USE OF EXCEL IN DATA HANDLING IN LABORATORIES

This Cook Book is not intended to provide comprehensive guidance on every thinkable use of Excel in laboratories, but to uncover some key sources of errors and provide good advice on how to avoid them – however, for a comprehensive understanding both require more extensive and detailed studies.

Areas of application - *lots of exciting and challenging possibilities, but be careful*

Calculations
Excel is used in most laboratories as the primary calculation tool, – especially for routine calculations, and integrated with application templates (although unfortunately the user often has not the faintest idea of what is actually behind it). It is important that the many functions and facilities in Excel (incl. more advanced calculations using macros) be used *consciously, systematically and in a controlled manner!*

Statistics
One special group of calculations that Excel can easily perform is statistical estimation and testing etc. There is a risk that simply selecting a tool or function from the toolbox without knowing if it is the right one for this purpose and without having the proper understanding of the correct interpretation of the results (unfortunately the HELP facilities for these functions are not very well elaborated or really helpful!).
The use of Excel for statistics therefore requires a basic understanding of statistical theories and ~~there~~their functions!

Data analysis
Excel is designed to handle and (re)arrange large amounts of data, not just for basic and statistical calculations, but also to organize or categorize the data, to highlight any coherence of data sets, trends or impacts of influence factors. The ultimate tool for this purpose in Excel ~~are~~is the Pivot Tables and related charts, but as always, *good and careful planning before entering the data is a prerequisite for a good analysis.*

Illustration
Excel has at least about 100 different types of charts that can be used to visualize data and demonstrate any possible coherence between sets of data – and that can easily be set up using integrated wizards (software wizards). However, most options are usually not suitable for illustrating laboratory data. Therefore, one should *be sure to choose the right chart for the purpose (e.g. a calibration curve) that represents the actual data and does ~~not~~ lead to better understanding rather than to misinterpretation ~~rather than better understanding~~*.

Storage
Excel can also be used as a database for collecting and storing even relatively large amounts of data. The program offers various functions/functionalities for entering, sorting, filtering, extracting, retrieving and viewing data in the database. In the laboratory, this is mainly relevant for data that are repeatedly generated/collected over time, along with information necessary for further treatment and evaluation. *But what at a first glance looks like a nice and simple table for collecting data and information is not necessarily the optimal way to organize data in Excel!*

Automation
Many employees in laboratories experience Excel as a kind of "black box" in which they enter data and just press "Calculate" to have the program automatically insert the data into other spreadsheets, do some calculations, display the results according to some predefined conditions, evaluate and comment etc. etc. Someone has (perhaps years ago) created an application that accomplishes the task using all the fancy functionalities of Excel, but *the daily user does not understand what actually happens, how he can change the procedure – and what happens when the conditions for the actual calculations and evaluation do not apply….?*

Presentation

Excel is not a great text editor, but it has almost endless tools for text formatting and orientation, backgrounds, and colors etc. for data in spreadsheets and charts, before the result can be printed as a nice report. However, laboratories that need to deliver clear messages and information to underpin key decisions need to use *these many options for fancy layouts with care to illustrate and underline the content of a report, rather than distorting the message itself.*

Excel IS a fantastic tool for data handling – even in a testing laboratory (though is it not specifically designed for that) – but it takes a conscious strategy in the laboratory to get the most out of it in an efficient, reliable and safe manner.

E.g. by implementing the following principles/advice:

Planning – *that pays off before you start.*

Some claim that the "80/20 rule" is also applicable to the creation of a new spreadsheet – i.e. spend 80% of the time planning with paper and pencil and then 20% for the actual implementation of the application on the computer, such as:…

- Which data have to be processed and how are they have to be entered in Excel (data transfer, format, import and entry)?
- Which calculation have to be performed, which formulas and inputs are necessary (incl. any constants, validation of input data (min/max values)?
- How should the result be presented and are there any special requirements in terms of QA (e.g. validation) and documentation?
- Is the application used routinely – or only once – and who uses it?
- If data is to be stored and organized, a suitable table ("list" in a separate tab) must be created before entering data….. i.a.

Setting up applications – *to be used by more than "one nerd" in the laboratory.*

Think of the usefulness of an Excel application by e.g. … -

- clarifying where data-input must be done and how
- making the use of an Excel application (e.g. Visual Basic forms) self-explanatory (on a level so that the daily user understands the actual functionality of the spreadsheet)
- using named cells and cell ranges, which makes it easier to read and understand formulas later
- hiding (if possible) background calculations in hidden cell areas of specific "calculation sheets"
- incorporating automatisms (by using logic functions, conditional formatting, data validation – and, if relevant, even array functions and macros), if this does not conflict with the users' logical judgements and the need to act individually
- keeping track of large amounts of data, by splitting them up into different sheets and files – and linking them in a convenient way (… which at the same time will probably reduce the time required for the calculations)
- backing up (protecting) cells/ranges/worksheets, before saving the application as a template for future use ….a. .o.

Documentation – *the irritating, time consuming but invaluable "paper work"*

Besides being a required part of any QA system that uses Excel applications, it can prove being of crucial importance for the daily use that the functionality of a given application is well documented. Not least when the colleague who has developed various formulas and other functionalities is no longer available. Excel offers several more or less elegant functions :

- Write an explanatory text in cells adjacent to the formulas (use e.g. the Equation Editor to show the actual formulas)
- Use the comment tool to explain the content of cells
- Print out of the worksheet in "Show Formulas" mode…i.a.

Quality Assurance – *also spreadsheets must be brought to follow the rules*

As an integrated – and very important – part of the entire testing process in the laboratory, the calculation tool in the form of Excel must also be included in the quality assurance activities, such as validation, documentation etc.

When using macros and complex functionality the software (EXCEL) is considered as Modified-Of-The-Shelf. Such application shall be suffieciently be verified or validated. Requirement hereto is given is ISO 17205 cl 6.4 and 7.11.2

The pitfalls are lack off:

- Identification
- The range of valid input values
- Test/ check of the functionality of the an application by manual check calculations with known input / known result
- Use (if necessary) built-in facilities for step by step calculations and error-tracking
- Audit trails of corrections in a spreadsheet are unfortunately not included in the functionality of Excel but must be established by creating special macros.
-

Note, that unmodified Off-the-shelf software is considered sufficiently validated by the software manufacturer

[Part of QA in using Excel, of course, will also be to formulate and obey policies and documented procedures; see the following paragraph]

Policies/procedures – *the laboratory must be in full control of the use of this powerful tool.*

Many employees may find working with Excel and developing new and better (!?) applications an interesting – and enjoyable – challenge, which, of course, can be seen positively, but at the same time in terms of controlling can create a lot of problems the actual and consistent use of Excel in the laboratory!

For all applications of Excel for data handling in a testing laboratory, it is therefore crucial that the laboratory has set up its own policies and relevant procedures to control (even limit) the temptation to develop its own applications in all kinds of fancy directions. For example, this must/could cover subjects such as

- Who is responsible for developing (and revising) the applications used for data handling in the laboratory?
- For which purposes is Excel used in the laboratory?
- How (at which level) are Excel workbooks backed up?
- How are statistics dealt with in the laboratory?

Training – *you can always learn something more about Excel*

Excel is often regarded as a tool which people can easily adopt and learn by themselves! However, it is often advantageous for the entire organisation in a laboratory to have staff receive specific training in using Excel for the laboratory applications (many of the specific functionalities, functions, and ways of developing a spreadsheet application relevant for laboratories are usually not taught in Excel courses).