Data formatting guidelines

# Linelist:

The linelist details the information on each individual. This is matched to the contacts sheet by “id”. The necessary columns are “id”, “report\_onset” and “death” and should not be renamed; other columns are optional.

**id**

This is the BEN code (if available) and should match **exactly**  the code used in the contacts sheet. If the person does not have a BEN code please use a **unique** code such as “NA100”. **This column is essential.**

**name**

This is the name of the person. The formatting is not critical but t help in other analyses, please enter in all capitals. **This column is optional.**

**report\_onset**

This is the date of reported debut des symptoms. It should be entered in “date” format in excel or in form “dd/mm/YYYY”. **This column is essential.**

**cte**

This is the date individuals were first admitted to cte. **This column is optional.**

**death**

The reported date of death. The same formatting requirements apply to all dates as those shown for “report\_onset”. If there is no date of death, leave a blank or type **NA. This column is essential.**

**status**

Whether an individual is deceased or surviving. Formatting is not critical should be consistent, ie. always type “deceased” or “Decede” but not a mixture of the two. **This column is optional.**

**sex**

Takes the values M, F or NA. **This column is optional.**

**age**

This should be a numeric value unless it is missing, in which case type **NA. This column is optional.**

**quartier**

Please enter place names in capitals separated only by a comma (no spaces). **This column is optional.**

**classification**

This should be entered all in lower case and takes values “confirmed” and “probable”. **This column is optional.**

**hcw**

Is the individual health care worker such as a nurse of doctor. Only enter TRUE, if the individual is not a health care worker leave blank or enter NA. **This column is optional.**

**traditional\_healer**

See above.

# Contacts:

This details the links between people. One line corresponds to one connection. “From” and “to” should **only** take id that are present in the linelist and the id should be in exactly the same format. The necessary columns are “from” and “to” and should not be renamed; the other columns are optional.

**from**

The source of infection. This should take the unique id of the individual. **This column is essential.**

**to**

This is the individual who becomes infected. This should take the unique id of the individual. **This column is essential.**

**plausible**

Is the link plausible? Enter TRUE is so, leave blank otherwise. **This column is optional.**

**investigating**

Is the link investigated? Enter TRUE is so, leave blank otherwise. **This column is optional.**

**known**

Is the link known? Enter TRUE is so, leave blank otherwise. **This column is optional.**

**in\_health\_structure**

Did the transmission occur in a health structure? Enter TRUE is so, leave blank otherwise. **This column is optional.**

**in\_traditional\_healing**

Did the transmission occur through traditional healing? Enter TRUE is so, leave blank otherwise. **This column is optional.**

**in\_funeral**

Did the transmission occur at a funeral? Enter TRUE is so, leave blank otherwise. **This column is optional.**

**in\_family**

Did the transmission occur between family? Enter TRUE is so, leave blank otherwise. **This column is optional.**

**in\_friends**

Did the transmission occur between friends? Enter TRUE is so, leave blank otherwise. **This column is optional.**

**in\_neighbourhood**

Did the transmission occur between neighbours? Enter TRUE is so, leave blank otherwise. **This column is optional.**

**in\_travel**

Did the transmission occur through travel? Enter TRUE is so, leave blank otherwise. **This column is optional.**

**to\_clarify**

Are there discrepancies in the dates of linked people. Enter TRUE is so, leave blank otherwise. **This column is optional.**