

# The Global Health Network Data Clinic Guidance Kit

## Table of Contents

<b><i>Purpose</i></b> .....	<b>2</b>
<b><i>Data Clinic Overview</i></b> .....	<b>2</b>
<b><i>Steps for organising a data clinic</i></b> .....	<b>2</b>
<b>Step 1: Identify the needs</b> .....	<b>2</b>
<b>Step 2: Prepare a plan or proposal for organising a Data Clinic</b> .....	<b>2</b>
<b>Step 3: Obtain institutional approval</b> .....	<b>5</b>
<b>Step 4: Promote the Data Clinic</b> .....	<b>5</b>
<b>Step 5: Pre meeting setup</b> .....	<b>5</b>
<b>Step 6: Scheduling</b> .....	<b>6</b>
<b>Step 7: Conduct the meeting</b> .....	<b>6</b>
<b><i>Measuring Results and Impact</i></b> .....	<b>7</b>
<b><i>Reporting</i></b> .....	<b>9</b>

## Purpose

The purpose of this guidance kit is to provide a practical resource to support the establishment and running of a Data Science Clinic. It is designed to guide institutions and practitioners in providing targeted assistance to address the challenges identified by researchers in the collection, management, analysis, interpretation, preservation, sharing, and reuse of health data.

## Data Clinic Overview

A Data Science Clinic is a platform for enhancing capacity in data science on topics such as analysis, interpretation, and presentation, focussed on solving real-world data challenges in individual or team research projects. Here, mentors or trainers can help individual researchers overcome data obstacles, such as data cleaning, analysis, and data sharing.

## Steps for organising a data clinic

### Step 1: Identify the needs

Identify the research training needs or gaps that need to be covered in the community through consultation with target audience or key stakeholders.

### Step 2: Prepare a plan or proposal for organising a Data Clinic

Planning is crucial for the organisation of a Data Science Clinic, since it guarantees clarity regarding the audience, objectives, resources and alignment with the goals of the organising institution, facilitating its approval and implementation. We provide a standard template to help you outline all the elements necessary to for your planning process.

Note: Preparing a plan or proposal for the Data Science clinic may be a long process and it takes shape over a period and changes to the plan can be made along the way.

### Plan or Proposal Template

1. Name of the Data Science Clinic
2. Organising institution/s – include organising partners
3. Primary coordinator or contact person
4. Objectives of the Data Clinic – this should clearly state **3-5 learning objectives** of the activity and the knowledge, or skills participants can expect to come away with after the clinic sessions

5. Target participants: This can be defined according to the objectives, area of interest, or scope of the Data Clinic.
6. Format – Online/in person/hybrid
7. Number of sessions
8. Registration – open access or by invitation
9. Location
10. Schedule – Dates, time and frequency
11. Duration of each session
12. Scope – local/national/regional
13. Number of participants – It would depend on format (online/in person) and the capacity of the venue
14. Facilitator – including their institution, role, contact and rationale for their participation
15. Tutor/s– This could be an expert on the identified topic. Include information on their institution, role, contact and rationale for their participation

#### Examples of Specialty Areas:

- ✓ Data Scientists: experience in data analysis and modeling.
- ✓ Data Engineers: specialists in the construction and maintenance of data infrastructures.
- ✓ Data Analysts: professionals who interpret and present data.
- ✓ Visualisation Experts: specialists in creating interactive graphs and reports.

16. Cost to participants – Attendance should be free for all participants
17. Tutor participation – Generally there is no remuneration and tutors contribute their time pro bono, they can be compensated for their time in kind or through mechanisms suitable to the organising team
18. Calendar with slot schedule: Create a calendar indicating the available slots for consultations. It is essential to choose a method for disseminating information about the services offered by the Data Science Clinic, including the dates and times available for booking consultation appointments.

Different approaches and the technologies involved in each are detailed here. You can choose the one that best suits your context:

Method	Description	Tool
<b>Manual Scheduling</b>	Use communication channels that require manual registration and confirmation of appointments, for example, a receptionist, notes appointments and fills in the necessary data for the consultation. It is the most basic method,	Email Telephone

	initially low cost but with limitations for its management.	
<b>Basic Electronic Scheduling</b>	Use software to create a digital calendar to schedule appointments. It has a low cost to use and improves management, although data entry is manual.	Google Calendar, Outlook Calendar, etc
<b>Online Scheduling</b>	It uses online booking platforms that allow the user to schedule appointments directly through the internet, with 24/7 access, reducing the work in receiving data.	Web platform
<b>Automated Scheduling with AI and Chatbots</b>	Use virtual assistants and chatbots to manage appointments on websites and messaging apps. It can provide immediate and personalized responses to patients, although advanced technological infrastructure is needed.	Google Assistant o Whatsapp chatbots
<b>Mobile Apps</b>	It uses an app to schedule appointments or an appointment scheduling program, it is a digital system that offers an online agenda, eliminating the need for you to be answering messages or calls, to write down appointments and to send reminders, all that is done by the application to schedule appointments.	Apps

19. Resources needed: Identify the necessary resources, according to the objectives, scope, modality and frequency of the meetings.

#### **Physical and Technological Infrastructure:**

- ✓ Adequate physical space for equipment and consultations.
- ✓ Computer.
- ✓ Servers or services for data storage and processing.
- ✓ Platform for videoconferencing, mainly for the clinic whose care will be carried out in the online modality.

#### **Software Tools for:**

- ✓ Data Management
- ✓ Data Analysis
- ✓ Data Visualization
- ✓ Big Data, Machine Learning y AI

#### Examples of open access resources:

1. TGHN resources of data science in global health research: <https://globalhealthdatascience.tghn.org/hub-resources/>
2. The “EpiR Handbook” is a free online R learning resource developed by Applied Epi: <https://epirhandbook.com/en/>
3. Using R for Data Science: [Welcome | R for Data Science \(had.co.nz\)](https://www.had.co.nz/)
4. Introductory Python course developed by University of Michigan: <https://www.coursera.org/specializations/python>
5. Introduction to the command line/Unix (used for bioinformatics or genomics applications of data science methods): [https://codethechange.stanford.edu/guides/guide\\_unix\\_commands.html](https://codethechange.stanford.edu/guides/guide_unix_commands.html)
6. Introduction to broad topics in statistics, textbook: [https://stats.libretexts.org/Bookshelves/Introductory\\_Statistics/Introductory\\_Statistics\\_\(Shafer\\_and\\_Zhang\)/01%3A\\_Introduction\\_to\\_Statistics](https://stats.libretexts.org/Bookshelves/Introductory_Statistics/Introductory_Statistics_(Shafer_and_Zhang)/01%3A_Introduction_to_Statistics)
7. GitHub, tool used by most data scientists - there is a great skill-development course on GitHub itself: <https://skills.github.com/>
8. Google machine learning course: <https://developers.google.com/machine-learning/crash-course/ml-intro>

#### Step 3: Obtain institutional approval

With a proposal for the establishment of the prepared clinic, it is suggested that you meet with the authorities of the institution from where it is intended to be coordinated, to obtain the due authorisation and necessary support. This can be verbal or in writing, according to the requirements of each place.

#### Step 4: Promote the Data Clinic

Publicising the existence of the Clinic is an important step to be considered. It is suggested to use all available communication channels (internal and external), such as posters, WhatsApp groups, social networks, email invitations, etc.

From TGHN we can support the dissemination of the Data Science Clinic, through our own channels such as the regional hubs or monthly newsletters. To find out how to spread the clinic's activities, please contact your regional manager.

#### Step 5: Pre meeting setup

- Schedule the meetings
- Set up the venue or online meeting including registration link, registration questions, follow up or feedback surveys, polls or breakout rooms, etc.

## Step 6: Scheduling

There are different ways to systematise the process of scheduling consultations, from the manual to the administrative management packages. You may use the one available at your institution.

The consultation is organised in a meeting format to facilitate an exchange between researchers seeking the help of data science experts to solve specific problems encountered during research data management. These consultations should have the potential to foster mutually beneficial collaborations between these researchers, such as joint publications or research projects. Data experts will act as short-term mentors, providing guidance on processes without taking responsibility for the final results, as execution depends on the participants.

A simple and easily accessible option is to use online forms, for example Google Forms, to manage the appointment and booking offers of the consultation.

Confirm the appointment reservation and the address for the consultation, whether in person or online (in this case the address is the link of the platform that will be used for the videoconference).

It is advisable to send a reminder 48 or 24 hours prior to the consultation, confirming the date, time, address and the name of the expert who is going to carry out the consultation.

### Example of items to create the scheduling form

1. Full Name:
2. Email:
3. City:
4. Institution:
5. Selection of Consultation Type: Individual/Team
6. Reason for Consultation:
7. Choose the Date of Consultation:
8. Choose Consultation Time:
9. What specific question or problem you would like to solve:
10. You would like to receive information about other TGHN: Yes/No
11. You authorise to be contacted in the future by TGHN: Yes/No

## Step 7: Conduct the meeting

The person responsible for the consultation must have the consultation response form, which contains the information on the scheduling form, and space to record the

actions carried out in the consultation. Likewise, it must have the minimum resources necessary to provide the service.

**Tips for conducting the consultation:**

- ✓ Start the meeting on time.
- ✓ Start the meeting by validating the information on the file with the consultant.
- ✓ Continue with the guidelines that collaborate in the resolution of the doubt or problem.
- ✓ Take note of the actions taken and/or suggested.
- ✓ Respect the end time of the match.

## Measuring Results and Impact

### Suggested KPI's for Data Clinic

Activities under Objective 1	Key Performance Indicators		
	Process Indicators	Outcome Indicators	Impact Indicators
<b>Data Clinic</b>	# of Data Clinic sessions held yearly  # of participants in each session  # of people satisfied by the Data Clinic	# of people with improved knowledge on Data Science topics	# of participants who have applied gained knowledge in their work

Here are a few examples of how to collect information on the above metrics:

## 1. Feedback form

If you would like to gather feedback from the participants for a data clinic, here is the example set of questions that can be utilised.

### Example feedback form questions:

1. Full name
2. Email
3. Your current role
4. Name of organisation
5. Country
6. Why did you attend this Data Clinic?
7. The Data Clinic was a good fit for my needs – Yes, No
8. The concepts and skills covered were explained well – Yes, No
9. Are you aiming to publish any materials (studies, guidelines, toolkits, etc) relevant to the topic covered in the Data Clinic? Yes, No
10. Will you be applying the knowledge gained from the Data Clinic in the work you are doing? Yes, No
11. Please explain your response above
12. What were the key takeaways for you from the sessions, anything that struck you or that you found most interesting?

## 2. Follow up questionnaire

A 6-month follow-up questionnaire can be used to measure the long-term impact of a Data Clinic by assessing how well participants have retained and applied the knowledge and skills gained. This questionnaire evaluates the lasting effects of the meetings and the sustained benefits and improvements in participants' practice or projects.

Here are example questions for a follow up questionnaire:

1. I believe that I learned new information/skills during the Data Clinic meetings. (Strongly agree, Agree, Neither agree nor disagree, Disagree, Strongly Disagree)
2. Please explain your answer
3. The skills I gained in the Data Clinic has been useful for my work (Strongly agree, Agree, Neither agree nor disagree, Disagree, Strongly Disagree)



4. The information I gained in the Data Clinic has been useful for my work (Strongly agree, Agree, Neither agree nor disagree, Disagree, Strongly Disagree)
5. I have noticed improvements/positive changes in my performance/produced work as a result of the Data Clinic (Strongly agree, Agree, Neither agree nor disagree, Disagree, Strongly Disagree)
6. Please explain your answer
7. I have applied the knowledge/skills gained from the Data Clinic in my work (Strongly agree, Agree, Neither agree nor disagree, Disagree, Strongly Disagree)
8. Please explain your answer
9. Are there any specific examples of how the Data Clinic has helped you in your work? (if they choose any of the first two options).
10. Are you working to publish any materials that are relevant to the topic covered by the Data Clinic? \*\* Follow up on this previous question (Yes, No, Not sure)
11. What is the type of work you want to publish? (if the answer is yes)
12. What is the tentative timeline until publication?
13. Do you consent to be contacted again by The Global Health Network to follow up regarding your progress in publishing your work?
14. Are there any collaborations that have emerged as a result of participating in the Data Clinic? (Yes, No, Not sure)
15. Please describe them (if they choose yes)

## Reporting

### 1. Reporting on the Regional Activity Tracker

The Regional Activity Tracker is designed to collect and track key information from research capacity strengthening activities. The information will be compiled to produce a Regional Activity Dashboard accessible and published on the TGHN Regional Knowledge Hubs to help The Global Health Network collaborators to stay informed of the ongoing and implemented activities in the regions. This information will further factor into The Global Health Network's impact metrics and progress evaluation.

Link to Regional Activity Tracker form: coming soon

### 2. Reporting on the TGHN Knowledge Hub

All regional activities should be reported on the relevant knowledge hub on the TGHN online platform using information on:

- Activity summary including location, background, aims and objectives
- Activity poster for comms and dissemination
- Activity programme or agenda
- Activity recording, if available
- Learnings from the activity and key takeaways