

**SHORT COMMUNICATION****Data collection using open access technology in multicentre operational research involving patient interviews**

Axshya SAMVAD Study Group: H. D. Shewade,¹ S. S. Chadha,¹ V. Gupta,^{1,2} J. P. Tripathy,¹ S. Satyanarayana,¹ K. Sagili,¹ S. Mohanty,¹ O. P. Bera,¹ P. Pandey,¹ P. Rajeswaran,³ G. Jayaraman,³ A. Santhappan,⁴ U. N. Bajpai,⁵ A. M. Mamatha,⁴ R. Maiser,⁴ A. J. Naqvi,⁶ S. Pandurangan,¹ S. Nath,¹ V. H. Ghule,¹ A. Das,¹ B. M. Prasad,¹ M. Biswas,¹ G. Singh,¹ G. Mallick,¹ A. J. Jeyakumar Jaisingh,³ R. Rao,⁷ A. M. V. Kumar^{1,8}

<http://dx.doi.org/10.5588/pha.15.0079>

Conducting multicentre operational research is challenging due to issues related to the logistics of travel, training, supervision, monitoring and troubleshooting support. This is even more burdensome in resource-constrained settings and if the research includes patient interviews. In this article, we describe an innovative model that uses open access tools such as Dropbox, TeamViewer and CamScanner for efficient, quality-assured data collection in an ongoing multicentre operational research study involving record review and patient interviews. The tools used for data collection have been shared for adaptation and use by other researchers.

Project Axshya, a flagship programme of the International Union Against Tuberculosis and Lung Disease South-East Asia Office (USEA), is being implemented in 300 districts across 21 states of India to enhance the visibility and reach of the Revised National Tuberculosis (TB) Control Programme.¹ Within this context, Axshya SAMVAD (Sensitization and Advocacy in Marginalised and Vulnerable Areas of the District), an active case-finding strategy,² was introduced in 2013. From April 2016 onwards, a multicentre operational research (OR) study was implemented in 18 randomly sampled districts across seven states of India (Figure 1) to determine the effectiveness of Axshya SAMVAD in 1) reducing diagnostic and treatment delays, 2) reducing patient costs and 3) reducing unfavourable treatment outcomes relative to passive case finding.

Quality-assured data collection in resource constrained settings is a challenge.^{3,4} The use of the open access technologies Dropbox™ (Dropbox, Inc, San Francisco, CA, USA) and EpiData (EpiData Association, Odense, Denmark) for quality-assured data capture in multicentre OR has previously been described, but only pertaining to record review.⁴ Our study, in contrast, involved programmatic record review as well as community-based patient interviews, the latter of which are more prone to error.⁵ To optimise the limited budget available for this study, we developed an innovative model of data collection/sharing using open access tools. In this paper, we describe and discuss the advantages and disadvantages of the model.

ASPECT OF INTEREST: INNOVATION IN DATA COLLECTION**Baseline training**

The principal investigator trained co-investigators and project staff involved in data collection (February–March 2016) across five cities in India. Three open access tools were installed: Dropbox for desktop and smartphones, TeamViewer (TeamViewer GmbH, Göttingen, Germany) for desktop and CamScanner (IntSig International Holding Ltd, Hong Kong, China), an application for smartphones (Figure 2). All of these tools are proprietary but offer free usage sufficient for most OR needs (Appendix Table A.1).

On the laptop computer of the principal investigator, Dropbox folders for each district* were created and shared with team members involved in data collection for that district. Any data saved by team members on their computers and smartphones thus automatically became available to the principal investigator stationed at the data control centre at the USEA in New Delhi. Dropbox had the dual advantage of offline use and online syncing. Dropbox provides additional free storage for the referee if a person creates a new Dropbox account based on a referral. The principal investigator ensured during training that all those involved in data collection created their own Dropbox account by sending them the Dropbox referral request.

The CamScanner application was linked to the district's Dropbox folder. The purpose was to scan all the completed forms rapidly in the field and generate quality (aligned, high resolution and in focus) jpg/pdf files for sharing with the project investigators.

Study participant enrolment

At the beginning of every month, and in every study district, the District Coordinator (DC) of Project Axshya prepared a list of study participants and updated their unique identifiers in the 'study participant monthly enrolment' Excel sheet (Microsoft, Redmond, WA, USA), a tool to enrol study participants on a monthly basis in Dropbox. Using a standard sampling technique, the principal investigator enrolled them

AFFILIATIONS

- 1 International Union Against Tuberculosis and Lung Disease (The Union), South-East Asia Office, New Delhi, India
- 2 All India Institute of Medical Sciences, New Delhi, India
- 3 Resource Group for Education and Advocacy for Community Health, Chennai, India
- 4 Catholic Health Association of India, Secunderabad, India
- 5 Voluntary Health Association of India, New Delhi, India
- 6 MAMTA Health Institute for Mother and Child, New Delhi, India
- 7 Central TB Division, Revised National Tuberculosis Control Programme, Ministry of Health and Family Welfare, New Delhi, India
- 8 The Union, Paris, France

CORRESPONDENCE

Hemant Deepak Shewade
International Union Against Tuberculosis and Lung Disease South-East Asia Office
C-6 Qutub Institutional Area
New Delhi 110016, India
e-mail: hemantjipmer@gmail.com

KEY WORDS

operational research; multicentre studies; longitudinal studies; data collection; India

Received 19 September 2016
Accepted 19 November 2016

PHA 2017; 7(1): 74–77
© 2017 The Union

*Available from the corresponding author upon request.

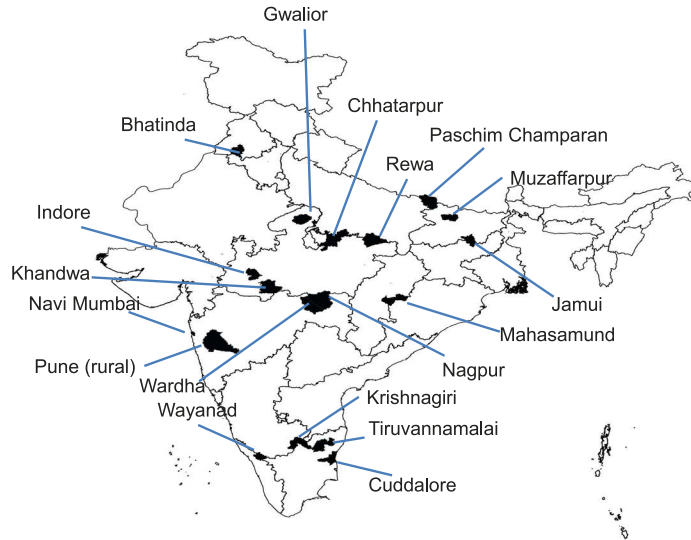


FIGURE 1 Map of India depicting the randomly sampled study districts ($n = 18$) in the Axshya SAMVAD study,* India, 2016–2017. *An active case-finding strategy implemented across 300 districts of India by the International Union Against Tuberculosis and Lung Disease, South-East Asia Office, New Delhi, India.

into the study and added them into the ‘Case-wise monitoring tool’ Excel sheet.[†] This activity was repeated every month until the desired sample size was enrolled (Appendix Table A.2).

Parts I and II of the questionnaire

The DCs and the district-level supervisors were expected to complete the data collection for Part I (a sin-

gle-page questionnaire filled out using record review) and for Part II (a three-page questionnaire filled out to read: filled out using structured close-ended patient interviews) of the questionnaire within respectively 1 and 2 months of the date of enrolment. Scanned copies of the paper forms were added to the Part I and Part II folders in Dropbox. Each interview was audio recorded using a smartphone, and the audio file was added to the Audio folder of Dropbox. All the scanned

[†] Available from the corresponding author on request.

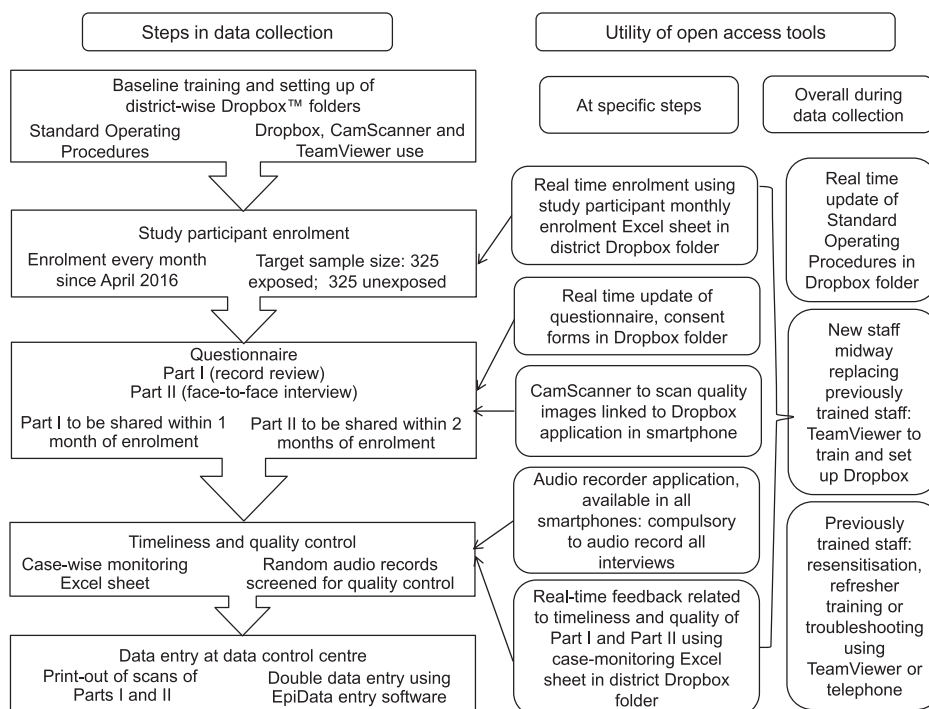


FIGURE 2 Utility of open-access tools at various steps of data collection during the Axshya SAMVAD study,* India, 2016–2017. *An active case-finding strategy implemented across 300 districts of India by the International Union Against Tuberculosis and Lung Disease, South-East Asia Office, New Delhi, India.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the contribution of sub-recipients of Project Axshya for their support: Population Services International, New Delhi; the Catholic Health Association of India, New Delhi; Voluntary Health Association of India (VHAI), New Delhi; Resource Group for Education and Advocacy for Community Health (REACH), Chennai; The Catholic Bishops' Conference of India-Coalition for AIDS and Related Diseases (CBCI-CARD), MAMTA Health Institute for Mother and Child, New Delhi and the Emmanuel Hospital Association (EHA), New Delhi, India. The authors thank the District Coordinators and the Interpersonal Communication Coordinators of Project Axshya in the 18 study districts for supporting the data collection: P Verma, M S Ansari, P Chandra Mishra, A K Sharma, P Singh, B K Srivastava, S Barik, V Mathew, R Robert, K Singh, M Ranjan, S Lohar, M Nema, Y Singh Rajput, R Singh, C S Gaarkhede, S Y Bale, G Parate, I P Koli, A K Bharadwaj, G Venkatraman, K Sathyanarayanan and G Sumesh. They thank B Chingsubam and J Lal for their support in data collection. They also thank other Project Axshya staff P K Singh, D Tigga, K J K Singh, K Kumar and R Singh, who participated in the initial training, and the Revised National Tuberculosis Control Programme staff in the study districts who supported the District Coordinators and the Interpersonal Communication Coordinators in study participant enrolment and record review. The authors thank the Department for International Development (DfID), London, UK for funding the Global Operational Research Fellowship Programme at the International Union Against Tuberculosis and Lung Disease, Paris, France, in which H D Shewade works as an operational research fellow. Project Axshya is being implemented by The Union South-East Asia Office and is supported by the Global Fund to Fight AIDS, Tuberculosis and Malaria (Geneva, Switzerland). This study is being conducted as an operational research under the project conditions using project staff. Therefore, no separate budget was required. The contents of this paper do not necessarily reflect the views of the government or non-governmental organisations or The Union. The authors thank the DfID for funding this open access publication. Conflicts of interest: none declared.

copies, available in Dropbox, were printed, and double data entry was undertaken centrally at the USEA using EpiData v. 3.1. A standard convention for naming the files was followed to maintain uniformity and avoid ambiguity.

Monitoring data collection and quality control

Indicators pertaining to the timeliness and quality of the data collection for Part I and Part II were followed up by the district and USEA supervisors, respectively, in the 'Case-wise monitoring tool' Excel sheet. Ten per cent of the audio recordings were randomly assessed for quality by the USEA supervisors. Repeat interviews were conducted if the data quality was suboptimal.

If any project staff identified that an enrolled study participant had to be excluded from the study (mostly due to initial misclassification), this was flagged in the comments column of the 'Case-wise monitoring tool' Excel sheet for review and decision by the principal investigator.

As it is a remote-control application, the TeamViewer software was helpful for trouble shooting, training new staff or retraining existing staff.

DISCUSSION

There were many advantages to this model (Figure 2). Each person knew immediately what the other person had done and could therefore plan activities accordingly. This methodology was resource-efficient, as travel and logistics costs were incurred only for baseline training. Existing project staff used the laptops provided by the project, internet and personal smartphones, and used their routine supervisory field visits to collect data for the study. Remote monitoring and quality control from the data control centre ensured the timeliness of data collection and the quality of the data, obviating the need for repeated re-training and the physical

presence of the supervisors. This could be undertaken even during international travel. Online backup in Dropbox (for 1 month maximum) helped with the retrieval of any files that were accidentally deleted. The principal investigator required a large amount of storage space in the Dropbox account, as the scanned files and audio records of all 18 districts were stored in his Dropbox. This potential problem was circumvented by using the Dropbox referral feature.

There are some limitations to this system. Without laptops, internet or smartphones, it would be difficult to implement the model. Telephone calls were also the principal mode of communication, especially when the project staff were in the field. As interviews were audio recorded, participants might have become self-conscious and responded accordingly.

In conclusion, we describe an innovative model for data collection and sharing in a multicentre OR study using open-access tools. We hope that this information is helpful to others who are undertaking or preparing to undertake similar ventures.

References

- 1 International Union Against Tuberculosis and Lung Disease. Project Axshya. Paris, France: The Union, 2016. <http://www.theunion.org/what-we-do/technical-assistance/tuberculosis-and-mdr-tb/project-axshya> Accessed December 2016.
- 2 Stop TB Partnership. Axshya SAMVAD goes door-to-door to fight TB [Internet]. 2014. Geneva, Switzerland: United Nations Office for Project Services, 2014. http://www.stoptb.org/news/frompartners/2014/fp14_085.asp Accessed December 2016.
- 3 Whitney C W, Lind B K, Wahl P W. Quality assurance and quality control in longitudinal studies. *Epidemiol Rev* 1998; 20: 71–80.
- 4 Kumar A M V, Naik B, Guddemane D K, et al. Efficient, quality-assured data capture in operational research through innovative use of open-access technology. *Public Health Action* 2013; 3: 60–62.
- 5 Rieder H L, Lauritsen J M. Quality assurance of data: ensuring that numbers reflect operational definitions and contain real measurements. *Int J Tuberc Lung Dis* 2011; 15: 296–304.

APPENDIX

TABLE A.1 Description of open access tools Dropbox™, TeamViewer and CamScanner

Tool	Description
Dropbox	Dropbox is a service that offers online cloud storage (backup on internet) of computer files and automatic file synchronisation across multiple computers and devices (tablets, smartphones), once a Dropbox account has been created by a user. The software creates a special 'Dropbox' folder on each computer, the contents of which are synchronised so that the same folder (with the same contents) is visible on each computer that it is viewed on by the user. Files and folders placed in this Dropbox folder are also accessible through a website and through smartphone applications. Sub-folders can be created within the main Dropbox folder and these sub-folders can be shared with other Dropbox users for mutual access. Dropbox uses a 'Freemium' business model, through which users are offered a free account with a set storage size (2 gigabytes in this case), and paid subscriptions for accounts with more capacity. (www.Dropbox.com)
TeamViewer	TeamViewer is a secure software package for remote control, desktop sharing, online meetings, web conferencing and file transfer between computers. TeamViewer is a tool that permits a computer user to take complete control of another computer via the internet. Upon running the TeamViewer application on a host computer, a username and password for that computer gets generated. Using this username and password, another TeamViewer user (using a different computer) can take complete control of the host computer via the internet. TeamViewer use is free for non-commercial purposes. (www.teamviewer.com)
CamScanner	CamScanner is an open-access application for taking pictures of documents, with a quality similar to scanning, and sharing them in pdf/jpg format directly via e-mail or other applications within the smartphone or tablet, including Dropbox. CamScanner converts a smartphone or tablet into a scanner for document sharing. (https://www.camscanner.com/)

TABLE A.2 Operational definition of study participants and sampling methodology in Axshya SAMVAD study,* India, 2016–2017

Participant	Definition
Study participant	New smear-positive patient with TB registered for treatment and belonging to marginalised population in the district
Study participant: exposed	Study participant diagnosed through Axshya SAMVAD, i.e., participant resides in a village/urban ward where Axshya SAMVAD was conducted before the date of diagnosis, and documentation of Axshya SAMVAD available in project records
Study participant: unexposed and eligible	Study participant not fitting the 'study participant: exposed' criteria and belonging to a village/urban ward where Axshya SAMVAD was not conducted (ever) before the date of diagnosis
Study participant: unexposed but excluded from the study	Study participant not fitting the 'study participant: exposed' criteria and belonging to a village where Axshya SAMVAD was conducted (ever) before date of diagnosis
Sampling	Once the DC updated the 'Study participant monthly enrolment tool' in the district Dropbox folder, the principal investigator followed the methodology to enrol study participants: all the 'exposed' participants were enrolled in the study; an equal number from the 'unexposed and eligible' list were randomly enrolled as 'unexposed' (1:1 ratio exposed:unexposed), and all the 'unexposed but excluded' participants were excluded from the study

*An active case-finding strategy for TB.

SAMVAD = Sensitization and Advocacy in Marginalised and Vulnerable Areas of the District; DC = district coordinator; TB = tuberculosis.

Conduire des recherches opérationnelles multicentriques est un défi, particulièrement dans les contextes de ressources limitées, en tenant compte des questions de logistique de déplacement, de formation, de supervision, de suivi et de soutien à la résolution des problèmes ; encore plus si cette recherche implique des entretiens avec des patients. Dans cet article, nous décrivons un modèle innovant qui

utilise des outils à accès ouvert comme Dropbox, TeamViewer et CamScanner pour un recueil de données efficace et de qualité assurée dans le cadre d'une recherche opérationnelle continue multicentrique impliquant des revues de dossiers et des entretiens avec des patients. Les outils utilisés pour le recueil de données ont été partagés pour l'adaptation et l'utilisation par d'autres chercheurs.

La realización de investigaciones operativas multicéntricas puede ser problemática, sobre todo en los entornos con restricción de los recursos, habida cuenta de las dificultades en la organización de los desplazamientos, la capacitación, la supervisión, el seguimiento y el apoyo a la resolución de problemas; más aun, cuando la investigación precisa entrevistas a los pacientes. En el presente artículo se describe un modelo innovador que utiliza herramientas de libre acceso como

las plataformas Dropbox, TeamViewer y CamScanner, con el fin de lograr una obtención de datos eficiente y de calidad garantizada, en una investigación operativa multicéntrica en curso que comporta el examen de las historias clínicas y entrevistas a los pacientes. Se comunican las herramientas utilizadas en la recogida de datos, con la finalidad de que otros investigadores puedan adaptarlas y las apliquen.