Introduction
Why are we here?

- Project management is complicated
- Academic project management is especially complicated
"This could be the discovery of the century. Depending, of course, on how far down it goes."
Wisdom comes from experience, and experience comes from making mistakes.

Omar Bradley, WWII General

Wisdom is the daughter of experience.

Leonardo da Vinci
Learn from the mistakes of others. You can’t live long enough to make them all yourself.

*Eleanor Roosevelt*

*US diplomat & reformer (1884 - 1962)*
To assure that you will never experience failure:

- don't take a risk,
- don't attempt anything new,
- don't expand on your ideas,
- don't set goals.

*Catherine Pulsifer, from Failure Is A Golden Opportunity*
Life Cycle of Research
Components of a Research Project
Management Fundamentals
Management = Answers to:

• Who?
• What?
• When?
• Where?
• Why?
• How?
What are we managing?

- Time
- Resources
- Direction
- People
- Data
Direction

If a man does not know to what port he is steering, no wind is favourable to him.

Seneca

Roman dramatist, philosopher, & politician
(5 BC - 65 AD)
People

• The Golden Rule
• Stewardship – replacing control with partnership and choice
• Fun
• Matching people with tasks so that they experience success
• Motivation
“If you want to build a ship, don't drum up people together to collect wood and don't assign them tasks and work, but rather teach them to long for the endless immensity of the sea.”

Antoine de Saint-Exupery
Key Components of Data Management Strategy

1. Planning
2. Creation of data
3. Documentation
4. Database system / Data storage mechanism
5. Naming conventions and codes / Data Dictionary
6. Security
7. Maintenance
Planning

• Deals with understanding the goal of the project, feasibility issues, deadlines, requirement analysis, assignment of responsibilities and establishment of data analysis mechanisms.

• The most important component of data management.

• Aids in better utilization of time and resources.

• Critical in designing the data capture process and data storage mechanism (data base schema).
Creation of data

There are three kinds of data collected or prepared during any research project:

• *Raw data:* Primary quantitative and empirical data collected from subjects / participants of the research. This is the data provided by the survey takers.

• *Derived data:* This is derived from the raw data. This could be based on some formula or logical reasoning or transitivity.

• *Analysis data:* This is the data used for the purpose of statistical analysis. This deals with representation suitable for data analysis etc.
Documentation

• Study specific documentation about data management principles, final signed protocols, clinical data handling conventions and data entry guidelines needs to be maintained.
• This will help in resolving issues related to data management and will also help in auditing and training new research personnel.
• Documentation process should be strictly followed from the planning phase to the maintenance phase.
Database system / Data storage mechanism

• Commercially available databases like MSSQL, MySQL or Access can be used.

• The advantage of using such systems is that they are reliable, scalable and already have built-in mechanisms for most of the data management tasks like data retrieval, data entry, optimization etc.

• We could also use static storage systems like spreadsheets.

• Such storage mechanism needs to be setup after proper planning so that the structure of the system is scalable and precise.
Naming conventions and codes / Data Dictionary

- When we build storage mechanisms, the data entry is done in a manner which is convenient for the data analysis phase.
- Hence proper coding schemes need to be agreed upon and followed and the process should be documented.
- Data dictionary is used to document how the database is structured and what each table represents. This will help in better understanding of the structure of the database and also aid in analyzing the data.
Security

• Mechanism to ensure that the data that was entered into the system is safe and consistent by controlling access to data

• Consistency, privacy and clarity of data can be maintained by having a very secure data repository

• Security is one of the most important aspects of data management and also the most vulnerable one
Maintenance

Maintenance of data includes the following:

• **Backing-up data:** To prevent loss of data due to system crash or any failure, data that is stored needs to be backed up. This is a very important aspect of data management. Back ups can be made to disks or another computer on the network or on a server.

• **Restoring data:** In case of a system failure, data that was lost needs to be restored from a backup. Mechanisms to do the same must be established.

• **Purging data:** This deals with deleting data that is/are no longer needed. This process helps in maintain a clean set up-to-date of data.
“It’s All About Team”
Quality, Transparency, Productivity Principles
Quality

• Protection of human subjects
• Professionalism – interactions with our team members
• Work products
Process quality principles for data life-cycle

1. Data Validation
2. Audit Trail
3. Quality Assurance (QA)
4. Archiving
Data Validation

- Data validation should be done right from when the data is captured to the point of data maintenance.
- This process helps in achieving an accurate, consistent and true representation of the captured data in the database.
- Source Data Verification (SDV) is an important step in data validation to ensure correctness in the data captured at the source and its representation in the database. This step is typically carried out as soon as the surveys/instruments are collected from the participants.
- During data entry, validation scripts should be used to ensure correctness of the data being entered. Also double entry of captured data will help in capturing any discrepancies in data entry.
Audit Trail

- Audit trails are used to record transactions on the system. This includes information on who accessed the system, what transaction was performed and when was it performed.
- They are useful for maintaining security as well as for recovering lost transactions by allowing reconstruction and examination of sequence of events.
- They can also be used to keep track of decisions made and procedures or rules introduced.
Quality Assurance (QA)

• It is the activity of ensuring quality in every phase by making sure that quality standards and procedures are being followed.

• This process is very critical in research because without QA there is no guarantee that the data that has been collected is consistent, reliable and precise.

• Usually a QA team is formed to check the records and make sure that the guidelines and protocols are being followed. This is done throughout the data life-cycle.

• The dimensions of Data Quality Assurance are precision, completeness, accuracy, non-duplication, portability and credibility.
Archiving

- Consists of records selected for permanent or long-term preservation.
- Consolidated data are moved from a primary online storage medium to a secondary offline storage medium.
- Data archiving should be done on a regular basis to prevent complete loss of data. Although back ups are done on a frequent basis, it is only a snap shot of data and not a preservation mechanism.
Transparency

- Full disclosure – principle of revealing all aspects of scientific method
- Allows team members to go on vacation
- Archive as you go
How to achieve transparency?

• Establish protocols for documentation
• Keep research notebook
• Generate technical reports for work products
• CONSORT checklist

www.consort-statement.org
Table of Contents for Research Notebook

- Team
- Work plan
- Literature Review
- Abstract
- Specific Aims and Hypotheses
- Background and Significance
- Study Design
- Measures
- Methods
- Human Subject Protection
- Description of Resources
Table of Contents for Research Notebook

- Budget
- Proposal
- IRB Protocol
- FERPA, HIPAA Certificates
- Data dictionary
- Results Templates
- Statistical analyses
- Analysis Log
- Results Tables
- Presentation
- Manuscript
- Appendix
## CONSORT CHECKLIST

<table>
<thead>
<tr>
<th>PAPER SECTION And topic</th>
<th>Item</th>
<th>Description</th>
<th>Reported on Page #</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TITLE &amp; ABSTRACT</strong></td>
<td>1</td>
<td>How participants were allocated to interventions (<em>e.g.</em>, &quot;random allocation&quot;, &quot;randomized&quot;, or &quot;randomly assigned&quot;).</td>
<td></td>
</tr>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td></td>
<td>Scientific background and explanation of rationale.</td>
<td></td>
</tr>
<tr>
<td>Background</td>
<td>2</td>
<td>Scientific background and explanation of rationale.</td>
<td></td>
</tr>
<tr>
<td><strong>METHODS</strong></td>
<td>3</td>
<td>Eligibility criteria for participants and the settings and locations where the data were collected.</td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td></td>
<td>Eligibility criteria for participants and the settings and locations where the data were collected.</td>
<td></td>
</tr>
<tr>
<td>Interventions</td>
<td>4</td>
<td>Precise details of the interventions intended for each group and how and when they were actually administered.</td>
<td></td>
</tr>
<tr>
<td>Objectives</td>
<td>5</td>
<td>Specific objectives and hypotheses.</td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td>6</td>
<td>Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (<em>e.g.</em>, multiple observations, training of assessors).</td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>7</td>
<td>How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules.</td>
<td></td>
</tr>
<tr>
<td>Randomization --</td>
<td>8</td>
<td>Method used to generate the random allocation sequence, including details of any restrictions (<em>e.g.</em>, blocking, stratification)</td>
<td></td>
</tr>
<tr>
<td>Sequence generation</td>
<td></td>
<td>Method used to generate the random allocation sequence, including details of any restrictions (<em>e.g.</em>, blocking, stratification)</td>
<td></td>
</tr>
<tr>
<td>Randomization --</td>
<td>9</td>
<td>Method used to implement the random allocation sequence (<em>e.g.</em>, numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned.</td>
<td></td>
</tr>
<tr>
<td>Allocation concealment</td>
<td></td>
<td>Method used to implement the random allocation sequence (<em>e.g.</em>, numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned.</td>
<td></td>
</tr>
<tr>
<td>Topic</td>
<td>Page</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Randomization -- Implementation</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blinding (masking)</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical methods</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RESULTS</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant flow</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recruitment</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline data</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numbers analyzed</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcomes and estimation</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ancillary analyses</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adverse events</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DISCUSSION Interpretation</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generalizability</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall evidence</td>
<td>22</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Randomization -- Implementation**: Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups.

**Blinding (masking)**: Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. When relevant, how the success of blinding was evaluated.

**Statistical methods**: Statistical methods used to compare groups for primary outcome(s); Methods for additional analyses, such as subgroup analyses and adjusted analyses.

**RESULTS**

**Participant flow**: Flow of participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome. Describe protocol deviations from study as planned, together with reasons.

**Recruitment**: Dates defining the periods of recruitment and follow-up.

**Baseline data**: Baseline demographic and clinical characteristics of each group.

**Numbers analyzed**: Number of participants (denominator) in each group included in each analysis and whether the analysis was by "intention-to-treat". State the results in absolute numbers when feasible (e.g., 10/20, not 50%).

**Outcomes and estimation**: For each primary and secondary outcome, a summary of results for each group, and the estimated effect size and its precision (e.g., 95% confidence interval).

**Ancillary analyses**: Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those pre-specified and those exploratory.

**Adverse events**: All important adverse events or side effects in each intervention group.

**DISCUSSION Interpretation**: Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision and the dangers associated with multiplicity of analyses and outcomes.

**Generalizability**: Generalizability (external validity) of the trial findings.

**Overall evidence**: General interpretation of the results in the context of current evidence.
Productivity

• Hallmark of a highly effective team
• Everyone benefits
• “You are not here merely to make a living. You are here to enable the world to live more amply, with greater vision, and with a finer spirit of hope and achievement. You are here to enrich the world. You impoverish yourself if you forget this errand.”

Woodrow Wilson

28th president of US (1856 - 1924)
How to be productive?

• Set S.M.A.R.T. goals
  Specific-Measurable-Attainable-Realistic-Timely

• Have regular team meetings
• Form action plans
• Establish ethic of mutual accountability
Thank you!