MRC and the UK Dept for International Development (DFID) under the MRC/DFID Concordat agreement (grant MR/J010723/1). G. Santis was supported in part by the Dept of Health *via* the National Institutes of Health Research comprehensive Biomedical Research Centre award to Guy's and St Thomas' National Health Service Foundation Trust in partnership with King's College London. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Funding information for this article has been deposited with FundRef.

Conflict of interest: Disclosures can be found alongside the online version of this article at erj.ersjournals.com

References

1 World Health Organization. Global tuberculosis report 2014. Geneva, World Health Organization, 2014. Available from: http://apps.who.int/iris/bitstream/10665/137094/1/9789241564809_eng.pdf
2 Blankley S, Berry MP, Graham CM, *et al.* The application of transcriptional blood signatures to enhance our understanding of the host response to infection: the example of tuberculosis. *Philos Trans R Soc Lond B Biol Sci* 2014; 369: 20130427.
3 Kaforou M, Wright VJ, Oni T, *et al.* Detection of tuberculosis in HIV-infected and -uninfected African adults using whole blood RNA expression signatures: a case–control study. *PLoS Med* 2013; 10: e1001538.
4 Berry MP, Graham CM, McNab FW, *et al.* An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* 2010; 466: 973–977.
5 Joosten SA, Fletcher HA, Ottenhoff TH. A helicopter perspective on TB biomarkers: pathway and process based analysis of gene expression data provides new insight into TB pathogenesis. *PLoS One* 2013; 8: e73230.
6 Bloom CI, Graham CM, Berry MP, *et al.* Detectable changes in the blood transcriptome are present after two weeks of antituberculosis therapy. *PLoS One* 2012; 7: e46191.
7 Cliff JM, Lee JS, Constantinou N, *et al.* Distinct phases of blood gene expression pattern through tuberculosis treatment reflect modulation of the humoral immune response. *J Infect Dis* 2013; 207: 18–29.
8 Smid M, Dorssers LC, Jenster G. Venn Mapping: clustering of heterologous microarray data based on the number of co-occurring differentially expressed genes. *Bioinformatics* 2003; 19: 2065–2071.
9 Rhodes DR, Yu J, Shanker K, *et al.* Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc Natl Acad Sci USA* 2004; 101: 9309–9314.
10 Chaussabel D, Quinn C, Shen J, *et al.* A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. *Immunity* 2008; 29: 150–164.
11 Saiga H, Kitada S, Shimada Y, *et al.* Critical role of AIM2 in *Mycobacterium tuberculosis* infection. *Int Immunol* 2012; 24: 637–644.
12 Tussiwand R, Lee WL, Murphy TL, *et al.* Compensatory dendritic cell development mediated by BATF-IRF interactions. *Nature* 2012; 490: 502–507.
13 Jacobsen M, Repsilber D, Gutschmidt A, *et al.* Candidate biomarkers for discrimination between infection and disease caused by *Mycobacterium tuberculosis*. *J Mol Med* 2007; 85: 613–621.
14 Kasvosve I, Gomo ZA, Mvundura E, *et al.* Haptoglobin polymorphism and mortality in patients with tuberculosis. *Int J Tuberc Lung Dis* 2000; 4: 771–775.
15 O'Garra A, Redford PS, McNab FW, *et al.* The immune response in tuberculosis. *Annu Rev Immunol* 2013; 31: 475–527.

Eur Respir J 2016; 47: 1873–1876 | DOI: 10.1183/13993003.02121-2015 | Copyright ©ERS 2016

Implementation of digital technology solutions for a lung health trial in rural Malawi

To the Editor:

There is a global epidemic of non-communicable disease including chronic lung disease [1–3]. In the case of chronic lung disease, this challenge plays out in the context of a high burden of communicable diseases and pulmonary tuberculosis in particular.

The World Health Organization Practical Approach to Lung Health (PAL) provides an approach to address the dual challenges of chronic lung disease and tuberculosis [4]. We designed the Triage Plus cluster randomised controlled trial to determine the effect of using an adaptation of the PAL strategy in combination with engagement of informal health providers on the case detection and treatment of chronic lung disease and tuberculosis in rural Malawi (trial registration number: PACTR201411000910192).

In this research letter, we describe the ways in which we used digital technology solutions to prepare for the Triage Plus trial including mapping the study area, obtaining a sample of participants and collecting and managing baseline data from study participants.

The study area was the Dowa and Ntchisi districts in Malawi covering 27 health facilities with an estimated population of 640 000.

We identified our study area using a Google Earth Pro (www.google.com/intl/en_uk/earth/; Google Inc., Mountain View, CA, USA) from which we demarcated the study area into clusters defined as catchment areas for health facilities. A buffer zone between neighbouring clusters was created to reduce contamination between clusters. Each cluster was demarcated into enumeration areas to enable enumeration of all dwelling units within each cluster. The size and shapes of clusters and enumeration areas were defined by housing density, natural geographical features like rivers and mountains, and man-made structures like roads and buildings. For the sampled enumeration areas, we used place mark pointers on Google Earth Pro to identify all potentially inhabited structures within each enumeration area on the map (an illustration of which is shown in figure 1). GPS coordinates for each potentially inhabited structure (dwelling unit) were extracted into an Excel document (Microsoft, Redmond, WA, USA) for sampling using KMLCSV Converter software (to convert Google Earth's KML file format to a comma separated value file format; http://kmlcsv-converter.soft112.com/).

When a cluster had been demarcated into enumeration areas and the total number of households in each enumeration area had been counted, 30 enumeration areas per cluster were selected based on a probability proportional to the number of households per enumeration area. At enumeration area level, eight dwelling units were selected for interviewing and an additional eight were selected as replacements to be used if the initially selected dwelling was found to be uninhabited.

Going by cluster, research assistants visited allocated dwelling units using their corresponding coordinates. The coordinates were entered into Samsung Galaxy S3 (Samsung Electronics, Suwon, South Korea) smartphones using Global Positioning Systems Essentials (www.gpsessentials.com) upon which the smartphones automatically showed the distance and the direction from the location of the research assistant to the target dwelling unit. Once dwelling units had been traced, consent for interview obtained, and consent forms filled, all interviews were conducted using an electronic questionnaire programmed onto the smartphones using ODK Collect (https://opendatakit.org). The electronic questionnaire had automatic skip patterns and data validation checks, to minimise physical and logical data collection errors.

At the end of each day, the saved forms were checked by supervisors and exported to an ODK Aggregate/ Briefcase database backed up offsite.

Mapping of enumeration areas on the Google Earth Pro map was straightforward as image resolution down to the required level of detail was excellent and geographical boundaries like mountains, forests and rivers were easy to identify. The only challenge we faced was clouds obscuring detail in one area which was overcome by importing satellite maps of the affected cluster from Bing maps (Microsoft) into Google Earth Pro.

Sampling was conducted at enumeration area and dwelling unit level as described in the methods with no challenges encountered.

Approximately one in 20 of the dwelling-like units sampled for interviewing were uninhabited. We overcame this challenge by further orientation for the mapping team on how to best identify uninhabited structures , such as churches, school blocks and tobacco barns. In the field, this was overcome by working through a pre-specified list of replacement households until the required number had been visited and interviewed. Another challenge was reaching remote and difficult-to-access households that were clearly visible on Google Earth Pro but difficult to get to because of natural barriers like forested areas that became impassable by vehicle during the rainy season. These challenges were overcome by using smartphones for guided household tracing and by enlisting the assistance of local guides.

FIGURE 1 Place marks on potentially inhabited structures, illustrative of the process undertaken using Google Earth Pro.

Data were transferred to the database using a wireless internet connection at our site office which enabled the data management team to monitor progress, and view and map data in real time.

We have reported on the successful use of digital technologies for mapping, sampling, data collection and management in preparation for a cluster level lung health trial in rural Malawi. We were able to demarcate cluster boundaries, buffer zones and enumeration areas using maps available from Google Earth Pro and then go on to sample individual dwelling units for inclusion in the study. This work was conducted without needing any fieldwork. Data collection in the field was efficient as we were able to plan fieldwork with contingency plans well in advance of field trips. Fieldwork was guided by digital technology that enabled fieldworkers to navigate to specific locations for data collection. Onward transfer of data to a central database was a simple process of automated data uploading from the smartphones to the database.

Our study adds to a growing evidence base that readily available and affordable digital technologies can be deployed effectively in challenging research settings like those seen in rural Malawi [5–9]. The software used for mapping, listing, household tracing and data collection and transfer (GPS Essentials, KMLCSV Converter and ODK) are open source, meaning that there were minimal software costs. There were costs in terms of human resources (information technology expertise and programming) and hardware costs (laptops and smartphones) compared with more traditional paper-based approaches. Although we did not set out to formally compare the costs of digital and traditional methodologies, our impression is that the former is more cost effective. The digital methodologies also provided potential advantages to the study conduct and minimisation of bias since sampling was performed objectively and independently from field conditions, and visits to dwelling units were both guided by the technology and tracked. It was possible to check both that specific dwelling units were visited and the location from which data were collected. The immediacy of data transfer to the study office meant that these quality control measures, together with query raising and resolution, could be implemented in real time. There were no reported data loss as data were directly transferred from the smartphones to the database and immediately backed up at the end of each day. Compared with what would have been required for a traditional approach, we needed very little paper which was another cost saving with environmental benefits too. The digital methodology adopted was not without challenges, although these were minimal and readily overcome as described.

The digital technology solutions we have reported on here worked well for mapping and sampling potential study participants in a difficult to reach area of rural Malawi without the need for fieldwork, allowed us to plan fieldwork, gave us immediate visibility of progress and immediate access to study data. We recommend the digital approaches we have reported on here to others planning similar kinds of fieldwork. The opportunities to expedite study set up and implementation, conduct quality control and assurance exercises in real time and do so in a cost effective way are substantial.

@ERSpublications
Digital technology solutions pave the way for a lung health trial in rural Malawi
http://ow.ly/107F5s

Blessings Chisunkha[1,2], Hastings Banda[1,2], Rachael Thomson[2], S. Bertel Squire[2] and Kevin Mortimer[2]
[1]REACH Trust, Lilongwe, Malawi. [2]Dept of Clinical Sciences and Centre for Applied Health Research and Delivery, Liverpool School of Tropical Medicine, Liverpool, UK.

Correspondence: Kevin Mortimer, Dept of Clinical Sciences and Centre for Applied Health Research and Delivery, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, UK. E-mail: kevin.mortimer@lstmed.ac.uk

Received: Dec 18 2015 | Accepted after revision: Feb 09 2016 | First published online: April 13 2016

This study is registered at www.pactr.org with identifier number PACTR201411000910192.

Support statement: This work was funded by NORAD and LHL International, Norway. Funding information for this manuscript has been deposited with FundRef.

Conflict of interest: None declared.

Acknowledgements: Sincere gratitude to the entire project team from both REACH Trust (Malawi) and Liverpool School of Tropical Medicine (UK), with expertise in Health and Medicine, Data Management, Computer Science, Research and Health Economics, for their insightful technical contributions and comments. Our thanks to the field team and the respondents.

References

1 World Health Organisation. Global Status Report on noncommunicable diseases, 2014. Geneva, Switzerland, World Health Organisation, 2014. Available from http://apps.who.int/iris/bitstream/10665/148114/1/9789241564854_eng.pdf Date last accessed: March 24, 2016.

2 World Health Organisation. Global Action Plan for the prevention and control of non-communicable diseases, 2013-2020. Geneva, Switzerland, World Health Organisation, 2013. http://apps.who.int/iris/bitstream/10665/94384/1/9789241506236_eng.pdf Date last accessed: March 24, 2016.
3 GBD 2013 Mortality and Cause of Death Collaborators. Global, regional, and national age–sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015; 385: 117–171.
4 Practical approach to lung health. Manual on initiating PAL implementation. WHO/HTM/TB/2008.410; WHO/NMH/CHP/CPM/08.02. http://apps.who.int/iris/bitstream/10665/69937/1/WHO_HTM_TB_2008.410_eng.pdf Date last accessed: March 24, 2016.
5 Escamilla V, Emch M, Dandalo L, *et al.* Sampling at community level by using satellite imagery and geographical analysis. *Bull World Health Organ* 2014; 92: 690–694.
6 Chang AY, Parrales ME, Jimenez J, *et al.* Combining Google Earth and GIS mapping technologies in a dengue surveillance system for developing countries. *Int J Health Geogr* 2009; 8: 49.
7 Lozano-Fuentes S, Elizondo-Quiroga D, Arturo Farfan-Ale J, *et al.* Use of Google Earth to facilitate GIS-based decision support systems for arthropod-borne diseases. *Adv Dis Surveill* 2007; 4: 91.
8 Wampler PJ, Rediske RR, Molla AR. Using ArcMap, Google Earth, and Global Positioning Systems to select and locate random households in rural Haiti. *Int J Health Geogr* 2013; 12: 3.
9 Kamadjeu R. Tracking the polio virus down the Congo River: a case study on the use of Google Earth in public health planning and mapping. *Int J Health Geogr* 2009; 8: 4.

Eur Respir J 2016; 47: 1876–1879 | DOI: 10.1183/13993003.00045-2016 | Copyright ©ERS 2016

Somatic *DICER1* mutations in adult-onset pulmonary blastoma

To the Editor:

Several rare lung tumours morphologically mimic embryonal structures of the developing human lung. Historically, these blastomatous tumours were described under the umbrella term of pulmonary blastoma. Subsequently, distinct entities were recognised, such as childhood pleuropulmonary blastoma (PPB) [1] (International Classification of Diseases for Oncology (ICD-O-3) code 8973/3). Later, adult-onset pulmonary blastoma was separated into well-differentiated fetal adenocarcinoma (WDFA) (ICD 8333/3) and pulmonary blastoma (ICD 8972/3) [2]. Pulmonary blastoma is a biphasic epithelial and mesenchymal malignancy, whereas PPB is purely sarcomatous and WDFA is characterised by a monophasic immature epithelium. Since 1988, the childhood PPB has received particular attention because of 1) its unique developmental progression from relatively indolent neonatal-onset lung cysts, to aggressive cystic-solid and solid sarcomas by age 72 months; 2) PPB's status heralding a newly recognised familial tumour predisposition syndrome and 3) its strong association with both germ-line and somatic *DICER1* mutations, which is not only true for PPB, but also for many other tumours in pleiotropic predisposition syndrome (now referred to as DICER1 syndrome). Until recently, neither WDFA nor pulmonary blastoma had been observed in families manifesting DICER1 syndrome [3]. However, in 2015, we identified a second somatic *DICER1* RNase IIIb mutation [4] in a WDFA that arose in a 16-year-old germ-line *DICER1* mutation carrier [5]. In addition to *DICER1* mutations in PPB and WDFA, somatic *CTNNB1* mutations (encoding β-catenin) appear to characterise WDFA and pulmonary blastoma [6], and are far less frequent in PPB [7, 8]. In contrast, *TP53* mutations are found in both PPB [8] and pulmonary blastoma [9], but not WDFA [9].

Given the original pathological grouping of WDFA, PPB and pulmonary blastoma and the existence of partially-overlapping molecular abnormalities in these lesions, as described above, we questioned whether pulmonary blastoma might also be characterised by *DICER1* mutations. We therefore analysed *DICER1* status in one infant and two adult pulmonary blastomas by Sanger sequencing and/or targeted capture followed by next-generation sequencing.

The study was approved by the Institutional Review Board of the Faculty of Medicine of McGill University (A12-M117-11A) and was performed with full informed patient or parental consent. Tumours were reviewed by pathologists at the referring institutions and by central reference pathologists (D. Bouron-Dal Soglio and V-H. Nguyen). We obtained peripheral blood DNA from cases 1 (adult onset) and 3 (infant onset); formalin-fixed paraffin embedded (FFPE) tumour tissue from all cases; FFPE non-tumourous lung tissue from case 2 (adult onset); and snap-frozen tumour tissue from case 3.

Sanger sequencing and/or Fluidigm Access Array-based next-generation sequencing (Fluidigm, San Francisco, CA, USA) was used to screen for coding mutations and mutations located near the exon–intron