

Molecular Epidemiological Information System to Support Management of Multidrug-Resistant Tuberculosis in Thailand

Areeya Disratthakit, Penpitcha Thawong, Pundharika Piboonsiri, Surakameth Mahasirimongkol*

Division of Genomic Medicine and Innovation Support, Department of Medical Sciences, Ministry of Public Health, Nonthaburi, Thailand

Abstract

Objective: To support the End TB strategy with an informatics system that integrates genomic data and the geographic information system (GIS) of *Mycobacterium tuberculosis* (MTB) clinical isolates. We aim to develop a system prototype for implementing genomic data to support multiple drug-resistant tuberculosis (MDR-TB) control.

Methods: A 12-step data value chain was applied to describe the information flow within the system. A prototyping-oriented system development method was utilized to test the feasibility of certain technical aspects of a system, and as specification tools to determine user requirements. A simulated dataset was entered as input for initial system testing.

Results: System prototype, namely Integrated MOL Outbreak detection and Joint investigation (iMoji), was established. The data entry modules consisted of (1) patient registration, (2) sample registration, (3) laboratory data entry and data analysis, and (4) verification and approval of the analyzed data. The initial system test demonstrated connectivity among modules without error. The system was able to report integrated genomic data and GIS information of MDR-TB for clustering analysis.

Conclusion: iMoji provides an interactive model for determining molecular epidemiological links of MDR-TB and corresponding spatial information to guide public health interventions for tuberculosis control.

Keywords: Molecular epidemiology, Tuberculosis, GIS information, Genomic, iMoji

Abbreviations: Geographic Information System (GIS), *Mycobacterium tuberculosis* (MTB), multidrug-resistant tuberculosis (MDR-TB), Integrated MOL Outbreak detection and Joint investigation (iMoji)

Correspondence:

Surakameth Mahasirimongkol*

Division of Genomic Medicine and Innovation Support

Department of Medical Sciences, Ministry of Public Health

88/7 Tiwanon Rd. Muang Nonthaburi 11000 Thailand

Tel. (66)2-951-0000 ext. 98095, 98096

e-mail: surakameth.m@dmsc.mail.go.th

DOI: 10.5210/ojphi.v12i1.10416

Copyright ©2020 the author(s)

This is an Open Access article. Authors own copyright of their articles appearing in the Online Journal of Public Health Informatics. Readers may copy articles without permission of the copyright owner(s), as long as the author and OJPHI are acknowledged in the copy and the copy is used for educational, not-for-profit purposes.

Introduction

Spatiotemporal information and genomics data have been shown to have high utility for guiding the public health interventions to control tuberculosis (TB) outbreaks [1]. A transmission network for an epidemic strain of multidrug-resistant tuberculosis (MDR-TB) is difficult to identify in areas of Thailand with a high MDR-TB burden [2,3]. Molecular epidemiology data of MDR-TB are of public health importance as they can be used to track the geographic distribution and spread of drug-resistant *Mycobacterium tuberculosis* (MTB) clones. It is vital to identify the MDR-TB transmission events so that appropriate interventions can be carried out for its control. This project focuses on the rapid detection of MDR-TB and on outbreak investigation utilizing spatiotemporal epidemiology coupled with classification of transmission events by molecular methods.

MTB genotyping is routinely performed in countries with a low incidence of TB. Genotyping is usually an extension of the national TB control program. In 2010, the US Centers for Disease control and Prevention (CDC) established a TB Genotyping Information Management System (TB-GIMS), which links surveillance, MTB genetics, and informatics data to control tuberculosis. The TB-GIMS enables the early detection of TB outbreaks in the United States [4]. The Health Protection Agency at Public Health England demonstrated that whole genome sequencing (WGS) combined with epidemiological information is a useful approach to prioritize public health resources towards tracking and responding to TB outbreaks [5]. Outbreak identification and investigation along with genotyping of *M. tuberculosis* are considered an important means of understanding the TB transmission network [6]. Despite the readiness of lab infrastructure and the availability of human resources for genotyping services, *M. tuberculosis* is not routinely genotyped in Thailand. Furthermore, currently there is no information system in place to collect spatiotemporal data for analyzing the TB molecular cluster in Thailand.

In order to develop an information system, a data value chain has been reviewed for processing data into a usable and desired format. The data value chain consists of four major stages: collection, publication, uptake, and impact [7]. To develop such an information system, a prototype is first used to collect user requirements for refined output. The prototyping aims to establish a link between user needs, designer's ideas, and the final system requirements [8]. There are three main prototyping approaches: exploratory, experimental, and evolutionary [8-11].

In this study, we aim to develop system for integrating genomic data and the GIS of MDR-TB using a prototyping-oriented system development method. The system development involved in three major steps: system requirements analysis, system design and system testing [12].

Methods

Data value chain analysis

System planning was performed to define the proposal and design concepts. The proposal established explicit terms for the objectives of the system. The design defined representations of the system [13].

Data value chain was applied for system planning. The data value chain has four main stages: collection, publication, uptake, and impact. These main stages are further divided into the following twelve steps: identify, collect, process, analyze, release, disseminate, connect, incentivize, influence, use, change, and reuse [7]. System modules were built on the basis of information derived from the value chain analysis. Role-based access control (RBAC) was adopted for module access control.

Prototyping-oriented system development

The prototyping-oriented methodology described an iterative paradigm of system development. The first step in system development was requirements analysis. An exploratory approach was used to clarify system requirements and potential solutions, which can be supported by information technology. System requirements included input, output, process, and data handled. Architecture and component design were the next step to define system architecture and module specifications. System prototype was produced to meet the system architecture [14]. Finally, system testing was performed to ensure that the simulated data entered exactly matches data source and the data processed successfully without error.

Seventy MDR-TB cases occurring between 2013 and 2017 were simple randomly selected from Tuberculosis Case Management database (TBCM) [15] for generating a simulated dataset including patient information, location of TB cases, drug resistance profile and genomic data.

The system prototype was developed under a collaborative project between the Department of Medical Sciences, Ministry of Public Health, Thailand, the CDC and the Thailand MOPH-US CDC Collaboration (TUC).

Initial system testing

The simulated data were entered manually into the system prototype for testing module connectivity. Patient addresses were added for processing data to provide GIS information. Genomic information inputs were sent to Linux system for clustering analysis.

Results

System requirements analysis

System planning through literature analysis showed that currently, there is no information system in place that combines epidemiological data and MTB genetic information to support the detection of MDR-TB outbreaks and the investigation of drug-resistant tuberculosis in Thailand. Therefore, we developed the iMoji system to efficiently collect both TB case information and MTB genotyping data in order to provide evidence of MDR-TB clustering. A design concept with a web-based system was used to collect essential data, including case-centric GIS obtained from patient registration in health facilities and MTB genotypes from laboratories.

The system prototyping scheme was built on a 12-step data value chain (Figure 1). As described above, cases of MDR-TB and genetic information of MTB were the essential inputs provided by health facilities, TB control program, and the *Mycobacterium* genetic laboratory. Considering data privacy, a TB registration number (TB no.) was used as the primary key for processing data related to MDR-TB cases, and a laboratory number (laboratory no.) was generated as an identifier for processing MTB genetics information. The first three steps of the collection phase provided inputs into the publication phase. Analysis of MTB sequencing data and the MDR-TB case location (XY coordinates) released MTB genotype and geospatial data. Data were disseminated in an easy-to-understand format as an MTB genotyping profile and a cluster map of MDR-TB cases. The uptake phase was the next step in system prototyping. A web-based system was used to connect four data sources to iMoji. Role-based access control (RBAC) was used to ensure data security. To increase the data value, geospatial and MTB genotype data were integrated to identify potential hotspots of MDR-TB transmission. It is expected that the iMoji system would facilitate case-area targeted interventions in response to MDR-TB outbreaks. However, the impact depends on how effectively the system could deliver this hotspot and MTB molecular linkage information to authorities in the TB control program.

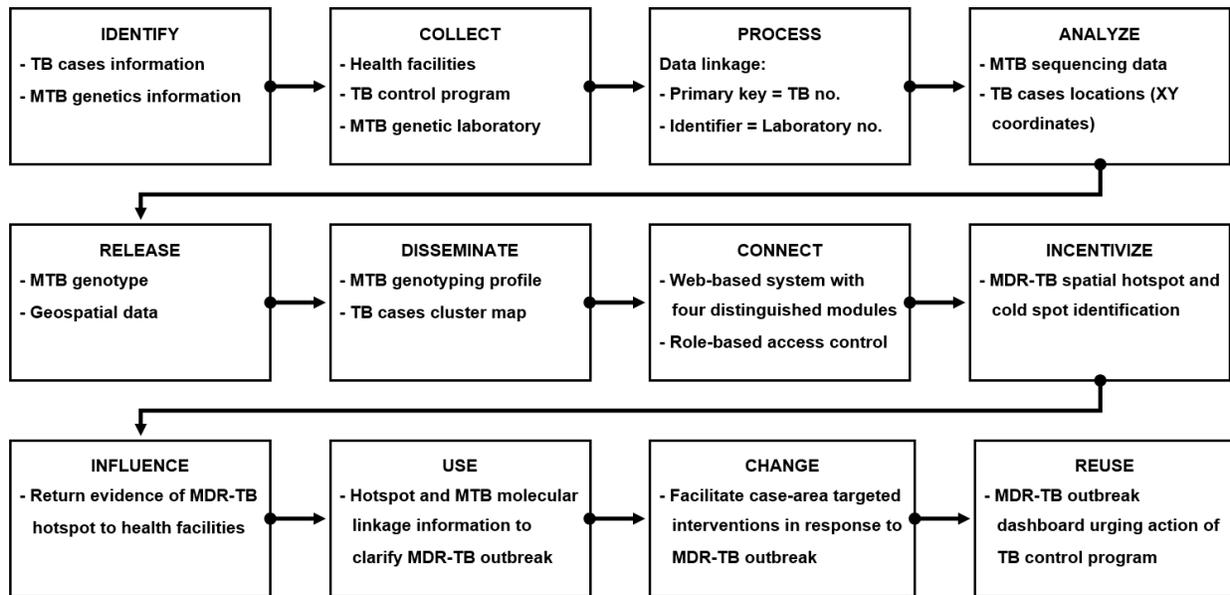


Figure 1: iMoji system prototyping scheme aligned with the 12-step data value chain.

System architecture and user flow diagram

Users were categorized on the basis of their usage of the four user interface modules:

- Patient registration module – User: Health facilities staff, e.g., TB clinic and the Office of Disease Prevention and Control (ODPC)
- Sample registration module – User: Sample inventory staff
- Laboratory data entry and data analysis module – User: MTB genetics laboratory analyst
- Verify and approve analysis data – User: Data reviewer

Concerning data privacy and security, RBAC was used to manage permissions. Each user group had its own set of access rights to a particular module and could view and edit only data accessible in that module (Figure 2). Visualization of the summary results from the data analysis is the only information shared across modules.

In this design, to protect patient privacy, the patient address was automatically transformed into XY coordinates at the sub-district level. This coordinate data was then integrated with MTB genotype data for clustering analysis.

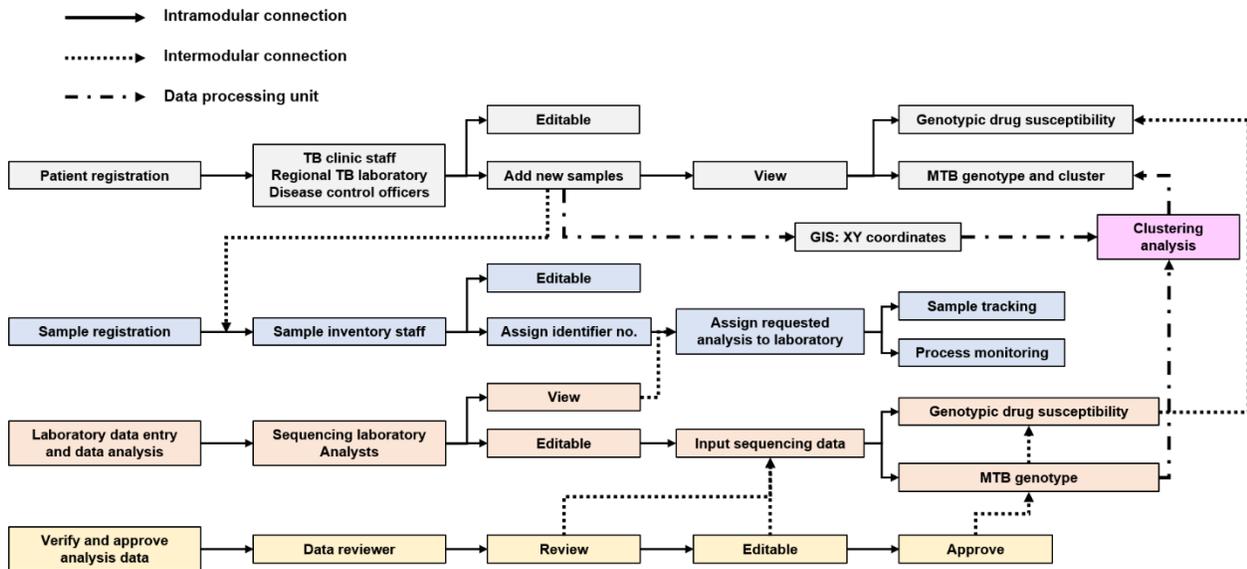


Figure 2: iMoji process model mapped to four application modules (patient registration, sample registration, laboratory data entry and analysis, verification and approval of analyzed data).

The iMoji system was designed according to the above prototyping scheme. Four distinguished information sources were linked to iMoji via the web-based system as show in Figure 3.

Initial system testing

Simulated TB cases information were entered into patient registration module as shown in Figure 4. The inputs were applied for generating XY coordinates and new sample record correctly as shown in Figure 5. These results demonstrated that an intramodular connectivity functioned properly without error.

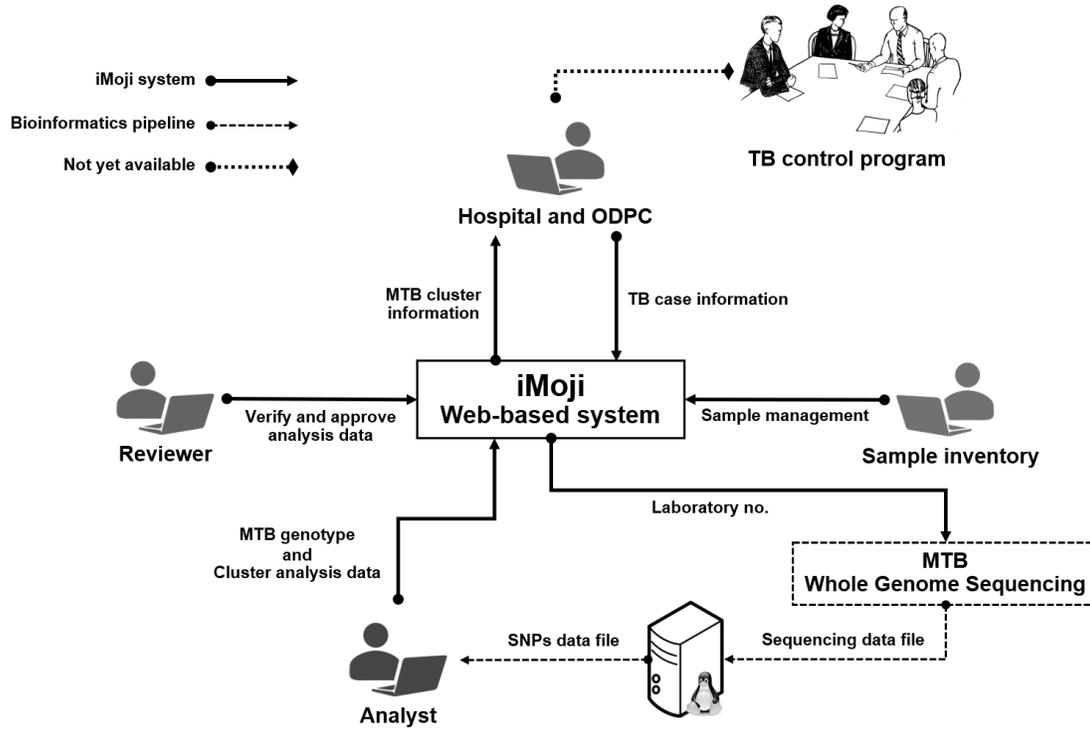


Figure 3: High-level iMoji system architecture.

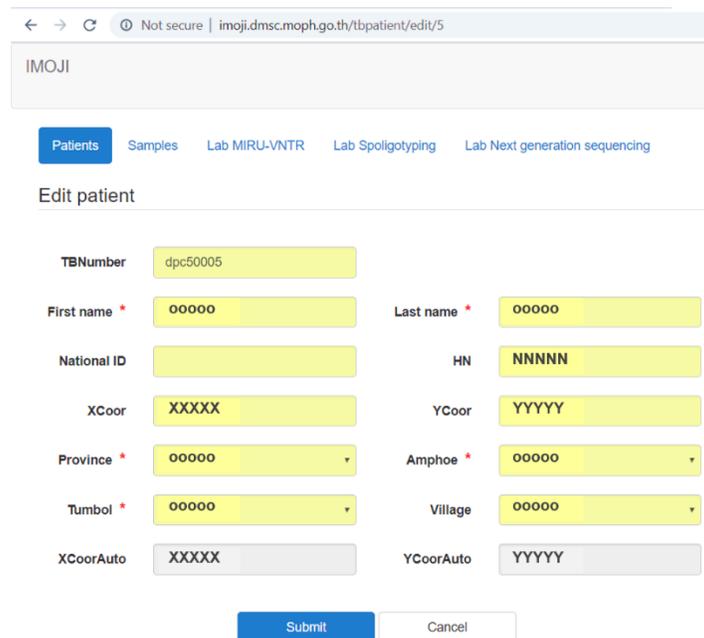


Figure 4: Patient registration module

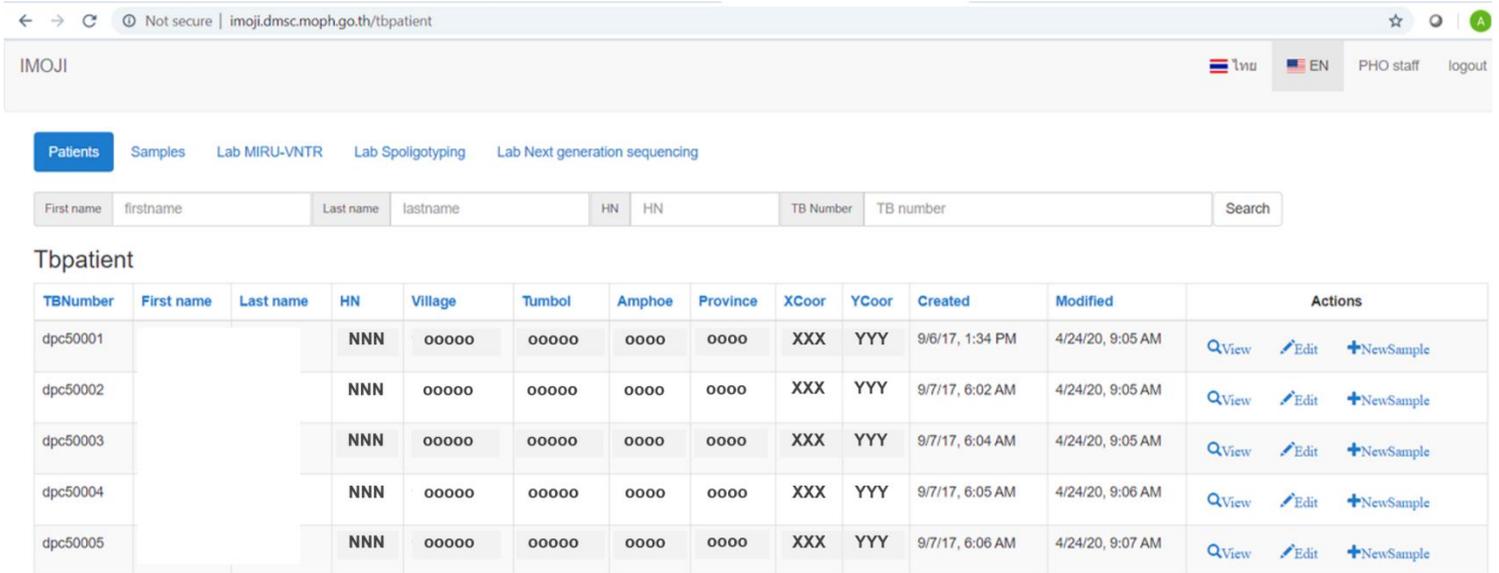


Figure 5: Data table with generated data from patient information module.

The sample records were sent to a sample registration module which facilitated communication between modules using sample ID and DMSc sample ID (Figure 6). Genomic data were entered into a laboratory data entry and data analysis module to integrate with Geographic Information System (GIS) data from the sample module for clustering analysis. These results indicated that intermodular connectivity was established successfully without error.

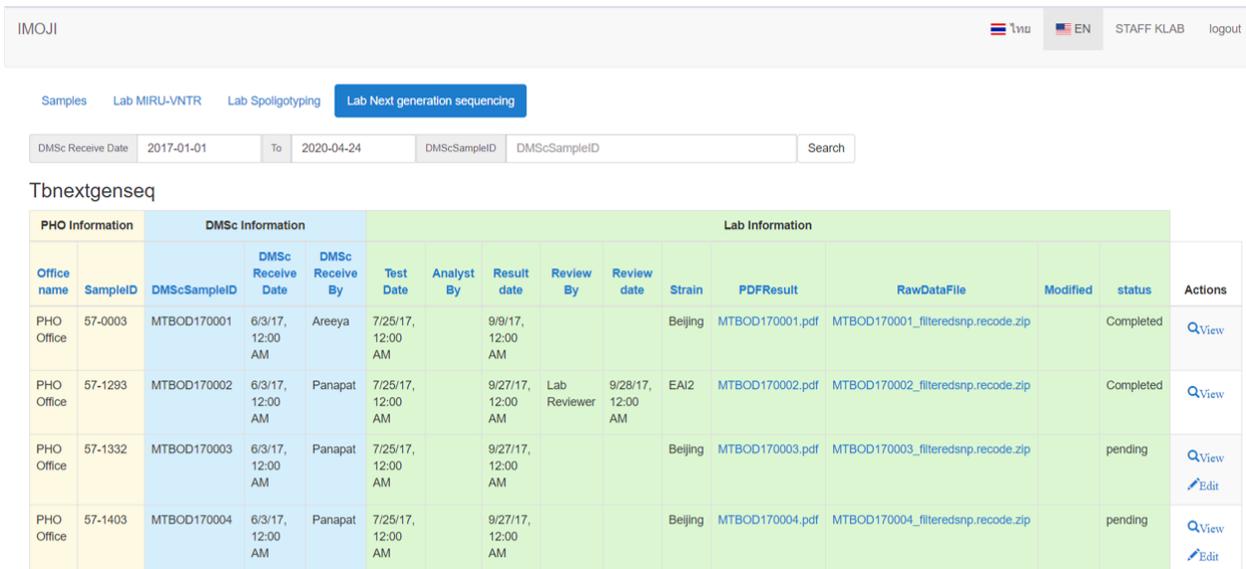


Figure 6: Data table with information links between modules.

Discussion

Establishing information systems to support the control of MDR-TB outbreaks is the cornerstone of public health practice. The iMoji system prototype was developed to meet the needs of diverse users related to molecular epidemiology and surveillance program in Thailand. Due to social stigma, person with TB disease always conscious about their privacy [16]. The system requirements analysis suggested that the use of TB no. and laboratory no. was applicable to personal data protection. In addition, RBAC was applied to restricting module access to authorized users.

The initial system testing results demonstrated that the iMoji prototype actually performed without error. However, the prototype system has limitations of data querying and visualization of the data in an easy-to-understand format for communication with the TB control program. Visualization of the information is an important issue that should be addressed to improve usability of the system [17].

The iMoji system prototype accomplished the first seven steps of 12-step data value chain. To achieve the last five steps, further system development is needed. For example, online and mobile devices for data collection are important tools for scale-up phase in TB outbreak control, involving forecasting using patient-level data scenario simulations for response planning. Furthermore, mandatory features such as response monitoring dashboards, statistical modeling of the intervention strategies, and risk assessment of the surrounding areas are required for TB surveillance [18]. Understanding and interpreting molecular cluster information of MDR-TB cases are crucial to leverage usability of the iMoji system prototype. Finally, TB control program's recognition of usefulness is a significant measure of effective application of the developed system.

Conclusion

In this study, we described “iMoji”, an information system, that was capable of integrating genomic data and GIS data to detect MDR-TB cluster on molecular level. Testing the system on the simulated dataset provided a proof of concept for determining molecular epidemiological links of MDR-TB and corresponding spatial information to guide public health interventions for tuberculosis control.

Acknowledgements

We are grateful to the Growing Expertise in E-Health Knowledge and Skills (GEEKS) training program for critical perspective on the important concepts such as the information value cycle, problem solving for data storage and retrieval processes (informatics) and the boundary spanning nature of informatics work. The program is launched and funded by the U.S. Government's Center for Disease Control and Prevention (CDC) in collaboration with Mahidol University and the Thailand Ministry of Public Health. The study was oversighted by the institutional review board of the Institute for the Development of Human Research Protections.

Financial Disclosure

This work has been co-funded by Centers of Disease Control and Prevention (CDRF agreement no. TBNX-17-63098-1) and Department of Medical Sciences, Ministry of Public Health, Thailand.

Competing Interests

No Competing Interests

Supplementary

Supplementary 1 User Manual: Integrated MOL Outbreak detection and Joint investigation (IMOJI) (Prototype)

References

1. Walker TM, Ip CLC, Harrell RH, Evans JT, Kapatai G, et al. 2013. Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: A retrospective observational study. *Lancet Infect Dis.* 13(2), 137-46 [PubMed https://doi.org/10.1016/S1473-3099\(12\)70277-3](https://doi.org/10.1016/S1473-3099(12)70277-3).
2. Jiraphongsa C, Wangteeraprasert T, Henpraserttae N, Sanguanwongse N, Panya L, Sukkasitvanichkul J, et al. Community outbreak of multidrug resistance tuberculosis, Kanchanaburi province, Thailand on 2002-June 2010. 2010;(June):261–71.
3. Boonthanapat N, Soontornmon K, Pungrassami P, Sukhasitwanichkul J, Mahasirimongkol S, et al. 2019. Use of network analysis multidrug-resistant tuberculosis contact investigation in Kanchanaburi, Thailand. *Trop Med Int Health.* 24(3), 320-27 [PubMed https://doi.org/10.1111/tmi.13190](https://doi.org/10.1111/tmi.13190).
4. Ghosh S, Moonan PK, Cowan L, Grant J, Kammerer S, et al. 2012. Tuberculosis Genotyping Information Management System: Enhancing Tuberculosis Surveillance in the United States. *Infect Genet Evol.* 12(4), 782-88 [PubMed https://doi.org/10.1016/j.meegid.2011.10.013](https://doi.org/10.1016/j.meegid.2011.10.013).
5. Black AT, Hamblion EL, Buttivant H, Anderson SR, Stone M, et al. 2018. Tracking and responding to an outbreak of tuberculosis using MIRU-VNTR genotyping and whole genome sequencing as epidemiological tools. *J Public Health (Oxf).* 40(2), e66-73 [PubMed https://doi.org/10.1093/pubmed/fox075](https://doi.org/10.1093/pubmed/fox075).
6. Nikolayevskyy V, Kranzer K, Niemann S, Drobniewski F. 2016. Whole genome sequencing of Mycobacterium tuberculosis for detection of recent transmission and tracing outbreaks: A systematic review. *Tuberculosis (Edinb).* 98, 77-85 [PubMed https://doi.org/10.1016/j.tube.2016.02.009](https://doi.org/10.1016/j.tube.2016.02.009).
7. 1985. Data2X. The Data Value Chain: Moving from Production to Impact. *Open Data Watch.* 2017, 1-8.

8. Carr M, Verner J. Prototyping and Software Development Approaches. *Prototyp Softw Dev Approaches*. 2004;(3):1–16.
9. Floyd C. A Systematic Look at Prototyping BT - Approaches to Prototyping. In: Budde R, Kuhlenkamp K, Mathiassen L, Züllighoven H, editors. Berlin, Heidelberg: Springer Berlin Heidelberg; 1984. p. 1–18.
10. Baeumer D, Bischofberger WR, Lichter H, Zuellighoven H. 1995. User interface prototyping - concepts, tools, and experience. *Proc Int Conf Softw Eng*. (March), 532-41.
11. Beaudouin-Lafon M, Mackay W. Prototyping Tools and Techniques. 2012;1081–104.
12. Detand J, Bastiaens R, Grimonprez B, Rysman O. The role of prototyping in product development. *Proc 4th Int PMI Conf*. 2010;(January):8.
13. House RW, Warfield JN. 1969. What is system planning? *Automatica*. 5(2), 151-57 [https://doi.org/10.1016/0005-1098\(69\)90008-9](https://doi.org/10.1016/0005-1098(69)90008-9).
14. Susanto A. 2019. Meiryani. System Development Method with The Prototype Method. *Int J Sci Technol Res*. 8(7), 141-44.
15. Bhatia V, Nunn P, Paramasivan CN, Wallengren K. Draft Report International Mini-Review of the Thailand National Tuberculosis Programme Table of Contents. 2016.
16. Atif M, Javaid S, Farooqui M, Sarwar MR. 2016. Rights and responsibilities of tuberculosis patients, and the global fund: A qualitative study. *PLoS One*. 11(3), 1-15 [PubMedhttps://doi.org/10.1371/journal.pone.0151321](https://doi.org/10.1371/journal.pone.0151321).
17. Ola O, Sedig K. Beyond simple charts : Design of visualizations for big health data. 2016;8(3).
18. Morgan O. 2019. How decision makers can use quantitative approaches to guide outbreak responses. *Philos Trans R Soc Lond B Biol Sci*. 374, 20180365 [PubMed https://doi.org/10.1098/rstb.2018.0365](https://doi.org/10.1098/rstb.2018.0365).
19. Jie Z, Ho M, Bin C, Chee E, Ong RT, et al. 2018. International Journal of Infectious Diseases Investigation of a cluster of multi-drug resistant tuberculosis in a high-rise apartment block in Singapore. *Int J Infect Dis*. 67, 46-51 [PubMed https://doi.org/10.1016/j.ijid.2017.12.010](https://doi.org/10.1016/j.ijid.2017.12.010).
20. Chee CBE, Gan S-H, Ong RT, Sng L-H, Wong CW, et al. 2015. Multidrug-resistant tuberculosis outbreak in gaming centers, Singapore, 2012. *Emerg Infect Dis*. 21(1), 179-80 [PubMed https://doi.org/10.3201/eid2101.141159](https://doi.org/10.3201/eid2101.141159).