International Journal of Infectious Diseases 116 (2022) 101–107

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

Optimizing cluster survey designs for estimating trachomatous inflammation–follicular within trachoma control programs

Julia W. Gallini[a], Eshetu Sata[b], Mulat Zerihun[b], Berhanu Melak[b], Mahteme Haile[c], Taye Zeru[c], Demelash Gessese[b], Zebene Ayele[b], Zerihun Tadesse[b], E. Kelly Callahan[d], Scott D. Nash[d,*], Paul S. Weiss[a]

[a] Department of Biostatistics and Bioinformatics, Emory University, Atlanta, Georgia, USA
[b] Trachoma Control Program, The Carter Center, Addis Ababa, Ethiopia
[c] Amhara Public Health Institute, Bahir Dar, Amhara, Ethiopia
[d] Trachoma Control Program, The Carter Center, Atlanta, Georgia, USA

ARTICLE INFO

Article history:
Received 23 September 2021
Revised 21 December 2021
Accepted 22 December 2021

Keywords:
trachoma
surveys
costs and cost analysis
Ethiopia

ABSTRACT

Objectives: The World Health Organization recommends mass drug administration (MDA) with azithromycin to eliminate trachoma as a public health problem. MDA decisions are based on prevalence estimates from two-stage cluster surveys. There is a need to mathematically evaluate current trachoma survey designs. Our study aimed to characterize the effects of the number of units sampled on the precision and cost of trachomatous inflammation–follicular (TF) estimates.

Methods: A population of 30 districts was simulated to represent the breadth of possible TF distributions in Amhara, Ethiopia. Samples of varying numbers of clusters (14–34) and households (10–60) were selected. Sampling schemes were evaluated based on precision, proportion of incorrect and low MDA decisions made, and estimated cost.

Results: The number of clusters sampled had a greater impact on precision than the number of households. The most efficient scheme depended on the underlying TF prevalence in a district. For lower prevalence areas (< 10%) the most cost-efficient design (providing adequate precision while minimizing cost) was 20 clusters of 20–30 households. For higher prevalence areas (> 10%), the most efficient design was 15–20 clusters of 20–30 households.

Conclusions: For longer-running programs, using context-specific survey designs would allow for practical precision while reducing survey costs. Sampling 15 clusters of 20–30 households in suspected moderate-to-high prevalence districts and 20 clusters of 20–30 households in districts suspected to be near the 5% threshold appears to be a balanced approach.

© 2021 The Authors. Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

INTRODUCTION

In countries where trachoma remains endemic, prevalence is monitored through periodic surveys in a random sample of villages. These surveys generate evidence for how long trachoma interventions, including mass drug administration (MDA) with azithromycin, are needed in each surveyed district. The World Health Organization (WHO) MDA guidelines recommend five annual rounds of MDA if the trachomatous inflammation–follicular (TF) prevalence among children aged 1–9 years in a district is ≥ 30%, three rounds for a prevalence of 10–29.9%, and one round for a prevalence of 5–9.9% (Solomon, 2016). A key threshold for eliminating trachoma as a public health problem is a district prevalence of TF < 5% among children aged 1–9 years.

Current survey sampling schemes for determining the district-level TF prevalence are relatively similar across trachoma-endemic countries, and are often based on published survey design recommendations (Solomon et al., 2018). The currently recommended sampling design is a two-stage cluster design whereby approximately 20–30 villages (clusters) are selected in the first stage of sampling, and approximately 25–30 households are selected within each cluster in the second stage (WHO, 2014). The design for surveys measuring the impact of trachoma interventions is based on an assumption of a 4% TF prevalence with ± 2% precision

* Corresponding author.
E-mail address: scott.nash@cartercenter.org (S.D. Nash).

https://doi.org/10.1016/j.ijid.2021.12.355
1201-9712/© 2021 The Authors. Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

among children aged 1–9 years (Solomon et al., 2018). In reality, many programs are surveying 25–30 clusters to reach the required number of children (Sata et al., 2021; Sanders et al., 2019). While analysis of optimal survey designs has been performed for other neglected tropical diseases (NTD) (Flueckiger et al., 2017; Knowles et al., 2017; Sturrock et al., 2010; Weiss et al., 2021), there have been fewer formal analyses conducted on trachoma survey design (Flueckiger et al., 2017; Mcleod et al., 2020) .

Although determining the district-level prevalence of trachoma is important because it allows for effective targeting of interventions, money spent by control programs on surveys is money unavailable for interventions such as MDA and surgical services. A recent study evaluating trachoma survey costs within the Amhara region of Ethiopia reported a median cost per survey cluster of $752, and thus $15,040 for a 20-cluster district-level survey and $22,560 for a 30-cluster survey (Slaven et al., 2020). Considering that Ethiopia had 673 trachoma-endemic districts in 2020, the cost of conducting country-wide surveys to gauge prevalence rates over the next 10 years will be substantial for the country's trachoma control program — ranging from $10–30 million if we extend these district-level calculations to the whole country (WHO, 2020b). Recent work has demonstrated the importance and cost-effectiveness of monitoring progress towards elimination as a public health problem globally (Solomon et al., 2020). In settings like Ethiopia, however, many districts are not reaching the elimination threshold at impact survey, with some districts requiring three or more rounds of surveys and 12 or more years of annual MDA before potentially reaching the threshold (Sata et al., 2021; Stewart et al., 2019). In the early phases of trachoma control programs, a 'one-size-fits-all' approach is warranted, since little is known about the trachoma burden in the area. However, as programs run longer, and serial survey data become available, a data-driven approach to guide sampling strategies allows for increased efficiency with respect to cost and time.

The objective of our study was to characterize the effect that varying the number of selected clusters and households has on the precision of TF estimates relative to the WHO-recommended MDA decision cut-points within a region with one of the highest burdens of trachoma. Additionally, this study aimed to analyze the cost-efficiency of various cluster sampling schemes.

METHODS

Ethical considerations

Survey protocols used in Amhara were approved by the Emory University Institutional Review Board (protocol 079-2006) and the Amhara Regional Health Bureau. Survey protocols are also reviewed by the Tropical Data Service (GTMP, 2021).

Survey methodology

In the first sampling stage of a district-level trachoma survey in Ethiopia, approximately 30 clusters (villages) are either selected from an enumerated list by simple random sample (SRS) or by using a method where selection probabilities are proportional to estimated village size (PPES). In the second sampling stage, one segment (parts of villages that are geographically close) is randomly selected from each sampled cluster. These segments normally consist of 25–30 households (approximately the number of households that a field team can reach in one day) (Missamou et al., 2018). In a typical survey, trachoma graders evaluate every individual aged $\geq$ 1 year old for TF, trachomatous inflammation–intense (TI), and trachomatous trichiasis (TT) (WHO, 2010).

Generation of simulated population dataset

To compare the precision and cost of different survey sampling schemes, a population database was created using SAS 9.4 (SAS Institute, Cary, NC, USA) simulating the population of the Amhara region, Ethiopia. The population TF distribution was characterized using 17 empirical surveys conducted by the Amhara Trachoma Control Program in 2017, using the Tropical Data system (https://tropicaldata.org). Our simulated dataset was based on the population observed in these 17 districts, with 30 districts created to represent the breadth of possible TF distributions within a setting such as Ethiopia. The simulated database represented individuals aged 1–9 years with their randomly assigned TF status, as this analysis focused only on the TF indicator. For 24 districts the default prevalence was set between 8% and 40%; for six districts it was set at 0–5%. These values represented probable district prevalence rates in Amhara, as of 2017 (Nash et al., 2018; Sata et al., 2021; Stewart et al., 2019) . The SAS macro written to create the dataset (found at https://github.com/jgallini/trachoma-survey-sample-macros) can be used to recreate the analysis performed in this study, and can be adapted to simulate different populations in Amhara or other areas.

Cluster and household sampling

Simulated samples were drawn with 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, and 34 clusters per district. This distribution of selected clusters was selected to observe the trends in prevalence precision relative to a specified default value of 30. In the second stage, one segment (30 households exactly for this simulation) was selected and all children aged 1–9 were used in the sample. TF prevalence estimates weighted for population were calculated for each of the 30 districts for all samples.

To analyze the effect of the number of second-stage units selected (households), the segment structure was temporarily ignored. Samples were drawn using 15, 20, 25, and 30 clusters in the first stage. Iterations for the second stage were 10, 20, 30, 40, 50, and 60 households, resulting in 24 possible sampling schemes. According to previous surveys carried out in Amhara, the average survey team can evaluate about 30 households in 1 day (Stewart et al., 2019). This analysis did not explore beyond 2 days of surveying (60 households), as any longer was deemed unrealistic.

Survey cost and cost waste

The cost of sampling a cluster is comprised of a fixed cost (the average cost of getting to a cluster) and a variable cost (based on the number of secondary units selected). To calculate an estimated cost for each sample design, two components were needed: the cost of measuring one household, and the cost of measuring a cluster (aside from the cost of measuring households within the cluster). Our basis for these two cost component estimates came from work similar to that conducted by Slaven et al. (2020). Based on our findings, the following function was derived to calculate overall cost (USD) of a given sampling design:

$$cost = 479 * clusters + 29 * households + 242(1 + I_{>30\ households})$$

The notation $I_{>30\ households}$ represents an indicator variable for sampling beyond the first 30 households in the segment.

Aside from the raw cost of each sampling scheme, a novel metric called cost waste was also developed to determine the cost efficiency of each sampling scheme relative to the efficiency of the TF prevalence estimate. Cost waste can be thought of as the funding used inefficiently toward the prevalence estimate relative to a simple random sample. Cost waste was calculated using the percentage of the sample efficiently used (calculated as the inverse

of the design effect multiplied by 100: $p_{used} = \frac{100}{D_{eff}}$) and the cost estimates derived above for each sampling scheme, using the following formula: $waste = cost * (1 - p_{used})$.

Data analysis

For all sampling schemes, the following metrics were used for analysis: the width of the 95% uncertainty interval (UI), the proportion of incorrect MDA decisions made relative to the WHO MDA decision cut-points, the proportion of low MDA decisions made, and the total cost wasted per sampling scheme.

To determine the relative precision of prevalence estimates, bootstrapping techniques were used. Bootstrapping (in this case with 1000 replicates) allows for estimation of parameters like the mean and variance, and consequently the calculation of empirical confidence (or uncertainty) intervals (UIs) (Gelman and Greenland, 2019). The widths of the intervals were compared across sampling schemes as the metric for comparing precision, and were examined relative to the number of clusters selected.

To compare sampling schemes with one another relative to MDA guidelines, the proportion of the 1000 samples that resulted in an incorrect MDA decision relative to the true district-level prevalence was calculated. The proportion of the 1000 samples in which a low incorrect MDA decision was made (i.e. one additional round of MDA instead of three rounds) was calculated for each sampling scheme as an additional metric. The last metric calculated for each design was cost waste, mathematically defined above. Figures were created using both SAS 9.4 and R 4.0.

RESULTS

Cluster level

When a true TF prevalence of 0.38 was assumed (Figure 1a), and when drawing between 16 and 18 clusters, the lower bound of the 95% UI crossed the MDA cut-point of 0.3 (below 0.3 warrants three rounds of MDA; above 0.3 warrants five rounds of MDA). In this theoretical district, if at least 18 clusters were sampled, 95% of samples would result in the correct decision for MDA according to the WHO guidelines. With an assumed prevalence rate of 0.31 (Figure 1b), it was observed that regardless of the number of clusters sampled, the incorrect decision (three rounds of MDA instead of five rounds of MDA) was possible given how close the true prevalence was to the MDA cut-point of 0.3.

At a true prevalence of about 0.20 (Figure 1c), the correct decision is made 95% of the time regardless of the number of clusters sampled. With an assumed true district prevalence of 0.08 (Figure 1d), one of three MDA decisions could be made depending on the sample: three rounds of MDA, one round of MDA (the correct decision according to the WHO guidelines), or the district moving into surveillance while pausing MDA programs. Regardless of the number of clusters sampled, any of these three decisions was possible, with a roughly 10% chance of incorrectly concluding that the district fell below the 5% threshold at both 20 and 30 clusters, and a 1% chance of this happening in two subsequent surveys (i.e. an impact and then a surveillance survey). With a district prevalence of 0.04 (Figure 1e), there was a risk of making an incorrect MDA decision regardless of the number of clusters sampled. Finally, when a district has a very low prevalence (0.01), the correct decision to terminate MDA programs will be made with 95% of samples, regardless of how many clusters are sampled.

Proportions of incorrect and low decisions

The proportion of incorrect MDA decisions out of the 1000 samples peaked around the treatment decision cut-points (5%, 10%, 30%), which is expected given that no amount of precision (unless the entire population is selected) will yield a consistently correct decision when the true prevalence itself is on the borderline (Figure 2). The proportion of wrong decisions dropped to its lowest point among all cluster levels midway between the treatment decision cut-points. When examining the proportion of low MDA decisions (Figure 3), peaks were observed near the treatment decision cut-points, with the proportion of wrong decisions dropping off between cut-points. For example, the probability of incorrectly estimating that a district was below the 5% threshold when in fact the district was in the 5–10% range hovered around 20%, regardless of the number of clusters sampled (Figure 3). Furthermore, the probability of this happening in two consecutive surveys (i.e., an impact survey followed by a surveillance survey) was this value squared, or 4%.

Household level

The second stage of sampling involved selecting either 10, 20, 30, 40, 50, or 60 households. When holding the number of clusters constant at 30 and varying the number of households selected, a similar trend to the cluster level results was observed (Figure 4). There was little separation across the numbers of households selected both for incorrect decisions made and low decisions made (Figures 4 and 5).

Cost waste

The cost waste per surveyed district was calculated for sampling schemes with all possible combinations of 15, 20, 25, and 30 clusters, and 10, 20, 30, 40, 50, and 60 households. Our simulations consistently reflected that the highest cost waste relative to a simple random sample of the same size was observed when more clusters were selected, regardless of the true TF prevalence in a district (Figure 6). For example, a 20 cluster by 30 household sample in a district with a true TF prevalence of 0.38 was estimated to waste $10,201, whereas a 15 cluster by 30 household sample was estimated to waste $7,826.

DISCUSSION

District-level TF prevalence plays a large role in the ability to obtain precise prevalence estimates, so prior knowledge of said prevalence should inform sample design decisions. Overall, it appeared that sampling 30 clusters per district did not achieve adequate precision to justify the cost wasted in many scenarios, particularly around treatment thresholds. In Ethiopia, where hundreds of district-level surveys will be required over the next 5–10 years, it is important for survey methodology to be an adaptive, data-driven process to best meet the needs of the country at a given time. Sampling 15 clusters of 20–30 households in suspected moderate-to-high prevalence districts and 20 clusters of 20–30 households in districts suspected to be near the 5% threshold appears to be a balanced approach. While this number of clusters will result in fewer children sampled, the increase from 20 to 30 clusters does not guarantee a level of precision that is worth the cost. Future operational research into alternative survey approaches would be useful for trachoma programs, and should be conducted with a focus on cost and sustainability (Weiss et al., 2021; Andrade-Pacheco et al., 2020).

Despite progress, there remain areas with persistently high levels of trachoma (Sata et al., 2021). For these districts, repeat surveys will likely be needed in the long term. Our models found that in districts that warranted five rounds of MDA before resurveying (> 30% prevalence), the 95% UIs remained as wide as 20%, even

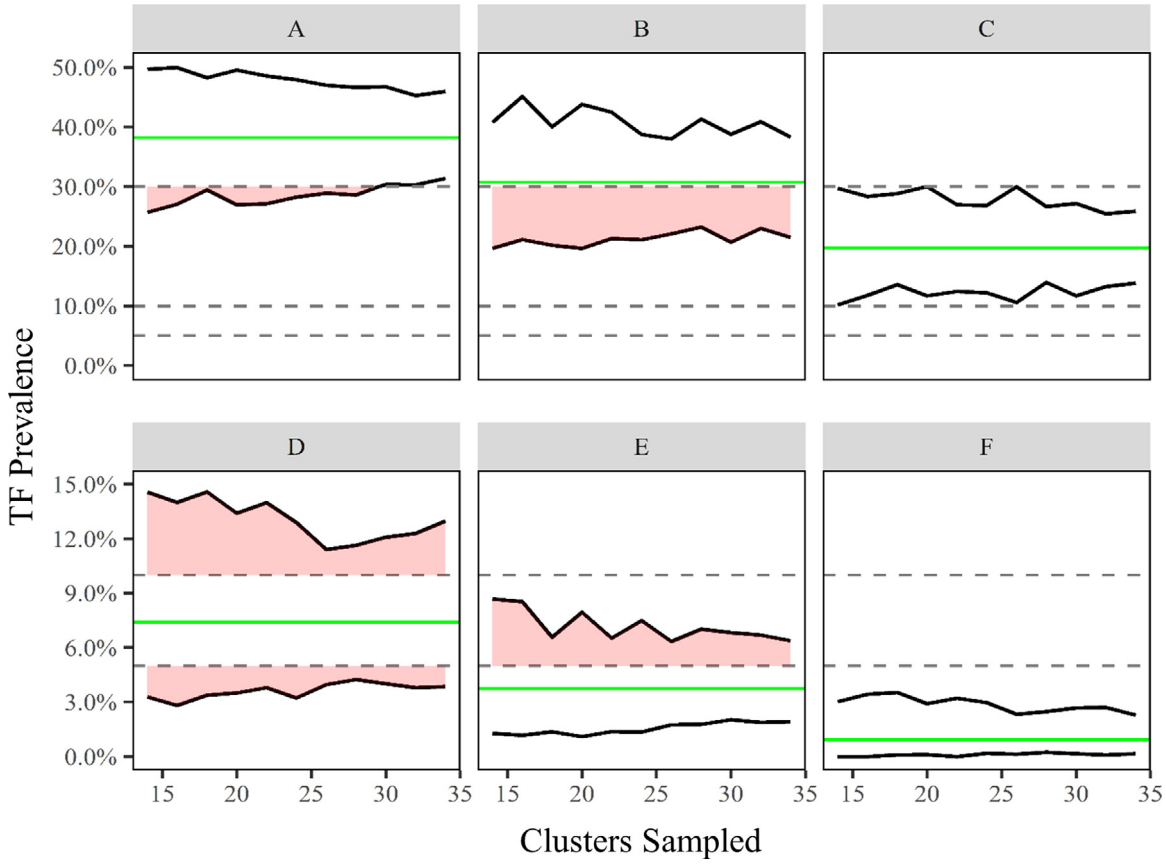

Figure 1. Precision of TF prevalence estimates by clusters sampled. True prevalence in order of panels: 38%, 31%, 20%, 8%, 4%, 1%. Regions of probability where an incorrect MDA decision is made relative to WHO MDA thresholds are shaded red. Green line: 'true prevalence', bold lines: 95% uncertainty intervals; dashed lines: WHO MDA cut-points.

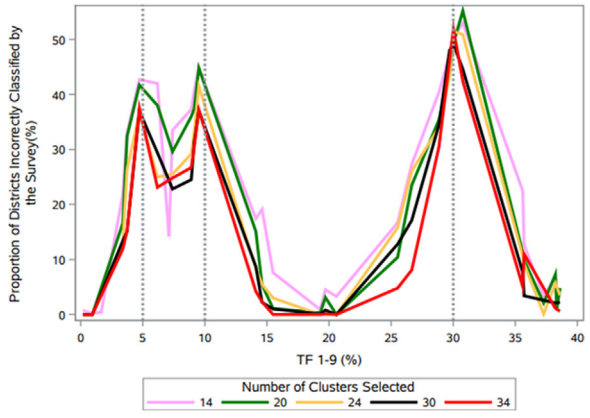

Figure 2. Proportions of wrong MDA decisions by cluster number and true prevalence. Peak probabilities of an incorrect decision occur at the MDA decision thresholds (vertical gray dashed lines). The lowest probabilities of an incorrect decision are found between two MDA thresholds.

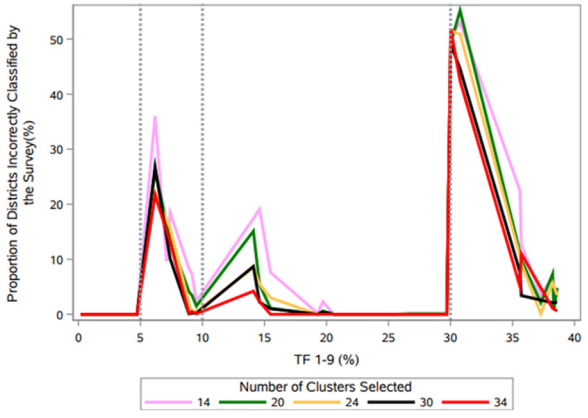

Figure 3. Proportions of low MDA decisions by cluster number and true prevalence. Peak probabilities of a low MDA decision occur at the MDA treatment thresholds (vertical gray dashed lines). The lowest probabilities of a low decision are found between two MDA thresholds.

when sampling an unrealistically high 34 clusters. While this observation may be initially frustrating to trachoma teams attempting to achieve 2% precision in all districts (Solomon et al., 2018), in practice, 2% precision is unnecessary with a TF prevalence as high as 35%, since the difference between 30% and 40% TF prevalence is programmatically negligible: both districts will need MDA for many years, as indicated by past trends (Ngondi et al., 2008; Sata et al,. 2021). Designing surveys for that level of precision in that setting is unrealistic to achieve, and will not substantially improve the quality of treatment decisions.

Assuming an a-priori TF prevalence of 4% in suspected high- and moderate-prevalence districts is probably unrealistic given the longitudinal trends in much of Amhara. Programs should consider basing sample size calculations on realistic assumptions for each district, using available data when possible. Given the long history of trachoma surveillance in Amhara, there is a significant amount of data available for approximating the prevalence in a district prior to the next survey. Previous reports have found that a continuation rate (rate of evaluation units requiring continued MDA after an impact survey) of greater than 71% implies that an impact sur-

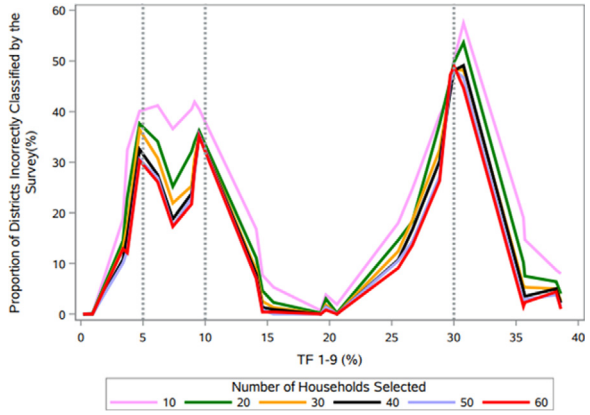

Figure 4. Proportions of incorrect decisions by household number and true prevalence. Peak probabilities of an incorrect decision occur at MDA decision thresholds (vertical gray dashed lines). The lowest probabilities of an incorrect decision are found between two MDA thresholds.

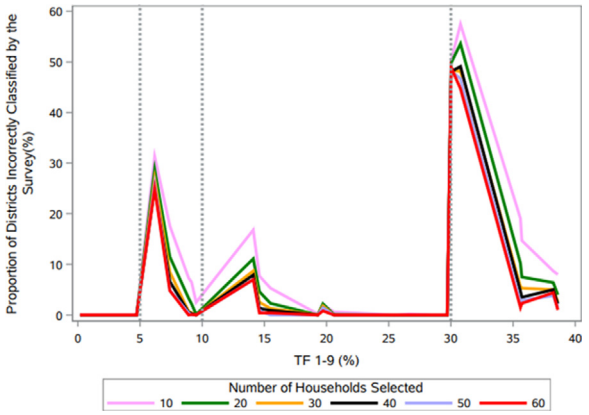

Figure 5. Proportions of low MDA decisions by household number and true prevalence. Peak probabilities of a low decision occur at MDA decision thresholds (vertical gray dashed lines). The lowest probabilities of a low decision are found between two MDA thresholds.

vey was an inefficient use of funds versus an additional round of MDA (Solomon et al., 2020). In Amhara many impact survey rounds are expected to yield continuation rates greater than 71%, suggesting cost inefficiencies in the current impact survey schedules.

In moderate-prevalence districts (10–30% TF prevalence), UIs become as narrow as 10–15% when sampling 34 clusters, which is still wide relative to treatment decisions. However, if MDA is to continue in a district, spending money to regularly resurvey in order to determine the exact prevalence is likely a high-cost, low-reward scenario. It has been determined that intraclass correlations (ICC) in Ethiopia are large and highly variable in comparison with ICCs in Nigeria and Mozambique, which leads to low precision relative to other countries (Macleod et al., 2020). The authors further state that precision of the prevalence estimate decreases

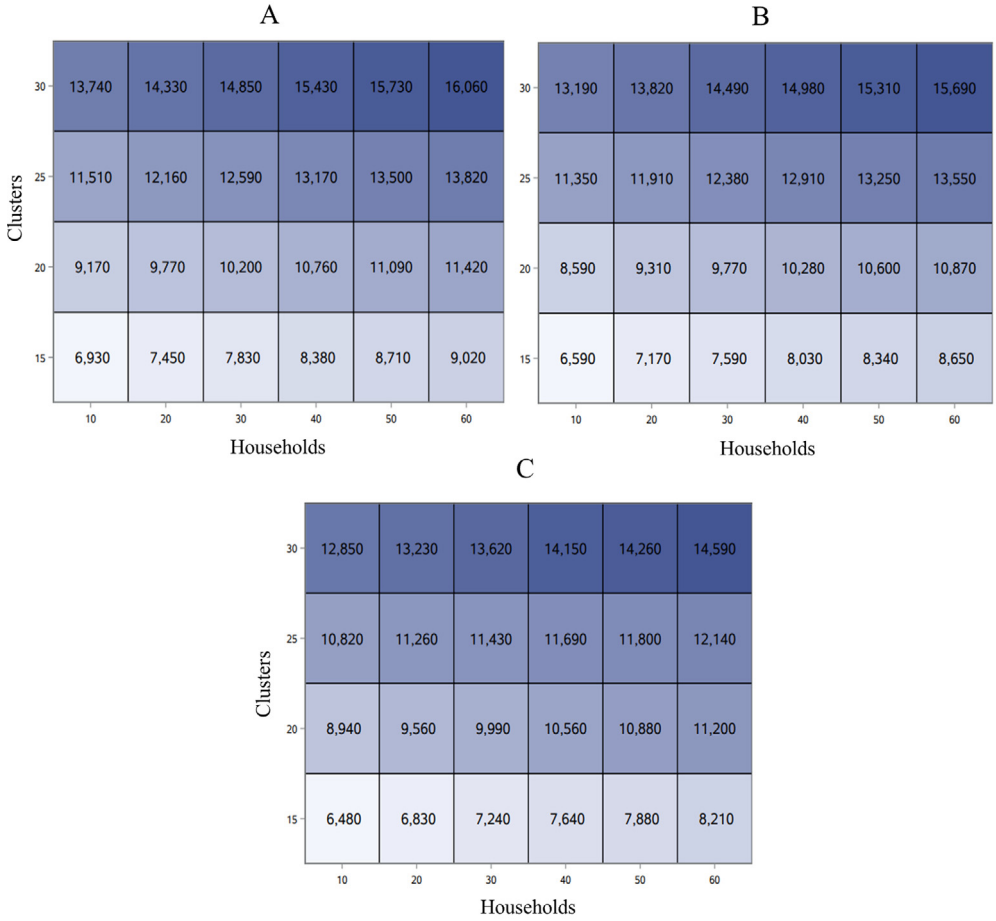

Figure 6. Cost waste by sampling design in districts of true prevalence of (a) 0.38, (b) 0.15, and (c) 0.04 by number of clusters and number of households, Amhara, Ethiopia (cost waste calculated per district in USD). The highest cost waste occurs at 30 clusters and 60 households for each prevalence.

as TF prevalence increases. Thus, sampling 15–20 clusters in suspected moderate- and high-prevalence districts should limit costs in areas unable to achieve precise estimates regardless of survey size.

Obtaining precise district-level estimates becomes most relevant when determining if a district is below the 5% TF threshold. In districts with assumed TF prevalence < 5%, despite achieving precision of 2%, precision barely improves between sampling 14 clusters and 34 clusters due to the homogeneity of prevalence among clusters. Our findings were in line with previous work using trachoma survey data from three countries, which showed that ICC decreased sharply at a low TF prevalence; thus, accurate estimates of TF can be made using smaller sample sizes (Macleod et al., 2020). While it is understandably tempting to sample as many clusters as possible in low-prevalence districts to ensure TF elimination, this practice does not result in substantially improved precision, and may not be worth the increased cost (Figure 2). Sampling 20 clusters in expected low-prevalence districts would allow for balance between precision and cost-effectiveness.

It is a common dilemma that trachoma seems to 'reappear' in districts where TF was below the 5% elimination threshold in a previous survey (Godwin et al., 2020; Weiss et al., 2021; Sata et al. 2021). In recent reports from Amhara region, 24% of surveillance surveys found a TF prevalence ≥ 5% (Sata et al., 2021). Our results provide strong evidence that concluding that TF has resurged in these districts may be misguided, since the simulations demonstrated that when the true TF prevalence was below the threshold (0.04), surveys were as likely to estimate a prevalence greater than or less than the 5% threshold. A surveillance survey prevalence over the threshold may simply represent the variability inherent in the surveys themselves (Godwin et al., 2020). Reappearance of trachoma observed in surveillance surveys is likely due to either: (1) the TF was never below the threshold and the previous impact survey underestimated the true prevalence; or (2) the TF is still below the threshold and the current survey overestimated true prevalence. Since restarting MDA is costly to elimination programs, an urgent need exists for increased operational research around the use of alternative indicators and timelines, and for an immediate review of survey approaches for trachoma surveillance.

The number of households selected had little impact on sample accuracy, suggesting that the current recommendation that teams survey the number of households that can be reached in one day (20–30) is sufficient for TF estimates. For all districts, cost waste increased as the number of clusters and households increased. The cost waste metric could help program managers better understand the tradeoff between precision and cost under a range of epidemiological settings. Furthermore, this metric could be useful for other NTDs that rely on population-based surveys as a monitoring tool.

A primary limitation of this study was the design of the population dataset and the associated assumptions. Only data from Amhara were used to approximate prevalence distributions. Sample size recommendations from the simulations depend on the ability to estimate the general endemicity of a given district prior to surveying — something that may be difficult, especially for younger programs. There has also been limited work on the cost of trachoma surveys (Chen et al., 2011; Slaven et al., 2020; Trotignon et al., 2017; Stelmach et al., 2019). Assumptions were made in deriving the cost formula, and all costs only reflected estimates. However, the relative cost of sampling designs is of interest for this study, not exact cost. Estimates were defined as 'correct' or 'incorrect' according to WHO treatment threshold guidelines (which are arbitrary to an extent), and not according to some truly optimal strategy. Lastly, sources of non-sampling error, such as measurement errors, were ignored. These errors play a role in TF estimation, and could also lead to incorrect MDA decisions.

There are many possible expansions of this study. The methods used for these analyses have been made publicly available to allow trachoma-endemic countries to develop their own sampling schemes. One could examine the sampling methodology trends in small districts with fewer than 30 villages. These methods could be also be extended to a re-evaluation of TT surveys, which call for approximately 30 clusters under most conditions (Flueckiger et al., 2017). Cost waste could be used to evaluate sampling schemes for other NTDs, such as schistosomiasis and soil-transmitted helminths. Lastly, as the TF prevalence in Amhara continues to drop over the coming years, the methods presented in this paper will be helpful in continually re-evaluating trachoma sampling methodology.

CONCLUSION

With a newly declared target date of 2030, trachoma programs serving the highest-burden areas need to be efficient with resources, balancing the need for high-quality monitoring with the widespread administration of interventions (WHO, 2020a). Efficient use of resources is required, and will promote improved progress towards the long-term elimination of trachoma as a public health problem.

Acknowledgements

This research was funded internally by The Carter Center.

Ethical approval

This research did not require statements of ethical consent or standards of animal care since the study did not use any human or animal subjects.

References

Andrade-Pacheco R, Rerolle F, Lemoine J, Hernandez L, Meïté A, Juziwelo L, et al. Finding hotspots: development of an adaptive spatial sampling approach. Sci Rep 2020;10(1):10939. doi:10.1038/s41598-020-67666-3.

Chen C, Cromwell EA, King JD, Mosher A, Harding-Esch EM, Ngondi JM, et al. Incremental cost of conducting population-based prevalence surveys for a neglected tropical disease: the example of trachoma in 8 national programs. PLoS Negl Trop Dis 2011;5(3):e979.

Flueckiger R, Courtright P, Mabey D, Pullan R, Solomon A. Design and validation of a trachomatous trichiasis-only survey. In: World Health Organization, editor. Strategic and Technical Advisory Group for Neglected Tropical Diseases 2017.

Gelman A, Greenland S. Are confidence intervals better termed 'uncertainty intervals'? BMJ 2019;366:l5381.

Godwin W, Prada JM, Emerson P, Hooper PJ, Bakhtiari A, Deiner M, et al. Trachoma prevalence after discontinuation of mass azithromycin distribution. J Infect Dis 2020;221(Suppl 5):S519 –S24.

GTMP. Tropical data; 2021. Available from: https://www.tropicaldata.org/. [Accessed 3/1/2021].

Knowles SCL, Sturrock HJW, Turner H, Whitton JM, Gower CM, Jemu S, et al. Optimising cluster survey design for planning schistosomiasis preventive chemotherapy. PLoS Negl Trop Dis 2017;11(5).

Macleod CK, Bailey RL, Dejene M, Shafi O, Kebede B, Negussu N, et al. Estimating the intracluster correlation coefficient for the clinical sign 'trachomatous inflammation–follicular' in population-based trachoma prevalence surveys: results from a meta-regression analysis of 261 standardized preintervention surveys carried out in Ethiopia, Mozambique, and Nigeria. Am J Epidemiol 2020;189(1):68–76.

Missamou F, Marlhand H, Dzabatou-Babeaux ASP, Sendzi S, Bernasconi J, D'Souza S, et al. A population-based trachoma prevalence survey covering seven districts of Sangha and Likouala Departments, Republic of the Congo. Ophthalmic Epidemiol 2018;25(sup1):155–61.

Nash SD, Stewart AEP, Astale T, Sata E, Zerihun M, Gessese D, et al. Trachoma prevalence remains below threshold in five districts after stopping mass drug administration: results of five surveillance surveys within a hyperendemic setting in Amhara, Ethiopia. Trans R Soc Trop Med Hyg 2018;112(12):538–45.

Ngondi J, Gebre T, Shargie EB, Graves PM, Ejigsemahu Y, Teferi T, et al. Risk factors for active trachoma in children and trichiasis in adults: a household survey in Amhara Regional State, Ethiopia. Trans R Soc Trop Med Hyg 2008;102(5):432–8.

Sanders AM, Abdalla Z, Elshafie BE, Elsanosi M, Nute AW, Aziz N, et al. Progress toward elimination of trachoma as a public health problem in seven localities in the republic of Sudan: results from population-based surveys. Am J Trop Med Hyg 2019;101(6):1296–302. doi:10.4269/ajtmh.19-0530.

Sata E, Nute AW, Astale T, Gessese D, Ayele Z, Zerihun M, et al. Twelve-year longitudinal trends in trachoma prevalence among children aged 1–9 years in Amhara, Ethiopia, 2007–2019. Am J Trop Med Hyg 2021;104(4):1278–89.
Slaven RP, Stewart AEP, Zerihun M, Sata E, Astale T, Melak B, et al. A cost-analysis of conducting population-based prevalence surveys for the validation of the elimination of trachoma as a public health problem in Amhara, Ethiopia. PLoS Negl Trop Dis 2020;14(9).
Solomon A. Validation of the elimination of trachoma as a public health problem. WHO; 2016.
Solomon AW, Hooper PJ, Bangert M, Mwingira UJ, Bakhtiari A, Brady MA, et al. The importance of failure: how doing impact surveys that fail saves trachoma programs money. Am J Trop Med Hyg 2020;103(6):2481–7.
Solomon AW, Macleod CK, Flueckiger RM, Al-Khatib T. Design parameters for population-based trachoma prevalence surveys. In: World Health Organization, editor. Strategic and Technical Advisory Group for Neglected Tropical Diseases 2018.
Stelmach RD, Flueckiger RM, Shutt J, Davide-Smith M, Solomon AW, Rotondo L, et al. The costs of monitoring trachoma elimination: impact, surveillance, and trachomatous trichiasis (TT)-only surveys. PLoS Neglected Tropical Diseases 2019;13(9) Epub 2019/09/06. doi:10.1371/journal.pntd.0007605.
Stewart AEP, Zerihun M, Gessese D, Melak B, Sata E, Nute AW, et al. Progress to eliminate trachoma as a public health problem in Amhara National Regional State, Ethiopia: results of 152 population-based surveys. Am J Trop Med Hyg 2019;101(6):1286–95.
Sturrock HJ, Gething PW, Clements AC, Brooker S. Optimal survey designs for targeting chemotherapy against soil-transmitted helminths: effect of spatial heterogeneity and cost-efficiency of sampling. Am J Trop Med Hyg 2010;82(6):1079–87.
Trotignon G, Jones E, Engels T, Schmidt E, McFarland DA, Macleod CK, et al. The cost of mapping trachoma: data from the Global Trachoma Mapping Project. PLoS Negl Trop Dis 2017;11(10).
Weiss PS, Michael E, Richards FO. Simulating a transmission assessment survey: an evaluation of current methods used in determining the elimination of the neglected tropical disease, lymphatic filariasis. Int J Infect Dis 2021;102:422–8.
WHO. Report of the Third Global Scientific Meeting on Trachoma. Baltimore, USA: World Health Organization; 2010.
WHO. Technical consultation on trachoma surveillance. Decatur, USA: World Health Organization; 2014.
WHO. Ending the neglect to attain the Sustainable Development Goals; 2020a Ntuli MM, editor.
WHO. Weekly Epidemiological Record; 2020b.