

Term sheet - Let's Experiment

Module 1

Neil's case study

Lateral branches/roots- Lateral roots extend horizontally from the primary root (radicle) and serve to anchor the plant securely into the soil. This branching of roots also contributes to water uptake, and facilitates the extraction of nutrients required for the growth and development of the plant.

(https://en.wikipedia.org/wiki/Lateral_root)

Polyethylene glycol (PEG)- A polyether compound with many applications, from industrial manufacturing to medicine. PEG is also known as polyethylene oxide (PEO) or polyoxyethylene (POE), depending on its molecular weight. The structure of PEG is commonly expressed as $H-(O-CH_2-CH_2)_n-OH$. (https://en.wikipedia.org/wiki/Polyethylene_glycol)

Ana's case study

Cancer- is a group of diseases involving abnormal cell growth with the potential to invade or spread to other parts of the body. These contrast with benign tumors, which do not spread to other parts of the body. (<https://en.wikipedia.org/wiki/Cancer>)

Oncogene- A gene which has the potential to cause cancer.

Overexpression- The process of artificially turning on the expression of a gene to increase the amount of mRNA and, therefore, usually the amount protein produced by that gene.

HER2- HER2 (human epidermal growth factor receptor 2) is a gene that can play a role in the development of breast cancer. HER2, also known as ERBB2 in humans, is the receptor tyrosine-protein kinase encoded by the *ERBB2* gene.

(<https://www.breastcancer.org/symptoms/diagnosis/her2>)

HER3- Receptor tyrosine-protein kinase erbB-3, also known as HER3 (human epidermal growth factor receptor 3), is a membrane bound protein that in humans is encoded by the ERBB3 gene. ErbB3 is a member of the epidermal growth factor receptor (EGFR/ERBB) family of receptor tyrosine kinases. The kinase-impaired ErbB3 is known to form active heterodimers with other members of the ErbB family, most notably the ligand binding-impaired ErbB2.

(<https://en.wikipedia.org/wiki/ERBB3>)

Druggable target- A biological gene, protein or mechanism that is targeted by a pharmaceutical compound.

Herceptin- Trastuzumab, sold under the brand name Herceptin, among others, is a monoclonal antibody used to treat breast cancer. Specifically, it is used for breast cancer that is HER2 receptor-positive. It may be used by itself or together with other chemotherapy medication. (<https://en.wikipedia.org/wiki/Trastuzumab>)

Module 2

Independent variable- The variable that you change in an experiment.

Dependent variable- The variable that you measure in an experiment.

Controlled variables- The variables that you ensure do not change between samples in an experiment or between repeat experiments.

Confounding variables- The unaccounted for variables that may differ between samples or repeat experiments, and affect both dependent and independent variables, causing an incorrect interpretation of results (or preventing an interpretation of the results).

Positive control- A validation control that demonstrates the presence of an effect in order to show that your experiment actually worked. Most of the time we get negative results, so you need a positive signal to show that the method worked even when there truly is no effect in your experimental group.

Negative control- A control that demonstrates the absence of an effect in order to show a baseline readout for the experiment. Negative controls are the unperturbed condition and are essential to determine the true effect of the perturbed condition.

Biological variation- The variability between individual samples that's naturally occurring, i.e., inherent to the biological system itself.

Technical variation- The variability between samples and/or repeat experiments that's caused by the experimenter's technique and/or their equipment or reagents.

Clones- A set of cells (or, in some cases, organisms) that are genetically identical.

Biological replicates- Independent samples in the same group of an experiment (e.g., the control or the experimental group) that are treated exactly the same way. They help you understand the degree of variability in the system you're studying.

Technical replicates- Multiple measurements made on the same sample. They help you understand the degree of variability that is caused by a certain aspect of the experimental technique (e.g., instrument error). They serve as internal quality checks of how the experiment performed.

Power analysis- Power analysis calculations are an approach to sample size estimation, which requires information from the literature or pilot experiments (for values of variables in the equation). This approach uses this information with a statistical model to explicitly calculate sample size, as opposed to “guesstimating” the best sample size.

Sample size- The number of biological replicates that you include in each group of an experiment, often called the ‘N’ of an experiment.

Minimal effect size- This is the smallest effect size you will be able to see in your experiment. Effect size is the size of the difference between the two groups you are comparing (e.g., drug vs. placebo).

False positive- A false positive is a result that indicates a given condition exists, when it does not. It is the rejection of a true null hypothesis (also known as a "type I error" finding). (https://en.wikipedia.org/wiki/Type_I_and_type_II_errors)

False negative- A false negative is a result that indicates, erroneously, no effect has been inferred when in truth there is an effect. It is the failure to reject a false null hypothesis (also known as a "type II error" finding). (https://en.wikipedia.org/wiki/Type_I_and_type_II_errors)

Probability of error- Alpha is the probability at which you will accept a false positive, which occurs when your experiment detects a difference between two groups, when no true difference exists. Say, if you choose an alpha value of 0.05 (very common), this means that you will have a 5% chance of seeing a false positive. Beta is the probability at which you will accept a false negative, which occurs when your experiment does not detect a difference between two groups when a true difference indeed exists.

Module 3

Rigor- The strict application of the scientific method to ensure robust and unbiased experimental design, methodology, analysis, interpretation and reporting of results.

Transparency- Scientific transparency refers to the act of conducting research without holding back data. Transparency also includes the practice of openly sharing complete and understandable methods for the research conducted, as well as sharing all relevant data with funding sources, publishers and collaborators.

Experimenter bias- Refers to designing, performing, or analyzing an experiment in a way that's skewed toward a particular conclusion. Bias runs the gamut from deliberate misconduct to unconsciously flawed experimental design. Here are some ways that these biases can surface, consciously or not:

- *Relaxed design:* not being rigorous in how you design the experiment (e.g., not doing all the proper controls, not taking sufficiently representative samples, or not using appropriate sample size).
- *Selective measurements:* consciously or unconsciously being selective in how you collect data to influence the outcome.
- *Selective analysis:* P-hacking (looking for an analytical test that will ultimately give you a statistically significant result).
- *Selective interpretation:* being selective in how you interpret data (e.g., cherry-picking or inappropriate inclusion or exclusion of data).
- *Reporting bias:* selective revealing or suppression of information.
- *Confirmation bias:* interpreting information or results in a way that confirms one's preexisting beliefs or hypotheses.

Reproducibility- The extent to which consistent results are obtained when an experiment is repeated using the same methodology described in corresponding scientific evidence (e.g. a publication in a peer-reviewed journal).

Randomization- A type of unbiased sampling, when experimental participants, cells, or animals have equal chance of being selected or assigned to a group (e.g., control vs. experimental).

Blinding- A blind or blinded-experiment is an experiment in which information about the test is masked (kept) from the participant, to reduce or eliminate bias, until after a trial outcome is known. It is understood that bias may be intentional or subconscious, thus no dishonesty is implied by blinding. If both tester and subject are blinded, the trial is called a double-blind experiment. (https://en.wikipedia.org/wiki/Blinded_experiment)

Measurement bias- Experimental errors that favor a particular result. (For example, if you were weighing mice and were favoring a hypothesis in which one group of mice gained weight, you might re-weigh a mouse thinking the scale was off if it lost weight and take the higher result). On the other hand, when the data matches your expectations you may not double check it. This systematic error would favor one result over the other.

Qualitative data- Data that is descriptive information that is not expressed numerically. One example of this kind of data is categorical data, such as the diagnostic stages of human cancers, or the distinct eye colors of fruit flies. Another example of qualitative data can be visual, such as the immunostaining of biological tissue on slides.

Categorical data- Categorical data, as the name implies, is grouped into some sort of category or multiple categories.

(<https://study.com/academy/lesson/categorical-data-definition-analysis-examples.html>)

Quantitative data- Data that is information that is expressed numerically. For example, measuring the concentration of a protein in blood serum would yield quantitative data.

Continuous data- Numerical data that can be any value across an infinite range, such as weight (e.g., 143.278 mg); it can also be a value over a defined range, such as percentages (e.g., 57.922%) where you can't have values below 0% or above 100%.

Discrete data- Numerical data that are discrete counts, so only a finite number of values is possible. For example, the number of pups per mouse litter or the number of children in a classroom (you can't have a fraction of a pup or a child).

Preregistration- When you preregister your research, you're simply specifying your research plan in advance, before you gather data. Preregistration separates hypothesis-generating (exploratory) from hypothesis-testing (confirmatory) research. (<https://cos.io/prereg/>)

P-value- When you perform a hypothesis test in statistics, a p-value helps you determine the significance of your results. Hypothesis tests are used to test the validity of a claim that is made about a population. The claim that's on trial, in essence, is called the null hypothesis. All hypothesis tests ultimately use a p-value to weigh the strength of the evidence (what the data are telling you about the population). The p-value is a number between 0 and 1 and interpreted in the following way:

- A small p-value (typically ≤ 0.05) indicates strong evidence against the null hypothesis, so you reject the null hypothesis.
- A large p-value (> 0.05) indicates weak evidence against the null hypothesis, so you fail to reject the null hypothesis.

(<https://www.dummies.com/education/math/statistics/what-a-p-value-tells-you-about-statistical-data/>)

Module 4

Protocol- An experimental protocol is a written procedural method for the design and implementation of a particular experiment. Protocols are written whenever it is desirable to standardize a laboratory method to ensure successful replication of results by others in the same laboratory or by other laboratories.

([https://en.wikipedia.org/wiki/Protocol_\(science\)](https://en.wikipedia.org/wiki/Protocol_(science)))

Key reagents- Experimental reagents that are essential to the success of the experiment.

Antibodies- A blood protein produced in response to and counteracting a specific antigen. Antibodies combine chemically with substances which the body recognizes as alien, such as bacteria, viruses, and foreign substances in the blood. These can be artificially designed to identify specific compounds in scientific experiments outside the body.

Cell lines- A cell culture developed from a single cell and therefore consisting of cells with a uniform genetic makeup.

Validate reagents- Refers to testing that the key experimental reagents work properly. For example, validating that an antibody is actually specific to the target molecule.

Authenticate reagents- Refers to verifying that the reagent is what it says it is. For example, if you order a cell line, it is important to verify they are actually the cells you ordered. Additionally, authentication refers to verifying that the cell line is derived from the stated cell-type and organism.

Module 5

Pilot experiment- A small-scale experiment that is done to help evaluate the feasibility, cost, and time of doing the experiment, with the goal of improving the design before doing it full-scale.

Troubleshooting- The process of changing, adjusting, and remaking an experiment so that it “works” (i.e., has an interpretable result).

Optimization- Systematically modifying and testing the components and variables of the experimental system to ensure the method of detection is specific and the results are accurate and robust.

Dose-response study- Describes the change in effect on an organism or biological process caused by differing levels of exposure (or doses) to a stressor (such as a chemical or an environmental factor) after a certain exposure time.

Orthogonal approaches- Using another experimental technique or another experimental system to address the same fundamental overarching research question.

Research meta-framework- The meta-framework starts with an overarching question. This project-based question seeds a hypothesis and a more focused experimental question to test this hypothesis that, when addressed effectively, helps to develop a predictive model about the underlying biology. From this model, you develop another hypothesis and experimental question that you test, and you repeat the process over and over again. In total, these sequential

experimental questions & results comprise a supporting network of orthogonal approaches that serve to address your fundamental research question and build a coherent story.