Instructions for using the Pathfinder tracking tool - Version 1.1

The Pathfinder tracking template has been developed to provide a standard format for use by pathfinder study teams to map and document the processes of host studies. It was designed as a simple method for tracking projects and can be adapted by each pathfinder team to meet their requirements.

The tracking template is available as an Excel file or Google Sheet. To use the Google Sheet, you will need to log into your Google account and select File -> Make a copy.

Below are instructions on how to use the Pathfinder Tracking Template – Version 1.0

Step 1 – add project details

Use the Project details sheet to add information about the host and pathfinder projects. This will include information on the study teams, datasets involved in the host study, and any resources used in the study, which could include study protocols or tools.

Step 2 – add project activities

Use the Project Tracking sheet to map the various stages of the host study. The first column 'Project Activity' defines the various steps that the host study has or will undertake. You can build this list of activities by adding rows and selecting from the dropdown option in column A.

- You can add as many of these study activity rows, including multiple of the same activity. E.g., you might have three rows for "DM: Data curation" if you have three datasets to curate.

- You don't need to include all activities. E.g., a host study that didn't collect data but used an existing dataset would not have a "DM: Data collection" activity.

- Each stage has two letters before it that define the data science lifecycle category that it falls under (RP - Research Planning, DM - Data Management, DA - Data Analysis, OI - Outputs and Impact, SE - Stakeholder Engagement). This is used to colour the cells based on the category.

- The sheet "Activities and descriptions" contains the list of different activity types and a description of these.

- You can add new activity types in the 'Activities and descriptions' tab if there is an activity in your project that doesn't meet the current options.

- Column B in the 'Project Tracking' tab allows you to add a description of the project activity

Step 3 – add quantitative metrics

Below are five columns used to record quantitative metrics for each project activity row. Some of these might not be applicable or known for each activity, and should be filled in if the information is available.

- **Dependencies on other activities:** Outline which other activities are dependent on this activity being completed
- **Started:** the date when the activity started
- **Completed:** the date when the activity ended
- **Weeks:** this is the number of weeks the activity took. This column is automatically generated

using the Started and Completed column dates
- **Number of human resources:** this is the number of people involved in the activity. Example, 1.0 = one full-time resource, 0.2 = one day a week

Step 4 – add qualitative responses

The 'Qualitative responses' column is used to capture free-text details of the project activity, which could include:

- What work was achieved during the activity?
- Who was involved? What skills and tools were needed? Were there costs involved in this?
- What were some of the challenges or barriers you faced? How were these resolved?
- What lessons have you learned from these experiences? What would you do differently next time?
- What tools, solutions, resources were developed during this activity? Please provide links or attachments.

These responses can be gathered through interviews, discussions or surveys of the host study teams. **Step 5 – add resources and links**

Add any resources or links that have been used or developed during the host study or pathfinder.

These could be links to information or resources on:
- **Research planning**, such as protocols, skills and team setup, process documents - **Information governance and ethics**, such as data access agreements, ethical assessment or consent forms
- **Data management**, such as the data management plan, data capture processes, data dictionaries and metadata, data curation processes, data infrastructure setup - Data Analysis, such as statistical analysis plans, tools and techniques used, shareable code and scripts
- Outputs, such as publications, articles, presentations, tools, dashboards
- Engagement activities, such as reports and findings from engagement, re-usable tools and assets used in engagement efforts

Next steps

We hope that this tracking template will provide a standard format for use in pathfinder studies.

This document can be used to:
- Identify challenges and barriers in the host study's health data science cycle - Provide initial requirements for identifying possible solutions to challenges and barriers to the research study or wider ecosystem
- Share and highlight key areas of learning and associated solutions and tools for the wider data science community
- Develop visual workflows, policy and procedure documents and
- Develop research publications that detail the methodology used
- Provide landscaping of the data science ecosystem associated with the research study and wider partnerships
- Contribute to wider knowledge of common challenge areas, and solutions to accelerate research studies

Feedback

The pathfinder work is intended to be iterative and will evolve as projects get underway. We welcome and encourage feedback on all aspects of the pathfinder work. We will actively seek this feedback, but please reach out if you have any suggestions or ideas on how we can improve the pathfinder project work.

Additional resources

Addition resources can be found on the Pathfinder Projects page on The Global Health Data Science Hub