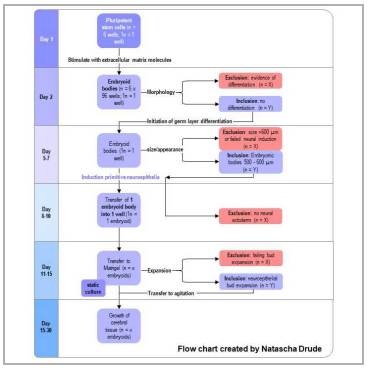


FEB 27, 2023

OPEN ACCESS

Protocol Citation: Tracey L Weissgerber, René Bernard, Friederike E Kohrs 2023. How to design a study flow chart. **protocols.io** https://protocols.io/view/how-to-design-a-study-flow-chart-cfzhtp36

Protocol status: Working
We use this protocol and it's working

Created: Aug 31, 2022

Last Modified: Feb 27, 2023

PROTOCOL integer ID: 69385

Keywords: flow chart, attrition, inclusion criteria, exclusion criteria, study design

How to design a study flow chart

Tracey L Weissgerber[1], René Bernard[2], Friederike E Kohrs[1]

[1]Berlin Institute of Health at Charité - Universitätsmedizin Berlin; [2]Charité - Universitätsmedizin Berlin

Tracey L Weissgerber
Berlin Institute of Health at Charité

DISCLAIMER

This protocol may not work for every study design or research area.

ABSTRACT

Flow charts can be used in human, animal and in vitro studies, as well as systematic reviews, to illustrate attrition or loss of observation. Flow charts provide a visual overview of the study design, and detailed information about the number of included and excluded observations at each stage of the experiment or study, as well as reasons for exclusion. This information is essential to evaluate the risk of bias, and also helps other investigators to estimate the resources that would be needed to perform similar studies. Despite these advantages, few studies have flow charts. Some templates for designing flow charts are available (e.g. CONSORT template for clinical trials, PRISMA template for systematic reviews), however these templates are designed for specific study designs.

Output: This protocol allows readers to efficiently design a flow chart for their study. The protocol is designed to work for many different study designs.

Expertise: Scientists using this protocol will need software for making flow charts, detailed knowledge of their study design and detailed study documentation (see last sentence of materials, below).

Materials: Scientists will need a software program that they are comfortable with, which allows them to design flow charts (e.g. Power Point, LucidChart). A calculator is helpful to confirm that the numbers in the flow chart add up. If the study is in progress or has been completed, scientists will need data on the number of included and excluded observations at each complete stage of the study to include in the flow chart. This information can be added later for studies that are in progress. Other specific expertise is not required.

IMAGE ATTRIBUTION

The thumbnail flowchart was created by Natascha Drude

GUIDELINES

This is designed to be a general protocol that will work for many types of study designs. Outlining your flow chart on paper to establish the structure before adding details in the software program can save you time.

Users who are conducting clinical trials may wish to consult the template in the CONSORT guidelines. Users who are conducting systematic reviews or meta-analyses may wish to use the template in the PRIMSA guidelines instead.

MATERIALS

Software for designing flow charts (e.g. power point, LucidChart, etc.)

- **Tip:** Outline your flowchart on paper first to set the structure, then fill in details using the software. This saves time.

Calculator (to confirm that numbers add up)

Detailed knowledge of study design

If the study is in progress or completed, users will need data for **the number of included and excluded observations at each completed stage of the experiment, and the reasons for exclusion**.

SAFETY WARNINGS

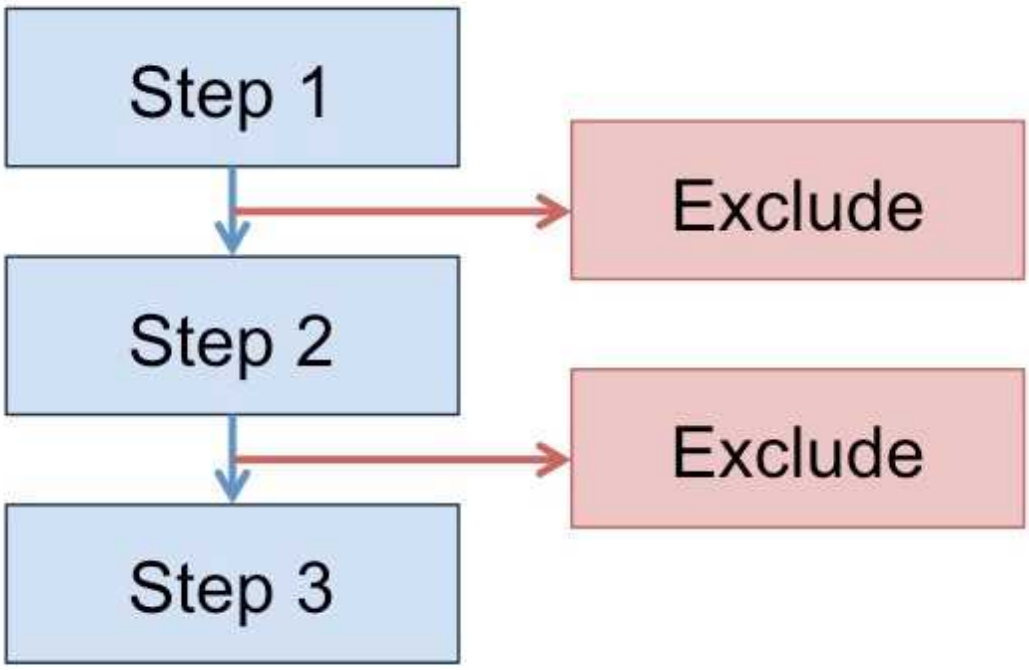
None

Choose the scenario that best fits your study design

1 **Which scenario describes your study?**

Scenario 1: Many steps (screening/selection, preparation, experimental) and one group. Common examples of this scenario include systematic reviews or meta-analyses and epidemiological cohort studies.

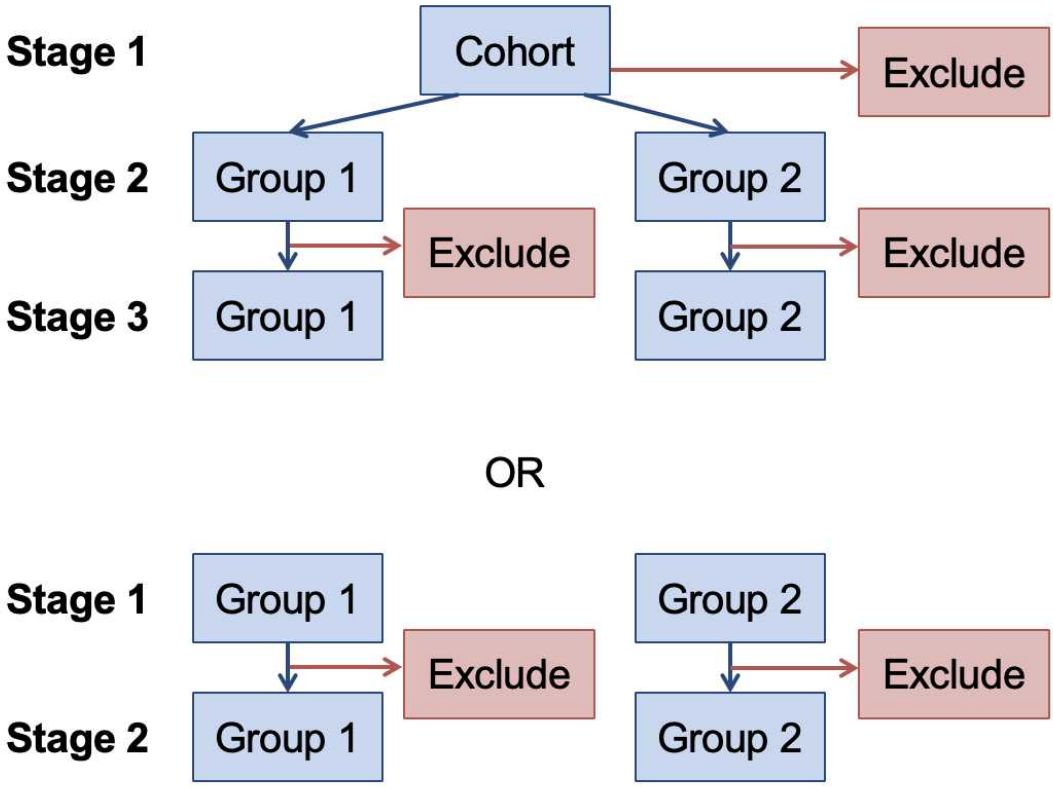
Go to Step 2 and follow the instructions for Scenario 1 (steps 2 - 11)



OR

Scenario 2: Many groups. Many study types use this scenario. Common examples include randomized controlled trials and animal studies.

Skip to Step 12 and follow the instructions for Scenario 2 (steps 12 - 20)



Scenario 1: Design a flow chart for a study with many steps..

2 **Define the stages of your study:** On the left side of the page, label the stages in your study (e.g. screening, randomization, etc.)

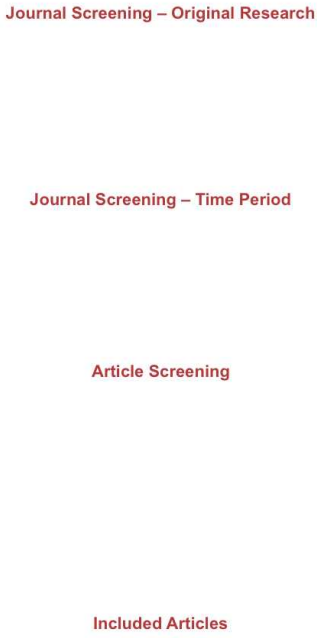
Tips

- The stage name should be specific – clearly specify what happened. You may want to add a short description below the stage name.
- Remember to include the screening phase, where participants, samples, or animals are evaluated for eligibility
- If you are doing in vitro work with samples derived from human participants or animals, your flow chart should start with the donors
- **If the timing or duration of the "Stage" is important, state this below the stage label.**

Example 1 – clinical study

Flow chart was modified from
DOI: 10.1053/j.ajkd.2016.07.034

Example 2: Literature survey



3 **Draw a box around each phase.** Use the stage name as the box title. Below each stage name, briefly specify the inclusion criteria.

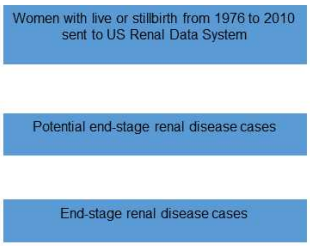
Tips

- One box per stage
- Boxes should appear in one column, so that readers can easily track participant, sample or animal flow by scanning down the page.

- Leave extra space to the right of each group for exclusions boxes – you'll add these in Step 3.

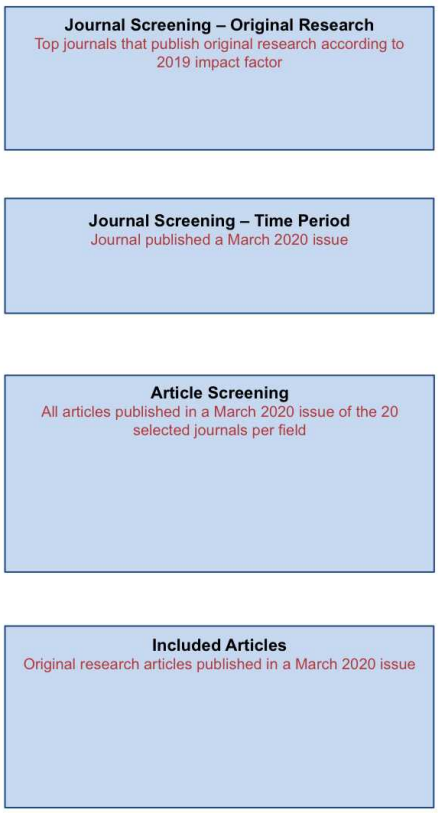
Advanced skills: Boxes in flow charts typically represent groups. If you have a box that represents a step (e.g. randomization) that all groups complete, try using a diamond shaped box to denote the step, instead of the rectangular boxes used to denote groups.

Example 1 – clinical study



Flow chart was modified from
DOI: 10.1053/j.ajkd.2016.07.034

Example 2: Literature survey



4 State the number of included observations in each box (n = x)

Tips

- **Clearly describe what x represents. Be specific.** Replace vague terms like “animals” or “participants” with more specific terms like “Sprague Dawley rats”, “brain slices from Bl6 mice”, or “healthy men and women”.
- **Don’t use the term “patients”.** This term can refer to humans in clinical studies, or animals in veterinary studies.
- **Don’t forget to include a sample size in the first box.**
- **Specify sex or gender:** After showing the total sample size, consider presenting a breakdown by sex or gender. **Example:** n = 12 Sprague Dawley rats (6 M, 6 F).
- **If your experimental unit differs from the number of participants, samples, or animals assessed, specify this.** Instead of “(n = x)”, write (N = x experimental units, n = x individuals). Replace the words “experimental unit” and “individual” with specific terms for your study.

Example 1: Cages of mice were randomized to a treatment or placebo. Describe your sample size as “(N = 5 cages, n = 10 mice).”

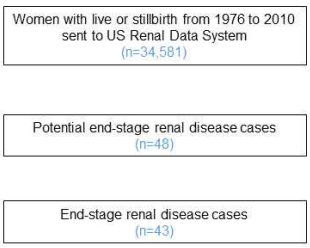
Example 2: Doctor’s offices were randomized to treatment or placebo. All patients within the practice were assigned to the same group. Describe your sample size as “(N = 10 practices, n = 5,623 patients).”

Example 3: 3-5 brain slices were selected from each mouse. Describe your sample size as “(N = 10 mice, n = 43 brain slices).”

- **If your observational or experimental unit changes at different phases of your study, clearly note this.**

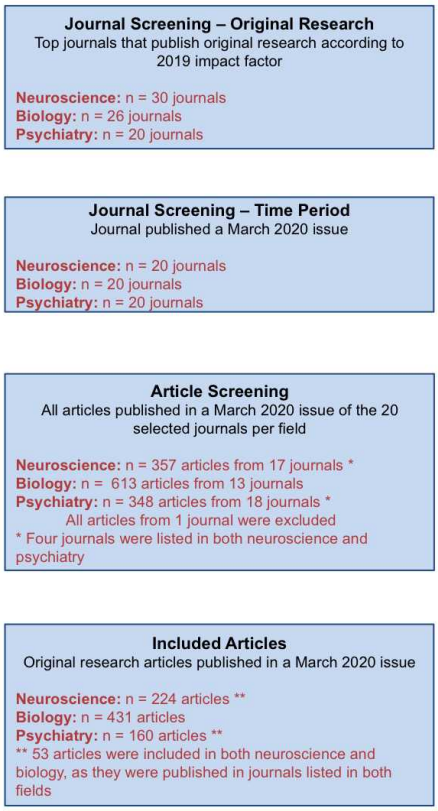
Example: You start by harvesting cells from human participants. You then conduct studies on cultured cells. At the sample collection phase, your units may be “n = 5 patients”. At later phases, your units may be “n = 5 6-well plates.”

Example 1 – clinical study



Flow chart was modified from DOI: 10.1053/j.ajkd.2016.07.034

Example 2: Literature survey

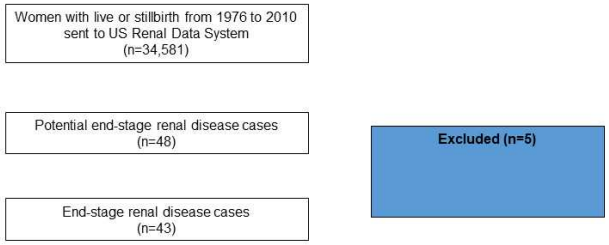


5 Add and label exclusion boxes

Tips

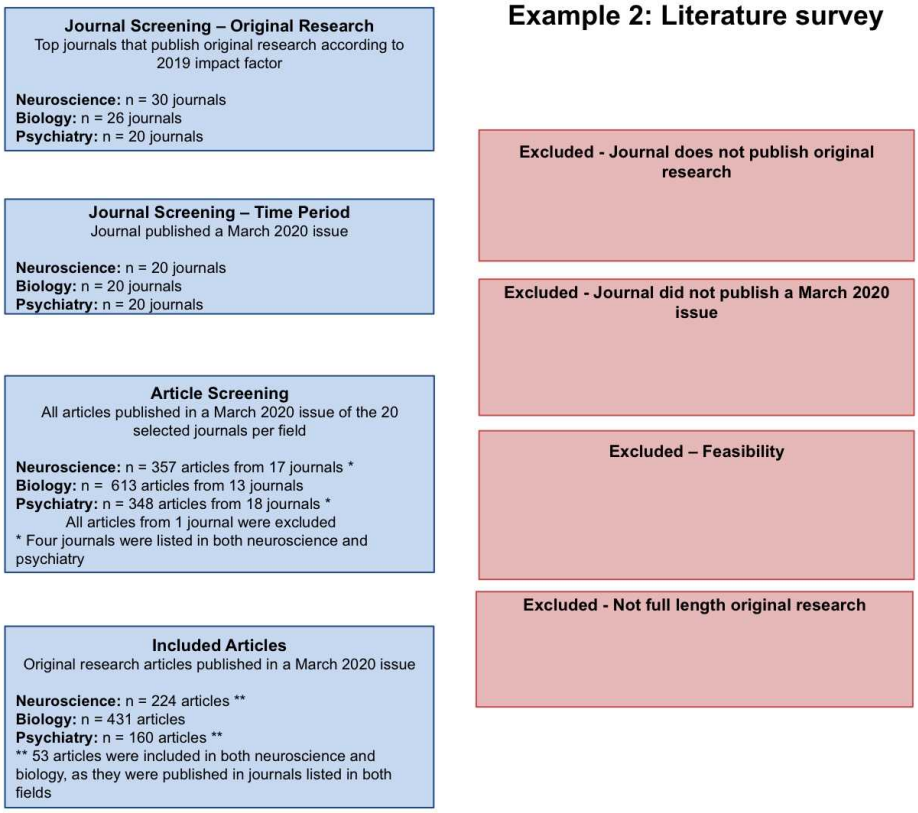
- Exclusion boxes go to the right of each group. Place all exclusions boxes in one column.
- Make exclusion boxes visually distinct from included boxes (e.g. different color or boxes with dashed line borders)
- Exclusions often happen during a phase, rather than at the start, so the exclusions boxes are often placed below the main box for the group at that phase.
- You can used dashed lines or a different color box outline to distinguish “excluded” boxes from “included” boxes.
- **Add a title to each exclusion box:** Write “Excluded (n = x)” at the top of each exclusions box.

Example 1 – clinical study



Flow chart was modified from DOI: 10.1053/j.ajkd.2016.07.034

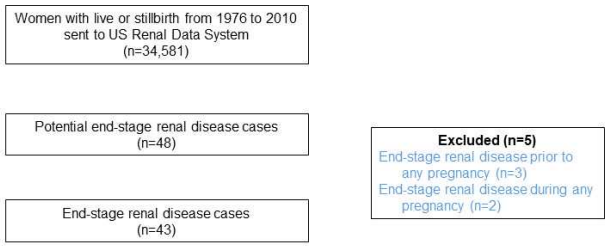
Example 2: Literature survey



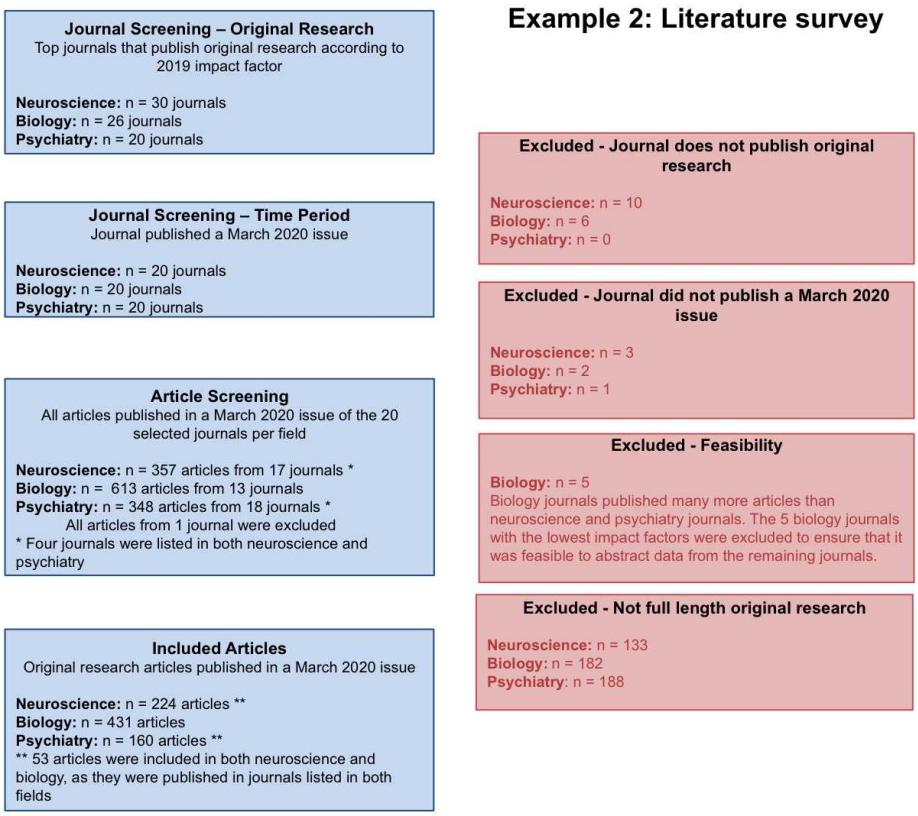
6 **Add numbers and reasons for exclusions:** Below the "Excluded" title, list all reasons for exclusion. Next to each reason, list the number of excluded observations during that phase. Example: "Lost to follow-up: n = 2", "Not enough sample collected: n = 1". This information will help readers to determine which exclusions may be a source of bias, and which are unlikely to contribute to bias.

- **Be specific** about your reasons for exclusion. Replace vague terms like “technical failure” with specific explanations like “FACS machine not working: n = 2 plasma samples”.
- **Avoid grouping participants into broad categories when more specific information is available**. For example, instead of stating “Withdrew: n = 10”, specify “Withdrew due to side effects: n = 3”, “Withdrew due to time commitment: n = 5”, “Withdrew for unknown reasons: n = 2.” **Caution:** This may not be possible for large studies, however you should avoid grouping reasons for exclusion that are a potential source of bias with those that are not.
- **Specify sex or gender:** After showing the total sample size, consider presenting a breakdown by sex or gender. **Example:** n = 12 Sprague Dawley rats (6 M, 6 F).

Example 1 – clinical study



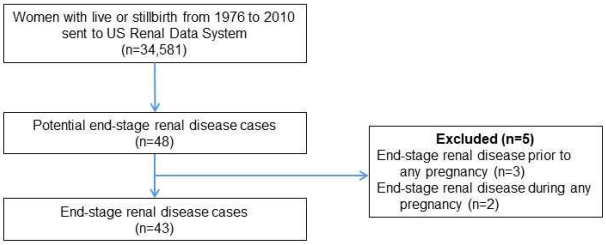
Flow chart was modified from DOI: 10.1053/j.ajkd.2016.07.034



7 Draw arrows

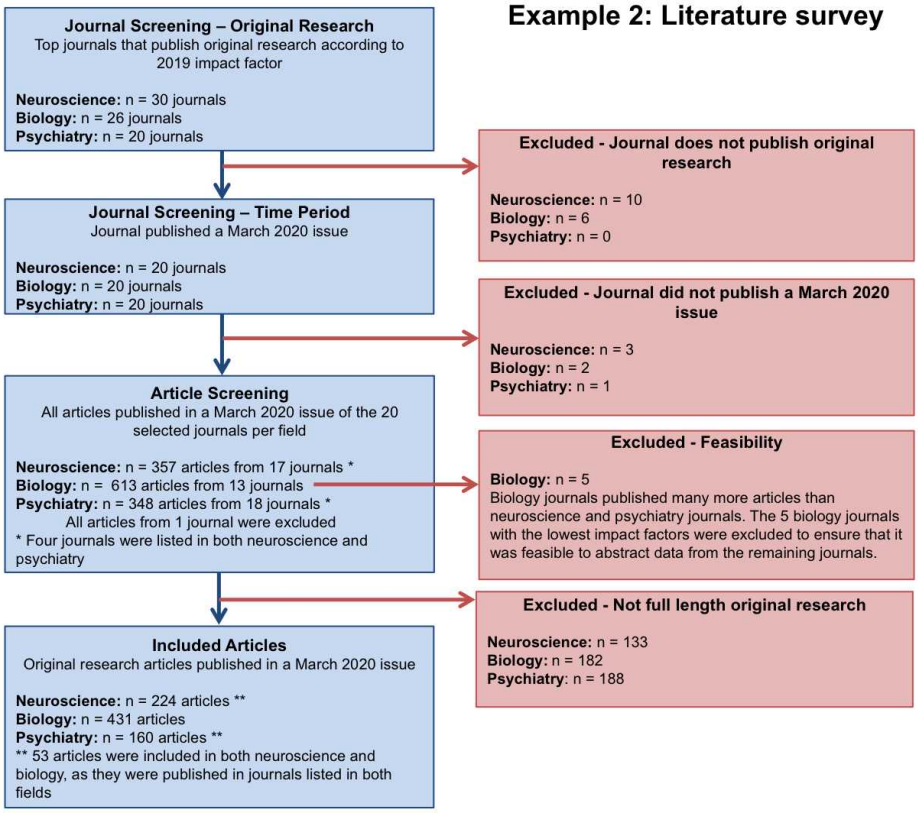
- **Connect boxes showing included observations with arrows pointing down.** This shows readers how participants, samples or animals flow through the stages of the experiment.
- **Draw arrows for each exclusion box.** Exclusion arrows should start on the “Inclusion” line, and point towards the excluded box to show that these observations are leaving the study. These arrows may be dashed or a different color to distinguish them from arrows tracking the flow of included observations (chose a format that matches the exclusion boxes).

Example 1 – clinical study



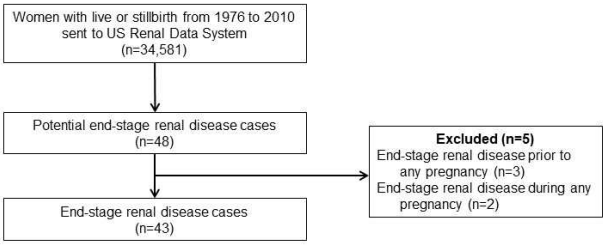
Flow chart was modified from DOI: 10.1053/j.ajkd.2016.07.034

Example 2: Literature survey



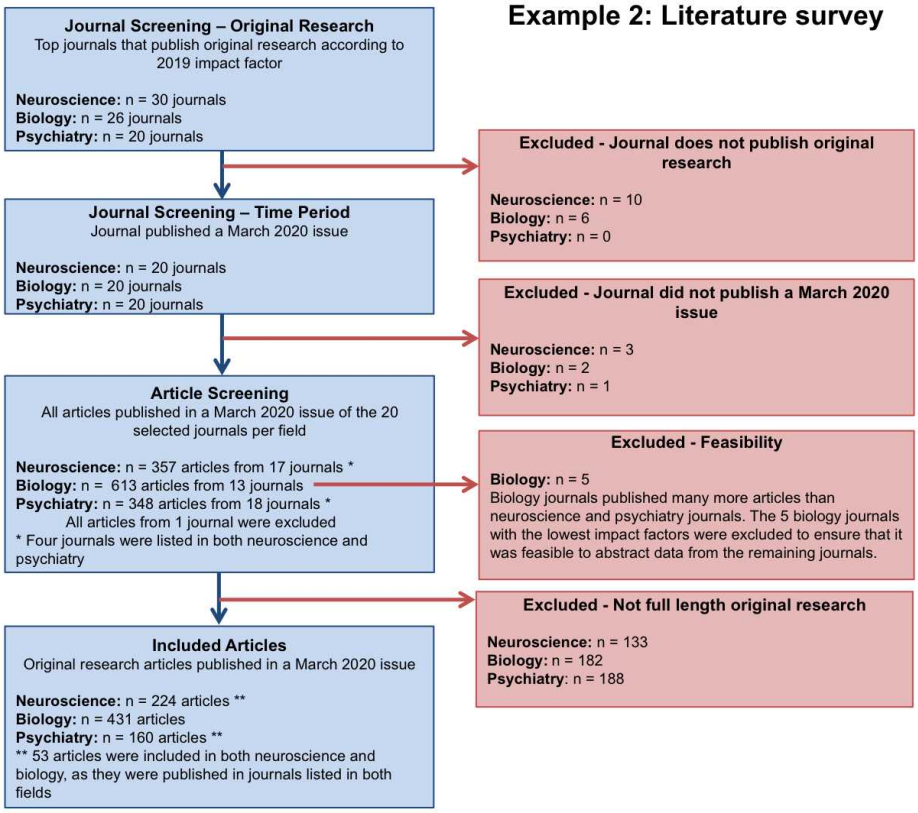
8 **Make sure all the numbers add up:** Check the data in each box to make sure that you've accounted for all observations.

Example 1 – clinical study



Flow chart was modified from DOI: 10.1053/j.ajkd.2016.07.034

Example 2: Literature survey



9 Prepare a legend

9.1 **Write a title** (e.g. Study flow chart). If your paper includes more than one study and flow chart, you may need a more specific title.

9.2 Write a legend

1. **Start with a general explanation**
   Examples:
   i. This flow chart illustrates the selection of papers for this cohort study.
   ii. This flow chart illustrates the selection of samples for this in vitro study.
2. **If needed, provide detailed explanation for things that wouldn't fit on the flow chart**.
3. **Define all abbreviations.** Note: There are tradeoffs to using abbreviations. Abbreviations shorten the text inside boxes, but make it challenging for readers to interpret the flow chart. The same abbreviation can mean very different things in different fields. Eye movements back and forth between the figure and legend are time consuming and frustrate readers, especially for large figures with many abbreviations or figures where the legend is placed on a different page. Consider these tradeoffs when deciding whether to use abbreviations and how many abbreviations to use. Abbreviations should generally be used sparingly for small flow charts, but may be necessary for large, complex flow charts.

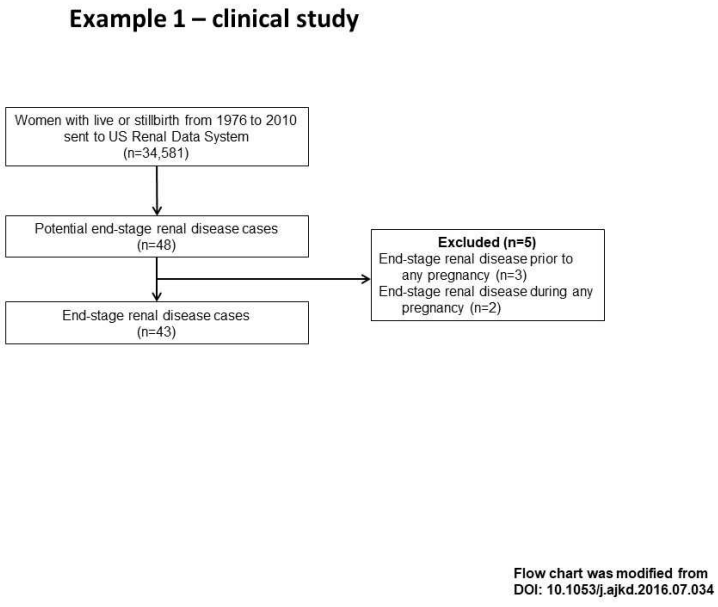


Figure 1: Identification of women with end stage renal disease through records linkage
Medical records from the Rochester Epidemiology Project were linked with records from the United States Renal Data System to identify women who had end stage renal disease. Pregnancy records from the Rochester Epidemiology Project were then reviewed to determine whether women developed end stage renal disease prior to, during or after pregnancy.
Abbreviations: US, United States

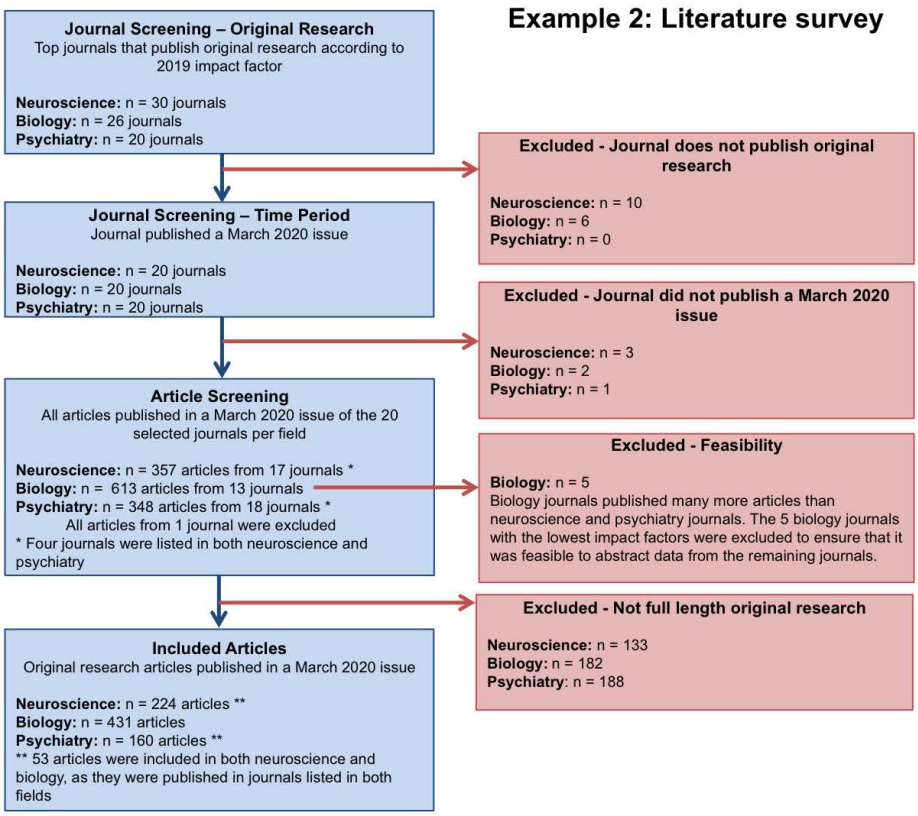


Figure X: Study flow chart
This flow chart illustrates the screening and selection process for journals and paper included in this meta-research study. Three different fields were examined to increase generalizability (neuroscience, biology, and psychiatry). Four journals were listed in both neuroscience and psychiatry, resulting in 53 articles being included in both samples.

10 Review your work

- Can someone who doesn't know about your study easily follow what you did?
- Does anything require additional explanation?
- Is the structure clear and easy to follow?
- Ask others for feedback

11 **Advanced techniques:** These techniques may be required depending on your study design.

11.1 Scenario 1: Add new observations after the first stage of the study

1. **Add boxes to the left of the group to indicate places where observations were added.**
2. **Use rightward pointing arrows** to connect the “Additional observations” box to the downward arrows illustrating the flow of participants through study stages.

3. **Title the box "Additional observations (n = x units)"**
4. **Specify the source of these new observations, or why they were added,** below the title. If different observations were added for different reasons, specify each reason and the number of observations added for that reason.

11.2 Scenario 2: You planned to make some measurements in a subset of participants, samples, or animals

1. In the stages where these measurements are performed, **list each measurement as a bullet point and specify the number of participants, samples or animals in which that measurement was completed**. Place an asterisk next to each measurement that you planned to perform in a subgroup.
2. **In the figure legend:**

- Note that you planned to perform this measurement in a subgroup of participants, samples or animals.
- Explain why you planned to do this
- Explain how you selected participants, samples or animals in which to perform the measurement.

Example: When designing the study, we planned to perform measurement 30 out of 60 randomly selected samples. This measurement is expensive and we did not have sufficient funding to analyze all samples.

Rationale: When assessing the risk of bias, it is important to differentiate between situations were you planned to perform measurements in a subset of participants, animals or samples, and situations where this decision was made after the study began or due to excluded samples.

11.3 Scenario 3: In vitro or other studies where...

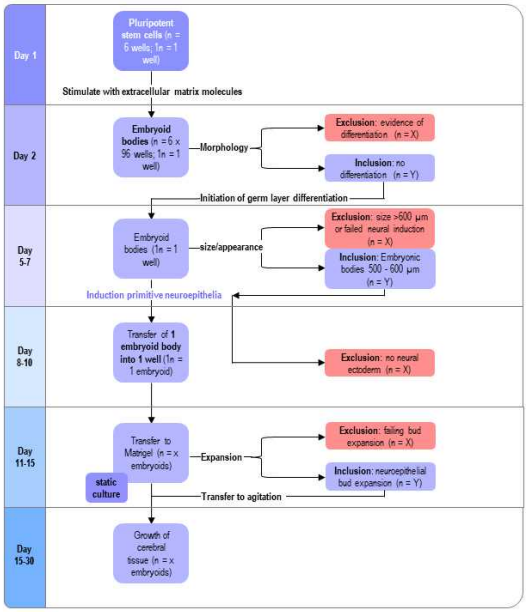
- **The unit of observation changes at almost every stage of the study**
- **Samples are evaluated at the end of each phase to identify those the meet the criteria to progress to the next phase**

The flowchart below shows an example for this type of study.

Tips

- For in vitro studies that use donor material (e.g. samples from animals or mice), the flow chart should start with screening and selection of the donors; then proceed through sample processing
- Show your timeline
- Specify what n represents at different stages. Distinguish between experimental and observational units.
- State essential details (cell passage, etc.)
- If your study includes multiple experiments, you may need separate flow charts for each experiments

Example 2 – *in vitro* study



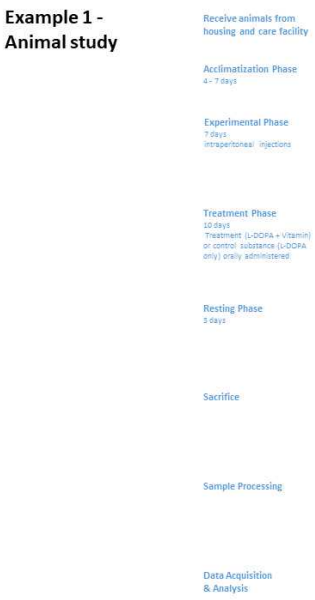
Flow chart created by Natascha Drude

Scenario 2: Design a flow chart for a study with many group..

12 **Define the stages of your study:** On the left side of the page, label the stages in your study (e.g. screening, randomization, drug application etc.)

Tips

- The stage name should be specific – clearly specify what happened.
- Remember to include the screening phase, where participants, samples, or animals are evaluated for eligibility
- If you are doing in vitro work with samples derived from human participants or animals, your flow chart should start with the donors If the timing or duration of the “Stage” is important, state this below the stage label.



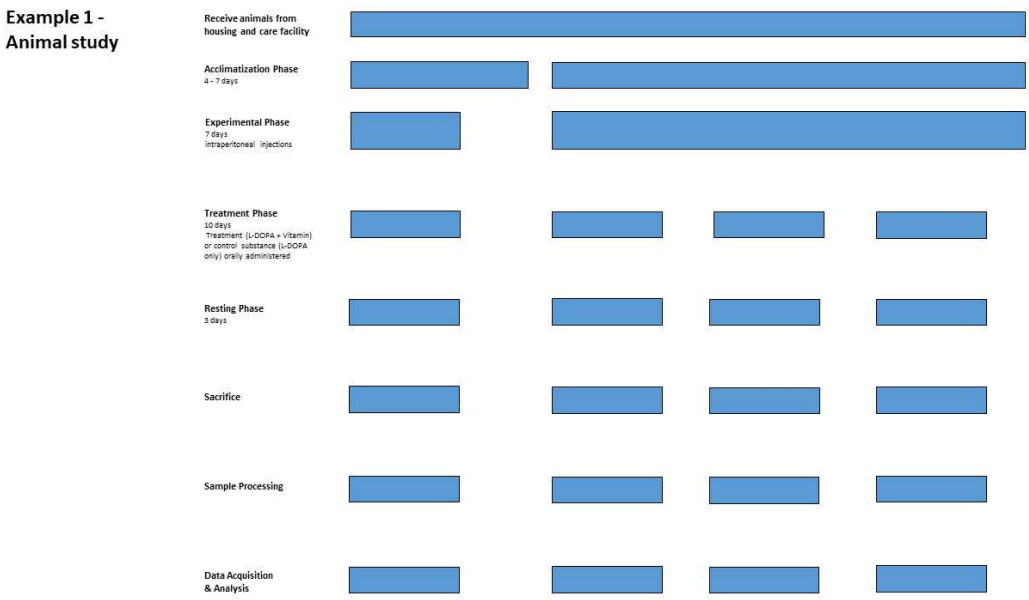
13 Draw boxes to illustrate the group structure at each phase of the experiment

Tips

- One box per group per stage
- Boxes for a single group should appear in one column, so that readers can easily track participant, sample or animal flow in that group by scanning down the page. The first four rows of the example flow chart below illustrate how to align your boxes if a group is divided into multiple groups in later phases of the study.
- Leave extra space to the right of each group for exclusions boxes – you'll add these in Step 3.
- The number of boxes (groups) may change across stages.

Example: If you screen patients, enroll them in the study, and then randomize them to two groups, you may have one box (group) for screening and enrollment, but two boxes (groups) for randomization and subsequent phases.

Advanced skills: Boxes in flow charts typically represent groups. **If you have a box that represents a step (e.g. randomization) that all groups pass through, try using a diamond shaped box to denote the step**, instead of the rectangular boxes used to denote groups.



14 **Label group boxes:** Title each box with the group name, followed by "(n = x)". Replace x with the number of patients, samples or animals included in that group at that phase of the study. If your group name doesn't clearly describe your unit of observation, describe your unit of observation after "(n = x)" (example: n = x Sprague Dawley rats).

Tips

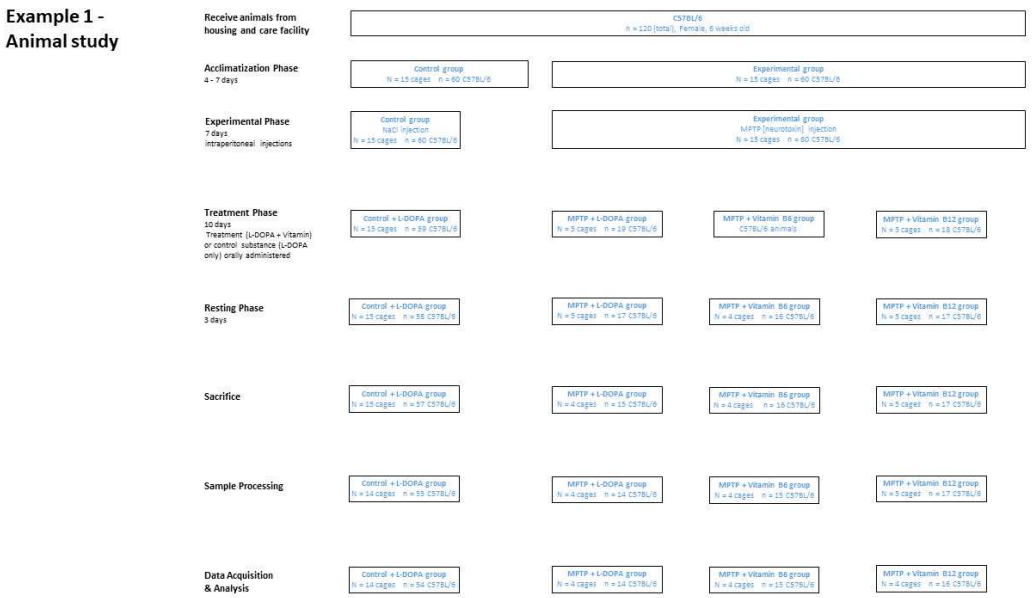
- **Don't forget to label the first box and include the sample size**. If your first stage is about screening and observations haven't been assigned to groups yet, the group name should specify the population studied (e.g. "Bl6 mice (n = 50)", "healthy pregnant women, <12 weeks gestation (n = 79)".
- **Group names should clearly specify what was done to the group, or inclusion criteria for the group**. Avoid vague names like "Group 1", "Control group", "Case group" or "Treatment group."
- **The "units" should clearly specify what was observed**. Replace vague terms like "animals" or "participants" with more specific terms like "Sprague Dawley rats", "brain slices from Bl6 mice", or "healthy men and women".
- **Don't use the term "patients"**. This term can refer to humans in clinical studies, or animals in veterinary studies.
- **Specify sex or gender:** After showing the total sample size, consider presenting a breakdown by sex or gender. **Example:** n = 12 Sprague Dawley rats (6 M, 6 F).
- **Avoid abbreviations when possible.** They make the flow chart hard to read and different abbreviations mean different things in different fields.
- **If your experimental unit differs from the number of participants, samples, or animals assessed, specify this**. Instead of "(n = x)", write (N = x experimental units, n = x individuals). Replace the words "experimental unit" and "individual" with specific terms for your study.

Example: Cages of mice were randomized to a treatment or placebo. Describe your sample size as "(N = 5 cages, n = 10 mice)."

Example: Doctor's offices were randomized to treatment or placebo. All patients within the practice were assigned to the same group. Describe your sample size as "(N = 10 practices, n = 5,623 patients)."

- **If your observational unit changes at different phases of your study, clearly note this in your flow chart**.

Example: You start by harvesting cells from human participants. You then conduct studies on cultured cells. At the sample collection phase, your units may be "n = 5 patients". At later phases, your units may be "n = 5 6-well plates."



15 Add and label exclusion boxes

1. **Place your exclusion boxes**

- Exclusion boxes go to the right of each group.
- Place all exclusions boxes for a group in one column.
- Make your exclusion boxes visually distinct from group boxes (e.g. different color or boxes with dashed line borders)
- Exclusions often happen during a phase, rather than at the start, so the exclusions boxes are often placed below the main box for the group at that phase.

2. **Add a title to each exclusion box:** Write "Excluded (n = x)" at the top of each exclusions box.

3. **Add reasons for exclusions:** Below the title, list all reasons for exclusion. Be specific. Next to each reason, list the number of excluded observations during that phase.

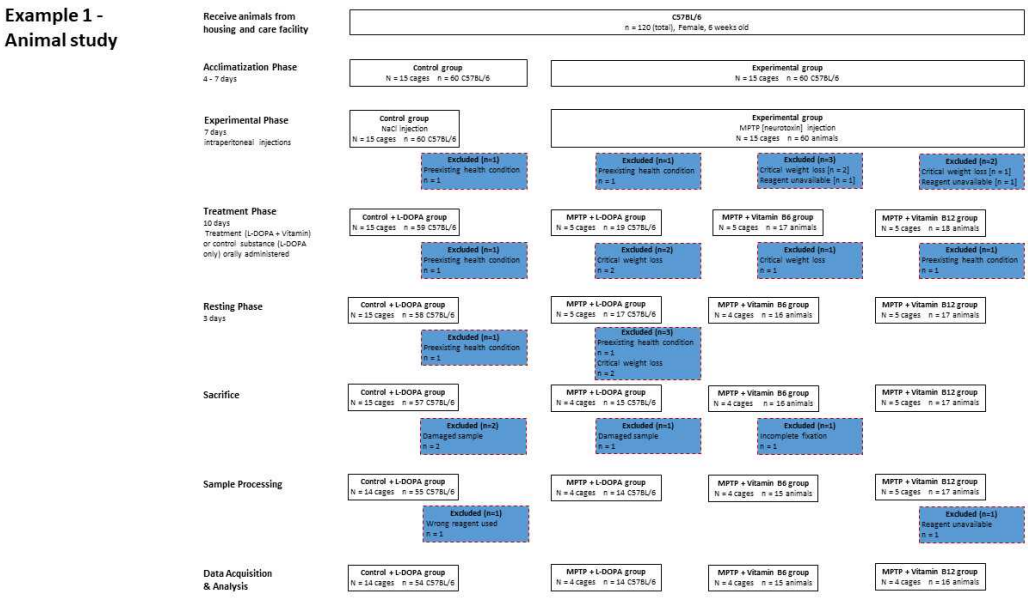
Example: "Lost to follow-up: n = 2", "Not enough sample collected: n = 1".

Tips

- **Be specific** about your reasons for exclusion.

Example: Replace vague terms like “technical failure” with specific explanations like “FACS machine not working: n = x”.

- **Avoid grouping participants into broad categories** when more specific information is available. For example, instead of stating “Withdrew: n = 10”, specify “Withdrew due to side effects: n = 3”, “Withdrew due to time commitment: n = 5”, “Withdrew for unknown reasons: n = 2.”
- **Specify sex or gender:** After showing the total sample size, consider presenting a breakdown by sex or gender. **Example:** n = 12 Sprague Dawley rats (6 M, 6 F).

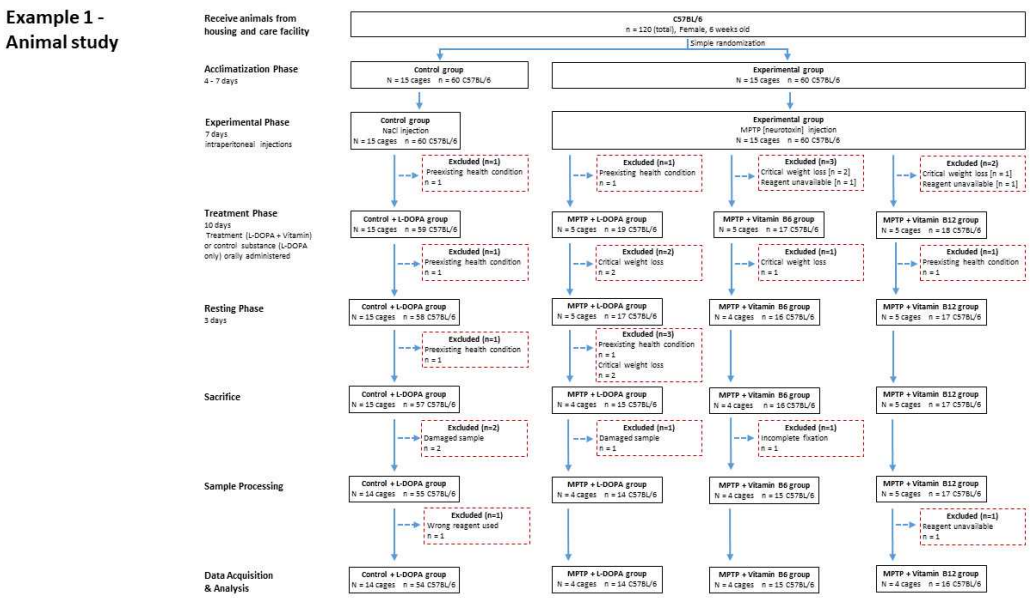


16 Draw arrows

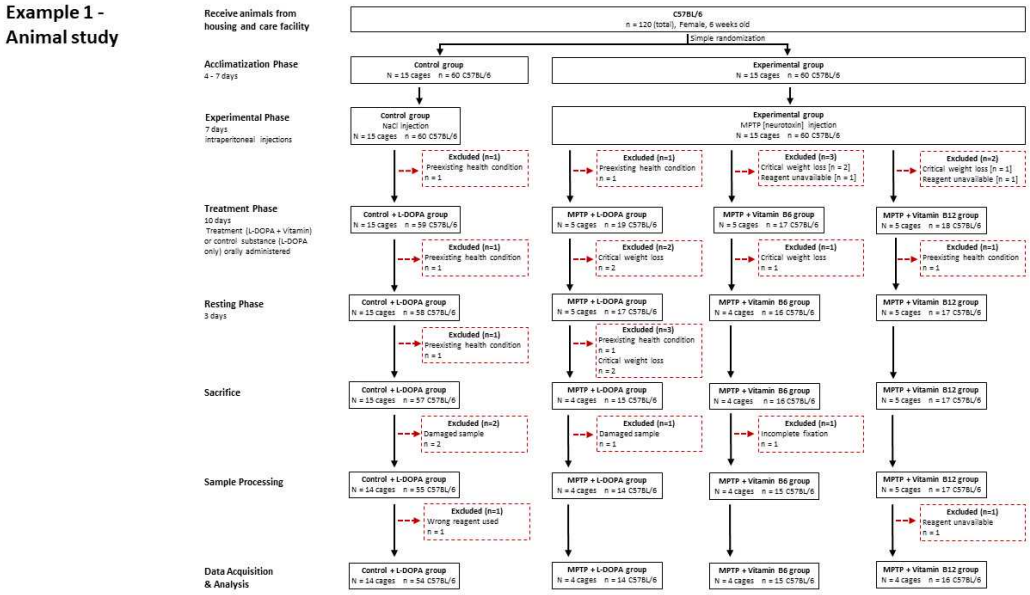
1. **Connect boxes showing included observations with arrows pointing down**. This shows readers how participants, samples or animals flow through the stages of the experiment.
2. **Draw rightward arrows for each exclusion box**. Exclusion arrows should start on the “Inclusion” line, and point towards the excluded box to show that these observations are leaving the study.

Tips

- Arrows should be straight whenever possible
- Avoid crossing arrows if you can
- You can differentiate excluded arrows and boxes from included arrows or boxes by using color or dashed outlines



17 **Make sure all the numbers add up:** Check the data in each box to make sure that you've accounted for all observations.



18 **Write a legend**

18.1 **Write a title** (e.g. Study flow chart). If your paper includes more than one study and/or flow chart, you may need a more specific title.

18.2 Write the legend

1. **Start with a general explanation**

Examples:

i. This flow chart provides an overview of the study design and attrition in this cohort study.

ii. This flow chart provides an overview of the study design and attrition in this animal study.

2. **If needed, provide detailed explanation for things that wouldn't fit on the flow chart**.

3. **Define all abbreviations.** Note: There are tradeoffs to using abbreviations. Abbreviations shorten the text inside boxes, but make it challenging for readers to interpret the flow chart. The same abbreviation can mean very different things in different fields (Google your abbreviations to make sure that they are commonly known). Eye movements back and forth between the figure and legend are time consuming and frustrate readers, especially for large figures with many abbreviations or figures where the legend is placed on a different page. Consider these tradeoffs when deciding whether to use abbreviations and how many abbreviations to use. Abbreviations should generally be used sparingly for small flow charts, but may be necessary for large, complex flow charts.

Example 1 - Animal study

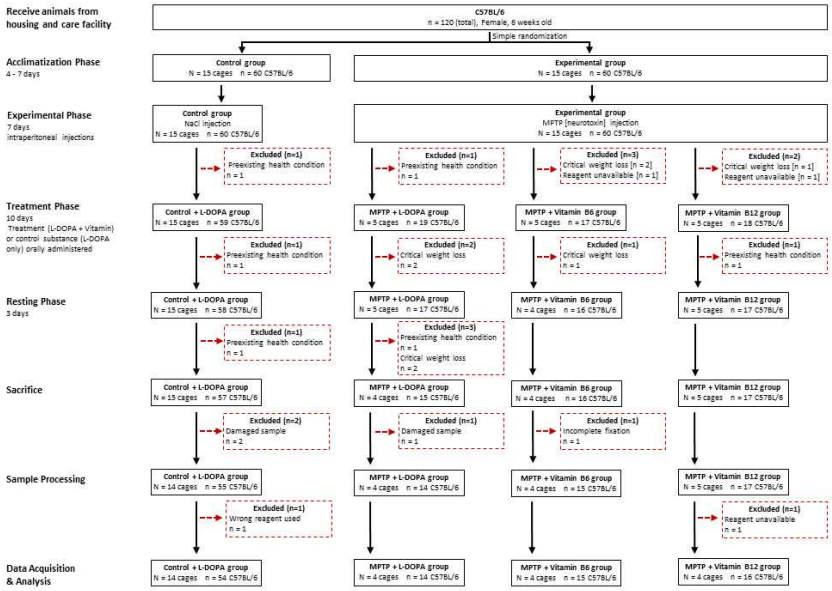


Figure x: Study flow chart
This flow chart illustrates the flow of mice trough this preclinical study. In the experimental phase control animals were injected intraperitoneal with NaCl. Experimental animals were injected with a MPTP, a neurotoxin acting selectively on dopaminergic neurons. In the treatment phase, L-DOPA, a standard symptomatic treatment substance for Parkinson's disease, was administered orally to all animals. Two experimental groups received Vitamin B6 or Vitamin B12 in addition to L-DOPA.

Abbreviations: C57BL/6 = Black 6 inbred mouse strain; L-DOPA = Levodopa; MPTP = 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine NaCl = sodium chloride.

19 Review your work

- Can someone who doesn't know about your study easily follow what you did?
- Does anything require additional explanation?
- Is the structure clear and easy to follow?
- Ask others for feedback

20 Advanced techniques: The techniques in this section may be required depending on your study design.

20.1 Scenario 1: Add new observations after the first stage of the study

Example 1 – clinical study

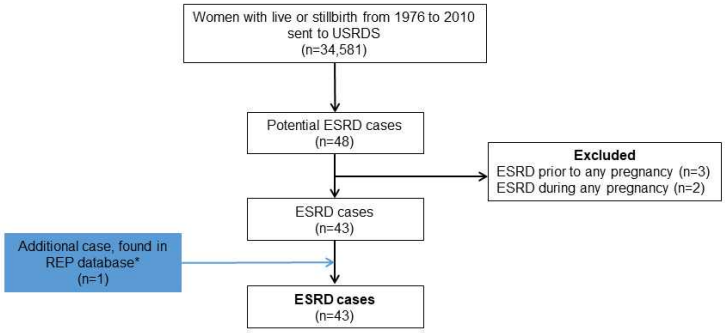


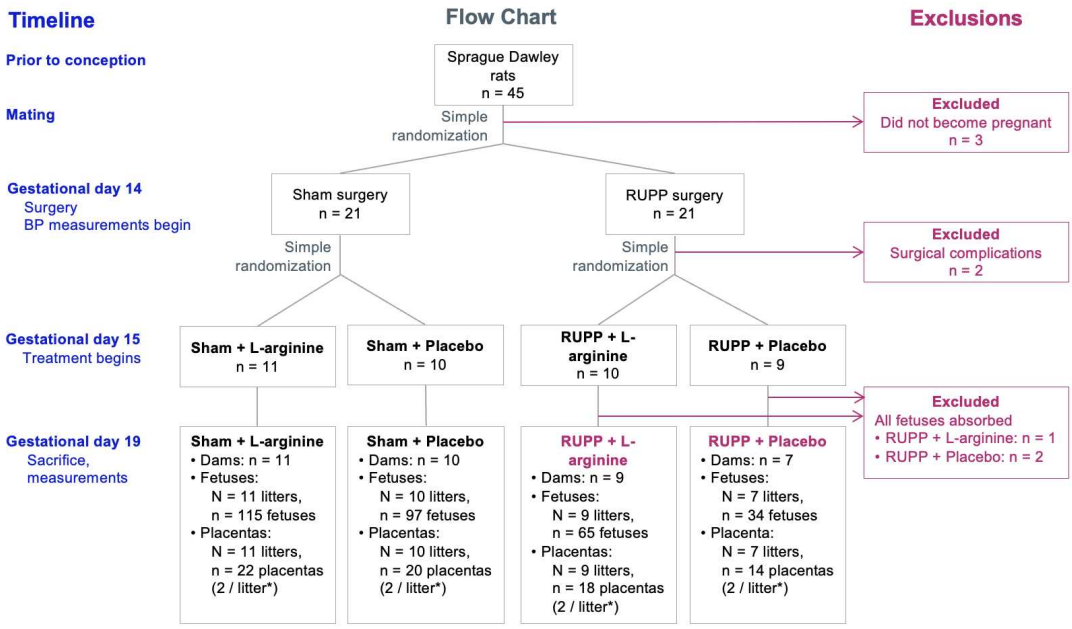
Figure 1. Identification of women with end-stage renal disease (ESRD) via linkage. *One case was not identified by US Renal Data System (USRDS) linkage, but was identified by the Rochester Epidemiology Project (REP). This women received a kidney transplant after the linkage with USRDS was performed.

Flow chart was previously published
DOI: 10.1053/j.ajkd.2016.07.034

1. **Add boxes to the left of the group to indicate places where observations were added.**
2. **Use rightward pointing arrows** to connect the "Additional observations" box to the downward arrows illustrating the flow of participants through study stages.
3. **Title the box "Additional observations (n = x units)"**
4. **Specify the source of these new observations, or why they were added,** below the title. If different observations were added for different reasons, specify each reason and the number of observations added for that reason.

20.2 Scenario 2: You planned to make some measurements in a subset of participants, samples, or animals

1. In the stages where these measurements are performed, **list each measurement as a bullet point and specify the number of participants, samples or animals in which that measurement was completed**. Place an asterisk next to each measurement that you planned to perform in a subgroup.



*Placental analyses were performed in two randomly selected placentas from each litter, in accordance with our pre-planned study protocol, to limit study costs.

2. **In the figure legend:**

- Note that you planned to perform this measurement in a subgroup of participants, samples or animals.
- Explain why you planned to do this
- Explain how you selected participants, samples or animals in which to perform the measurement.

Example: When designing the study, we planned to perform measurement 30 out of 60 randomly selected samples. This measurement is expensive and we did not have sufficient funding to analyze all samples.

Rationale: When assessing the risk of bias, it is important to differentiate between situations were you planned to perform measurements in a subset of participants, animals or samples, and situations where this decision was made after the study began or due to excluded samples.

20.3 Scenario 3: Your study includes many different time points and you remove some animals, tissues or samples from each group at each time point

The removed animals, samples or tissues are excluded from further phases of the study, but they don't fit in the exclusions boxes because they were analyzed.

In the stages where these removals occur, create a separate row of "**Measurements**" or "**Analysis**" boxes above the exclusion boxes. Use a different color or shading to distinguish measurement or analysis boxes from inclusion and exclusion boxes. Follow the same procedures as you would for an exclusion box (e.g. labelling, placement, arrows).

20.4 Scenario 4: You have too many groups to show in a flow chart

Use a table instead, showing groups in columns and study stages in rows. Use sub-columns for each group to enter the number of included and excluded observations and reasons for exclusion. Consider color coding cells for excluded observations. The colors won't be machine readable, but they will allow human readers to quickly visually identify groups and stages with higher exclusion rates.