THOUGHTFUL DATA COLLECTION: GETTING THE BEST DATA FOR YOUR STUDY, BY DESIGN

ITHS | Institute of Translational Health Sciences
Accelerating Research. Improving Health.
Learning objectives

Describe how to match the appropriate data collection method to your study populations

What type of data collection should you pick for your study?

Apply best practices to maximize data quality and quantity to real life scenarios

Learn about best practices about data collection and how they relate to data analysis

Practice evaluating your data needs

Put all this new knowledge to work
Een goed begin is het halve werk!
- Dutch Proverb -

A good start is half the job!
- Rough translation -
Your data collection does not live in a vacuum

► Principal investigators
► IRB’s
► Funding sources
► Participants
► Your employees or coworkers
► University / Departments

Your data is the lynch pin in your study
So where to start?

*Let me tell you how I work*...
What’s your ultimate goal?

Answer a specific set of research questions

*The classic research model*

Creating a repository for future studies

*Collecting simple contact information and/or consent might be beneficially to future studies*

To get funding for my next research project

*Larger grants sometimes want to see pilot data*

“Fishing expedition”

*Collecting as much data as possible in the hope that something will be relevant and/or significant*
Determine some basic aspects of your study

- Target population
- Available resources
- Target sample size
- Variables
- Sites

Ballpark estimates are the goal during this phase
Target population

- General age
- Willingness to participate
- (Computer) literacy
- Tolerance of contact
  - Number of encounters
  - Timing of encounters
- Location
  - Where do your subjects live?
Available resources

► People
► Time
► Money
► Tools
► Internet availability
Resources

People
- Available FTE
- Expertise
  - Study design
  - People skills
  - Computer savviness
- Compensation
- Participants can be resources!
### Resources

<table>
<thead>
<tr>
<th>People</th>
<th>Time</th>
<th>Money</th>
<th>Tools</th>
<th>Internet Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Time

- Timeline of your study
  - When are you collecting data?
- Design
- Testing
- Deadlines
  - Data entry
  - Data analysis
Resources

- People
- Time
- Money
- Tools
- Internet Availability

Money

- Operating costs
  - Hardware
  - Licensing
  - Office supplies
  - Paper
  - University “Cut”
  - Salaries
  - Consulting fees
- Participant compensation
Resources

People

Time

Money

Tools

Data capture tools
- REDCap, Catalyst, Excel or Access
- Paper
- Industry SaaS options

Data analysis
- Statistics
- Visualization

Hardware
- Tablets
- Desktops or laptops
- B.Y.O.D. (Bring Your Own Device)
## Resources

<table>
<thead>
<tr>
<th>People</th>
<th>Time</th>
<th>Money</th>
<th>Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Internet availability

- **Online**
  - Reliability
  - Speed
  - Email capability

- **Offline**

- **Hybrid**
  - Access depends on:
    - Location
    - Time
    - Costs
Target sample size

- Quantity
- Quality
- Ease of recruitment
- Enrollment ratios
- “Finishers” ratios
- Statistical power
Variables

- Essential vs Optional
  - Crucial for hypothesis
  - Legally mandatory
- Anonymous vs identified
- Nature of variables
  - Low risk
  - Personal Health Information (PHI)
  - “Super” PHI
Sites

- Single site
- Multiple sites
  - Stakeholders
    - Local politics
    - Multiple IRB’s
  - Location
    - Local, National, International
  - Language
  - Data model
    - Federated
    - Spider in a web
Choosing your workflow

- What?
- When?
- Where?
- Who?
- How?
Choosing your workflow

What are you collecting?

- Variables
  - Number
  - Type
  - Nature
- Time commitment (average)
- What are your targets?
  - People
  - Events
  - Things (diagnoses, medications, devices)

What?

When?

Where?

Who?

How?
Choosing your workflow

When are you collecting it?

- Collection points
  - Total number
  - Intervals
  - Triggers
    - Time
      (Daily, weekly, monthly)
    - Event based
      (When participant visits ED)

What?

When?

Where?

Who?

How?
## Choosing your workflow

### Where are you collecting it?

<table>
<thead>
<tr>
<th>Where?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online</td>
</tr>
<tr>
<td>Desktop</td>
</tr>
<tr>
<td>Mobile</td>
</tr>
<tr>
<td>Text message</td>
</tr>
<tr>
<td>Browser</td>
</tr>
<tr>
<td>Phone call</td>
</tr>
<tr>
<td>Offline</td>
</tr>
<tr>
<td>In-person</td>
</tr>
<tr>
<td>Paper</td>
</tr>
<tr>
<td>REDCap Mobile App</td>
</tr>
<tr>
<td>Extraction from a medical system</td>
</tr>
</tbody>
</table>

### Questions

<table>
<thead>
<tr>
<th>What?</th>
</tr>
</thead>
<tbody>
<tr>
<td>When?</td>
</tr>
<tr>
<td>Where?</td>
</tr>
<tr>
<td>Who?</td>
</tr>
<tr>
<td>How?</td>
</tr>
</tbody>
</table>
Choosing your workflow

Who is doing the collecting?
- Participants
- Study personnel
- “Hearsay” people
  - Clinicians
  - Parents
  - Social workers
  - Teachers

What?
When?
Where?
Who?
How?
Choosing your workflow

How are you doing your collection?

► Electronic Data Capture (REDCap)
  ► Data entry
  ► Survey
  ► Data import
► Paper forms
► Machine generated
  ► Images (MRI, CAT, Ultrasounds)
  ► “Fitness trackers”
  ► Data dump from a database (EMR)
► All of the above
Example Cases

- Long Acting Reversible Contraception (LARC)
- Soccer Traumatic Brain Injury (TBI)
- Headache Clinic
- Autism Center
Long Acting Reversible Contraception (LARC)

**What?**
- EMR Medical data (~500 variables)
- Targets: Women with IUD’s

**When?**
- Retrospective over the last 3 years

**Where?**
- Extraction from the EMR data warehouse, the EMR itself

**Who?**
- Data analyst & research coordinators

**How?**
- EMR data dump & manual abstraction
# Soccer Traumatic Brain Injury (TBI)

<table>
<thead>
<tr>
<th>What?</th>
<th>Mental state, trauma data, scoring tools (150 variables)</th>
<th>Targets: 7 to 14 year olds that play soccer</th>
</tr>
</thead>
<tbody>
<tr>
<td>When?</td>
<td>Single time point. At various soccer tournaments</td>
<td></td>
</tr>
<tr>
<td>Where?</td>
<td>On location, Online via mobile devices</td>
<td></td>
</tr>
<tr>
<td>Who?</td>
<td>Research coordinators, parents, children</td>
<td></td>
</tr>
<tr>
<td>How?</td>
<td>Questionnaires on tablets or phones</td>
<td></td>
</tr>
</tbody>
</table>
### Headache Clinic

| What?                  | Headache symptoms, scoring tools (100~200 variables)  
                        | Targets: Adult patients with chronic headaches |
|------------------------|------------------------------------------------------|
| When?                  | Multiple time points. Generally before a clinic visit |
| Where?                 | Online                                               |
| Who?                   | Study participants                                  |
| How?                   | Questionnaires via email invitation, choice of desktop or mobile |
### Autism Center

| **What?** | Social history, scoring tools (200~300 variables)  
|           | Targets: Parents, patients |
| **When?** | Multiple time points. During clinic visits, online follow-up surveys |
| **Where?** | In clinic, online |
| **Who?**  | Study coordinators, parents |
| **How?**  | Questionnaires via email invitation, choice of desktop or mobile, manual data entry online |
Questions?
Workshop Case (Marathon)

**Goal**

Track the wellbeing of the Marathon runners

**Target Population**

Marathon runners doing the 2017 Seattle marathon

**Available Resources**

UW Medicine MD, First aid workers, mobile laboratories, tablets

**Target Sample Size**

~7,000 runners

**Variables**

Injury details, medical history, demographics

**Sites**

14 first aid tents spread out over the course
Workshop Case (Marathon) example solution

| What?    | • Medical history, injury history, local lab results, demographics  
 |          | • Targets: Marathon runners |
| When?    | • Pre-race, during the race, post race |
| Where?   | • 14 First aid tents, online |
| Who?     | • Marathon runners, first aid workers |
| How?     | • Questionnaires via email invitation, choice of desktop or mobile, REDCap mobile app during the race on tablets |
Thank You
Questions?
CONNECT WITH ITHS

www.iths.org

@ITHS_UW

/ithsuw

/InstituteofTranslationalHealthSciences
Visit ITHS.org to Become an ITHS Member

Join a unique catalyst that accelerates discoveries to practice.

Access

Members gain access the different research services, resources, and tools offered by ITHS, including the ITHS Research Navigator.

Education and Training

Members can access a variety of workforce development and mentoring programs and apply for formal training programs.

Funding

Members can apply for local and national pilot grants and other funding opportunities. ITHS also offers letters of support for grant submissions.

Collaboration

Members can connect with collaborators across the CTSA consortium.