Developing a Research Question and Testable Hypothesis

1. Key considerations in planning a research study
   a. Ask a question worth answering
   b. How will you answer the question?
   c. Share ideas with colleagues

2. Preparing a Proposal
   a. Short and focused: state key points and basic concepts directly, avoid jargon
   b. Use pictures and graphics to enhance presentation

3. Hypothesis
   a. Choose an important, focused, testable hypothesis that increases understanding of biologic process, disease pathology, treatment or prevention, resource utilization
   b. Must be provable

4. Study Proposal Organization:
   a. **Abstract:** *what, why and how.*
      i. Introduce research question and provide brief background and significance
      ii. State hypothesis and specific aims
      iii. Describe approach/methods
      iv. Write this last.
   b. **Specific Aims:** *what* you are doing
      i. Describe long-term goals
      ii. Propose research question
      iii. Clearly state hypothesis
      iv. List specific aims necessary to prove (or disprove) hypothesis
   c. **Background and Significance:** *why* you are doing it
   d. **Preliminary Studies and Results:** *what* you have already done and *how* you did it
   e. **Research Design and Methods:** *how* you plan to do the research, *why* you are doing it this way, *when* and *where* you are doing it
Basic Study Design

1. **Causal inference and association**
   a. What we want is causation but what we can see is association
      i. Causation ≠ Association (e.g., correlation)
      ii. Association: co-occurrence of the exposure and outcome
      iii. Sole plausible explanation: after one considers all the alternative explanations for the observed association between the exposure and the outcome, the only plausible explanation is that the exposure has caused the outcome
   b. Five possible explanations for association between intervention and outcome
      i. Causal effect of exposure of interest (what we want)
      ii. Bias
         1. Selection bias
         2. Measurement bias
      iii. Confounding: a common cause of exposure and outcome
      iv. Chance
         1. 95% confidence interval
         2. p-value
            a. Power
            b. statistical significance versus minimally clinical important difference
      v. Reverse causation
         1. Temporal priority: the exposure is present before the outcome

2. **Basic methodology**
   a. Basic Measures
      i. Risk: The number of outcome-free people in a fixed cohort who obtain the outcome over specified time period
      ii. Rate (Incidence rate): The number of new cases in a fixed cohort who obtain the outcome per unit of time (e.g. 1-person-year) over specified time period
      iii. Odds: A probability divided by its complement
      iv. Prevalence: Proportion of people with an outcome during a designated time period
      v. Incidence: Proportion of new cases of an outcome during a designated time period
   b. Basic Quantification in Clinical Research
      i. Ratio: 1=null value
         1. Risk ratio: Risk in exposed ÷ Risk in unexposed
         2. Rate ratio: Rate in exposed ÷ Rate in unexposed
         3. Odds ratio: Odds in exposed ÷ Odds in unexposed
ii. Difference: 0=null value
   1. Risk difference: Risk in exposed - Risk in unexposed
   2. Rate difference: Rate in exposed - Rate in unexposed
   3. Odds difference: Odds in exposed - Odds in unexposed

c. Types of data (variables)
   i. Quantitative
      1. Continuous (ex. Height, weight)
      2. Discrete (Ex. # of siblings)
   ii. Qualitative
      1. Nominal (ex. Race)
      2. Ordinal (ex. Sizing of clothes)

d. Basic Statistics

<table>
<thead>
<tr>
<th>Predicators</th>
<th>Outcome variable</th>
<th>Continuous (Discrete)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Categorical</td>
<td>Dichotomous</td>
<td>&gt;2 Categories</td>
</tr>
<tr>
<td></td>
<td>(2x2 table) Chi square tests; 2-sample proportion z-test</td>
<td>Chi square test 2-sample means t-test; ANOVA</td>
</tr>
<tr>
<td>&gt; 2 Categories</td>
<td>Chi square test</td>
<td>ANOVA</td>
</tr>
<tr>
<td>Continuous (or continuous &amp; Categorical)</td>
<td>Logistic regression</td>
<td>Polytomous regression</td>
</tr>
</tbody>
</table>