# Project Outline 02 - Data analysis for cancer drug discovery

## Background to the project

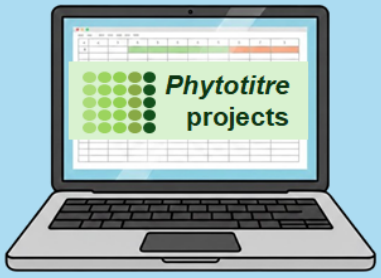
Cancer is a leading cause of mortality worldwide. It is characterized by the uncontrolled proliferation of abnormal cells that evade the body's normal defence mechanisms. Despite significant advances in immunotherapy and targeted biologics, there remains a pressing need for the discovery of new classes of small-molecule drugs that can selectively induce apoptosis or arrest the cell cycle in cancer cells.

Historically, the plant kingdom has provided a rich repository of anti-cancer agents. Iconic chemotherapy drugs such as paclitaxel, vincristine, and etoposide, were all originally derived from natural sources. The large-scale screening of natural products has also given rise to successful and widely used drugs in many other categories.

These discoveries were enabled by the vast chemical diversity present in the Natural world, a uniquely powerful resource which retains great potential for further drug discovery. By systematically testing hundreds or thousands of different natural extracts or compounds, researchers aim to discover and isolate a small handful of rare 'hits' which possess the desired biological activity against a target of interest. This method, called **high throughput screening** (HTS), is a critical step in the modern drug discovery process. Your project will involve the analysis of datasets arising from screens of a large natural product library for novel molecules with potential to inhibit the growth of four model tumour cell-lines.

### Project at a glance

- Data analysis project
- 8 weeks of analysis
- BSc / MSc level
- Cancer drug discovery




## Benefits of choosing this project

Through this project, you will gain experience of several key data analysis and statistical techniques which are in high-demand from life-science employers. These include management of large datasets, background correction, normalisation, plate mapping, correlation analyses, ANOVA, and four parameter logistic (4PL) curve fitting. Beyond gaining key industry-relevant skills, your analyses will contribute to the search for new antibiotic leads.

## How long will it take?

The basic project outlined below can be achieved typically within an **8 week** period of analytical time (not including the time necessary for writing of the dissertation). However, this can be extended to over **16 weeks** if necessary by following the suggestions for further, more in-depth analyses given in the later sections. Please discuss with your supervisor how to modify the project to achieve your aims while fitting within the available timeframe.

## What will I do on this project?

The core activities you need to complete the basic version of this project are listed below:

1) Perform background correction and normalisation of a large HTS dataset
2) Calculate Z'-factor scores for assay plates from a large screen for anti-cancer extracts
3) Map assay values to extract IDs using plate maps
4) Identify the top "hits" from the screen for further analysis
5) Perform correlation analyses to seek insight into the specificity of hit extracts
6) Use 4 parameter curve fitting to calculate $IC_{50}$ values for top hits
7) Use these analyses to prepare your final report or dissertation

## What skills will I gain through this project?

Completing this project should give you experience of the following core employability skills:

| **Data handling skills** | | **Statistical analyses** | |
|---|---|---|---|
| Standard Microsoft Excel techniques | ✓ | Calculation of assay Z' factor | ✓ |
| Data Management and Integrity | ✓ | One way ANOVA | ✓ |
| HTS data handling | ✓ | Two way ANOVA | ✓ |
| Background correction & normalisation | ✓ | 4-parameter dose response curve fitting | ✓ |
| Mapping assay values to plate maps | ✓ | Linear regression for correlation analysis | ✓ |
| Creation of appropriate charts | ✓ | Applying corrections for multiple testing | ✓ |

## What support is available?

Comprehensive method sheets, which describe how to perform each step of the data analysis protocols are available from the downloads page on our website. These are free to download without registration. Troubleshooting advice to help solve common problems that may arise during the project is also given towards the end of this project guide. Your university supervisor will be your primary point of contact and will be responsible for your day-to-day supervision.

# Instruction list

Complete these tasks in the following order for a basic computer-based data-analysis project in Antibiotic Discovery. These should be completed in the order shown below.

# The preparation phase

## Step 1: Read the introduction to the *Phytotitre* library

The *Phytotitre* library is a collection of natural extracts that has been developed to support drug discovery projects in academia and industry. The full-size kit comprises both polar and non-polar extracts of 400 plants, for a total of 800 extracts provided in ten 96-well plates. This collection comprises almost entirely of traditional herbs or medicines with a history of oral use in humans. By focussing on such plants, the aim is to increase the likelihood of identifying leads with both a high hit rate for biomedical targets, and a favourable safety profile. A smaller version of the *Phytotitre* collection, comprising 400 extracts, is also available to support student laboratory research projects.

To learn more about the background to the *Phytotitre* library, please read **Method sheet 100**.

## Step 2: Familiarise yourself with relevant Health and Safety information ⚠

Although you will not be performing any laboratory work during a data analysis projects, you should still be aware of the risks associated with the collection and analysis of data. Please read our '**Safety First**' document to understand the primary risks associated with a computer-based data analysis project in this area.

Every institution has a slightly different way of training and recording student health and safety inductions. However, the following are some common steps that you may have to complete before starting your data analysis project:

1) Read the project risk assessment

Every new project should always begin by writing a **Risk Assessment** document to cover the anticipated risks associated with the project. You should discuss with your supervisor who will prepare this, then read and sign the risk assessment associated with your project. Take note in particular of what the key risks are, and how to mitigate them (e.g. adjusting your seating height and keyboard placement, etc.).

2) Be aware of and follow your local health and safety rules

In addition to reading the relevant risk assessment documents, there may be other procedures necessary to complete before you are able to work in the laboratory. Please ask your supervisor for advice on what else may be necessary to complete before starting.

## Step 3: If necessary, complete a Research Ethics application

Some universities require students to submit an Ethics Application to address any potential ethical issues the project may raise before you start your project. The format of such applications varies greatly between institutions. However, if you are required to submit an ethics application for your project, we advise that you cover at least the following points in your submission:

1: No use of human samples or personal data

It will be helpful to state near the start of your application that the project will be based on the analysis of publicly available data only. Mention that there will be no laboratory work performed, which means there will be no biological, genetic or chemical risks associated with the project. You should also state that you will not be collecting any data from volunteers, or analysing any

personal or identifiable data. It may help to mention that as your project be *in silico* only, with no use of animals or animal products, your project aligns well with the modern movement towards reducing animal use in research (3Rs - Refine, Reduce, Replace).

3: Compliance with the Nagoya Protocol

The Nagoya Protocol is an international agreement under the Convention on Biological Diversity (CBD) that ensures the fair and equitable sharing of benefits arising from the use of genetic resources (such as plants) or related traditional knowledge in certain ways. Your ethics application should acknowledge that (i) the Nagoya Protocol is relevant as the experiments from which the data were obtained involved the screening of plant extracts from diverse countries and (ii) the experiments as designed are compliant with the Nagoya Protocol, as discussed in more depth on our website.

*Reminder*: Always show your draft ethics application to your supervisor before submission, as they are the 'Principal Investigator' (PI) responsible for your supervision.

## Step 4: Make sure the necessary IT resources are in place

Our data analysis projects are designed in such a way that students should be able to complete them without requirement for expensive specialist software or subscriptions. We recommend that you use Microsoft Excel for the majority of the work, but any equivalent spreadsheet software should work as well, including freeware alternatives. The main requirements for equipment are as follows:

Equipment

- Any PC or laptop capable of running standard spreadsheet software

Software

- *Recommended*: Microsoft Excel (any version)
- *Alternative*: Any equivalent spreadsheet software

## Step 5: Plan strategies for Good Laboratory Practice, Data Management and Integrity

A key employability skill for those seeking work in the life-sciences is knowledge and experience of working to the standards of **Good Laboratory Practice (GLP)** and **Data Management and Integrity**. You should therefore aim to practise these principles throughout your project. Begin by reading **Method Sheet 17**. This document will give you an overview of how to maintain your laboratory notebook and data files to comply with GLP and Data Management standards, which are essential for workers in the life science industries.

# The data analysis phase

Please make sure you have completed all necessary activities in sections 1-5 before beginning the analysis phase of your project.

## Step 6: Download the necessary Excel files and Method sheets

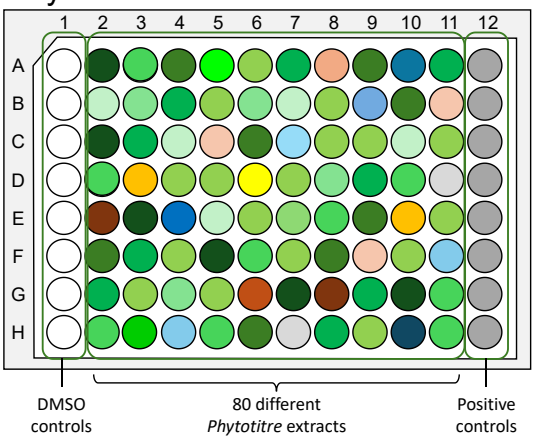
You should begin by downloading the necessary Excel files from the downloads page on our website. These files contain the raw data which you will analyse. As soon as you have downloaded the files, place them in a folder marked 'Primary data files', and remember to never alter these files. Instead, you should copy the Excel files into a new folder, marked 'Data analysis', and give each file a new name (e.g. "Phytotitre analysis 01"). In this way, you keep the original data safe, minimising the risk of accidentally altering it during the analysis. This is an essential principle of Data Management and Integrity. It may also be helpful to download all of the relevant Method sheets, and save them in a separate folder.

At this stage you should also read **Method sheet 101**, which gives advice on the basic data manipulation methods you will need to be familiar with to be able to complete the project. Practise some of the examples yourself until you feel confident performing all of the main data processing functions described in this method sheet.

## Step 7: Perform background correction of plate data

Open the working version of the *Phytotitre* dataset relating to the screen for inhibitors of growth of the tumour cell-line A2780. The first tab explains the methods used during the screen, the second gives the plate maps, and the third gives information on each plant contained within the library. The following tabs contain the raw data from the screening experiments. As each stock plate contains 80 different extracts, and there are 10 different stock plates in the *Phytotitre* library, there are 10 tabs containing data, each obtained from assays of 80 different extracts.

In this screen, every 96-well plate containing cell cultures was challenged with 80 extracts per plate, a positive control (the chemotherapy drug cisplatin) and a negative control (DMSO, as shown in the image at right). 24 hours later, cell biomass remaining in each well was measured by staining each plate with the dye crystal violet. Lower absorbance values in this assay indicate cell loss, mostly arising from either cell killing or slower growth. The raw absorbance values from these assays are arranged in 96-well plate format as they were screened (i.e. in 12 x 8 cell format). Each stock plate was assayed once only.




Your first task will be to perform **background correction** of the absorbance values for every plate shown in each tab. This is necessary to account for background absorbance arising from non-specific staining of the dye on the plastic of the plates that could otherwise confound interpretation of the results. To do this, follow the advice given in **Method sheet 102**.

## Step 8: Perform normalisation of the data

Once all the data are background-corrected, you should proceed to **normalising** the data from every plate. This means transforming the raw absorbance values into a percentage of the negative control values (i.e. the DMSO controls in column 1 of the plate). This is necessary to enable correction for day to day variation as different plates may have been screened on different days, and the extent of staining of the control cells may have been different. You can do this by following the advice given in **Method sheet 103**.

## Step 9: Calculate Z' factor values for each of your plates

Now that you have normalised the data from each plate, the next step is to test how reliable and reproducible the results of the screening assay were. This is a necessary step in the drug discovery process, as it helps to establish whether the observed results reflect real differences, and are not just artefacts or "chance events". The industry standard approach is to calculate a metric called the Z' factor (pronounced "Z-prime factor") for every plate tested during the primary screen. Z' factor values greater than 0.5 indicate that the positive and negative controls are reproducible and well-separated, raising confidence that any hits discovered on that plate are likely to be real. Values between 0 and 0.5 indicate a less reproducible assay and less confidence in the hits. Values below zero indicate that the assay has far too much variability and cannot be relied upon to identify hits.

You should therefore calculate the Z' factor value for every plate measured in the screen, using the normalised absorbance data. Do this by following **Method sheet 104**.

## Step 10: Mapping experimental data to plate maps

Now that you have gauged how reproducible the results of the screening experiments were, you should map the normalised data to the respective extract ID and plant name. To do this, insert a new tab in your Excel file (e.g. 'Normalised results') and collate all of the results from the preceding 10 tabs for the 8 hour timepoint to yield a table in which you align every normalised absorbance value (i.e. the percentages) to the corresponding extract ID number and plant name, according to the plate maps shown in the second tab of the Excel file. Please take great care with the placement and copying of your data at this stage, as any error in the alignment will prevent the correct identification of "hit" extracts. Once you have aligned all 800 ID numbers with the normalised absorbance at the 8 hour timepoint, you should prepare two scatter plots, one for all the polar extracts (IDs 1-400) and one for all the non-polar extracts (IDs 401-800). Place the extract IDs on the x-axis and the normalised absorbance values on the y-axis. For advice on how to complete these tasks, please follow **Method sheet 105**.

## Step 11: Identify hits from the primary screen

Now that you have aligned the data for normalised cell-growth against the extract IDs, you are ready to identify the "hits" from the screening experiment. For your project, these will be the extracts that show the greatest potential to inhibit growth of each tumour cell-line. To do this, insert a new tab (e.g. 'Hit extracts') into your working Excel file for this cell-line. Now copy the whole table of normalised values, including the corresponding extract ID and plant names, from the previous 'Normalised results' tab. Paste the values only of this table, with no equations or cell-references, using the Excel function: Paste-special → Values. Now select the whole table of values, including all columns containing data or labels, and use the Excel 'Sort' function to order the table by ascending absorbance values. Those extracts that appear near the top of the table will be those with the lowest adherent cell count at the time of crystal violet staining. These are the "hits" from this screen, as they have likely either inhibited cell growth or promoted cell death.

Take a note of the identities of your top 7 hit extracts, including the extract ID number, the genus and species names, and the common name of the plant. For advice on how to complete these tasks, please follow **Method sheet 106**.

## Step 12: Repeat the above steps for the other cell-lines

Now that you have completed an analysis of the *Phytotitre* screen for inhibitors of growth of A2780 cells, you should repeat steps 6 to 11 using the *Phytotitre* datasets for the HepG2, MCF-7 and PC-3 cell-lines. Prepare a fresh Excel file for analysis of these new data.

## Step 13: Explore correlations within and between the HTS datasets

Insert a new tab into each of your working Excel files (you should have four by now, one for each cell-line), and name it 'Correlation analyses'. Paste as values the normalised data from the whole screen, alongside the extract IDs, into a new data table. The first analysis you should perform is to test whether there is any correlation between the results observed in response to the polar extracts (numbered 1-400) and those of the non-polar extracts (numbered 401-800) of the same plants. To do this, paste as values the first 400 normalised absorbance measurements for one cell-line in one column, then paste as values the next 400 normalised absorbance (i.e. relating to extracts 401 to 800) into the next column along. Make sure to align the data carefully so that Extract 1 is paired with Extract 401, and Extract 2 with Extract 402, etc. In this way, the growth values for polar extracts of each plant are paired with those of the non-polar extracts of the same plant. Prepare a scatter plot from these two columns and insert a linear trendline, checking the option to show the $r^2$ value on the chart. Take the square root of this value to calculate the coefficient of correlation, r. Then follow the advice given in **Method sheet 109** to test whether the coefficient reaches statistical significance for this particular test.

A key question to answer for your report will be, are there any correlations between the extent of growth inhibition achieved by the extracts for the different cell-lines? To test this, prepare a new Excel file and name it 'All cell-lines". Now paste as values the normalised cell growth data from each of the four separate analysis Excel files into the same tab of the new file. Fill the first column of the sheet with the extract IDs in numerical order (i.e. 1 to 800). Then, one at a time, paste the corresponding values for each of the four different cell-lines into the neighbouring four columns. Insert the name of each cell-line in the row just above the first entry for each column. Use this table to calculate the coefficient of correlation and p-value for each cell-line against every other cell-line, as shown in **Method sheet 109**. You should end up with 6 different pair-wise correlations from this analysis.

## Step 14: Plot a bar chart of hit extracts

The next step in drug discovery after completion of a screening experiment is to attempt **replication** of the top hits. This is the process where the highest scoring extracts only, not the full collection, are tested in further experiments to confirm that they have biological activity against the target of interest. These tests aim to validate that the hits are genuine, and not false positives, which may arise from artefacts in the screening process, statistical error arising from variability in the assay, or mis-calculation during analysis of the original dataset.

Five *Phytotitre* extracts have been tested as part of this replication process for their ability to inhibit growth of the MCF-7 tumour cell-line. In these tests, the extracts were tested at a single concentration of 128 µg/ml in four separate experiments, rather than the single measurement performed during the primary screen. You should download the replication dataset (`mcf7_replication.xlsx`) from the 'Downloads' page and perform the background correction and normalisation processes as before on the data from all four experimental plates. Then you should prepare a data table in a new tab that collates the data from all 4 experiments. Calculate the mean and standard deviation of normalised growth values for every extract and the negative control (DMSO alone). Then plot a bar chart with the plant names as the bar labels, and error bars indicating the standard deviation of each set of four experiments, as shown in **Method sheet 107**.

Next, perform a one-way ANOVA with Dunnet's post-hoc test (or Tukey's post-hoc test if Dunnet's post-hoc test is not available in your software) to determine if cell growth was significantly lower than the DMSO control in the presence of any of the five tested extracts. For advice on how to do this, please follow **Method sheet 108**.

## Step 15: Examine dose response assays of individual hits

One of the top hits from the *Phytotitre* screen was taken forward to test whether its ability to inhibit the growth of MCF-7 cells is dose-dependent. This is achieved using a **dose-response experiment**, which simply involves measuring the effects of multiple different doses of the extract on the phenotype of interest. Four separate dose response experiments were conducted using doubling dilutions of herb extract from a maximum dose of 128 µg/ml, down to a lowest dose of 0.5 µg/ml. A series of dilutions of DMSO alone (the vehicle) were also prepared to match the concentration of DMSO present in each dilution of extract. The effects of these preparations on growth or viability of MCF-7 cells after 24 hours of treatment was then measured using the crystal violet assay. Download the `mcf7_dose_response.xlsx` file to obtain the raw data.

Your next task is to analyse the data from these four experiments, by first performing background subtraction of the positive control (SDS treatment for complete cell killing) and normalising the absorbance values to a percentage of the negative control (medium alone). Then use a two way ANOVA to test if any differences in cell growth seen between the extract dose curve and the DMSO dose curve are statistically significant, as explained in **Method sheet 108**.

Finally you should calculate the concentration of extract necessary to achieve a 50% inhibition of cell growth. This value, referred to as the $IC_{50}$ (inhibitory concentration 50%), is a key metric in the drug discovery process as it gives an indication of how potent your hit extract or compound is. $IC_{50}$ can be calculated by fitting a 4 parameter logistic (4PL) curve to the dose response data. This can be easily achieved using specialised data analysis tools such as GraphPad Prism or R. However, it can also be calculated using Microsoft Excel, as explained in **Method sheet 112**.

Got this far and still have time left in the analysis phase of your project? Please see the following section for suggestions of optional, additional analyses you can try to extend your project if you'd like to do so.

# Writing up: Advice for your project dissertation

Congratulations on the completion of your data collection phase! Now comes the writing up. Giving as much care and attention to this process as your experimental phase is necessary to help you score the highest possible marks in your dissertation. If you would like some advice on how to begin writing up your project report or dissertation, please refer to our notes in **Method Sheet 114**.

# Suggestions for project 'Extension' or upgrading to 'Level 7' (MSc) work

It should be possible to complete the analyses suggested above for the basic antibiotic discovery data project within about **8-10 weeks**. However, if you have managed to complete all of these analyses, and still have more time available, please discuss with your supervisor whether it would be appropriate for you to attempt one or more of the additional analyses shown below to further explore your hits and extend your project.

If you are completing a project at Masters level (i.e. Level 7), or have much more time available, you should aim to complete at least two of the additional analyses shown below. These additional analyses do not have to be performed in any particular order, and you are of course free to modify them with the approval of your supervisor to better fit your own project aims.

## Project variation 1 - Test if the hit extracts are also toxic to bacterial cells

Some molecules are toxic to both eukaryotic and prokaryotic organisms. To test whether your top hits are, in addition to being able to kill cancer cells, also capable of killing bacteria, you could explore the first part of Project 05. This data-analysis project explores the potential of the same 800 *Phytotitre* extracts to kill *Escherichia coli* (a Gram-negative bacterial species) or *Micrococcus luteus* (a Gram-positive bacterial species). Testing this hypothesis would be helpful from a drug discovery perspective, since it would be preferable for a new chemotherapy drug to cause no harm or imbalances in the resident bacteria of the human gut, since these organisms are essential for health and digestion.

## Project variation 2 - Align the cell viability data with traditional medicinal use and epidemiological data of hit extracts

Most of the plants that comprise the *Phytotitre* collection are either edible plants, or plants that have been used historically as a traditional medicine. If any of the hits from your analyses are common dietary constituents, perform a literature search to establish whether dietary intake of this plant is associated with lower risk of cancer in epidemiological studies. If the plant is used as a traditional medicine, perform a literature search to explore whether it has been used in the past to treat cancer, or any other relevant health condition. Combine the insight you gain from these searches with your interpretation of the assay data when writing your discussion section.

# Troubleshooting FAQ

Science does not progress without challenges! It is very unusual for a student project to reach completion without facing some difficulties along the way. A key element of your training is learning how to identify problems and apply solutions as you encounter them. The FAQ section below lists some issues that are commonly seen in such projects, and approaches you can try to remedy them.

## 1) Some of the growth values are greater than 100% after normalisation of the data - why is this?

**Probable Cause:** This is commonly seen in natural extract library screens. In this example, while some extracts may inhibit cell growth, others might encourage it by acting as a rich source of nutrients to support their metabolism.

**Suggested Action:** Do not discard or attempt to modify these data, keep them within your dataset and include them as they are in the following analytical steps.

## 2) Some of the growth values are below zero after normalization - is that possible?

**Probable Cause:** These values do not mean there are less than zero cells remaining in the well, it just means the test well had a lower absorbance than the mean of your positive (100% inhibition) controls.

**Suggested Action:** Values only slightly below zero are common and statistically normal due to minor pipetting variations. Values that are substantially below zero typically arise because of loss of extract colour over time as certain cell-lines can metabolise some plant pigments. You should retain the negative values within the dataset for your mean/ANOVA calculations to maintain statistical integrity, but for your final IC50 graphs, you can treat them as “0% growth”, for example by setting the minimum value of the y-axis on your chart to zero.

## 3) I can’t get the sort tool to work in Excel, the values are still in the same order as before or not sorting properly - why is this?

**Probable Cause:** Excel can only sort numerical values, not cell references.

**Suggested Action:** Ensure you paste the data table as ‘Values’ (Paste special - Values), rather than a simple “Paste”, or Ctrl-V alone before you attempt a sort function. If you paste without this ‘Values’ distinction, the cell references will be copied and pasted into the destination cells. Click on one of the cells in your table. If it says something like “=AVERAGE(C8:C16)”, or any other cell-reference, you have to delete the table and start again, pasting the table as values only. Once every well has a value (e.g. “0.1158”), then you will be able to sort the values appropriately. Remember also when sorting to select the whole table, not just individual columns, before sorting. Failing to do this can also result in an incorrect sorting result.

## 4) My “hits” are often different for the different cell-lines - why is this?

**Probable Cause:** Genetic variation between tumour cell-lines.

**Suggested Action:** This is a common observation in tumour cell-line screening. Every different tumour cell-line has evolved via a different route, accumulating a different set of mutations. This means different tumours escape regulation of proliferation via different mechanisms. A compound or extract may work well against one target of cellular transformation, but not another. In this way, extracts or compounds can inhibit the growth of some types of cancer cell, but not others.

## 5) My ANOVA gives “significant” results, but the bar chart error bars overlap heavily - is this possible?

**Probable Cause:** ANOVA tests analyse the variance across the whole experiment, not just between two means.

**Suggested Action:** Overlapping standard deviation error bars do not always mean “not significant”. If you have used ANOVA with Dunnett’s test (comparing each value to a control) or Tukey’s test (comparing each treatment to every other treatment), it is quite common to find statistically significant differences between treatments even if there is some overlap between the error bars.

## 6) My IC50 curve is a flat line or looks like a “V” shape - what happened?

**Probable Cause:** Incorrect data pairing or non-logarithmic scaling.

**Suggested Action:** First, ensure your x-axis is set to a Logarithmic scale (the data will not plot well on a linear scale). Second, check that you haven't accidentally included the "0 concentration" (Control) point in the log-plot, since zero values cannot be plotted on a log-scale axis. Finally, ensure your concentrations are paired with the correct absorbance values; a "V" shape often means the data was sorted incorrectly.

## 7) I can't get the 4-parameter logistic (4PL) formula to work in Excel - why is this?

**Probable Cause:** Poor initial “guesses” for the slope and IC50 values.

**Suggested Action:** When using Excel to calculate IC50 using a 4PL model, it is necessary to begin by estimating the IC50 and slope values. Try to make sure you make a good first guess with the IC50 value by looking at where the 50% inhibition value lies on your x-axis. Then, try various values for the slope between -1.0 and -3.0. Keep adjusting these until the curve fits reasonably well between the points. After that, adjust the values more finely by adjusting the slope and IC50 estimates in a stepwise fashion until the sum of residuals is as low as you can make it. More advice on these functions are available in **Method sheet 112**.

# Project Progress Checklist

The following checklist table will enable you to keep track off the tasks you should aim to complete for the basic project before entering the writing-up phase.

| Step | Task | ☑ |
|---|---|---|
| 1 | Read the introduction to the *Phytotitre* library (Method sheet 100) | ☐ |
| 2 | Familiarise yourself with relevant Health and Safety information | ☐ |
| 3 | If necessary, complete a Research Ethics application | ☐ |
| 4 | Make sure the necessary IT resources are in place | ☐ |
| 5 | Plan strategies for Good Laboratory Practice, Data Management and Integrity | ☐ |
| 6 | Download the necessary Excel files and Method sheets (Method sheet 101) | ☐ |
| 7 | Perform background correction of plate data (Method sheet 102) | ☐ |
| 8 | Perform normalisation of the data (Method sheet 103) | ☐ |
| 9 | Calculate Z' factor values for each of your plates (Method sheet 104) | ☐ |
| 10 | Map experimental data to plate maps (Method sheet 105) | ☐ |
| 11 | Identify hits from the primary screen (Method sheet 106) | ☐ |
| 12 | Repeat the above steps for the other cell-lines | ☐ |
| 13 | Explore correlations within and between the HTS datasets (Method sheet 109) | ☐ |
| 14 | Plot a bar chart of your top hit extracts (Method sheets 107 & 108) | ☐ |
| 15 | Examine dose response assays of individual hits (Method sheets 108 & 112) | ☐ |
| 16 | Write the dissertation - hints, tips and advice available in Method sheet 114 | ☐ |

**Found a Bug?**
We care greatly about the utility of our resources for customers, and are very responsive to user feedback. If you spot an error, incorrectly numbered method sheet, or a broken link, please feel free to report the issue by sending an email to: contact@caithnessbiotechnologies.com.