

Data Collection and Analysis: Available Resources



Nicklaus
Children's
Hospital

Research Institute

Apeksha Gupta,
MPH

Qualities of Good Data Collection

- Consistency of methodology
- Reliable and error free
- Utility (understandable and interpretable variables)
- Completeness
- If an exempt study, no PHI is collected



Advantages of using REDCap

- **Features** - Electronic questionnaires, public survey, and data repository (much like SurveyMonkey)
- **Fast setup** - A simplified methodology for building databases quickly and easily
- **Secure and web-based** - Secure data input within MCHS network
- **Multi-site access** - REDCap projects can be used by researchers from multiple sites and institutions
- **Flexible** - Fully customizable. You are in total control of shaping your database
- **Mid-study modifications** - You may modify the database at any time during the course of your study
- **Easy in** - Data may be imported from external data sources to begin a study or to provide mid-study data uploads
- **Easy out** - Export data to common data analysis packages
- **Audit trails** - To track data manipulation and export procedures

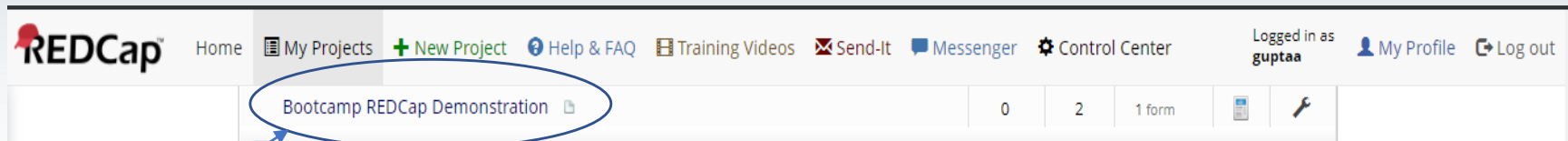
REDCap Best Practices

- Plan your data collection
- Document your data collection needs in a study protocol
- Group your data fields/questions based on how they are being collected.
 - Data source (chart review, patient assessment, lab report, etc.)
 - Time point (baseline clinic visit, follow-up visit, etc).
- Use categorical field types (yes/no, multiple choice, etc.) whenever possible
- Keep forms fairly short to minimize risk of data loss
- Involve a statistician early during the development of the database
- Test and Retest the Project
- Do not change variable names or values of categorical field types once you begin collecting real data

REDCap should be used when your study has...

- More than 1 individual collecting data
- Many variables
- Large number of subjects/participants
- Providing/distributing a survey

REDCap Demonstration Data Collection Sheet



First page in REDCap when you login with your project already created by RAD

Project status: Development

Completed steps 0 of 7

Main project settings

Not started

Enable Use surveys in this project? [?](#) [VIDEO: How to create and manage a survey](#)

Enable Use longitudinal data collection with defined events? [?](#)

[I'm done!](#)

[Modify project title, purpose, etc.](#)

Design your data collection instruments

Not started

Add or edit fields on your data collection instruments. This may be done by either using the Online Designer (online method) or by uploading a Data Dictionary (offline method). Quick links: [Download PDF of all instruments](#) OR [Download the current Data Dictionary](#)

Go to [Online Designer](#) or [Data Dictionary](#)

You may also browse for pre-built data collection instruments in the [REDCap Shared Library](#)

Have you checked the [Check For Identifiers](#) page to ensure all identifier fields have been tagged?

Enable optional modules and customizations

Optional

Enable Repeating instruments [?](#)

Disable Auto-numbering for records [?](#)

Enable Scheduling module (longitudinal only) [?](#)

Enable Randomization module [?](#)

Enable Designate an email field to use for invitations to survey participants [?](#)

[Additional customizations](#)

Click on Online designer

REDCap Demonstration Data Collection Sheet

[Project Home](#) [Project Setup](#) [Online Designer](#) [Data Dictionary](#)

[Create snapshot of instruments](#)

Last snapshot: never ?

[VIDEO: How to use this page](#)

The Online Designer will allow you to make project modifications to fields and data collection instruments very easily using only your web browser. NOTE: While in development status, all field changes will take effect immediately in real time.

Data Collection Instruments

Add new instrument:

[Create](#)

a new instrument from scratch

[Import](#)

a new instrument from the official [REDCap Shared Library](#)

[Upload](#)

Instrument ZIP file from another project/user or [external libraries](#)

Instrument name

My First Instrument

Fields

1

View PDF



Instrument actions

[Choose action](#)

[Return to list of instruments](#)

Current instrument: **My First Instrument**

[Preview instrument](#)

Variable: record_id

Record ID

NOTE: The field above is the record ID field and thus cannot be deleted or moved. It can only be edited.

[Add Field](#)

[Add Matrix of Fields](#)

Create new data collection instruments if you have more than 1

Click on My First Instrument. You can also rename the instrument


Record ID is populated automatically when you start collecting the data

Click on add field to add data collection variables

REDCap Demonstration Data Collection Sheet

Add New Field



You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the  [Field Types video \(4 min\)](#).

Field Type ✓ ---- Select a Type of Field ----

- Text Box (Short Text, Number, Date/Time, ...)
- Notes Box (Paragraph Text)
- Calculated Field
- Multiple Choice - Drop-down List (Single Answer)
- Multiple Choice - Radio Buttons (Single Answer)
- Checkboxes (Multiple Answers)
- Yes - No
- True - False
- Signature (draw signature with mouse or finger)
- File Upload (for users to upload files)
- Slider / Visual Analog Scale
- Descriptive Text (with optional Image/Video/Audio/File Attachment)
- Begin New Section (with optional text)
- Dynamic Query (SQL)

REDCap Demonstration Text Box

Add New Field

You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the [Field Types video \(4 min\)](#).

Field Type: Text Box (Short Text, Number, Date/Time, ...)

Field Label [How to use Piping](#)

Age:

Action Tags / Field Annotation (optional)

[Learn about Action Tags](#) or [using Field Annotation](#)

Variable Name (utilized during data export)

age

☐ Enable auto naming of variable based upon its Field Label?

ONLY letters, numbers, and underscores

Validation? (optional) Number

Minimum: 0

Maximum:

- or -

Enable searching within a biomedical ontology ?

-- choose ontology to search --

Required?* ☐ No ☒ Yes

* Prompt if field is blank

Identifier? ☒ No ☐ Yes

Does the field contain identifying information (e.g., name, SSN, address)?

Custom Alignment Right / Horizontal (RH)

Align the position of the field on the page

Field Note (optional) years

Small reminder text displayed underneath field

REDCap Demonstration Single Answer

Add New Field

You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the [Field Types video \(4 min\)](#).

Field Type: Multiple Choice - Drop-down List (Single Answer)

Field Label [How to use Piping](#)

Sex:

Choices (one choice per line) [Copy existing choices](#)

1, Female
2, Male

☒ **Enable auto-complete for this drop-down** [How do I manually code the choices?](#)

Action Tags / Field Annotation (optional)

☐ [Learn about Action Tags](#) or [using Field Annotation](#)

Variable Name (utilized during data export)

sex ☐ **Enable auto naming of variable based upon its Field Label?**

ONLY letters, numbers, and underscores

Required?* ☐ No ☒ Yes
* Prompt if field is blank

Identifier? ☒ No ☐ Yes
Does the field contain identifying information (e.g., name, SSN, address)?

Custom Alignment Right / Horizontal (RH)

Align the position of the field on the page

Field Note (optional)

Small reminder text displayed underneath field

REDCap Demonstration Multiple Answer

Edit Field

You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the [Field Types video \(4 min\)](#).

Field Type: Checkboxes (Multiple Answers)

Question Number (optional)
Displayed only on the survey page

Field Label [How to use Piping](#)
Symptoms that the patient had:

Choices (one choice per line) [Copy existing choices](#)
0, Fever
1, Nausea
2, Headache
3, Other
[How do I manually code the choices?](#)

Variable Name (utilized during data export)
symptoms
ONLY letters, numbers, and underscores

☐ Enable auto naming of variable based upon its Field Label?

Required?* ☐ No ☒ Yes
* Prompt if field is blank

Identifier? ☒ No ☐ Yes
Does the field contain identifying information (e.g., name, SSN, address)?

Custom Alignment Right / Vertical (RV)
Align the position of the field on the page

Field Note (optional)
Small reminder text displayed underneath field

Save Cancel

REDCap Demonstration Calculated Field

Create date of admit and date of discharge to calculate LOS.

The screenshot displays two variable creation forms in REDCap. The top form is for 'Variable: date_of_admit' and the bottom is for 'Variable: date_of_discharge'. Both forms have a title bar with icons for edit, copy, paste, and delete. The 'Date of admit' form includes a text input field, a calendar icon, a 'Today' button, and a date format dropdown set to 'M-D-Y'. A red asterisk and the text '* must provide value' are shown below the input field. The 'Date of discharge' form has identical elements. Below each form are buttons for 'Add Field' and 'Add Matrix of Fields'. A blue arrow originates from the text 'Create date of admit and date of discharge to calculate LOS.' and points to the 'Date of admit' form.

Variable: date_of_admit

Date of admit:

* must provide value

Add Field Add Matrix of Fields

Variable: date_of_discharge

Date of discharge:

* must provide value

Add Field Add Matrix of Fields

REDCap Demonstration Calculated Field

Edit Field

You may add a new project field to this data collection instrument by completing the fields below and clicking the Save button at the bottom. When you add a new field, it will be added to the form on this page. For an overview of the different field types available, you may view the [Field Types video \(4 min\)](#).

Field Type: Calculated Field

Field Label [How to use Piping](#)

LOS:

Calculation Equation [How do I format the equation?](#)

`datediff([date_of_discharge], [date_of_admit], "days")`

✓ Valid [Clear calculation](#)

Test calculation with a record: -- select record --

Action Tags / Field Annotation (optional)

@HIDDEN

[Learn about Action Tags](#) or [using Field Annotation](#)

Variable Name (utilized during data export)

los ☐ Enable auto naming of variable based upon its Field Label?

ONLY letters, numbers, and underscores

Required?* ☒ No ☐ Yes
* Prompt if field is blank

Identifier? ☒ No ☐ Yes
Does the field contain identifying information (e.g., name, SSN, address)?

Custom Alignment Right / Vertical (RV)

Align the position of the field on the page

Field Note (optional)

Small reminder text displayed underneath field

Save Cancel

Date difference
formula to create LOS

There are many action tags.
@HIDDEN hides LOS from the
survey link. So the respondent
will not be able to see LOS

REDCap Demonstration Branching Logic

Branching Logic: may be employed when fields/questions need to be hidden for data entry under certain conditions. For instance, you may want to hide the question “If yes for surgery, then please specify?” until a “Yes” answer is checked for a previous question, “Had surgery?”

The screenshot displays the REDCap branching logic configuration interface. It consists of two stacked question blocks. The top block is for the variable 'had_surgery' and contains a radio button question: 'Had surgery:' with options 'Yes' and 'No'. A red asterisk indicates that a value must be provided. A 'reset' link is located at the bottom right of this block. Below the question block are two buttons: 'Add Field' and 'Add Matrix of Fields'. The bottom block is for the variable 'surgery_yes' and contains a text input field with the label 'If yes for surgery, then please specify:'. A red asterisk indicates that a value must be provided. Similar to the top block, it has 'Add Field' and 'Add Matrix of Fields' buttons at the bottom.

Variable: had_surgery

Had surgery: ☐ Yes ☐ No

* must provide value

reset

Add Field Add Matrix of Fields

Variable: surgery_yes

If yes for surgery, then please specify:

* must provide value

Add Field Add Matrix of Fields

REDCap Demonstration Branching Logic

Add/Edit Branching Logic ✕

Choose method below for the following field: **surgery_yes** If yes for surgery, then please specify.

☐ **Advanced Branching Logic Syntax** [\(How do I use the advanced syntax?\)](#)

Show the field ONLY if...

[had_surgery] = '1'

Clear logic

Test logic with a record: -- select record -- ▾

— OR —

☒ **Drag-N-Drop Logic Builder**

Field choices from other fields
(drag a choice below to box on right)

date_of_admit = (define criteria)

date_of_discharge = (define criteria)

los = (define criteria)

had_surgery = Yes (1)

had_surgery = No (0)

my_first_instrument_complete = Incomplete (0)

my_first_instrument_complete = Unverified (1)

➡

Drag and Drop

➡

Show the field ONLY if...

☒ ALL below are true

☐ ANY below are true


had_surgery = Yes (1) ✕

Clear logic

Save


Cancel

REDCap Demonstration Data Collection Sheet

Project status:  Development

Completed steps 0 of 7



Enable the survey and click on I am done once the data collection sheet is final






Not started


I'm done!

Main project settings

 Use surveys in this project? 

 Use longitudinal data collection with defined events? 



 [VIDEO: How to create and manage a survey.](#)






Complete!


Not complete?

Main project settings

 Use surveys in this project? 

 Use longitudinal data collection with defined events? 

 [VIDEO: How to create and manage a survey.](#)



Not started

I'm done!

Design your data collection instruments & enable your surveys

Add or edit fields on your data collection instruments (survey and forms). This may be done by either using the Online Designer (online method) or by uploading a Data Dictionary (offline method). You may then enable your instruments to be used as surveys in the Online Designer. Quick links: [Download PDF of all instruments](#) OR [Download the current Data Dictionary](#).

Go to or

You may also browse for pre-built data collection instruments in the

Have you checked the [Check For Identifiers](#) page to ensure all identifier fields have been tagged?

REDCap Demonstration Data Collection Sheet

Data Collection Instruments

Survey options:

Survey Queue

Survey Login

Survey Notifications

Add new instrument:

Create

a new instrument from scratch

Import

a new instrument from the official [REDCap Shared Library](#)

Upload

instrument ZIP file from another project/user or [external libraries](#)

Instrument name	Fields	View PDF	Enabled a survey	Instrument actions	Survey-related options
My First Instrument	8			<div>Choose action</div>	<div><div>Survey settings</div><div>+ Automated Invitations</div></div>

Enable the survey under
Online Designer

REDCap Demonstration Move Project to Production

Move Project To Production Status?

Are you sure you wish to leave the DEVELOPMENT stage? If you proceed, the project will be moved to PRODUCTION status so that real data may be collected. If you select the 'Delete ALL data' option below, all current collected data, calendar events, and uploaded documents will be deleted, otherwise all will remain untouched as the project is moved to production.

★ Have you checked the [Check For Identifiers](#) page to ensure all identifier fields have been tagged?

Keep existing data or delete?

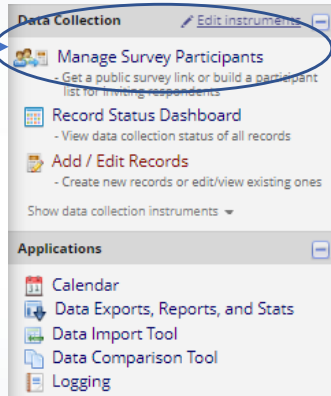
☒ Keep ALL data saved so far.

☐ Delete ALL data, calendar events, documents uploaded for records/responses, survey responses (if applicable), and any logging events pertaining to data collection.

Once in production, you will not be able to edit the project fields in real time anymore. However, you can make edits in Draft Mode, which will be auto-approved or else might need to be approved by a REDCap administrator before taking effect.

YES, Move to Production Status Cancel

REDCap Demonstration Manage Survey Participant



Data Collection [Edit instruments](#)

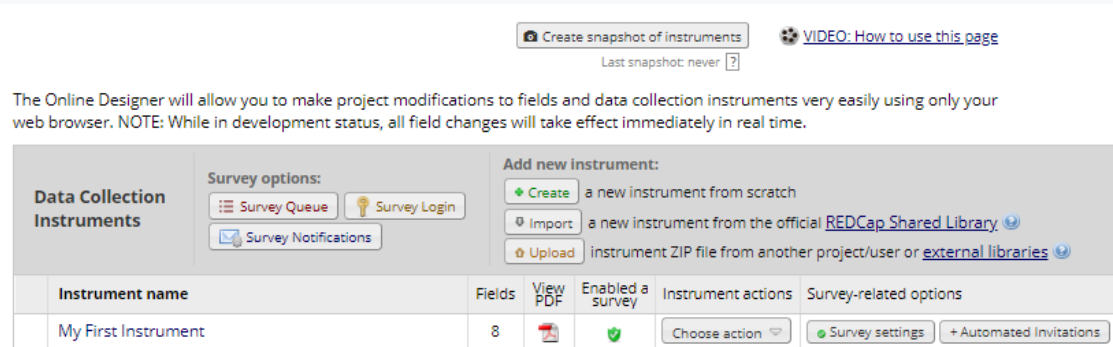
- Manage Survey Participants**
- Get a public survey link or build a participant list for inviting respondents
- Record Status Dashboard**
- View data collection status of all records
- Add / Edit Records**
- Create new records or edit/view existing ones

Show data collection instruments ▾

Applications

- Calendar
- Data Exports, Reports, and Stats
- Data Import Tool
- Data Comparison Tool
- Logging

Once the survey is enabled there will be a survey participant link generated by REDCap



Create snapshot of instruments
Last snapshot: never ?

[VIDEO: How to use this page](#)

The Online Designer will allow you to make project modifications to fields and data collection instruments very easily using only your web browser. NOTE: While in development status, all field changes will take effect immediately in real time.

Data Collection Instruments

Survey options:

- [Survey Queue](#)
- [Survey Login](#)
- [Survey Notifications](#)


Add new instrument:


- [Create](#) a new instrument from scratch
- [Import](#) a new instrument from the official [REDCap Shared Library](#)
- [Upload](#) instrument ZIP file from another project/user or [external libraries](#)


Instrument name	Fields	View PDF	Enabled a survey	Instrument actions	Survey-related options
My First Instrument	8			Choose action	Survey settings + Automated Invitations

REDCap Demonstration Manage Survey Participant

Manage Survey Participants


 Public Survey Link

 Participant List


 Survey Invitation Log


Using a public survey link is the simplest and fastest way to collect responses for your survey. You may obtain the survey link below to email it to your participants. Responses will be collected anonymously (unless the survey contains questions asking for identifying data from the participant). **NOTE:** Since this method uses a single survey link for all participants, it allows for the possibility of participants taking the survey multiple times, which may be necessary in some cases.

To obtain the survey link, copy the URL below and paste it into the body of an email message in your own email client. Your email recipient(s) can then click the link to begin taking your survey.

Public Survey URL: 


Link Actions


 Open public survey


 Send me URL via email

 Survey Access Code or  QR Code

Link Customizations

 Get Short Survey Link

 Create Custom Survey Link


 Get Embed Code


<https://redcap.mch.com/surveys/?s=EEXAWMEEDX>


REDCap Demonstration Codebook


[Project Home](#) [Project Setup](#) [Other Functionality](#) [Project Revision History](#) [Edit project settings](#)


Quick Tasks


 Codebook


 Manage Survey Participants


 Export data


 Create a report

 Check data quality

 User Rights

 Online Designer and Data Dictionary Upload

 Copy this project

 Data Access Groups

☆ Project is not used as a template [Add](#)

The Codebook is a human-readable, read-only version of the project's Data Dictionary and serves as a quick reference for viewing field attributes.

Invite participants to complete your survey by emailing a public survey link or building a participant list for batch notification.

Export your data from REDCap to open or view in Excel or various stats packages.

Build custom reports for quick views of your data, and export reports to Excel/CSV.

Build or execute data quality rules to find discrepancies and errors in your project data.

Grant new users access to this project or modify user privileges for current users.

Create new fields/questions on your data collection instruments or modify existing ones using the Online Designer or by uploading a Data Dictionary. Quick link: [Download the current Data Dictionary](#) OR [Download Data Dictionary with drafted changes](#)

Create an exact duplicate of this project, which copies over all data collection instruments, any surveys that exist, as well as the option to copy all users and reports to the new project.

Create groups of users to limit user access to certain records/responses, in which only users within a given Data Access Group can access records created by users within that group.

REDCap Demonstration Data Stats Feature

Data Collection

Manage Survey Participants

Record Status Dashboard

Add / Edit Records

Show data collection instruments ▼

Applications

Calendar

Data Exports, Reports, and Stats

Data Import Tool

Data Comparison Tool

Logging

Field Comment Log

File Repository

User Rights and DAGs

Record Locking Customization

E-signature and Locking Mgmt

Data Quality

API and API Playground

REDCap Mobile App

Help & Information

Create New Report

My Reports & Exports

Other Export Options

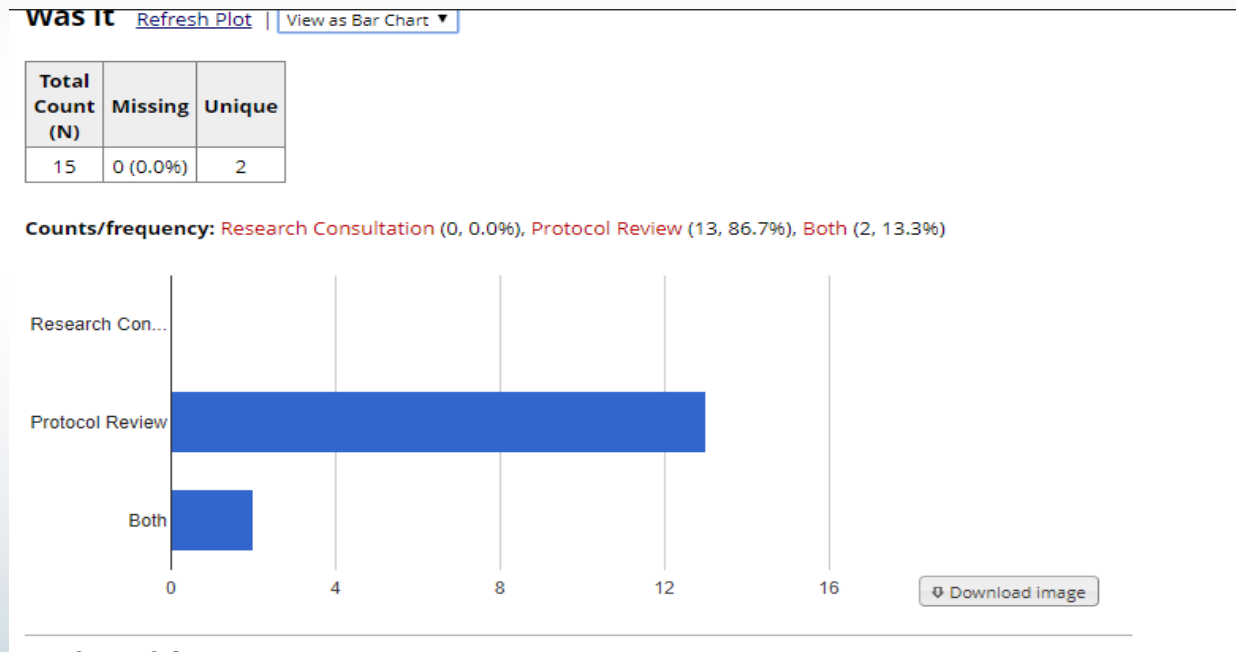
This module allows you to easily view reports of your data, inspect plots and descriptive statistics of your data, as well as export your data to Microsoft Excel, SAS, Stata, R, or SPSS for analysis (if you have such privileges). If you wish to export your *entire* data set or view it as a report, then Report A is the best and quickest way. However, if you want to view or export data from only specific instruments (or events) on the fly, then Report B is the best choice. You may also create your own custom reports below (if you have such privileges) in which you can filter the report to specific fields, records, or events using a vast array of filtering tools to make sure you get the exact data you want. Once you have created a report, you may view it as a webpage, export it out of REDCap in a specified format (Excel, SAS, Stata, SPSS, R), or view the plots and descriptive statistics for that report.

My Reports & Exports

	Report name	View/Export Options	Management Options	Report ID (auto-generated)
A	All data (all records and fields)	<div><div>View Report</div><div>Export Data</div><div>Stats & Charts</div></div>		
B	Selected instruments (all records)	<div>Make custom selections</div>		
	<div>+ Create New Report</div>			

REDCap Demonstration Data Stats Feature

Data Stats will provide you the below statistics for each variable.



REDCap Demonstration Data Quality Feature

Show data collection instruments ▼

Applications

Calendar

Data Exports, Reports, and Stats

Data Import Tool

Data Comparison Tool

Logging

Field Comment Log

File Repository

User Rights and DAGs

Record Locking Customization

E-signature and Locking Mgmt

Data Quality

API and API Playground

REDCap Mobile App

Reports

Edit reports

Processing Complete! Execute rules: All All except A&B Clear

Apply to: -- All records --

Rule #	Rule Name	Rule Logic (Show discrepancy only if...)	Real-time execution ?	Total Discrepancies	Delete rule?
A	Missing values*	-		4,750 view	
B	Missing values* (required fields only)	-		<div>Execute</div>	
C	Field validation errors (incorrect data type)	-		<div>Execute</div>	
D	Field validation errors (out of range)	-		<div>Execute</div>	
E	Outliers for numerical fields (numbers, integers, sliders, calc fields)**	-		<div>Execute</div>	
F	Hidden fields that contain values***	-		<div>Execute</div>	
G	Multiple choice fields with invalid values	-		<div>Execute</div>	
H	Incorrect values for calculated fields	-		<div>Execute</div>	

REDCap References

1. <https://projectredcap.org/resources/videos/>
2. https://www.unmc.edu/vcr/documents/unmc_redcap_usage.pdf
3. <https://www.ctsi.ufl.edu/wordpress/files/2019/02/Project-Creation-User-Guide.pdf>
4. <http://cri.uchicago.edu/wp-content/uploads/2015/12/REDCap-Beginners-Guide.pdf>

REDCap Access Request Form

Once the study is approved by regulatory team, please inform RAD at rad@nicklaushealth.org to get access to REDCap to design the survey



Simultaneously please complete the REDCap access request form survey <http://bit.ly/REDCapAccessRequestForm> which needs to be signed by PI



After RAD team receives signed REDCap Access form along with approval letter from regulatory team, you will get access to design your project in REDCap



Once the survey is ready in REDCap, notify RAD at rad@nicklaushealth.org



RAD team will make sure that the survey is compliant with Data Collection Sheet



RAD team will move the survey to Production and change your access so you can start collecting the data. No further modifications will be possible



Once the data collection is over you can e-mail RAD team at rad@nicklaushealth.org to export the data



DATA ANALYSIS IN EXCEL

Descriptive Statistics

Mean

- The mean (or average) can be used with both discrete and continuous data.
- Mean is particularly susceptible to the influence of outliers so should not be used if data has outliers.

Median

- The median is the middle score for a set of data that has been arranged in order of magnitude.
- The median is less affected by outliers and skewed data.

Standard Deviation

- The standard deviation, like the mean, is appropriate when the continuous data is not significantly skewed or has outliers.

Quartile & Interquartile Range

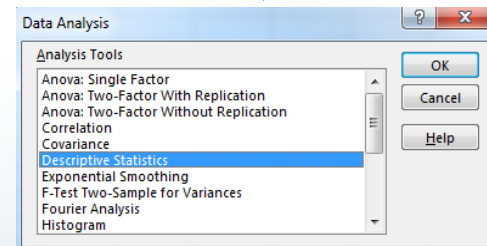
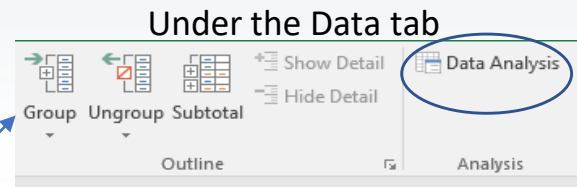
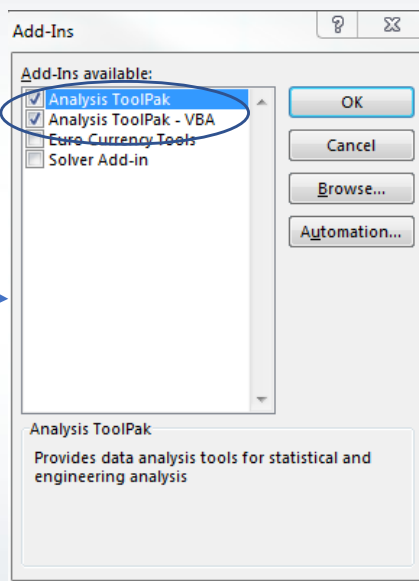
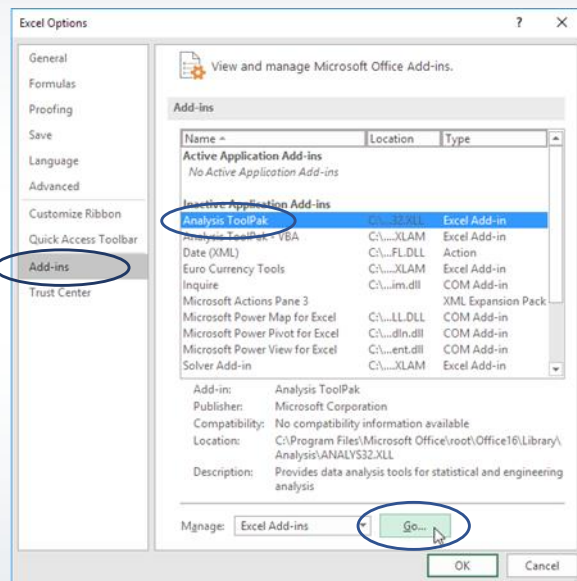
- Quartiles are often reported along with the median when dealing with skewed and/or data with outliers.

Poll Question

1. Which one of these statistics is not affected by outliers?

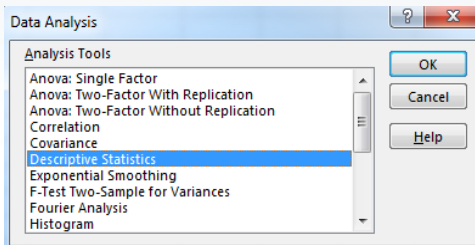
- a. Mean
- b. Interquartile range
- c. Standard deviation
- d. Range

Downloading Data Analysis ToolPak in Excel



Data Analysis – Descriptive statistics using Data Analysis ToolPak

Descriptive statistics (Overall Mean, Median, Standard Deviation, Range, Sum, and Count)



Data selected in Input Range

Name	Sex	Age	Height	Weight
Alfred	M	14	69	112.5
Alice	F	13	56.5	84
Barbara	F	13	65.3	99
Carol	F	14	62.8	102.5
Henry	M	14	63.5	102.5
James	M	12	57.3	83
Jane	F	12	59.8	84.5
Janet	F	15	62.5	112.5
Jeffrey	M	13	62.5	84
John	M	12	59	99.5
Joyce	F	11	51.3	50.5
Judy	F	14	64.3	90
Louise	F	12	56.3	77
Mary	F	15	66.5	112
Philip	M	16	72	150
Robert	M	12	64.8	128
Ronald	M	15	67	133
Thomas	M	11	57.5	85
William	M	15	66.5	112

Descriptive Statistics

Input Range: \$C\$1:\$E\$20

Grouped By: Columns

Labels in first row: ☒

Output options:

Output Range:

New Worksheet Ply: ☒

New Workbook: ☐

Summary statistics: ☒

Confidence Level for Means: 95 %

kth Largest: 1

kth Smallest: 1

	Age		Height		Weight	
Mean	13.31578947	Mean	62.33684211	Mean	100.026316	
Standard Error	0.342442479	Standard Error	1.17623173	Standard Error	5.22469867	
Median	13	Median	62.8	Median	99.5	
Mode	12	Mode	62.5	Mode	112.5	
Standard Deviation	1.492672159	Standard Deviation	5.127075247	Standard Deviation	22.773935	
Sample Variance	2.228070175	Sample Variance	26.28690058	Sample Variance	518.652047	
Kurtosis	-1.11092552	Kurtosis	-0.138969241	Kurtosis	0.68336484	
Skewness	0.063611668	Skewness	-0.259669589	Skewness	0.18335097	
Range	5	Range	20.7	Range	99.5	
Minimum	11	Minimum	51.3	Minimum	50.5	
Maximum	16	Maximum	72	Maximum	150	
Sum	253	Sum	1184.4	Sum	1900.5	
Count	19	Count	19	Count	19	
Largest(1)	16	Largest(1)	72	Largest(1)	150	
Smallest(1)	11	Smallest(1)	51.3	Smallest(1)	50.5	
Confidence Level(95.0%)	0.719444951	Confidence Level(95.0%)	2.471171167	Confidence Level(95.0%)	10.9766846	

Poll Question

2. A list of 5 pulse rates is: 72, 80, 84, 75, 92. What is the median for this list?

- a. 74
- b. 76
- c. 77
- d. 80

Data Analysis – Descriptive statistics using Excel

Descriptive statistics (Quartiles and Interquartile)

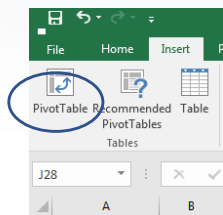
	A	B	C	D	E	F	G
1	Name	Race	Sex	Age	Height	Weight	To Calculate Quartile for Weight
2	Alice	White	F	13	56.5	84	50.5
3	Barbara	American African	F	13	65.3	98	84.25
4	Carol	Asian	F	14	62.8	102.5	=QUARTILE(F2:F20,2)
5	Jane	White	F	12	59.8	84.5	QUARTILE(array, quart)
6	Janet	American African	F	15	62.5	112.5	150
7	Joyce	Asian	F	11	51.3	50.5	
8	Judy	White	F	14	64.3	90	
9	Louise	American African	F	12	56.3	77	
10	Mary	Asian	F	15	66.5	112	
11	Alfred	American African	M	14	69	112.5	
12	Henry	American African	M	14	63.5	102.5	
13	James	White	M	12	57.3	83	
14	Jeffrey	American African	M	13	62.5	84	
15	John	White	M	12	59	99.5	
16	Philip	American African	M	16	72	150	
17	Robert	Asian	M	12	64.8	128	
18	Ronald	White	M	15	67	133	
19	Thomas	American African	M	11	57.5	85	
20	William	American African	M	15	66.5	112	

IF QUART EQUALS	QUARTILE RETURNS
0	Minimum value
1	First quartile (25th percentile)
2	Median value (50th percentile)
3	Third quartile (75th percentile)
4	Maximum value

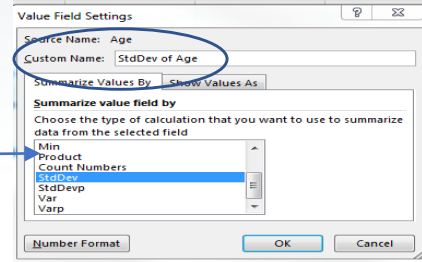
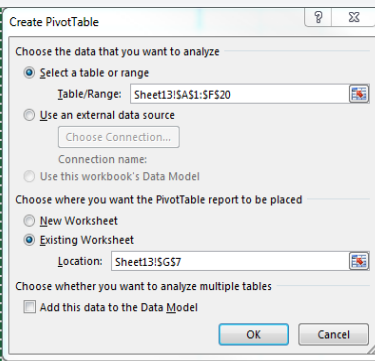
Interquartile range
Third quartile – First Quartile
= 112.25-84.25
=28

Data Analysis – Basic statistics using PivotTable

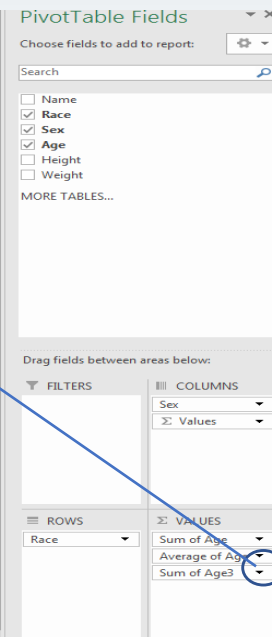
Alternative way: Descriptive statistics (Mean and Standard Deviation)



Name	Race	Sex	Age	Height	Weight
Alice	White	F	13	56.5	84
Barbara	American African	F	13	65.3	98
Carol	Asian	F	14	62.8	102.5
Jane	White	F	12	59.8	84.5
Janet	American African	F	15	62.5	112.5
Joyce	Asian	F	11	51.3	50.5
Judy	White	F	14	64.3	90
Louise	American African	F	12	56.3	77
Mary	Asian	F	15	66.5	112
Alfred	American African	M	14	69	112.5
Henry	American African	M	14	63.5	102.5
James	White	M	12	57.3	83
Jeffrey	American African	M	13	62.5	84
John	White	M	12	59	99.5
Philip	American African	M	16	72	150
Robert	Asian	M	12	64.8	128
Ronald	White	M	15	67	133
Thomas	American African	M	11	57.5	85
William	American African	M	15	66.5	112



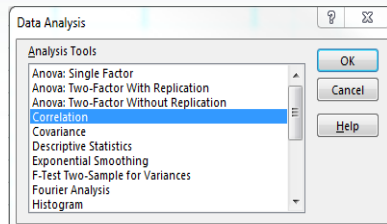
Click on the arrow next to "Sum of Age 3" you will get an option for "Value Field Setting". Select "Value Field Setting" to get the above box and select StdDev to get StdDev for Age. Similar process has been done for "Average of Age".



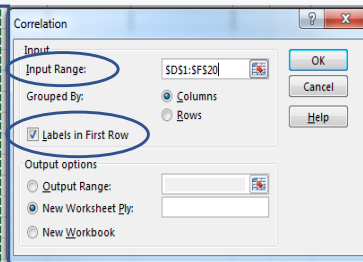
Column Labels										
F		M						Total Sum of Age	Total Average of Age	Total StdDev of Age
Row Labels	Sum of Age	Average of Age	StdDev of Age	Sum of Age	Average of Age	StdDev of Age				
American African	40	13.33333333	1.527525232	83	13.83333333	1.722401424		123	13.66666667	1.58113883
Asian	40	13.33333333	2.081665999	12	12	#DIV/0!	Sample size of 1	52	13	1.825741858
White	39	13	1	39	13	1.732050808		78	13	1.264911064
Grand Total	119	13.22222222	1.394433378	134	13.4	1.646545205		253	13.31578947	1.492672159

Data Analysis – Descriptive statistics using Data Analysis ToolPak

Descriptive statistics (Correlation)



Name	Race	Sex	Age	Height	Weight
Alice	White	F	13	56.5	84
Barbara	American African	F	13	65.3	98
Carol	Asian	F	14	62.8	102.5
Jane	White	F	12	59.8	84.5
Janet	American African	F	15	62.5	112.5
Joyce	Asian	F	11	51.3	50.5
Judy	White	F	14	64.3	90
Louise	American African	F	12	56.3	77
Mary	Asian	F	15	66.5	112
Alfred	American African	M	14	69	112.5
Henry	American African	M	14	63.5	102.5
James	White	M	12	57.3	83
Jeffrey	American African	M	13	62.5	84
John	White	M	12	59	99.5
Philip	American African	M	16	72	150
Robert	Asian	M	12	64.8	128
Ronald	White	M	15	67	133
Thomas	American African	M	11	57.5	85
William	American African	M	15	66.5	112



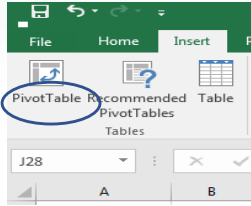
	Age	Height	Weight
Age	1		
Height	0.811434	1	
Weight	0.740885	0.877785	1

Size of Correlation	Interpretation
.90 to 1.00 (-.90 to -.100)	Very high positive (negative) correlation
.70 to .90 (-.70 to -.90)	High positive (negative) correlation
.50 to .70 (-.50 to -.70)	Moderate positive (negative) correlation
.30 to .50 (-.30 to -.50)	Low positive (negative) correlation
.00 to .30 (.00 to -.30)	negligible correlation

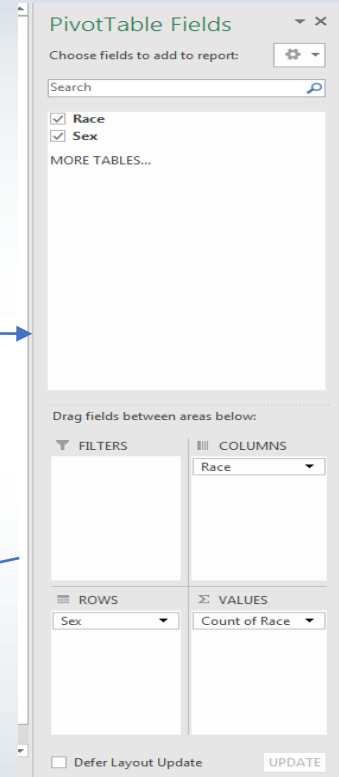
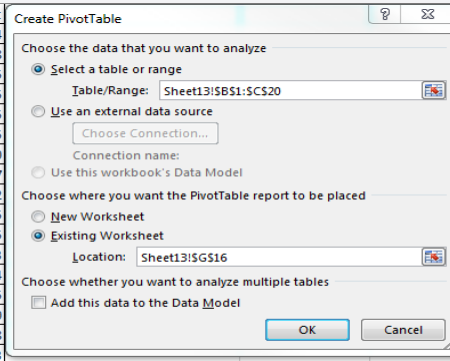
Reference: Diener E, Wirtz D, Tov W, Kim-Prieto C, Choi DW, Oishi S, Biswas-Diener R. New well-being measures: Short scales to assess flourishing and positive and negative feelings. Social Indicators Research. 2010 Jun 1;97(2):143-56.

Data Analysis – Descriptive statistics using Pivot Table

Descriptive statistics (Frequencies)



Name	Race	Sex	Age	Height	Weight
Alice	White	F	13	56.5	84
Barbara	American African	F	13	65.3	98
Carol	Asian	F	14	62.8	102.5
Jane	White	F	12	59.8	84.5
Janet	American African	F	15	62.5	112.5
Joyce	Asian	F	11	51.3	50.5
Judy	White	F	14	64.3	90
Louise	American African	F	12	56.3	77
Mary	Asian	F	15	66.5	112
Alfred	American African	M	14	69	112.5
Henry	American African	M	14	63.5	102.5
James	White	M	12	57.3	83
Jeffrey	American African	M	13	62.5	84
John	White	M	12	59	99.5
Philip	American African	M	16	72	150
Robert	Asian	M	12	64.8	128
Ronald	White	M	15	67	133
Thomas	American African	M	11	57.5	85
William	American African	M	15	66.5	112



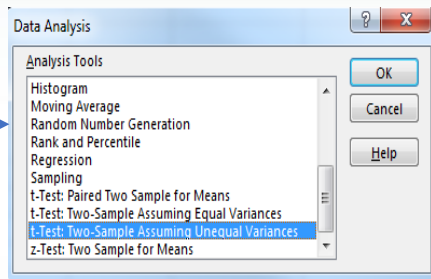
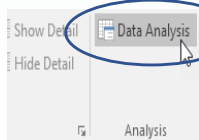
Count of Race		Column Labels			
Row Labels		American African	Asian	White	Grand Total
F		3	3	3	9
M		6	1	3	10
Grand Total		9	4	6	19

Data Analysis – Basic statistics using Data Analysis ToolPak

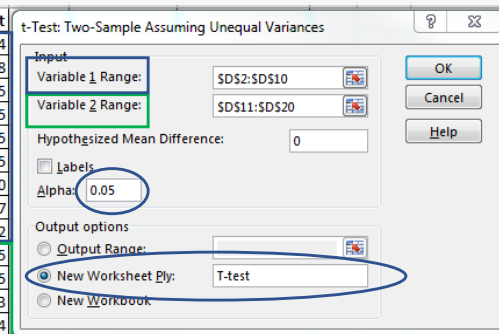
Testing a difference in means between groups

(Null Hypothesis [H_0]: There no difference in the Weight between Males and Females)

Under the Data tab



Name	Race	Sex	Age	Height	Weight
Alice	White	F	13	56.5	84
Barbara	American African	F	13	65.3	98
Carol	Asian	F	14	62.8	102.5
Jane	White	F	12	59.8	84.5
Janet	American African	F	15	62.5	112.5
Joyce	Asian	F	11	51.3	50.5
Judy	White	F	14	64.3	90
Louise	American African	F	12	56.3	77
Mary	Asian	F	15	66.5	112
Alfred	American African	M	14	69	112.5
Henry	American African	M	14	63.5	102.5
James	White	M	12	57.3	83
Jeffrey	American African	M	13	62.5	84
John	White	M	12	59	99.5
Philip	American African	M	16	72	150
Robert	Asian	M	12	64.8	128
Ronald	White	M	15	67	133
Thomas	American African	M	11	57.5	85
William	American African	M	15	66.5	112



t-Test: Two-Sample Assuming Unequal Variances		
	Female	Males
Mean	90.111111	108.95
Variance	375.73611	516.525
Observations	9	10
Hypothesized Mean Difference	0	
df	17	
t Stat	-1.949304	
P(T<=t) one-tail	0.0339772	
t Critical one-tail	1.7396067	
P(T<=t) two-tail	0.0679545	
t Critical two-tail	2.1098156	

We conclude that there is no significant difference in Weight between males and females. ($p=0.0679545 > \text{Alpha } 0.05$)

Note: Reject H_0 when P is <0.05

Poll Question

3. A result is called “statistically significant” whenever:

- a. The null hypothesis is true.
- b. The alternative hypothesis is true.
- c. The p-value is less or equal to the significance level.
- d. The p-value is larger than the significance level.

Data Analysis – Basic statistics using Excel

Calculating Odds Ratio (OR) with confidence interval

Exposure	Lung Cancer	Without Lung Cancer	Grand Total
Smokers	a 17	b 5	22
Non-Smokers	c 9	d 69	78
Grand Total	26	74	100

$$OR = (a/c)/(b/d) \text{ or } (a*d)/(b*c) = 17*69/5*9 = 1,173/45 = 26.07$$

$$\text{Upper} = e^{\ln(OR) + 1.96*\sqrt{1/a + 1/b + 1/c + 1/d}} = \text{EXP}(\text{LN}(26.07) + (1.96*(\text{SQRT}((1/17)+(1/5)+(1/9)+(1/69)))))) = 87.87$$

$$\text{Lower} = e^{\ln(OR) - 1.96*\sqrt{1/a + 1/b + 1/c + 1/d}} = \text{EXP}(\text{LN}(26.07) - (1.96*(\text{SQRT}((1/17)+(1/5)+(1/9)+(1/69)))))) = 7.73$$

You can directly copy paste the formula in excel and change numbers to get Upper and Lower Confidence Interval.

Interpretation: Patients who smoked were 26.07 times more likely to have lung cancer than patients who did not smoke.

Note: If confidence interval includes 1 then the result is not statistically significant

Poll Question

4. A case-control study of 1700 participants looked at the association between Tamoxifen and uterine cancer. The study included 689 cases. There were 139 cases and 58 controls taking Tamoxifen.

Calculate the odds ratio of the above study

	Uterine Cancer	
	Yes	No
Tamoxifen		
Yes	139	58
No	550	953

- a. 3.20
- b. 4.15
- c. 4.20
- d. 4.00

Poll Question

5. Odds Ratio of Tamoxifen and risk of developing uterine cancer is 4.15. Choose the correct interpretation:

- a. Tamoxifen has higher risk of developing uterine cancer
- b. Tamoxifen has lower risk of developing uterine cancer

Effect Size

- Measures the strength of the relationship between two variables on a numeric scale.
- Statistic effect size helps us in determining if the difference is real.
- In hypothesis testing, effect size, power, sample size, and critical significance level are related to each other.

Effect size calculation

Cohen'd:

$$d = M_1 - M_2 / s_{\text{pooled}}$$

M1 = mean of group 1

M2 = mean of group 2

S_{pooled} = pooled standard deviations for the two groups

Effect size of 0.2 is a small effect, 0.5 is a medium effect, and 0.8 is a large effect.

Pooled standard deviation equation:

$$\sqrt{[(s_1^2 + s_2^2) / 2]}$$

Where S_1 and S_2 are the standard deviation of group 1 and 2, respectively.

Effect size calculation

Phi (ϕ): 2 x 2 Contingency table

$$\text{Phi} = \sqrt{X^2 / n}$$

X^2 is the Chi-Square test statistic
 n = total number of observations

How to Interpret

A value of $\phi = 0.1$ is considered to be a small effect, 0.3 a medium effect, and 0.5 a large effect.

Other Statistical Assistance

- If your study requires more complicated statistical approach please contact the RAD team for additional support.
- If the above is the case, please adhere to the research project deadlines so that we can best provide assistance.

2 years research project deadlines

October
12th, 2020

Present research concept:
https://is.gd/research_consultation

December
11th, 2020

Submit draft protocol in protocol builder for Regulatory and RAD review and let us know about it by emailing to: rad@nicklaushealth.org and RegulatoryAffairs@nicklaushealth.org

January
11th, 2021

Submit application to conduct research with final Research Protocol and all supporting documents https://rebrand.ly/research_eApp

July 12th,
2021

Complete all data collection in REDCap and contact biostatistician to start data analysis (analysis can take up to 6 weeks)

October 1st,
2021

Complete manuscript draft and send to rad@nicklaushealth.org and all co-authors for review.

October
2021 - May
2022

Finalize and submit/re-submit manuscript for publication

