Data Sharing Toolkit

Module 4

Sharing Personal Data

Case Study 4

Credit: Chhor Sokunthea / World Bank, WWARN site

1. (2020) Elizabeth Pisani, Stella Botchway, 'Learning from the pioneers: lessons about data platforms drawn from the WWARN experience.' Accessed October 2020. https://wellcome.figshare.com/articles/Learning_from_the_pioneers_lessons_about_data_platforms_drawn_from_the_WWARN_experience/4476308/1

Sharing personal data

The Worldwide Antimalarial Resistance Network (WWARN) turns 10 this year (2020). Over the past decade, they have navigated the transition from 'data sharing [being] a swear word' to having gathered, curated and enabled access to trial data from over 180,000 patients.[1] By allowing researchers to access this data, WWARN has made a significant contribution to the fight against drug resistance for one of the world's deadliest infectious diseases – one which disproportionately affects the world's poorest people.

The challenge

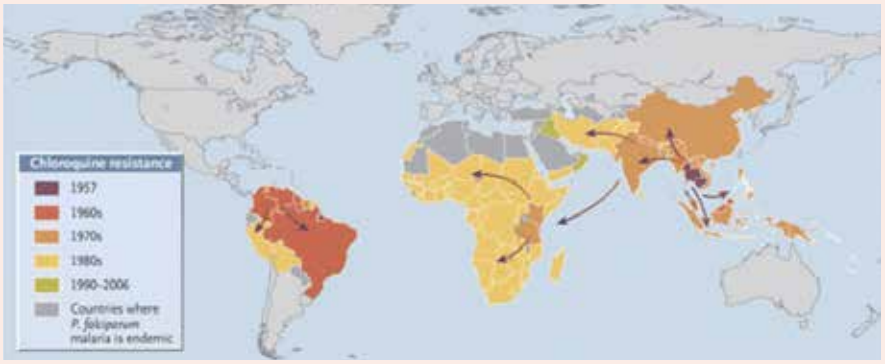
In 2018, Malaria infected 228 million people and killed 405,000. Antimalarials developed in the 1920s (chloroquine) were highly effective against the disease up until the 1960s, when malaria developed widespread resistance. In the 1970s, new artemisinin-based combination therapies (ACTs) developed in China were effective at combating the disease, however, there are now concerns that malaria is again developing

resistance. Resistance is the most likely explanation for a doubling of malaria-attributable child mortality in eastern and southern Africa between 1990 and 1998.[2] Concerns over an increasing lack of response to new forms of treatment led to the formation of WWARN in 2010.

Access to data is key to understanding how malaria, its effects on people and its resistance to drugs is developing. By taking and combining patient data from lots of different sources it is possible to produce new insights.

However, the sharing of malarial patient data has long been a contentious issue, raising logistical and ethical concerns related to the management and use of this type of personal data. WWARN needed to establish mechanisms which would make malaria researchers feel comfortable and able to share patient data, whilst also protecting the subjects of patient trials from harm.

[2] (2003) Eline L Korenromp, Brian G Williams, Eleanor Gouws, Christopher Dye, Robert W Snow. 'Measurement of trends in childhood malaria mortality in Africa: an assessment of progress toward targets based on verbal autopsy'. Accessed October 2020. https://www.sciencedirect.com/science/article/abs/pii/S1473309903006571



The solution

WWARN is a collaborative research network made up of 282 partners from a global range of research institutions. Hosted at the University of Oxford, the network's core team comprises an internationally based Secretariat, six Scientific Groups and four Regional Centres. The network is connected to the broader malaria research community through a Board and Scientific Advisory Committee. The network is funded by a number of institutions including: the Bill & Melinda Gates Foundation, Exxon Mobil, Initiative 5%, Expertise France, UKAID, The Wellcome Trust and The Global Fund.

Logistical issues and solutions

At the start of the project, WWARN had three options for where the management and location of the database would sit: the World Health Organization (WHO), a stand-alone non-profit organisation or an academic institution.[3]
An academic institution was selected due to a perceived lack of potential baggage to the funder and founders and because it would not be tied to the WHO's recruitment process or come with a large bill in creating a governance structure and organisation from scratch. Of academic institutions, Oxford was chosen due to its relative geographic accessibility and pre-existing links to malaria research. The repercussions of this choice are discussed in the lessons learnt section of this case study.

[3] (2020) Elizabeth Pisani, Stella Botchway, 'Learning from the pioneers: lessons about data platforms drawn from the WWARN experience.' Accessed October 2020. https://wellcome.figshare.com/articles/Learning_from_the_pioneers_lessons_about_data_platforms_drawn_from_the_WWARN_experience/4476308/1

The Terms of Submission for data into WWARN are articulated in full on their website.[4] A sample Case Record Form, which complies with baseline standards developed by the network, is also published on the site.[5] This includes variables such as:

- Date of birth
- Race
- Pregnancy and HIV status
- Previous medical history
- Current symptoms and their frequency
- Results of diagnostic tests (such as blood and urine tests)

Data must be collected in accordance with any laws or regulations, including 'without limitation in the country of origin', however, it is not clear how this is monitored or supported from the information provided in the Terms of Submission. Data contributors, those researchers and organisations contributing data to the network as identified in Terms of Submission, are responsible for submitting anonymised personal data – where the definition for personal data comes from the UK Data Protection Act 2018[6]. WWARN offers support and guidance for this process via an email address provided on a different part of their site, and monitors data contributions to ensure compliance. According to the Terms of Submission, the network moderators 'take appropriate technical and organisational measures to protect the security and confidentiality of Data'. The limitations of this are discussed below.

[4] (2020) University of Oxford, 'WWARN Data Platform Terms of Submission'. Accessed October 2020. https://www.wwarn.org/sites/default/files/attachments/documents/wwarn_terms_of_submission05may20_-_final.pdf

[5] (2018) Worldwide Antimalarial Resistance Network, ' Malaria Case Record Form'. Accessed October 2020. https://www.wwarn.org/tools-resources/procedures/malaria-case-record-form-crf

[6] (2018) UK Government, 'Data Protection Act 2018',. Accessed October 2020. https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted

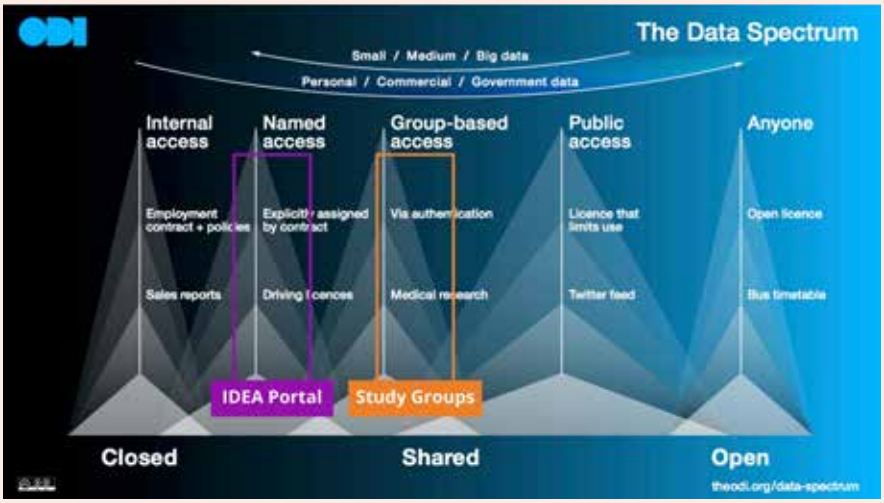
By mandating anonymisation, WWARN helps to protect patients. However, support in this process, and the standard to which data needs to be anonymised, could be actively provided to help contributors more readily understand what is necessary when contributing data. As the research necessitates some personal data, it should be shown by the network to what degree and how contributors should omit or change data without compromising aims. The network does make offers of support broadly, both at the end of the Terms of Submission and on the site itself, however, someone who is contributing data for the first time may struggle to readily understand what is expected of them in terms of anonymising data. This means those who are contributing data for the first time may have to spend time independently researching how to manage the data they are contributing to ensure it is in the right format for the network. However, as the same contributors return over time and become more familiar with the needs of the network and how to manage their data to comply, this problem has the potential to gradually resolve itself.

Ethical considerations and data misuse

WWARN developed two mechanisms to minimise the potential for misuse when sharing personal data. The first, the WWARN portal, has a FAIR compliant access protocol. Although the data is not openly published for anyone to access, the protocol for potential use is clearly articulated within the Inventory Data Explorer Application (IDEA), in brief: An inventory of data on the portal is viewable by anyone, a request for specific data is made via a form, which is then approved by an independent data access committee or by the data contributors themselves.[7] The data contributors dictate whether it is themselves of the data access committee who makes this decision.

[7] Worldwide Antimalarial Resistance Network, 'WWARN Data Inventory'. Accessed October 2020. https://app-live.wwarn.org/DataInventoryExplorer/#0

Data contributors are notified of data requests and within a set time period can opt out, request that the potential reuser contact them directly, or delegate the decision to the committee. By making the data available upon request, WWARN can vet who is using the data and what for. This may help contributors feel more comfortable sharing data with the network, however, depending on to what degree the data has been anonymised, it may not be necessary to protect its subjects. This is a mechanism for 'Named Access' in the data spectrum below, and has a relatively high administrative burden for the network.



The data spectrum demonstrates that the openness of data is not binary – data is not just 'fully open' or 'fully closed'; there are a host of intermediary gradations to the ways data can be shared. For example, Named Access is access explicitly granted by a contract, usually on an ad hoc basis, while 'Group-based Access', is where data can be accessed by a number of individuals or organisations according to their membership to a certain group. WWARN makes use of both types of data access.

The second method for access to patient-level data is through Study Groups. These are WWARN facilitated collaborations between researchers, centred around a specific topic of focus, that provide access to pooled data relevant to the aims of the research. This eliminates the need for ad hoc data requests via the portal, thereby reducing the logistical challenges to data access encountered by both the researcher and the network.

As well as from being less administratively intensive, this mechanism has further advantages. Researchers are more likely to contribute data if they are also receiving access to relevant data around their research topic. There is an increased incentive when collaborations and all data contributors are guaranteed acknowledgment in subsequent publications. Additionally, by data contributors from Malaria endemic countries becoming part of Study Groups with those from high, collaboration between researchers with acute and relevant knowledge is fostered in historically low resourced settings, where they are needed most.

Lessons learned

WWARN was originally developed as a surveillance portal for drug resistance, where researchers would share live data from areas where malaria was showing resistance to treatment. However, as those developing the network conducted more research into the data ecosystem, they found that where the main contributors of data would be researchers. Researchers who would be unlikely to share ‘live’ data which has not yet been mined for publications. Additionally, as the WHO, a key stakeholder in the project, already had a project which they felt occupied this niche, WWARN

instead became a network for researchers to share clinical trial patient-level data. The network increased the success of the initiative by taking steps to understand their key partners and those contributing data, and being flexible enough to adjust the scope to suit their wants and needs.

During some initial work with external organisations, the network had problems with materials being broadcast which were difficult for potential contributors to understand. There was very little engagement from the people who were sent these materials, and most refused to share data when greeted with such 'forbidding language', further reinforcing the message that, where possible, simple language is best to foster trust.[8]

From the start, data use agreements were clearly worded and not overly legalistic. The network hired communications specialists, who were initially seen by researchers to be unnecessary. However, their impact became clear as the network established itself. As one employee said 'The scientists want the data, the policy-makers need the research evidence, but neither spend as much time worrying about how they will receive it. Communications is the glue that binds the whole enterprise together.' The communications team were integral to the simplification of the Terms of Submission, which clearly state that 'Neither WWARN, IDDO nor the University of Oxford will claim any rights of ownership (including intellectual property rights) over the Data submitted. Data Contributors will continue to own and have full control over the Data they submit'. By keeping the language simple and clear, the network ensured that their vision for the rights of the data contributors was readily understood.

[8] '..will accept no liability or responsibility for any loss or damage arising out of any failure on the part of Oxford to comply with the restrictions on access to data set forth in Annex 2 hereto in respect of data (including data generated from blood samples) and/or the provisions contained in Annex 2…' (2020) Elizabeth Pisani, Stella Botchway, 'Learning from the pioneers: lessons about data platforms drawn from the WWARN experience.' Accessed October 2020. https://wellcome.figshare.com/articles/Learning_from_the_pioneers_lessons_about_data_platforms_drawn_from_the_WWARN_experience/4476308/1

What isn't clear from the Terms of Submission, however, is the process around the consent of trial participants and what the legal rights are of the researchers and network to share this data. Although there exists some guidance on the site for this, it isn't necessarily clear to which trials this applies and is not signposted in the Terms of Submission.[9]

WWARN felt that the governance of the network must be transparent, equitable and flexible. Initially, the choice of Oxford had a negative perception for researchers, particularly in endemic countries; 'We certainly thought of it as just one big Oxford data-grab,'[10] said one. A former WWARN employee also reflected: 'For the national programmes, the question of data ownership is very real. More than once I heard people say: why are we giving African data to England?'. WWARN strove to mitigate this through a number of strategies described below.

The openly published access protocol was an attempt to demonstrate the transparency that is attractive to researchers, whilst keeping patient level data sufficiently closed to protect its subjects and those who contribute. The development of Study Groups also helped to minimise distrust and incentivise data contributions. There was the clear gain of access to pooled data around research topics which wouldn't have ordinarily been available, which was again, balanced against the potential risks to the patients. The presence of the Study Groups also minimised the administrative burden on WWARN staff by reducing ad hoc requests for individuals who are all researching the same topic.

[9] (2018) Worldwide Antimicrobial Resistance Network, 'Informed consent' Accessed October 2020. https://www.wwarn.org/tools-resources/procedures/informed-consent

[10] (2020) Elizabeth Pisani, Stella Botchway, 'Learning from the pioneers: lessons about data platforms drawn from the WWARN experience.' Accessed October 2020. https://wellcome.figshare.com/articles/Learning_from_the_pioneers_lessons_about_data_platforms_drawn_from_the_WWARN_experience/4476308/1

That being said, some researchers, particularly those from low resource, malaria endemic countries, feel WWARN does not engage with them enough, especially when it comes to clarity around how to become involved in a Study Group. One endemic country researcher articulated his frustration with the process: 'My own feeling is that the communication has not been clear, how to be part of a Study Group other than just contributing data. What is missing is information, and when information is missing… people become suspicious…. Then because of a simple lack of information, a good thing gets killed'[11]. As these researchers are on the frontline of the fight against malaria, with the highest degree of understanding of the context, they may be able to provide Study Group focuses with more relevance to the situation on the ground than those in Oxford, for example.

WWARN is not in a position to provide financial incentives for researchers – which some feel is the the most appropriate way to be rewarded for their data – however, by articulating how Study Groups are formed with further clarity, and more clearly demonstrating it as an option available to all, WWARN may be able to engage more thoroughly with individuals and institutions essential for advancement in the field.

[11] (2020) Elizabeth Pisani, Stella Botchway, 'Learning from the pioneers: lessons about data platforms drawn from the WWARN experience.' Accessed October 2020. https://wellcome.figshare.com/articles/Learning_from_the_pioneers_lessons_about_data_platforms_drawn_from_the_WWARN_experience/4476308/1

Impact

More than 120,000 individual patient records have been contributed to the WWARN data repository so far. As the network itself is predominantly researchers, their promoted impact is via numbers of publications, of which there have been 137 since 2016 (April, 2020). Through these studies, WWARN have developed tools to track medicine quality, and contributed to an increased understanding of what factors affect an anti-malarial working, often leading to policy change. For example, in 2016, following a meta-analysis produced by WWARN, researchers found that one group of drugs were not as effective in children. Subsequently, a Study Group rapidly changed focus to identify what would happen if children were given the drug in higher doses. Following their research, the WHO altered the dosage recommendations. This is evidence of pooled analysis of well-curated, shared data leading rapidly to a targeted clinical trial. An urgent question was answered, and policy quickly changed. The result will be fewer treatment failures and more healthy children. Through their quality assurance programmes and Study Groups, WWARN have developed tools and strategies to monitor drug quality, reducing the probability that patients are taking ineffective medicines, thereby getting healthier, faster. To date there have been 15 Study Groups successfully run and closed, all grouped around essential topics such as these.

Future Plans

There are 20 active Study Groups currently conducting life saving research, such as into genetic factors that might help malarial resistance, thereby identifying groups most likely to not be able to find drugs that work, and into the impact of malaria on infants during pregnancy, at birth and into infancy.

Additional future plans include:

Continuing to improve and refine their approach to data sharing. New tools will support research partners to create good quality data that can be more easily shared and analysed by the global health community.

Continuing to release tools and deliver workshops to build capacity in effective data collection and publication.

Adapting the WWARN model for other diseases, through the IDDO. WWARN managers said the development of functional pilot platforms for schistosomiasis, Ebola and visceral leishmaniasis has cost a great deal less than the original malaria platform.

Further resources

If you are planning to share personal data in a similar way to this case study there are resources in the Data Sharing Toolkit to help you do this and avoid the same challenges:

- Module 4 – Protecting individual's rights when sharing data
- Guide: Protecting people – managing risk when handling personal data
- Guide: Anonymising data in agriculture
- Guide: Deciding how to provide access to data

Data Sharing Toolkit

ACKNOWLEDGEMENTS

This document was authored by the Open Data Institute and CABI as part of a Bill & Melinda Gates Foundation funded investment.

The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the Bill & Melinda Gates Foundation or CABI.

datasharingtoolkit.org

DOI: 10.21955/gatesopenres.1116765.1

cabi.org | theodi.org | gatesfoundation.org

CABI Data Sharing Toolkit

open data institute

BILL & MELINDA GATES foundation