

Seeing is Believing

Final Report for Technology Project entitled:

Development and pilot testing of an Image Capture and Processing System for Trachoma Control (ICAPS)

January 25, 2020

Project Leaders: Sheila West, PhD
Christopher Brady, MD
Christopher Bradley, PhD
Robert Massof, PhD

BACKGROUND

Problem Statement

The Alliance for the Global Elimination of Trachoma (GET 2020) is progressing towards their goal, with countries conducting surveys for their application to WHO for validation of elimination. The mainstay of determining that districts have achieved their goal are population-based prevalence surveys, which measure program impact and confirm that re-emergence has not occurred (pre-validation or surveillance surveys). These surveys require assessment by standardized graders of trachoma signs in the upper conjunctiva. Preliminary estimates suggest that a minimum of 2,560 such surveys will be required between now and global elimination of trachoma.

However, the low rates of trachoma have created a significant problem with obtaining and standardizing trachoma graders, as there are rapidly becoming few places with enough trachoma to undertake training to ensure valid results. As new countries map potential trachoma areas, the grader trainees are traveling farther away to access the remaining areas with enough trachoma for training. Where trachoma may have been gone for years, graders may not be available or may be out of practice, and need re-standardization. Overall, there are insufficient funds to ensure quality standardization for such an effort. Yet WHO requires high-quality data for validation, so an innovative solution for this problem must be found.

Project Statement

We propose to develop and test an image capture and processing system that uses smartphone technology, cloud storage and crowdsourcing to acquire good quality images of the upper conjunctiva that can be stored and graded in a virtual reading center for signs of trachoma. This system could ease the burden of relying on quality trained trachoma graders and ensure that the thousands of planned trachoma surveys in low prevalence areas can be done.

OBJECTIVES and RESULTS

There are three key project objectives, and the results of the development and testing described below

1. Development of the image acquisition system for ICAPS.

We have completed the development of the hands free image acquisition system for the Samsung and VR headgear hardware set. This involved creation of a headmounted unit that would do the following:

1. Allow the operator to visualize her hands to enable flipping the lid
2. Present a magnified image to the operator to enable detection of when the image was in focus
3. Assign bar codes to the image of interest
4. Signal when the image was to be taken to the shutter embedded in a foot pedal.
5. Design and manufacture a durable foot pedal which would sync with the headmounted display.

The System was tested in multiple field trials, detailed in each of the progress reports, and summarized below:

We tested over 42 eyes during the development stage, to be certain the hands free component worked, and that the focus was appropriate. The first system test was carried out in 6 eyes of 4 children, and all but one image was acceptable. We tested 10 eyes of five subjects, and all but one image was acceptable for grading. In this basis, we undertook the first field test in Tanzania

Field test 1: 60 children were enrolled, and after 31 children, when we had unacceptable images for 30% of eyes, the field test was halted. The field test identified a problem with the magnified image in the headset that did not allow ideal determination of focus. This was not evident at Wilmer, but a problem for the field graders. This problem was resolved after further development.

Second Wilmer test: 41 eyes of 22 subjects were imaged, and all were imaged successfully. The ophthalmologist who did the image acquisition approved the new magnification window, and we proceeded to the second field test.

Field Test 2: 60 children were enrolled in the field test, and all 60 children went through the test. A single grader graded the eyelid for trachoma, and took an image using the ICAPS system. The field grade was considered the gold standard. The grader then independently graded the image and the grades were compared. As noted, the kappa was not above 0.7 and 50%-60% of the TF cases were missed.

Grader 1 field grade versus Grader 1 grade for ICAPS image, Right Eye*

	Grader 1 Image grade: 0	Grader 1: TF	Total
Grader 1: Field grade=0	48	0	48
TF	6	4	10
Total	54	4	58

Kappa=0.52 (0.21-0.84) *Two images were ungradeable

Reviewing the images, and compared to the images taken with our research camera system, the primary problem was the resolution of the images to enable follicles to be graded, a problem not detected during testing at Wilmer because there is no trachoma in Baltimore. We worked to develop an optimization between size of the image to send to the Virtual Reading Center, and resolution, and maximized the size at 2.5 MB. The application was modified and again underwent a survey in Tanzania,

Field Test 3: The study ophthalmologist was also part of this final field test. Following the same protocol, we enrolled 60 children in the test. Agreement between the ophthalmologist and Tanzania field grader on the field grades for trachoma was excellent, kappa=0.78. However, when both graded the ICAPS

images and compared the grades to the field grades, the agreement scores improved but remained below 0.70 due to continued difficulties grading follicles in the images. (Kappa=0.62 and 0.58). Between 45% and 50% of the cases of follicles were missed on images. At this point, the app developers migrated to a different phone system where a larger size image can be downloaded more quickly. This was the last modification that could be made to the system without an entire re-design. The modification worked, 100% of images were gradeable and deemed of sufficient quality to proceed to the field trial.

2. Development of the image processing system for ICAPS that encapsulates storage, delivery to the Virtual Reading Center, and grading images.

We have also completed the development and testing of the Virtual Reading Center. Prior to the field test, we deployed of the test set of images, obtaining satisfactory agreement for the crowd source graders. A system was to allow for local compression of images for transmission to a cloud-server to minimize the need for broadband internet connections. For crowdsourcing validation, images (n=47) were grades as for TF status (normal, TF, or abnormal without TF). The images were then posted to Amazon Mechanical Turk (AMT) for grading by untrained volunteers as either "TF" or "not-TF." Each image was graded by 10 unique Turkers in all trials for \$0.10 per image. The mode of AMT grades for each image was compared to the adjudicated expert grade to determine AMT grading accuracy as well as sensitivity and specificity for detecting TF. To date, over 700 images have been successfully uploaded using our system in 2 different district surveys in Tanzania. Compressing images locally allowed for a 95% reduction in file size with attendant reductions in upload times. All normal images (n=16) were correctly identified, and only 1 of the TF images (n=11) was graded incorrectly. 4 abnormal images without TF (n=20) were graded as "TF." Overall sensitivity was 91% and specificity was 83%. Area under the ROC was 0.87. With the expansion of the image size, it is clear that a local broadband network will be needed to avoid long delays in transmission. This is discussed further in the section under the field test.

This component was added to the manual of operations for ICAPS, and has been tested by Dr Solomon at World Health Organization who signed in as a master grader and also graded images for the field test.

3. Field test the ICAPS system in conjunction with a district survey.

Several groups were interested in having us field test the ICAPS system in upcoming district surveys. The Carter Center was interested, but because they'd not use Tropical Data, we could not integrate our findings as we wished. Project ORBIS was interested, but the surveys in Ethiopia were in areas deemed unsafe for our field team who would be doing the training and on site supervision. The Tanzania National Program was interested and we worked with them to select Chamwino North as the district to test ICAPS. Below we describe highlights of the field test:

1. TRAINING: The training for ICAPS was one full day: a morning of orientation to the system, and an afternoon of taking images in a village for certification (See Figure 1 for an image during training). The National Program planned to use three field graders in Chamwino district, so we provided 3 trainers and each trainer supervised one field team for the test Four trachoma graders were assigned for training, one was unable to understand the system and the other three graduated with 5 consecutive photos of good quality each. The post training survey indicated that the graders were happy with the system but wanted a second day of practice using the system in the field.



Figure 1: Image taken during training using ICAPS system

2. SURVEY: 30 villages were randomly selected and 50 children ages 1-9 were randomly selected by the National Program for the survey in Chamwino district. (see Figure 2 for an image of the survey) Due to an unforeseen “holiday” in Tanzania, our team was unable to participate in the final day of the survey as they had flights home and had to take the ICAPS systems with them. Therefore, the data are based on the days of ICAPS participation. Several field issues developed during the course of the survey:

a. The phones and their extra battery had to be charged each night, which necessitated staying in locations with electricity. In the future, extra batteries may be needed to accommodate surveys where electricity is not available.

b. Despite worry about overheating, the phones did not overheat except on one afternoon in one system. As long as the grader removes the headset between households, this does not appear to be an issue. In the one instance, several houses were grouped together providing no respite for the phone and the phone did heat up. We were unable to collect images for 11 children who were surveyed while the system was down. In future surveys, allowing respite for the system will alleviate the problem.

c. The bar codes continue to be an issue. Placing bar codes on the children results in loss when they pull them off. The graders often neglected to check the bar codes, and we have some images with bar codes that do not match what was scanned by Tropical Data. Our images always have bar codes, but Tropical Data phones do not have high quality bar code scanners and if they are not checked, could result in loss of fidelity. In some cases we were able to infer the match. In the future, training will include careful assessment of the match in Tropical Data as well as in our ICAPS system.

d. the bandwidth for transmission of the images to the reading center was not great enough in the local environment. One of our trainers had to use his 4G data plan with hotspot, and bought a \$20 worth of 12gigabytes of data to transmit, which was sufficient. It still took several hours every night to transmit over 300 images taken each day (more if multiple images per eye were taken).

e. The initial plan to have 10 graders crowd source grade each of the images was not feasible for a district survey. This is a cost of \$1 per image, and with a minimum of 3,000 images, is not feasible for scale up. It is also a volume that –if multiple surveys world-wide are ongoing– would quickly overwhelm the capacity of the reading center. We are using the district survey to estimate the minimum number of readers per image, but it is likely that multiple reading centers will be involved in a world wide scale up.



Figure 2: Image of use of ICAPS during the survey

3. RESULTS

There were 3014 images uploaded for the survey. Of these, 21.9% were ungradeable. The two primary problems were the darkness of the image and the movement of the child resulting in a blurred image. All of the ungradable images were of eyes that were judged in the field to be no trachoma, so there is no evidence that there are preferentially missing eyes with disease. A total of 10% of persons were not able to be graded due to ungradeable images, as we were able to use the other eye if it had gradeable images. This is clearly a training issue, not a systems issue, as most images were gradeable, and speaks to the need to train the graders not to save images that are not of high quality.

The prevalence of trachoma according to the field grade (excluding those with ungradeable images) was 6.7% and according to the image grade was 6.7%, (95% confidence interval=2.8-12.5). The kappa value for agreement was 0.65 (95% CI=.56-.74), which was less than the 0.7 hoped for, but still rated as evidence for substantial agreement. There was no evidence of bias in grading, with roughly equal numbers of those graded no trachoma in the field, but trachoma present on images, and those graded trachoma in the field and not on the images.

PUBLICIZING ICAPS

The development of ICAPS was presented at the Association for Research in Vision and Ophthalmology in 2019, and two other presentations (of the field survey results and of the results of testing the Virtual Reading Center) are accepted for presentation at the same meeting in 2020. There is a conference on Global Health in Washington DC in March, and a presentation on the development and testing of ICAPS has been selected for that meeting. Two manuscripts are in progress on this system.

Of note, the COR-NTD meeting, which is a major meeting of NTD funders, program managers, and researchers was held in November 2019. There is a section on Innovation, and ICAPS was selected among many applicants for presentation. We had a table and a 2 hour session to present and enable participants to try out our system. As a result of the session at present, there is interest from Fred Hollows Foundation in using the system in Australia and we are in the process of sending a prototype to try there. The program manager for Malawi, Dr Kumbo, together with SightSavers are planning to use the system in surveys in Malawi. The Carter Center has approached us for training and use of the system for surveys in Sudan, where trained graders are unable to travel due to political restrictions. Finally, PAHO is interested in using the system in Brazil where several surveys are needed to determine the magnitude of trachoma in the country.

Together with Dr Anthony Solomon, we are planning to cohost a meeting on Image acquisition in connection with Tropical Data. The meeting will likely be held in advance of the 2020 WHO GET2020 meeting, the last week of April. We will also present the results of the ICAPS system testing at the Trachoma Scientific Informal Workshop which is also held in connection with GET2020 meeting.

NEXT STEPS

We anticipate the development of more demand for our system than we can support at present. While we can provide training and initial supervision, we cannot provide sufficient ICAPS systems for country needs. We have stockpiled 30 headsets and are in negotiations to create the foot pedals, but the outlay for the phone systems and back up technical support is outside of our current finances. We are hoping to reach out to Samsung through other contacts to see of their interest in donation of phones that are no longer their newest product. We estimate \$1million will be needed to support creation of sufficient systems and three years of technical support and back up. Because we do not want to price this system beyond what programs can afford-which we estimate at \$500 per system plus \$3000 for a training session of up to 5 persons-we will need partners going forward. As Seeing is Believing has closed their Foundation, we will not be able to apply for a scale up grant to attract other partners, so we are currently re-thinking our scale up strategy.