Aggregation of tracker data values

Approach & technical solution for pushing DHIS2 tracker data to aggregate datasets

### Outstanding issues

* There is a bug in 2.33.3/2.33.4 that prevents us from creating this custom attribute; not exposed in the UI ([JIRA 8755](https://jira.dhis2.org/browse/DHIS2-8755)): for now must insert the attribute via the API.
* There is a bug ([Jira 8868](https://jira.dhis2.org/browse/DHIS2-8868)) that causes metadata-dependency-export tool to fail when custom attribute assigned to program indicator
* Decide on use of codes or UIDs for linking PIs to data elements and categoryoptioncombos. The example script uses codes for data elements, uids for categoryoptioncombos.
* Improving the script so that it will at a minimum notify administrators if there is an error and/or sending a DHIS2 message with the import summary.

## *There could also be other and more advanced scripts out in the community that can be considered.*

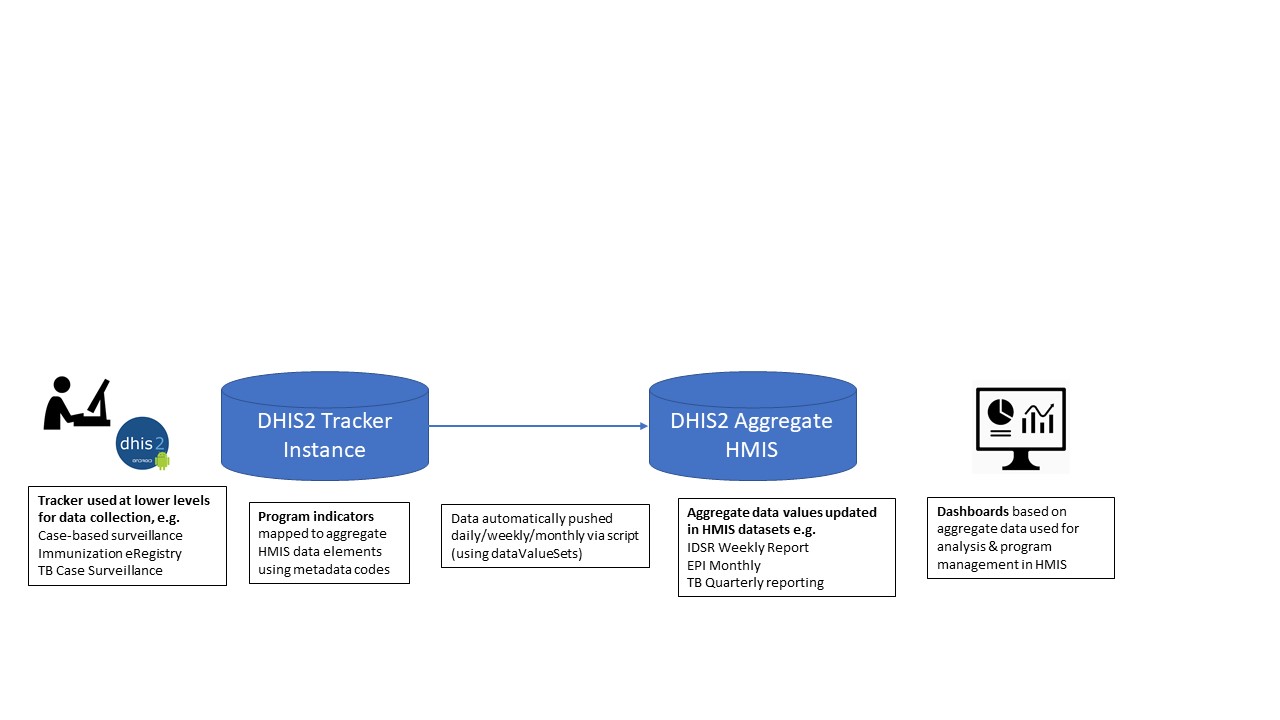
## Use Case

For countries implementing DHIS2 tracker for individual-level data collection, a separate dedicated DHIS2 instance is recommended for tracker deployment. Many countries have a mature, stable HMIS used primarily for capturing aggregate data across health programs in an integrated environment. By maintaining separate tracker and aggregate DHIS2 instances, performance can be better managed by system admins and data governance principles can be applied to ensure personally identifiable data captured by Tracker can be protected according to national policies and governance frameworks.

There is a clear benefit in being able to leverage individual data collection through DHIS2 tracker to automatically ‘report’ aggregated data to the routine HMIS (e.g. to aggregate datasets, often monthly/weekly/quarterly reports from facility level). Capturing individual level data through DHIS2 tracker can improve the quality of the data reported into the routine aggregate HMIS, while also enabling ad hoc analysis of the tracker data in the Tracker Instance as required.

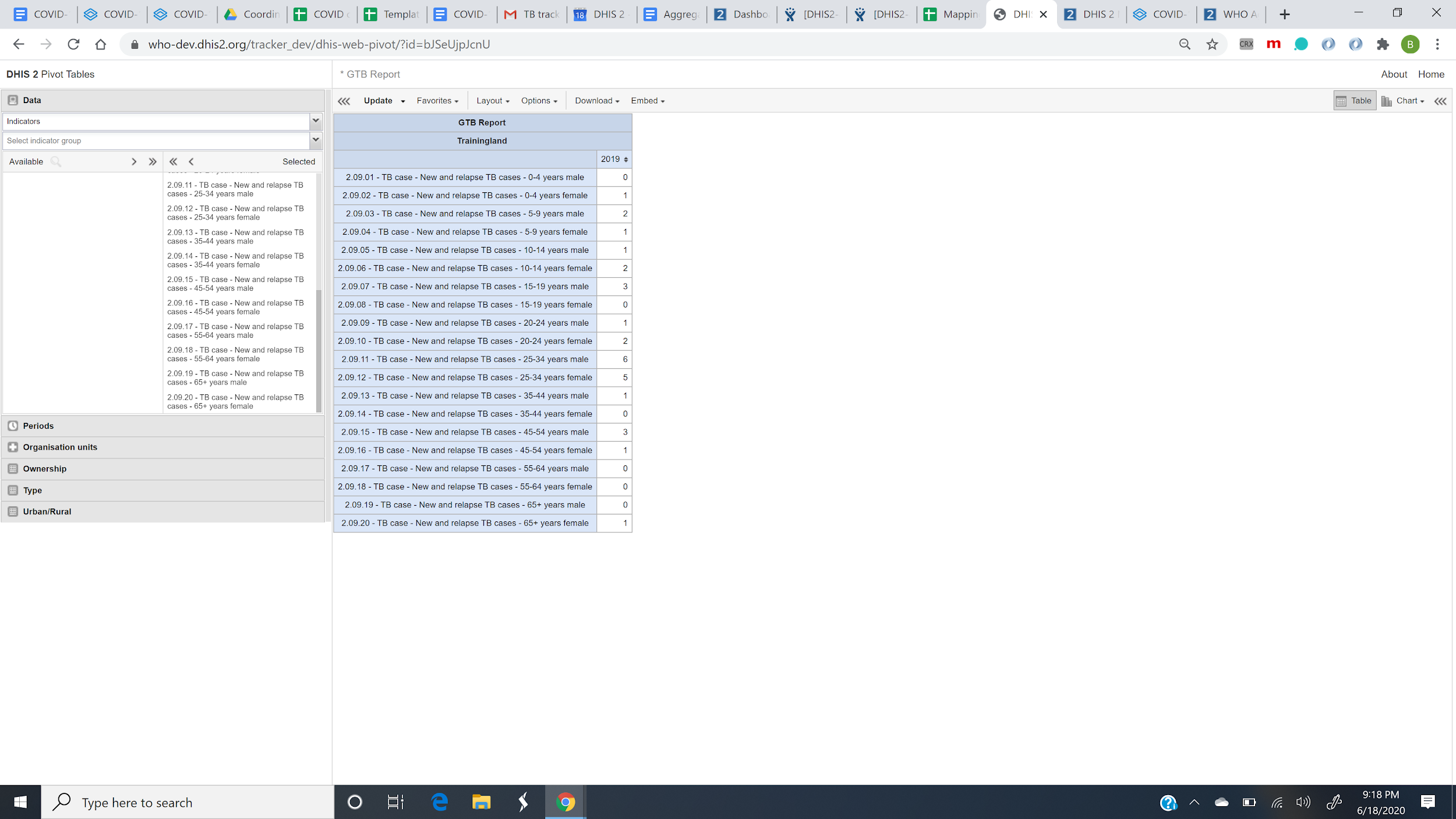
Approach

Individual-level data collected in DHIS2 tracker in one instance (or within the same instance) is aggregated via Program Indicators. Each program indicator is mapped to a corresponding aggregate data element (and category option set, as applicable) via a code. The program indicator value should represent the same value as the aggregate (e.g. <<number of new and relapsed TB cases notified>> or <<number of BCG doses given to children under 1>>). At a frequency agreed upon by health programmes and national HIS units, a script is used to push data value from the program indicators to complete data elements in corresponding aggregate datasets in the HMIS.

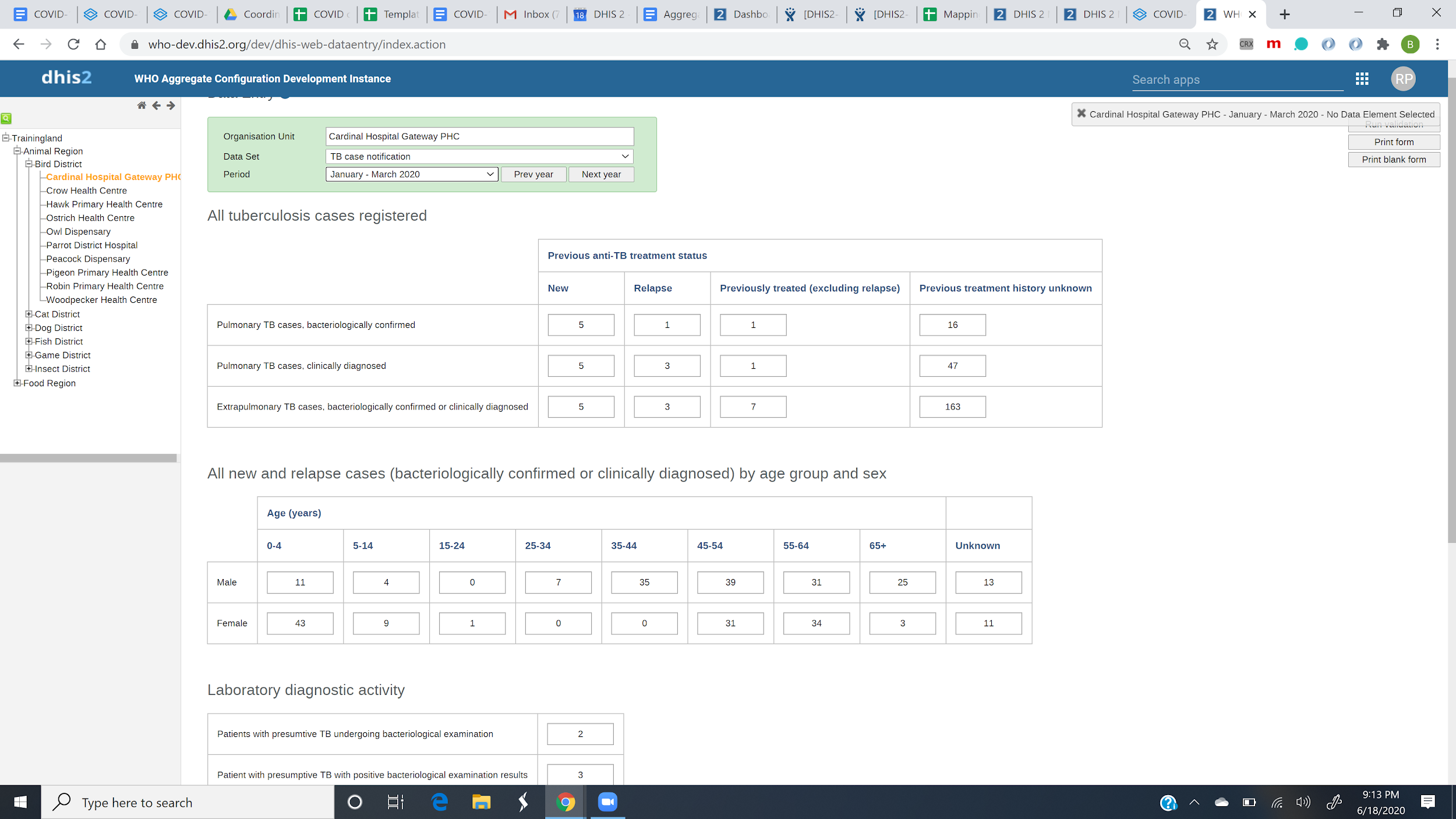


Example: TB Case Surveillance to aggregate quarterly TB report for TB notifications

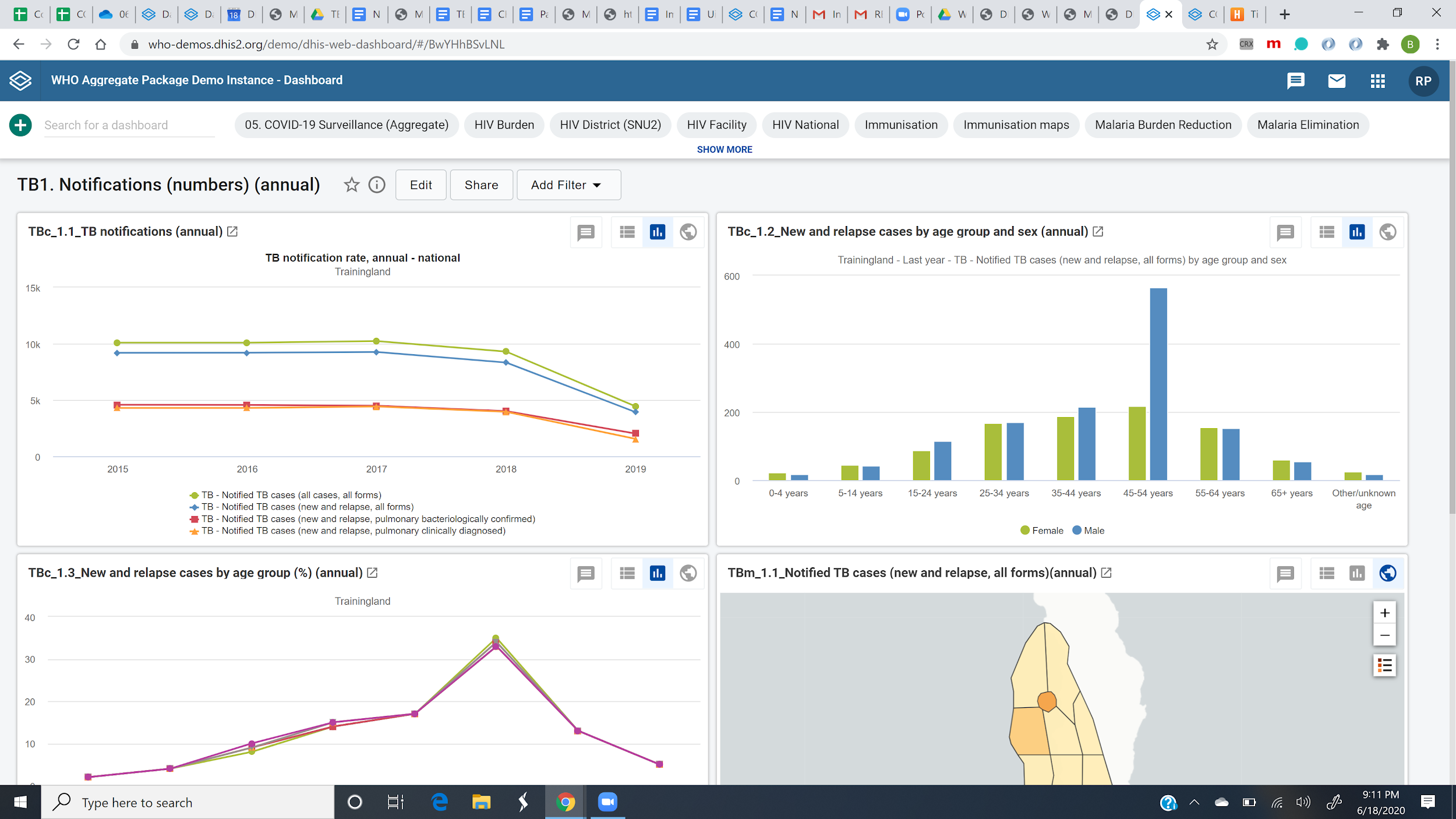
*Program indicator data values from tracker data (data elements/disaggregation can be produced, but not pivoted by dimensions as gender, age group)*



*Aggregate data entry (automatically filled by the data pushed from tracker program indicators)*



*Aggregate dashboard output (with ability to pivot male/female as a data dimension)*

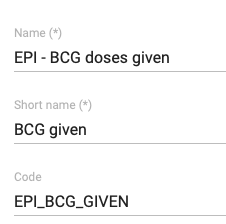
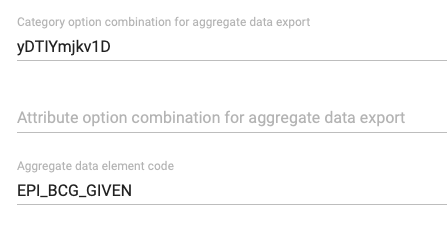


Technical Solution

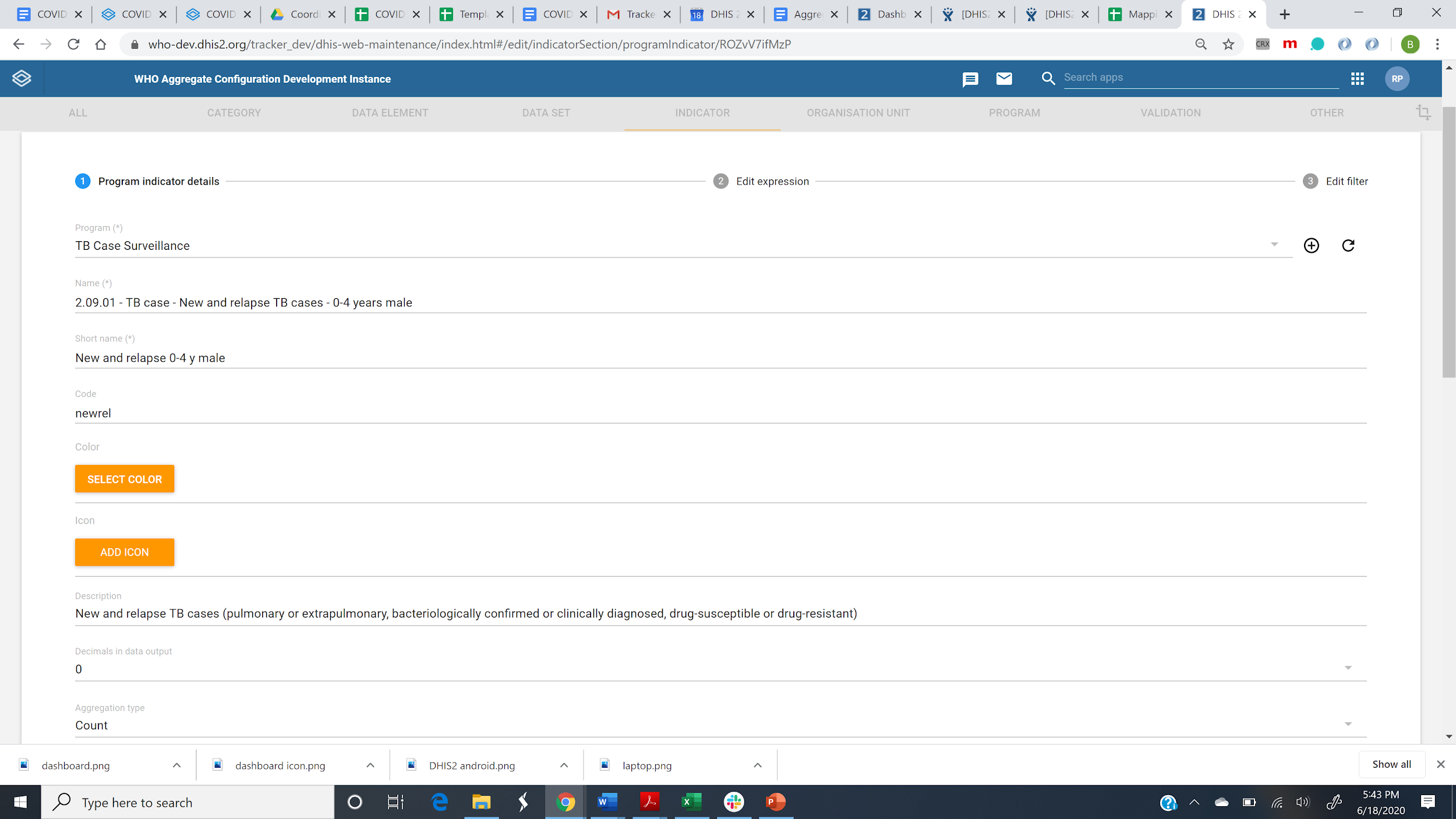
A solution for aggregation of the tracker data into periodic aggregated values has been created and initially tested using an EPI Tracker instance in Rwanda (using tracker as an immunization eRegistry at national scale) to send aggregated data to the monthly EPI dataset in the HMIS instance. The following elements are needed:

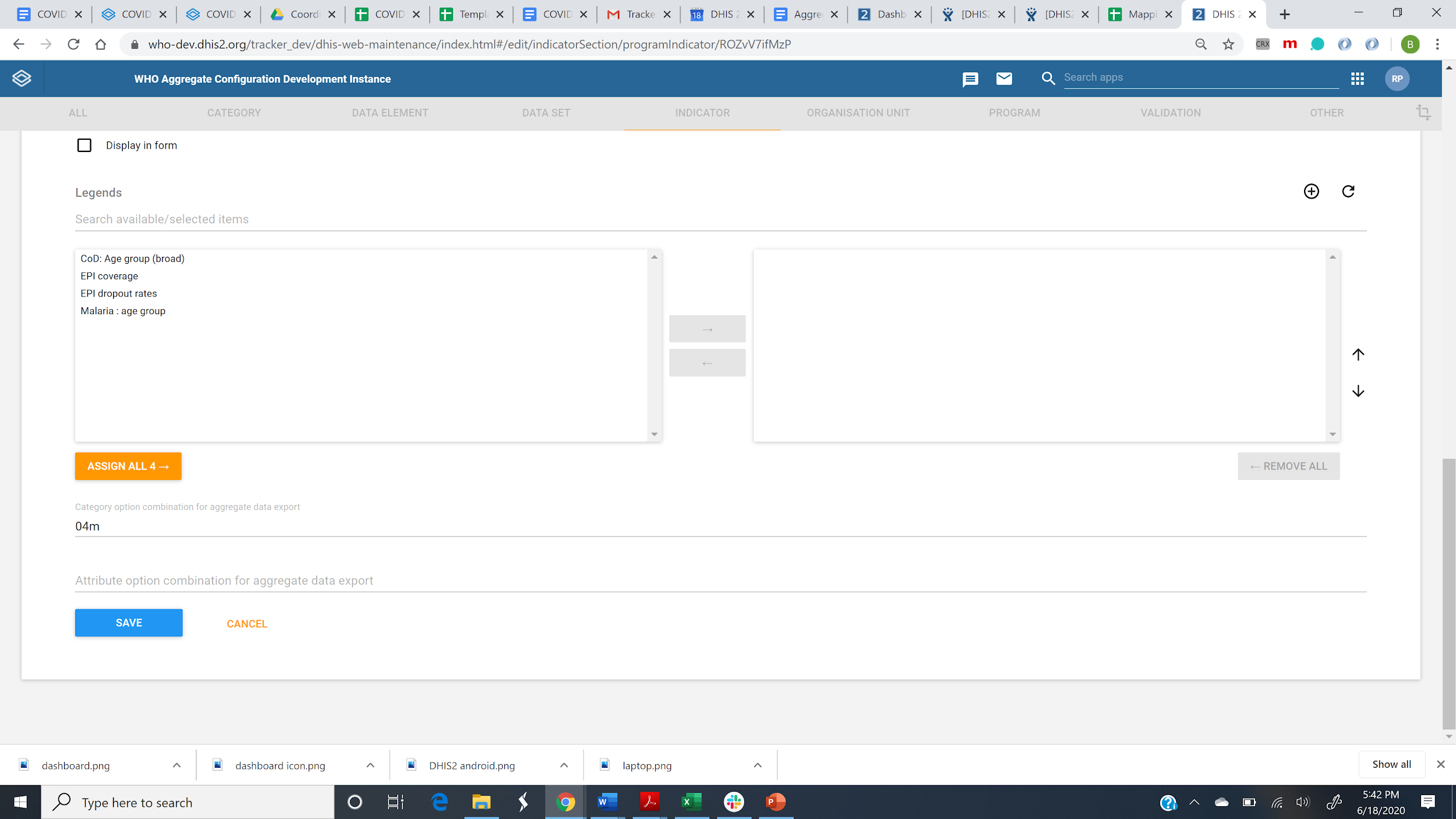
* A data set with data elements which aggregate value from tracker can be saved to.
* Program indicators for each of the data elements (including disaggregations), configured to correspond to the definition of target data element/disaggregation
* A custom attribute called “Aggregate data element code”, which is applied to the program indicators.
* A script that pulls (for example) relative (e.g. monthly, quarterly) values from the program indicators and pushes them back to the server as a (for example) monthly data value set.

Every aggregate data element and categoryOptionCombo must be assigned a code. The “Aggregate data element code” attribute is used to associate the data element code with the corresponding program indicator. For categoryOptionCombo, this field is already built in to the Program Indicator user interface.

*Data element (left) and program indicator (right).*

**

**

Using the analytics API *dataValueSet* option, program indicator data can be exported as aggregate data values and imported back in using the *dataValueSets* endpoint. A simple script for doing this is shown below. [replace with a link to script on GH that we can maintain?] 

## DHIS2 Digital Data Packages Support for Tracker → Aggregate

DHIS2 digital data packages have been developed to support both aggregate reporting & analysis, as well as tracker data capture and facility-level analysis.

Aggregate digital data packages (inclusive of standard aggregate dashboards) are available for health programmes such as TB, HIV, malaria, RMNCAH and disease surveillance. Aggregate packages include:

1. Data set, data elements and category option sets (‘target’ for sending tracker data)
2. Metadata codes that are AWX compliant and enable the mapping of data values from tracker to the aggregate ‘target’

In addition, tracker packages are being developed for a growing number of use cases such as immunizaion eRegistries and case-based surveillance for TB, HIV and integrated disease reporting. Where tracker data packages are designed to capture data that can be aggregated and submitted to the corresponding aggregate dataset, we have included the following in the tracker digital data packages:

1. Program indicators configured to produce data values corresponding to the data elements and disaggregations included in the aggregate digital data package data
2. Custom attribute for ‘Aggregate data element code’
3. Attributes per program indicator populated with the data element codes and category option combination code from the aggregate digital data package

A script for pushing the

## Implementation Considerations

Implementation considerations must be taken into account when an approach for automating the submission of tracker data to routine HMIS aggregate data sets. These include:

1. Accurate mapping of tracker program indicators to aggregate data elements
   1. Adherence to common, precise definitions
   2. Proper mapping of metadata codes between program indicators and aggregate data elements and category option combinations
2. Timing and frequency of the data transfer
   1. How often to do the aggregation, for example daily, or on day of HMIS reporting deadline etc. How often the aggregate data is used for analysis and decision making should also be considered.
   2. How many months back in time to update aggregate data for, for example last month, last 3 months, current and last month etc.
   3. Governance principles and national HMIS policies for locking datasets, timely reporting; How can values for ‘late reporting’ be updated in the aggregate dataset; managing discrepancies between tracker and aggregate data set based on the timing of data submission
3. Parallel and hybrid reporting flows
   1. Phased scale up of tracker: some facilities/districts adopting tracker sooner while others may retain the paper based aggregate reporting
   2. Parallel reporting: some countries necessitate a period of parallel reporting of paper-based and tracker data collection for a pilot period
4. Revisions to reporting flows
   1. Often need to be updated in HIS governance documents
   2. If not well communicated, the ‘source of truth’ can be obscured
5. Maintenance
   1. If metadata in the tracker instance or in the aggregate HMIS instance are altered, the mapping of program indicators to aggregate data elements may need to be updated