



Designed by Charlie Barclay, Emma Griffiths and Rhiannon Cameron as part of PHA4GECon 2025 pre-conference workshops

## PHA4GECon 2025: Needs assessment role cards

The following role cards are designed to support the Needs Assessment activity in the PHA4GECon Developing Data Standards Workshop. In this exercise, participants practice conducting short, structured interviews to identify what contextual data are needed for surveillance and research, mirroring the early “scoping and needs assessment” phase of data specification design.

Each card represents a fictional stakeholder involved in the response to the emerging *E. griffithsi* outbreak. The goal is for interviewers to uncover each stakeholder’s data requirements, gaps, and priorities through open-ended questioning.

### Interviewee Card 1 – Public Health Epidemiologist (Clinical Surveillance Focus)

#### Role Summary:

You coordinate *E. griffithsi* case reporting for a provincial public health lab. You oversee sample submissions from hospitals and clinics and need data that help you track outbreaks, link cases, and assess risk to the community.

#### Talking Points / Priorities:

- You receive **clinical swab samples** (nasal/throat) from patients with respiratory symptoms.
- Data you currently collect: patient demographics (age, gender, location, travel), collection date, test type and severity of symptoms.
- Data are captured in a spreadsheet and are meant to be uploaded weekly to a regional database, however there are often delays in upload as data doesn’t often get sent to the data management team in an upload ready format
- You often lack:
  - Exposure information (travel, contacts).
  - Collection method details or who collected the sample.
  - Linkages between diagnostic data and sequencing results.

- Concerned about timeliness and duplication — labs use different formats.
- Would like standardized fields for: **symptom onset**, **test date**, **hospitalization status**, **sequencing assay type**, and **lineage calls**.

## Interviewee Card 2: Environmental Scientist (Wastewater surveillance focus)

### Role Summary:

You lead a wastewater monitoring project trying to detect *E. griffithsyi* in municipal systems. You collaborate with labs that process and sequence environmental samples. The monitoring project collects raw wastewater and influent samples from treatment plants and sewer sites weekly.

### Talking Points / Priorities:

- You support several sites, some of which use an autonomous sampler, which continuously collects a sample over a 24 hour period, whereas other treatment plants just take a single grab sample first thing in the morning.
- Data you collect: site ID, GPS coordinates, sample volume, flow rate, presampling weather, and collection date/time.
- Data stored in Excel; shared by email with sequencing partners.
- Problems:
  - Inconsistent **site naming conventions** across regions.
  - No standard way to capture **catchment population** or **environmental conditions** (e.g. temperature, rainfall).
  - Downstream labs don't return sequencing metadata (e.g. primer scheme, dehosting method).
- Needs better integration of **sampling and sequencing contextual data** for environmental interpretation.
- Interested in linking environmental and clinical datasets.

## Interviewee Card 3: Academic Researcher (Genomic and Phylogenetic focus)

### Role Summary:

You study *E. griffithsyi* genomics to understand its evolution and potential zoonotic origins. You integrate sequence data from multiple partners.

### Talking Points / Priorities:

- Receive isolate and metagenomic sequences from multiple labs.
- Need contextual data to interpret results: sampling date, host species, geographic origin, sequencing platform, library prep, and QC metrics.
- Frustrated by missing or inconsistent contextual data (e.g. "sample site: BC", "sample site: London hospital", "sample site: unknown").
- Store data in a research database, export to spreadsheets for analysis.
- Would like:
  - Clear links between sample IDs and sequence accessions.
  - Consistent ontology-based terms for sampling context and methodology.
  - Capture of bioinformatics methods (assembly software, version, reference database).

- Advocates for **machine-readable standards** and contextual data completeness to enable downstream comparative analysis.

