris Brady (University of Vermont), Emily Gower (University of North Carolina (UNC))	- Discussed criteria for photographers, adapting ICAPS training program. Finalization awaits better understanding of camera systems to use (see item 5 below). - Manuscript published on proof of principle for use of photography; agreement between field grading and image grading, and between field graders.⁹ - Manuscript published on use of hands free camera system for field surveys.¹⁰ - Quality metrics for images developed, tested for reproducibility, and submitted for publication.
Collecting data on the optimal characteristics of different photography systems, particularly the smartphones. This will help to have recommendations on best phone systems to field test using expert photographers - Minimum specifications for good quality images,	Jeremy Keenan (UCSF), Sheila West, Fahd Naufal (JHU), Harran Mkocha (Kongwa Trachoma Project)	- Trial of 4 Smart phones available in Africa, Latin America, compared to SLR using expert photographers; to start in December 2021 in Tanzania.

ease of use for photographers, cost. - Robustness, storage.		
Collection of information on current image grading platforms and possibilities of combination systems. This will help to have recommendations on models of platforms to test in operational research with thousands of images of varying quality - Data includes ease of getting images to system, volume of grading assessments that can be handled, cost, experiences, approaches to consensus gradings. - Specifications for photo upload in different settings of internet connectivity; photo compression opportunities on front end.	Christopher Brady (University of Vermont)	Ongoing.
Workshop on use of photography for grading TT - Organize a virtual workshop with appropriate partners on use of photography to grade TT.	Emily Gower (UNC)	To be confirmed.

Set up a directory to outline minimum specs for devices, tools, and technology	Anthony Solomon WHO (Lead), Tropical Data, Fahd Naufal, Sheila West (JHU), Chris Brady (University of Vermont)	To be confirmed.
9. Further research needed on regulations around 'medical devices'	Anthony Solomon WHO (Lead), Tropical Data	Ongoing.

8. CONCLUSIONS

The workshops have identified an urgent need to move forward with development and validation of photography for field surveys for trachoma, from training and standardizing field graders to replacement of field grading altogether. While many opportunities exist, key challenges impeding progress have been identified and operational research that will enable this transition has been outlined. While funding in the current environment for these initiatives is also a substantial challenge, the long-term value is clear, as trachoma prevalence declines and field grading standardization proves more difficult and expensive to assure.

9. REFERENCES

1. World Health Organization. *Design Parameters for Population-Based Trachoma Prevalence Surveys*. Geneva, Switzerland; 2018.
2. Solomon AW, Pavluck AL, Courtright P, et al. The Global Trachoma Mapping Project: Methodology of a 34-Country Population-Based Study. *Ophthalmic Epidemiol*. 2015;22(3):214-225. doi:10.3109/09286586.2015.1037401
3. Stelmach RD, Flueckiger RM, Shutt J, et al. The costs of monitoring trachoma elimination: Impact, surveillance, and trachomatous trichiasis (TT)-only surveys. *Glob Heal Div*. doi:10.1371/journal.pntd.0007605
4. Southisombath K, Sisalerm Sak S, Chansan P, et al. National Trachoma Assessment in the Lao People's Democratic Republic in 2013–2014. *Ophthalmic Epidemiol*. 2016;23(sup1):8-14. doi:10.1080/09286586.2016.1236973
5. Courtright P, MacArthur C, Macleod CK, et al. Training system for trachoma prevalence surveys. *Trop Data Train Syst Trach Preval Surv Int Coalit Trach Control London*. 2019.
6. West SK, Taylor HR. Reliability of photographs for grading trachoma in field studies. *Br J Ophthalmol*. 1990;74(1):12-13. doi:10.1136/bjo.74.1.12
7. Solomon AW, Bowman RJC, Yorston D, et al. Operational evaluation of the use of photographs for grading active trachoma. *Am J Trop Med Hyg*. 2006;74(3):505-508.
8. Prasad BP, Bhatta RC, Chaudhary JSP, et al. Agreement between novice and experienced trachoma graders improves after a single day of didactic training. *Br J Ophthalmol*. 2016;100(6):762-765. doi:10.1136/bjophthalmol-2015-307224
9. Naufal F, West SK, Brady CJ. Utility of photography for trachoma surveys: A systematic review. *Surv Ophthalmol*. 2021. doi:10.1016/j.survophthal.2021.08.005
10. Naufal F, Brady CJ, Wolle MA, et al. Evaluation of photography using head-mounted display technology (ICAPS) for district Trachoma surveys. Wanji S, ed. *PLoS Negl Trop Dis*. 2021;15(11):e0009928. doi:10.1371/JOURNAL.PNTD.0009928

APPENDICES

Appendix 1: Participants in one or more of the workshops and subsequent meetings, and Agendas for the workshops

Workshop Participants

Name	Institution
Aemero Abateneh	Kimma University Ethiopia/Tropical Data
Yilikal Adamu	Ethiopia, Addis Ababa University
Wondu Alemayehu	Ethiopia, Berhan Public Health and Eye Care Consultancy
Tawfik Al-khatib	Yemen, Ministry of Health/Tropical Data
Chris Bradley	Johns Hopkins University
Chris Brady	University of Vermont
Robert Butcher	London School of Hygiene & Tropical Medicine/Tropical Data
Dennis Cain	Johns Hopkins Medicine
Ana Cama	Fred Hollows Foundation/ Tropical Data
Paul Courtright	Sightsavers
Michael Dejene	Sightsavers/Tropical Data
Luca Della Santina	University of California San Francisco
Donal Bisanzio	RTI International
Bilghis Elkhair	Sudan, Ministry of Health/Tropical Data
Emily Gower	University of North Carolina at Chapel Hill
Emma Harding-Esch	London School of Hygiene & Tropical Medicine/Tropical Data
Carlos Hernandez	Eyre and Far North Local Health Network SA
Cristina Jimenez	Sightsavers/Tropical Data
Fikreab Kebede	Ethiopia, Ministry of Health
Jeremy Keenan	University of California San Francisco
Amir Kello	World Health Organization, Regional Office for Africa/ESPEN
Tom Lietman	University of California San Francisco
Patrick Massae	Tropical Data
Mariam Mbofana	Mozambique, Ministerio da Saude
Caleb Mpyet	Sightsavers/Tropical Data
Scott Nash	The Carter Center
Fahd Naufal	Johns Hopkins University
John Neemann	University of California San Francisco
Jeremiah Ngondi	RTI International/Tropical Data
Catie Oldenburg	University of California San Francisco
Martha Saboya	Pan American Health Organization
Fikre Seife	Federal Ministry of Health, Disease Prevention and Control Directorate, Addis Ababa, Ethiopia
Anthony Solomon	World Health Organization
Sandra Talero	Instituto Barraquer de America
Turbe Valerian	University College London
Sheila West	Johns Hopkins University
Cassidy Whitson	Emory University
Dionna Wittberg	University of California San Francisco
Meraf Wolle	Johns Hopkins Medicine

Workshop Agenda: July 2020

Day 1: Thursday 23rd July 2020 / Facilitator & Chair: Martha Saboya (PAHO)

Time (UK)	Session title	Speaker
2.00-2.10	Welcome	Martha Saboya (PAHO)
2.10-2.15	Introductions (name, role & organisation, involvement in trachoma surveys)	All
2.15-2.25	Introduction to the problem	Emma Harding-Esch (LSHTM/Tropical Data)
2.25-2.35	Literature review on photography as an image tool	Fahd Naufal (JHU)
2.35-2.45	iCAPS	Sheila West (JHU)
2.45-2.55	The use of photographic grading for trachoma diagnosis within trachoma impact surveys: A pilot study and introduction to the Gondar Grading Center, Amhara region, Ethiopia	Scott Nash/Cassidy Whitson (TCC)
2.55-3.05	BREAK	
3.05-3.15	Acceptability & feasibility of photography for trachoma diagnosis	Robert Butcher (LSHTM)
3.15-3.25	Using Machine Learning to Identify TT	Emily Gower (UNC)
3.25-3.35	Developing an automated approach to TF grading based on deep learning	Luca Della-Santina (UCSF)
3.35-3.45	How do we grade images once taken?	Chris Brady (JHU)
3.45-4.00	Wrap-up	Martha Saboya (PAHO)

Day 2: Friday 24th July 2020 / Facilitator & Chair: Martha Saboya (PAHO)

Time (UK)	Session title	Speaker
2.00-2.15	Specific questions for speakers	Presenters
2.15-2.25	Summary of key presentations points, review of spreadsheet: https://lshtm-my.sharepoint.com/:x:/g/personal/sdblehar_lshtm_ac_uk/EbmoyAxkRbpDrIS_GJQFZeABPJeKc383P3lrJ6s4gV1Bsg?e=jWWx53	Martha Saboya (PAHO)
2.25-2.30	Comments from participants	All
2.30-2.55	Fun Retro - challenges & opportunities for: - Training - Supervision - Replacing field grading https://funretro.io/publicboard/cnxMydCWfHXx9HGnuhuAmm7yKrC2/b3c43da1-b76e-4870-99fc-611cb6223e36	Groupwork (introduced by Cristina Jimenez)
2.55-3.05	BREAK	
3.05-3.55	Next steps: - Training - Supervision - Replacing field grading	Group discussion
3.55-4.00	Wrap-up & close	Martha Saboya (PAHO)

Workshop Agenda: November 2020

DAY ONE: November 16th (Note that all times are Eastern Standard (EST) Daylight Savings time)

Time (EST)	Session description	Speaker
10:00-10:15	Welcome and brief introduction	Dr Anthony Solomon (WHO) and Dr Martha Saboya (PAHO)
10:15-10:45	<u>The Case for Need:</u> What are the circumstances and contexts where the potential for using photographs with subsequent grading offers distinct advantages over trained field graders? <u>Short term:</u> Conflict Zones? Regions with low TF prevalence (Latin America, Pacific?) <u>Long Term:</u> Surveillance surveys where re-emergence is suspect? <i>Discussion</i>	Dr Martha Saboya (PAHO)
10:45-10:55	Quality Images and the Camera systems that provide them.	Dr Fahd Naufal (JHU)
10:55-11:30	What are the best characteristics and configuration of camera system to achieve the best quality images for grading? <i>Discussion:</i> Does it need to be hands free? Balance of robustness in the field, ease of use in the field, replacement, cost, etc.	Dr Jeremy Keenan (UCSF)
11:30-11:40	BREAK	
11:40-12:25	Assuring quality Grading of Images and processing of data: Presentation: (retinal specialist) <i>Discussion:</i> How would images be transferred to a remote site efficiently? What about areas with internet or bandwidth problems? What is the best way to link the data, say from Tropical Data, to the image and grades of the image? Is a team of trained graders (then how does one obtain a consensus grade?) or crowd sourcing, or using artificial intelligence best for grading TF-or perhaps a combination? What evidence do we have? What is the likely cost of upload and final grading liable to be?	Dr Chris Brady (Retina Specialist, UVM)
12:25-12:30	Quick summary and wrap up for Day One	Dr Martha Saboya (PAHO)

DAY TWO: November 17th (Note that all times are Eastern Standard (EST) Daylight Savings time)

Time (EST)	Session description	Speaker
10:00-10:10	Opening and quick summary of topics from Day One	Dr Martha Saboya (PAHO)
10:10-10:30	<u>The photographer:</u> Essence of training ophthalmic photographers	Dennis Cain (Director, Ophthalmic Photography, JHU)
10:30-10:45	Experience in training for the field using ICAPS. <i>Discussion:</i> What would a training session for photographers look like? What are the qualifications? What would be a certification process? Does a photographer need to be able to flip lids, or/and grade TT, or is a dedicated photographer the best option as a 3rd person and just use a TT grader who also knows how to flip lids?	Dr Sheila West (JHU)
10:45-11:45	Evidence needed to consider replacing trained field graders with photo images and grading. <i>Discussion:</i> Do we need district surveys or are village surveys sufficient? What about TT assessment?	Dr Anthony Solomon (WHO)
11:45-12:00	Concluding summary and action items https://funretro.io/publicboard/9RED114RVZPYoXWZs4xGOPZKUF01/3b90a1f1-8d5e-4044-a229-9c734348fe52	Dr Martha Saboya (PAHO)

Appendix 2: Timeline of workshops and meetings

Event	Date
Trachoma Photography Workshop 1: Future-proofing trachoma clinical diagnosis training and quality assurance during trachoma prevalence surveys	23-24 th July 2020
Working Group Meetings: Open-access set of consensus photos	3 rd November 2020 & 2 nd March 2021
Working Group Meeting: Research/Evidence required to train graders using photos	19 th November 2020
Working Group Meeting: Developing a consensus grading system	21 st June 2021 & 5 th October 2021
Trachoma Photography Workshop 2: Photography and image grading as potential substitutes for TF grading during trachoma prevalence surveys	16-17 th November 2020
Working Group Meetings: Progress update and next steps	26 th April 2021 & 20 th September 2021

Appendix 3: Summaries of research study presentations from the workshops

A summary of the discussions and presentations are described below from the first (July 2020) and second (November 2020) workshops.

First Workshop

a. Introduction to the problem

Dr Harding-Esch outlined that as TF becomes rarer, identifying locations with live TF cases for Tropical Data trainings is becoming more complex and expensive. Survey supervisors should also be able to accurately grade trachoma, but if supervisors cannot be trained, or have their training refreshed, it is not possible to verify field grading accuracy, especially “borderline” cases. The solution must enable standardized, accurate grading to maintain the ability to compare intra- & inter-country prevalences; have quality management processes in place throughout system; be logistically feasible (and cost-effective) to integrate into prevalence surveys at-scale; and be acceptable to all stakeholders.

b. Literature review on photography as an image tool

Dr Naufal outlined studies that have been undertaken using photography to support both training and as a substitute for field grading. When image quality is good, the levels of agreement between image grading and field grading are good for TF and TI. Agreement between field graders of the same children and between image graders of the same images are also reportedly good.

c. Image capture and Processing system (ICAPS): hands free camera system

Dr West outlined the goals, development, testing and preliminary results of the Image Capture and Processing System (ICAPS) head-mounted technology. Selecting and training photographers was key to the successful use of the system.

d. The use of photographic grading for trachoma diagnosis within trachoma impact surveys: A pilot study and introduction to the Gondar Grading Center, Amhara region, Ethiopia

Dr Whitson outlined a pilot study where photos were taken alongside normal field grading and a photography grading center was developed. Photo and field gradings were compared, to provide proof of principle that the use of photography and image assessment is feasible within routine monitoring.

e. Acceptability & feasibility of photography for trachoma diagnosis

Dr Butcher described the acceptability and feasibility of the use of photography as reported from survey staff and communities during a training workshop where photos were taken of individuals in the field during training.

f. Successes and challenges in using Smartphones to document TT

Dr Gower described studies and their experience, where photography has been used to support grading of trichomatous trichiasis (TT) in the field, as well as the training and tools used to support these studies.

g. Developing an automated approach to TF grading based on deep learning

Dr Della-Santina outlined ongoing efforts to use deep learning for trachoma grading, including the benefits and progress.

h. How do we grade images once taken?

Dr Brady described the processes involved in grading images taken during the ICAPS work and the use of crowdsourcing to support the grading of TF, something that has been successfully used for images taken for diabetic retinopathy and glaucoma grading.

Second Workshop

a. The Case for Need of replacing field grading with image capture and grading

Dr Saboya discussed the current requirement that graders be trained, and have training refreshed, which incurs challenges of cost and effort to send graders to areas where trachoma still is present. In addition, in the future when TF prevalence is low, refresher training will be very difficult. Finally, regardless of TF prevalence there are other situations where the use of trained graders may not be ideal or sustainable: these include areas with limited numbers of graders and high attrition, such as The Americas; conflict zones that need surveys where graders may not be allowed in; and hard to reach areas also present challenges. Dr Saboya described how photography and image grading can potentially overcome these challenges in the future.

b. Quality Images and the Camera systems that provide them:

i. A Grading System for assessment of quality of the images

High quality images that allow TF to be graded are essential to the success of the use of photography and image grading and are key for training photographers. Dr Naufal outlined one system at Dana Center that used the following metrics to assess quality, which was graded as high, medium, or low: lighting/shadow, glare/tear film, focus, eyelid eversion, blanching, finger obstruction, debris in grading zone. Dr Naufal stressed the importance of standardizing the assessment of image quality when evaluating camera systems and using examples for training.

ii. Camera characteristics/configuration/logistics

Dr Keenan contrasted SLR cameras, used primarily for research in trachoma, with smartphones, which are just starting to be used. Both require batteries and spares in the field, and both have challenges to integrate with data collection in the field. Both likely require a person who is just a photographer, in addition to a lid flipper and a recorder, although this is testable. Of note, a hands-free system could integrate both lid flipping and photography. Smartphones appear to be more robust, more acceptable to

community members, and easier to replace versus costly repairs for SLR systems (the latter usually must be shipped to specialty shops). The smartphone lens appears easier to smudge, and the images were not as sensitive for detecting TF in one study, although analyses using newer model smart phones is needed. The problem of autofocus in smartphones may also be an issue. The use of a smart phone with attachments (cellscope) or a hands-free system (ICAPS) with proper training may alleviate some problems. Operational research to identify a robust, easy to use system, that with training can take high quality images is needed.

c. Assuring Quality Grading of Images and Processing of Data

Dr Brady discussed the need to design systems for grading and data processing that considers the users and the scale involved. Sending images to reading centers is straightforward: Images captured by a camera system can be manually offloaded onto flash drives or into computers for upload, or an integrated phone-based software system that automatically sends images could be developed. Design challenges with upload and sending via the internet are the size of the images; using ICAPS system, 9 MB images could be compressed on the upload platform prior to sending to 300KB, which reduced the internet bandwidth required to send. Others report success with 100KB JPEGs. Other challenges include linkage of the photo to the field data during a survey; ideally, one phone system for the entire process would be used, but this will likely require newer smartphones than currently available in many areas that use Tropical Data. The number of photos per eye is also a cost issue - some have used two, others three images per eye, with all two or three presented for grading at the same time and one grade returned. Dr Brady described his crowdsourcing platform for grading, where optimally up to 7 minimally trained graders review images and render a grade, which then ranges from all negative for TF to all positive for TF. The cost is \$.06/grader/image or \$0.42 per eye. The reading center in Ethiopia processes close to 1,000 images with two graders per eye, comprising of ophthalmology residents and medical students. The cost is \$.09/grader/image. Ideally, artificial intelligence approaches to grading can be developed, which if installed in the phone, could provide grades as the image was taken. Hybrid approaches could also be considered, especially in low prevalence settings. For example, using crowdsourcing to screen out images without TF or clearly with TF, and expert graders review images when there is uncertainty. Several key operations research questions arose from the discussion.

d. The Photographer: Essence of Training Ophthalmic Photographers

Mr. Cain (director, the Division of Ophthalmic Photography at Wilmer Eye Institute) provided key points to consider in training photographers. First, although there are formal training programs available, he has had most success with on-the-job training, with applicants who have people skills and some familiarity with photography. Trainees are taught the anatomy of the eye, specifics of each camera system, optimal lighting, and settings, and then practice with one another. He certifies each trainee before they can see patients, based on written test and practicum. The group also recommended

screening applicants for near vision acuity and for basic dexterity skills, like pinch and zoom using a smart phone.

e. Experience in Training for the Field using ICAPS

Dr West discussed experiences training trachoma graders to be photographers using the ICAPS system. First, the length of training should be at least two days, with ½ day on gaining familiarity with the equipment and procedures. The rest should be training in the field, practicing with children and especially those ages 1-4 years, where most of the poor-quality images occurred. Second, ideally a large cadre of applicants would be screened (experience using a camera system essential) and trained to identify the best candidates, rather than training only a small group of trachoma graders, as skills needed for trachoma grading may not translate to being a photographer. Finally, a certification process should be developed, including submission of several consecutive good quality images.

f. Evidence Needed to Consider Replacing Trained Field Graders with Photos and Grading from the WHO perspective

Dr Solomon reviewed key components needed to accept using photos and grading in place of field work, as the evidence base must be strong to permit WHO to consider photos and grading as a tool for surveys. Any system must be able to be used at scale with program level application, including the training of personnel. The system should be tested in multiple locales, especially low prevalence areas. Field grading versus photo grading must be assessed, recognizing there is variance in field grading as well. The question of assessment of TT in surveys must be addressed, and the adequacy of images. WHO will be working with the NTD community to develop an R&D blueprint for neglected tropical diseases next year, which will hopefully help leverage funds to answer priority operational research questions going forward.