Overview of an Outbreak Tools linelist features

Yves Amevoin   Mathilde Mousset

Introduction

Outbreak-tools is a set of tools to facilitate the creation and maintainance of Excel linelists. In particular, you can easily:

- Define different types of variables (dropdown lists, Excel formulas) their format and data validation rules.
- Print a register book that perfectly match your linelist.
- Translate your linelist into several languages.
- Use pre-defined geo databases for patient origin.
- Display custom analyses (univariate, bivariate, temporal, spatial) as tables and graphs.

- Export the entire data or selected variables/lines.
- Migrate the data to an updated version of the linelist.

The linelist is defined in a configuration file called *setup*. This *setup* file is then loaded into another file that designs the linelist (the *designer*). The designer creates the linelist, based on the configurations of the *setup*.

The goal of this exercise is to explore the functionality and use of a measles linelist containing fictitious data.

> **!** Functionalities
>
> The features are described in a user guide sent with this exercise. You can to get an idea of all the features available by reading the user guide. We'll just go through a few of them in the linelist. The main ones are presented in the following two diagrams:
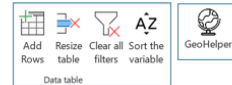
Getting started

Setup your linelist

1 **Fill in the information about your site.**
These information's are to be entered *once* in the *Info* sheet of the linelist.

2 **Adapt the linelist to your needs.**
- Decide **which variables to show or hide**; name "Extra variables" that you want to add
- Import geobase (optional)
- Import previous data (optional)

3 **Print a register** and/or a **case report form** (optional)
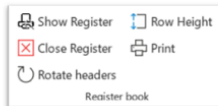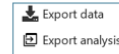
Enter your data

4 **Enter data in the data entry sheet(s)**

5 **Update and inspect analyses**

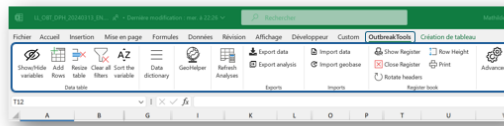**6** **Export (anonymous) data** to MSF or MOH

(a) Getting started

Linelists features

Dedicated buttons in the OutbreakTool ribbon.

Hide or Show optional variables, and then print the corresponding **register/CRF**

Special column types:

* **Custom dropdowns**: define the categories you want in the dropdown (click on the name of the variable to go to the table in the *Custom dropdowns* sheet)

* Columns to enter **geodata** from an embarked geobase. Use the **GeoHelper button** or the cascading dropdown menus

* **"Extra" variables**: for variables not in the linelist, customize their names

Extensive analyses sheets, including on filtered data (don't forget to **refresh** before looking at them!).

Predefined specific exports (ex: anonymous data, or for MoH)

Migrate data from one linelist file to another version of this file (export + import features)
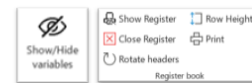
(b) Features

Figure 1: Linelist Features

Exercise: Using linelist features

You have a linelist named `linelist_measles_usage_en.xlsb`, and an excel s sheet containing data called `import_linelist_en.xlsb`. We're going to explore the linelist after importing the data from the sheet, and answer some questions.

1- Import data, discover the data entry sheet.

Import data

Open the `linelist_measles_usage_en.xlsb` sheet.

On the `OutbreakTools` tab, click on the `Import data` button. Browse and select the `import_linelist_en.xlsb` file, then click OK.

At the end of the import, you will be asked if you want to see a report. Select 'Yes'. Which variables are not available in the imported data?

Sorting data

Sort the data in the 'ID number' column.

To do this, Click on the `$A$9` cell.

On the `OutbreakTools` tab, click on the `Sort Variable` button.

What are your thoughts?

Now sort the data on the column `Identification number padded`.

Show/Hide Variables

On the `OutbreakTools` tab, click on `Show/Hide Variables` button.

Hide the first three custom variables.

2- Entering data

You have data for the following two patients:

```
Patient 1:

Notification date: 2023-10-22
Name: John Doe
Sex: Male
Age: 3 years
Origin: Lualabala, Kapanga, Kanampumb, Kasankara
Date of onset of symptoms: 2023-10-18
```

```
Date hospitalised: 2023-10-25
Vaccinated against measles: no
Date of discharge: 2023-11-01 (Died)

Patient 2:

Identification number: 44
Date of notification: 2023-11-05
Name: Jane Doe
Sex: Female
Age: 7 Months
Origin: Lualabala, Kapanga, Mulambu, Mulambu
Date of onset of symptoms: 2023-11-04
Date of hospitalisation: 2023-11-10
Vaccination status unknown
Date of discharge: 2023-11-12 (Cured)
```

Enter this data in the linelist.

3- Browse analyses

Click on the sheet `Uni and bi-variate analyses`.

Global overview, univariate analyses.

In the `OutbreakTools` tab, click on `Refresh analyses`. How many patients do you have in the database? What is the case fatality rate? What is the proportion of children under 5?

Analysis on filtered data.

Click on the sheet `Linelist patients` and filter the `age group` column to keep only children under 5.

Click on the sheet `Uni and bi-variate analyses` then click on `Refresh analyses` in the `OutbreakTools` tab.

You now have analyses only on the data you have filtered. What is the CFR of children under 5?

Bivariate analysis

Click on the sheet `Linelist patients`.

In the `OutbreakTools` tab click on `Remove all filters`.

Click on the `Uni and bi-variate analyses` sheet.

Using the `Go to Section` button on the `$C$2` tab, select the the section `Patient Age Characteristics`.

Among the unvaccinated patients, what is the most represented age group?

Have any adult patients (over 15 years of age) died? How many?

Temporal analysis

Click on the `Temporal analyses` sheet, and update the analyses as before.

In cell `$G$11$`, choose the month as the time unit.

In cell `$E$11`, set the start date as 2019-01-01.

During 2019, in which month were there the most cases?

During this month of 2019, in which epidemiological week were there the most cases?

Overall during the epidemic, in which epidemiological week were there the most cases?

During this epidemiological week, what was the case fatality rate?

You can use the filters in the `Patient list` sheet to answer the following questions.

What is the case-fatality rate for under-5 during the most affected epidemiological week?

In which epidemiological week were the most deaths recorded?

Spatial analysis

Click on the `Spatial analyses` sheet and refresh the analyses as before.

In cell `$C$14$`, select the option `Health-area`.

How many patients have no information about their origin?

Which health area is the most affected? We will keep its name for future use.

Where do most of the patients who die come from?

Spacio-temporal analysis.

Remember that we had to keep in mind the health area most affected.

Click on the `Spatio-temporal analysis` sheet.

In cell `$C$14`, click twice. A box for selecting locations is displayed.

Select the most affected health area. You can stop at level 3 and click OK.

Then update the analyses as before.

During the epidemic, this health area had the most cases in which week?

You can choose up to 10 health areas and display them on a graph if you wish.

4- Exporting data

In the `OutbreakTools` tab, click on `Export Data` button.

Click on `Anonymous export for MoH`. Choose where you want to save the export.

Do the same and make an `Anonymous export for MSF`. This is the export that is sent to the dashboard. Don't rename it. You can open it to explore the data and see the differences between the two exports. The password is 1234.