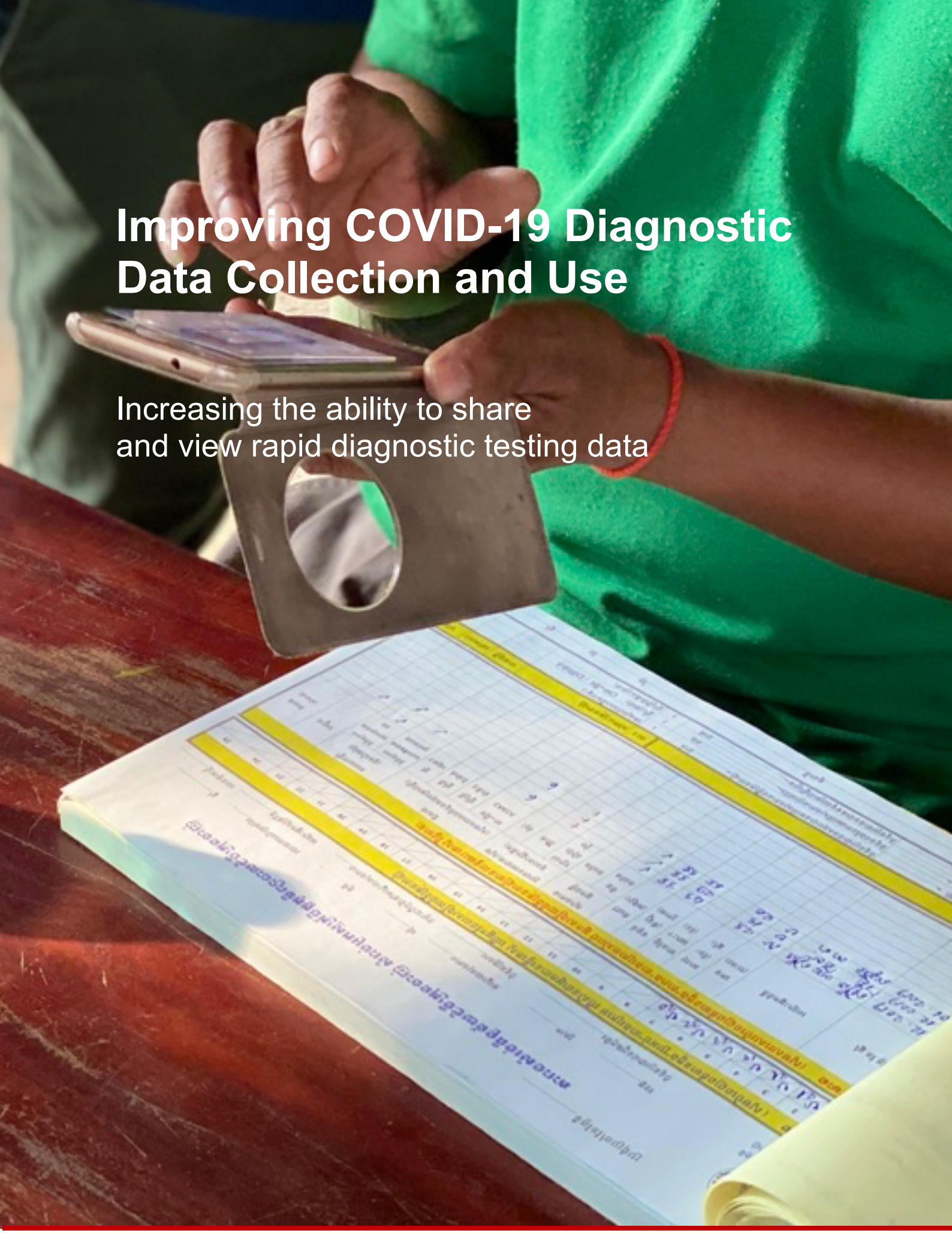


Improving COVID-19 Diagnostic Data Collection and Use

Increasing the ability to share
and view rapid diagnostic testing data

FIND
Diagnosis for all

PATH

Introduction

There are over **400 COVID-19 based digital health tools** available for country partners to choose from for varying use cases across the COVID-19 continuum of care including diagnosis, prevention, delivery of care, and surveillance.

Country implementors need to determine the most appropriate ways to deliver digital health interventions to collect diagnostic data that will support programme needs. In parallel, these solutions should provide evidence for more informed decision making to provide maximum global health impact, but implementors often lack a globally standardized approach to data capture and usage. This has become increasingly clear as COVID-19 has uncovered the need for cross-country data analysis, which is currently nonviable given different standards across countries for data collection.

FIND and PATH have been working together to support development and implementation of appropriate digital tools for global COVID-19 response. As part of its role as a co-convener of the Access to COVID-19 Tools-Accelerator (ACT-A) Diagnostics pillar[1], FIND has partnered with low- and middle-income countries (LMICs) to help implement digital tools in their public health responses. PATH is working across multiple ACT-A pillars to support the global COVID-19 response. Within the digital health space, PATH has designed a strategic framework for governments, investors, implementing organizations, and the broader digital health community to adapt existing digital tools during different phases of an outbreak. The Digital Square initiative within PATH's Center of Digital Excellence has convened a network of global collaborators to identify any redundant efforts, duplicate solutions, and piecemeal approaches to promote an efficient, collaborative approach to digital solutions for COVID-19. To kick-off this work, FIND conducted a desktop-based review of different digital tool repositories that included information on the current availability, maturity and status of any digital tools used by countries in their COVID-19 responses.

Background and Rationale

Poor-quality data can cause inefficiencies in patients' treatment cascades and affect the monitoring and improvement of programmes.

When global health implementors and programmes use differing data collection schemas, it is difficult to draw comparisons between the datasets and make informed decisions that are comparable with other programmes. For example, a country may be using a single, unified reporting system but has multiple digital tools across programmes sourcing different pieces of information to be transmitted. Therefore, analytics and aggregation of data do not match up, resulting in global health implementors not utilizing their data to its full potential.

The COVID-19 pandemic has presented an immediate need with focused opportunities for the development and adoption of digital health interventions. With this comes the challenge of ensuring that solutions are forward thinking and not unnecessarily fragmenting health systems. The use of standardised digital tools to screen and triage patients, track the number of tests being used in the field, and transmit those results allows for rapid clinical decision-making throughout the patient

[1] ACT-Accelerator diagnostics pillar: progress
https://www.finddx.org/covid-19/act-accelerator-progress/

care cascade. At a policy level, accurate real-time data strengthen surveillance systems and help guide the design and implementation of targeted testing strategies.[2] To combat "short-lived" pandemic-focused interventions, consistent thinking and a design pattern around system architecture and data are needed to sustainably be part of the larger health system. Even given a standardized basic approach, it is crucial to ensure that the functionalities and usability of digital health tools are developed specifically for a country, facility-type, or user's context which can be ensured by starting with a common data model and committing to development based on a human-centred design approach.

Approach and Methodology

To this end, FIND and PATH created a template of data fields specifically relevant to the COVID-19 RDT (rapid diagnostic test) workflow, called the "ACT-A COVID-19 ANTIGEN RDT DATA MODEL".[3] The intention was to accelerate the development and deployment of a set of minimum functionalities with a minimum dataset (designed specifically for the project), incorporating data exchange and format standards, for the collection of COVID-19 data.

During a health emergency, it is vital that stakeholders at various levels of the health system can leverage accurate and quality data of relevant diagnostics being used to indicate disease burden and resource mobilization. Our data model and pre-built vendor solutions based on the data model, enable software developers to create usable and interoperable digital health tools to support rapid deployment by global health implementors.

Using the ACT-A COVID-19 Antigen RDT Data Model or resulting pre-built vendor solutions, aids the speed of deployment and usability of COVID-19 digital health tools

This effort was aimed at supporting LMIC (Low- and Middle-Income Countries) partners in the collection of COVID-19 antigen RDT results at a decentralised level for use in patient care, disease surveillance, supply chain management, and as relevant, other use-cases to related to pandemic responses. The creation of the data model was based on the rationale that it is better to adapt or onboard an existing data capture module for an existing digital solution that is currently being used, if one already exists, as this allows rapid deployment with lower costs and minimal training for end users, strengthening the sustainability of the COVID-19 data collection solution.

The use of data exchange and format standards (e.g. HL7 FHIR (fast healthcare interoperability resource)) and data standards for codification (e.g. ICD-11) are key building blocks and design patterns that pave the way to a continuous care record for patients. This ensures data will be curated in a format that is more conducive to inclusion in a central health record and as part of the national population data set for analysis and reporting. These tools need to be carefully implemented to ensure the

[2] Surveillance in emergencies. (2021). Retrieved 6 August 2021, from https://www.who.int/emergencies/surveillance

[3] Request for Proposals (RFP) from digital health solution providers to enable COVID RDT data capture. (2021). Retrieved 6 August 2021, from https://www.finddx.org/wp-content/uploads/2021/04/RFP-Digital-health-solution-COVID-RDT-07APR2021.pdf

sustainability of the digital tool and its alignment with global and national IT and enterprise architecture policies.

A Common Data Model

What is a common data model, what do you do with it and why is it important?

A common data model helps to define the terminology that digital health programmes use globally for all their data sources, in this case for COVID-19 Antigen RDT results, as well as the relationships that exist between the different data items.

Using the ACT-A COVID-19 Antigen RDT Data Model or resulting pre-built vendor solutions, reduces development time for software developers over building a digital solution from scratch

Based on this conceptual standpoint, programmes can map their application data definitions to the common data model, which provides equivalent fields between different disease programmes, allowing both interoperability of systems, data, and the analysis of consistent metrics.

While the data model contains Protected Health Information these data can easily be deidentified, allowing the sharing of aggregate results in a safe and efficient manner and allowing them to be used for programmatic analysis without the risk of re-identification. This can help to reduce programmatic costs and improve workflow and day-to-day efficiencies.

Using a standards-based approach to transfer data between systems strengthens the ability to report diagnostic results and make informed programmatic decisions

The goal of harmonized data models and exchange formats is to see systems continuing to build towards strengthening a holistic Health Information System and avoiding fragmentation of health systems.

What does the Data Model Look Like?

The data model is freely available to download, adapt to context, and use here.

Diagram showing the data model which allows for easy customisation:

What Fields are Collected and Why?

The following data fields are available within the common data model to evaluate and optimise programmatic performance as well as inform decision making and allow for maximum impact:

Field	Description ***Reason for collection***
Patient_ID	Unique, system generated patient ID *To identify a specific patient easily*
Patient_First_Name	Patient's first or given name *Patient identifiers*
Patient_Family_ Name	Patient's last or family name *Patient identifiers*
Patient_Birth_Date	Patient's date of birth capturing day, month, and year of birth *Patient identifiers*
Age	Estimated age in years of the patient, captured if patient_birth_date is unknown *Patient identifiers*
Sex	Patient's biological sex at birth, either male or female *Patient identifiers*
Contact_number	Number where patient may be reached *Patient identifiers*
Address	Patient's address including street name, district/county, city and region *Patient identifiers*
Session_ID	Unique, system generated case ID to capture instance of patient's visit; each patient could have multiple Session IDs *Session IDs allow the number of patient visits to be counted rather than the total number of RDTs run*
Administrator_ID	Unique ID of the health worker or practitioner who is conducting the test *To identify the healthcare worker (HCW) performing the test*
Administrator_Name	Name of the health worker or practitioner who is conducting the test *To identify the HCW performing the test*
Test_ID	Unique, system generated instance of a test; each Session_ID could have multiple test numbers *Session IDs allow the number RDTs run to be counted*
Test_Location	Geocodes of where test is being conducted, collected from device *Allows mapping of tests by location, using a map*
Test_Setting	Type of facility where the test is being performed *Allows disaggregation of tests settings, by type, e.g., Health Centre, Community, etc.*
Test_Reason	Indicates reason for performing test *Allows disaggregation by reason for testing, including symptomatic or non-symptomatic*
Symptoms	Details of symptoms patient is exhibiting *Allows comparison of the most reported symptoms associated with COVID-19*
Days_Since_ Symptoms_Began	Duration for which symptoms have persisted

	Provides an indication of how far into the symptomatic or isolation period the patient may be
Specimen_Type	Indicates how biological sample was collected *Allows disaggregation of different specimen types and also confirmation that the right specimen has been used for the specific RDT test used*
Specimen_Type_ Other	Indicates how biological sample was collected if "other" was selected in "Specimen Type" *Allows disaggregation of different specimen types, if the brand is not entered in the 'Specimen_Type" dropdown field*
RDT_Name	Device/RDT test model name as assigned by the manufacturer *Allows disaggregation of different RDT brands*
RDT_Name_Other	Device/RDT test model name as assigned by the manufacturer if "other" was selected in "RDT_Name" *Allows disaggregation of different RDT brands, if the brand is not entered in the 'RDT_Name" dropdown field*
Syndrome_Test_ Type	Indicates type of test being conducted / what the RDT is testing for - HIDDEN Field - Pre-populated - (COVID-19) *Allows easy disaggregation of data, when stored with similar data from other disease states*
RDT_LOINC_Code	HIDDEN Field - Pre-populated - (94558-4) *Allows easy disaggregation of data, when stored with similar data from other disease states*
RDT_LOINC_Long_ Common_Name	HIDDEN Field - ("SARS-CoV-2 (COVID-19) Ag [Presence] in Respiratory specimen by Rapid immunoassay") - Allows easy disaggregation of data, when stored with similar data from other disease states
RDT_Lot	Lot or batch number to which the individual device belongs- Allows easy identification of quality control issues, if a high number of invalids or indeterminates were seen, the reports would show they all came from the same Lot of RDTs
RDT_Expiry_Date	Expiry date of the device with respect to the lot number *This field allows for reports to show if RDTs are being used past their expiry date*
RDT_Expired	Message Popup Screen/Message only - No data to be filled *This field is just used to show the user a warning is the date of the RDT expiry is in the past. The warning would tell the user that they should use another RDT which is in date*
RDT_Start_Time	Time at which test was begun *This field is used to calculate the total running time of the RDT, to ensure users read the RDT in the reading window, not too soon or not too late*
RDT_End_Time	Time at which the result was read *This field is used to calculate the total running time of the RDT, to ensure users read the RDT in the reading window, not too soon or not too late*
RDT_Result	Final observation as evidenced by the test result *Positive, Negative, Invalid or Indeterminate are included as standard, allowing dissemination by RDT Result type to get an accurate indication of disease positivity of the patients presenting for testing*

Image_of_RDT	Image of the RDT device that clearly shows test result from which the result can be inferred *This field is useful if you wish to collect the images for quality control issues, or user training, if a high number of invalids are seen, the images can be checked. There are multiple uses for this field*
Repeat_Test	Indicates whether the test being conducted immediately follows another, in the event of a failed or inconclusive test *Allows an HCW to quickly process a new RDT without re-entering patient details, setting context or symptoms*
Repeat_Test_ Reason	Indicates reason for repeating test *A repeated test should have a clinical reason for the repeat test reason*
Repeat_Test_ Reason_Other	Allows user to capture additional details for test being repeated *Allows an HCW to enter additional details regarding the need to perform a second RDT*
Facility_ID	Unique ID assigned to the testing facility
Facility_Name	Name by which the facility is locally identified
Facility_Address	Address of the facility capturing district/county, city, region
Additional_Notes	Any other information of relevance - open field

What do I gain by adopting this data model?

Adopting this data model will save you time and ensure you are collecting data consistent with other programmes, reducing costs and development stress.

How would you integrate two different data sources, e.g., two different programme databases, if the two databases have different columns or headers, or even different data types?

Using a common data model alleviates these problems by simplifying the incorporation of data from different sources into a single, unified source in a single database which can then be aggregated as required.

How might we bridge the gap between different functional domains where data is being collected to accurately triangulate information?

Clinical, programmatic, laboratory, and supply chain reporting often have divergent or incongruent views of the data. Being able to support semantic interoperability between systems improves exploration and better insight. Overall, electronic registries have been shown to improve overall system performance and automation of aggregate data reporting, therefore standardized data collection and reporting mechanisms will incrementally add value throughout the health system organizational structure.

What if we don't want to build our own solution or want an already validated solution?

Easy, use one of the of existing data collection and reporting applications or the modules that have already adopted the data model. This allows you to concentrate your efforts on integrating a comprehensive digital health tool into your workflow.

We created COVID-19 RDT Data collection modules for well-established vendor applications in use in LMICs, which utilise the ACT-A COVID-19 ANTIGEN RDT DATA

MODEL in their respective parent applications and allow users of those vendor applications to easily collect COVID-19 test data.

The five developers and platforms are:

1. COMMCARE
2. DHIS2
3. RDT DATA CAPTURE MEDPLAT
4. ODK – OPEN DATA KIT
5. COMMUNITY HEALTH TOOLKIT

Links to vendor solutions:

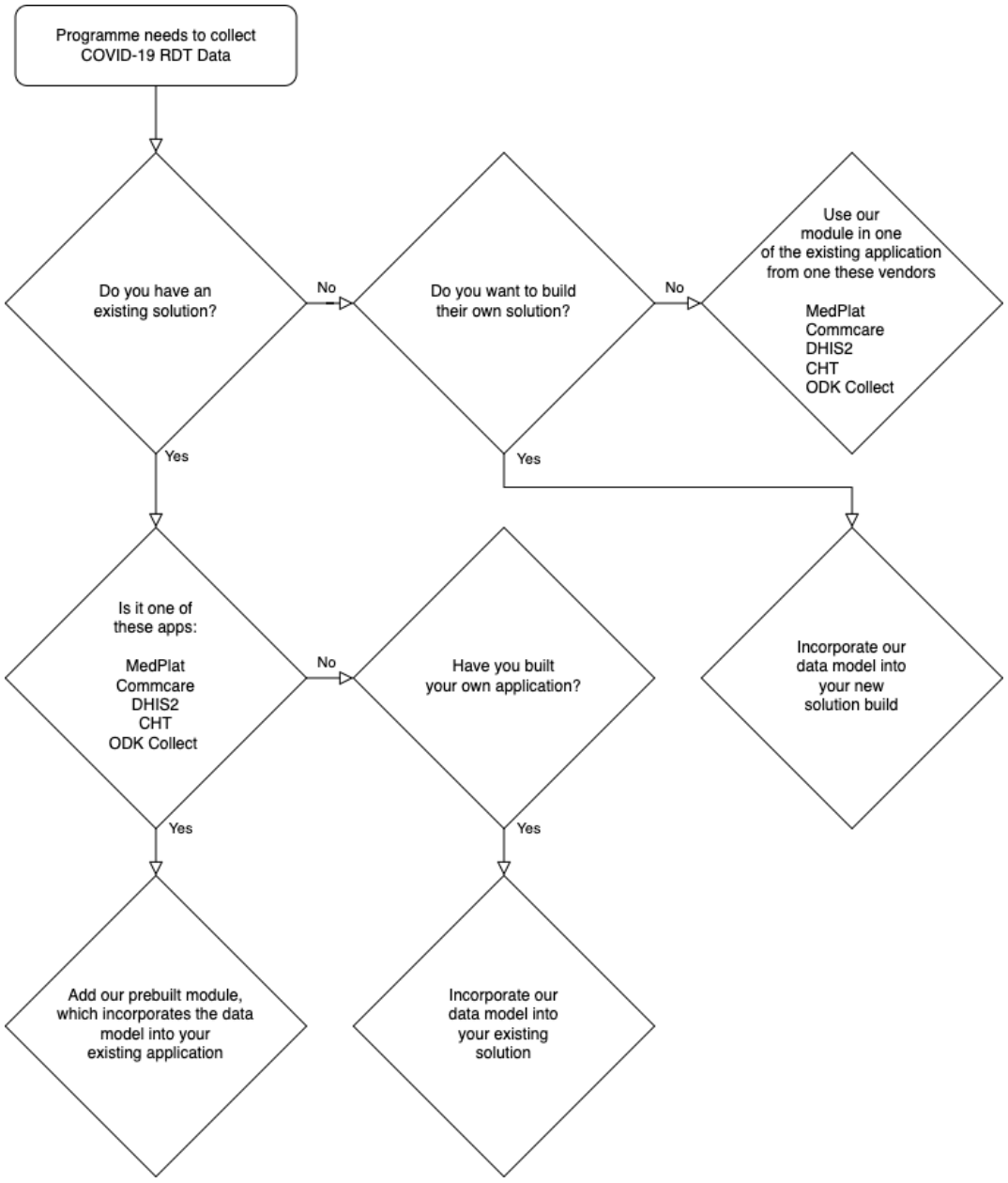
- Dimagi
- HISP India
- Argusoft
- ODK
- Medic

By adopting the ACT-A COVID-19 ANTIGEN RDT DATA MODEL or one of these applications, you ensure that the data you collect is consistent with other programmes and systems in addition to working to make the most use out of it in a global context.

What if you have an existing data collection application and haven't already developed your own COVID-19 RDT data model?

Using our common data model within your own application, or as a template to start the development of your own application, will ensure you cover all the bases before you introduce additional data fields to satisfy your own programmatic needs. Our solution is specifically built to ensure consistency of mapping for future data interoperability. Of course, if you have an existing data model you can instead align any aspects of our model to fill any gaps you may have.

Where are you in your COVID-19 Data collection journey?



Benefits of using the data model

Using the ACT-A COVID ANTIGEN RDT DATA MODEL or resulting pre-build vendor solutions will: save you time and resources, ensure you are collecting data in the same format as other programmes, and allow you to share and visualise high-level aggregate data.

The data model and range of pre-built vendor solutions developed by FIND and PATH are available to any software developers, global health implementors, or other interested stakeholders. Using this inceptive work, you can choose to onboard our common data model to integrate it into your own programme or have us assist you in building a bespoke application, or even onboarding one of the pre-built vendor solutions to enable your programme to start collecting COVID-19 RDT data today.

Reduce development expenses and the need for technical resources by repurposing one of our existing solutions to ensure rapid deployment and maximum impact from your COVID-19 diagnostics data collection programme.

While the data model is COVID-19-centric in its current form, it offers a framework which is flexible and reusable. This allows it to be disease-agnostic in the future through simple adaptation with minimal changes to suit different health areas, including Malaria, HIV, Dengue, or any other disease where RDT data collection is required.

Start on your Path Today

The ACT-A COVID-19 ANTIGEN RDT DATA MODEL was designed by FIND with contributions from PATH to align with existing systems and fast healthcare interoperability resource codes, with additional contributions from Global Fund to ensure that the model was truly reusable and disease agnostic.

The impact of the data model will only be fully unlocked with thoughtful implementation and integration. To ensure the appropriateness and sustainability of the data model, PATH and FIND will be incorporating the data model into more comprehensive health care worker facing applications and conducting human-cantered design to inform expanded functionality and design improvements. PATH and FIND will then implement the resulting tools in two geographies. The journey to implementation and the lessons learned along the way will provide a blueprint for others interested in doing the same. As such, we aim to create an implementation guide to capture our experiences, best practices, and lessons learned in the hopes of supporting other implementers. To stay informed or to provide feedback, please email **nick.banks@finddx.org** or **abpatil@path.org.**

If you are interested in using the COVID-19 Antigen RDT common data model or resulting pre-built vendor solutions, then a discussion with the FIND and/or PATH people would be the best place to start.

Investigate the inclusion of the data model in your application or platform and provide us feedback on what works and doesn't. As we look ahead, let us know how your organization could use this data model to support work across the health system.

The ACT-A COVID-19 ANTIGEN RDT DATA MODEL was designed by FIND with contributions from PATH to align with existing systems and FHIR to ensure that it was truly reusable and disease-agnostic.