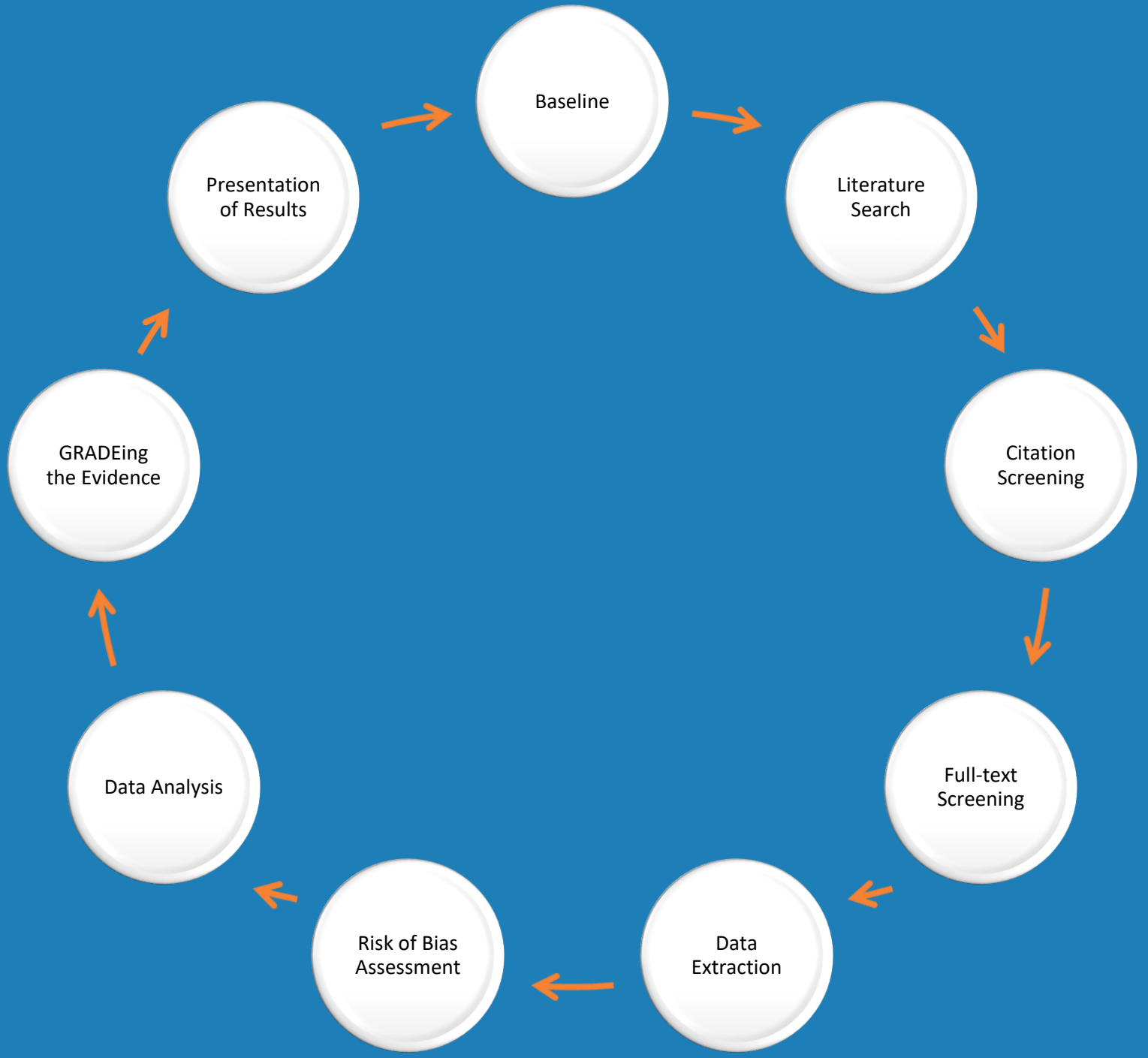
Process issues and data abstraction

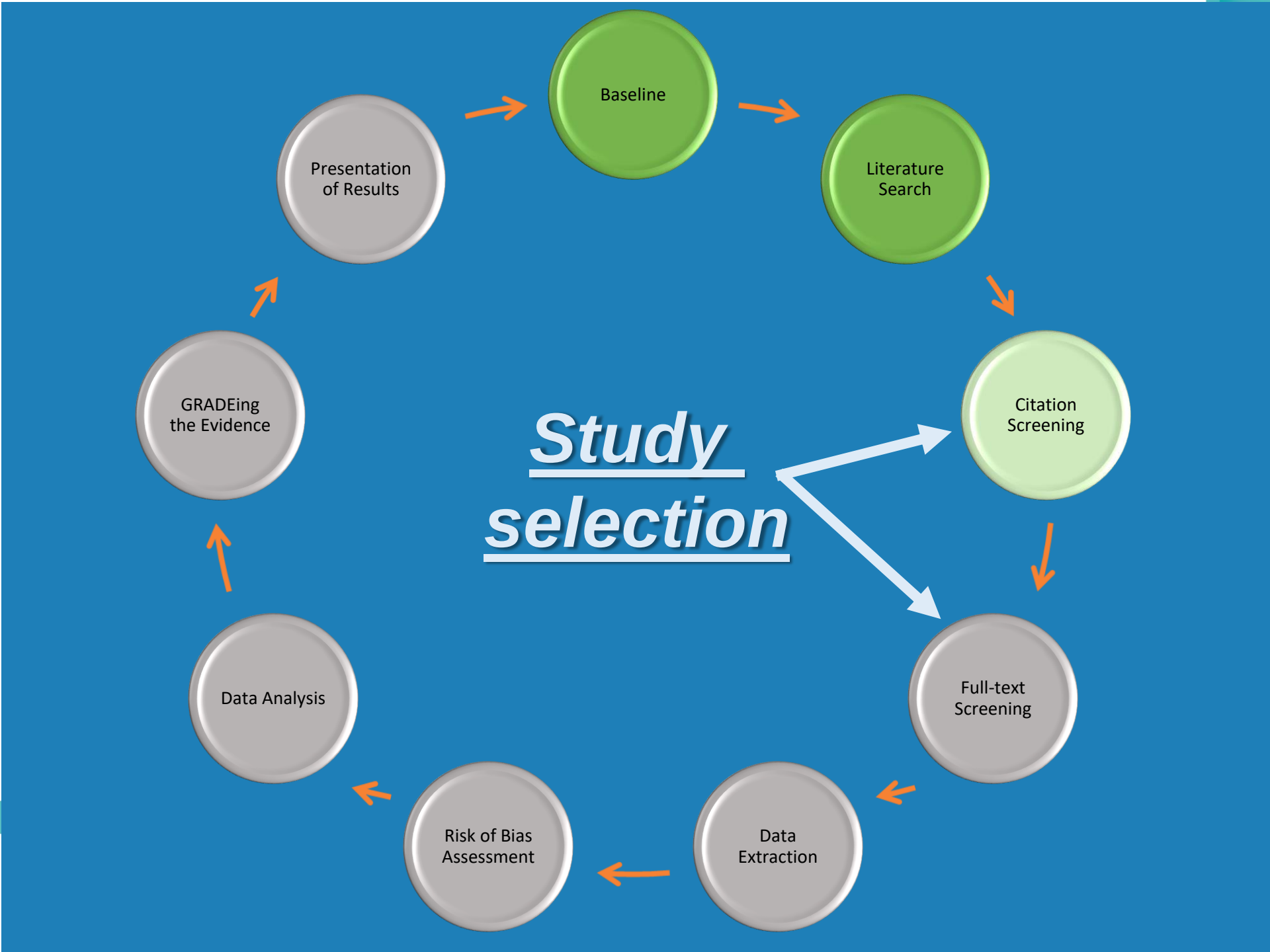
Lecture #5

CHSC 7362 A01
Systematic reviews and meta-analysis


Baseline
Literature Search
Citation Screening
Full-text Screening
Data Extraction
Risk of Bias Assessment
Data Analysis
GRADEing the Evidence
Presentation of Results

Literature search

- Goal of search strategy:
  - **sensitive** enough to capture all relevant citations
  - **specific** enough to be feasible


Baseline
Literature Search
Citation Screening
Full-text Screening
Data Extraction
Risk of Bias Assessment
Data Analysis
GRADEing the Evidence
Presentation of Results
Study selection

Study selection

- Goal of Title/ Abstract screening phase (Level I Screening):
  - Exclude all irrelevant citations (e.g. obviously doesn't meet the PICOS criteria)
  - Include any citations that may be relevant
  - When in doubt or it is unclear then include for full-text screening
  - Note:
    - citations may present little details pertaining to the PICOS (e.g. details of randomization, age range, severity of disease, all outcomes in the study, etc.).
    - not have abstracts or key words (e.g. title only)
    - be published in a non-English language

Study selection

- Goal of Full-text screening phase (Level II Screening):
  - Exclude all citations of studies that don't meet all the PICOS elements (e.g. doesn't report outcomes of interest)
  - Exclusions must be classified (preferably using a hierarchy)
  - Example of reasons for exclusion:
    - **Population(s)** not of interest to this review
    - **Intervention(s)** not of interest to this review
    - **Comparator(s)** not of interest to this review
    - No **outcome(s)** related to this review reported
    - **Setting(s)** not of interest to this review
    - **Study design(s)** not of interest to this review
    - **Publication type(s)** not of interest to this review
    - **Language** of publication not English

Study selection

- **Literature Search Results (Database)**
  - 01 - Medline (Ovid) (0)
  - 02 - Embase (Ovid) (0)
  - 03 - Central (Wiley) (0)
- **Level I Screening**
  - 01 - Duplicates (0)
  - 02 - Titles/ Abstracts - Deduped (0)
- **Level II Screening**
  - 01 - Full-text screening - All (0)
  - 02 - Awaiting assessment - English (0)
  - 03 - Awaiting assessment - Non-English (0)
  - 04 - Unavailable via library services (0)
- **Excluded Studies**
  - 01 - Excluded - All (0)
  - 02 - Excluded - Population (0)
  - 03 - Excluded - Intervention/ Control (0)
  - 04 - Excluded - Outcomes (0)
  - 05 - Excluded - Study design (0)
  - 06 - Excluded - Settings (0)
  - 07 - Excluded - Timing (0)
- **Included Studies**
  - 01 - Included - All (0)
  - 02 - Included - Primary reports (0)
  - 03 - Included - Secondary reports (0)
- **Background Papers**
  - 01 - Background citations (0)
  - 02 - Systematic review citations (0)

Questions

How important is the data extraction stage?

- How often do mistakes occur in data extraction in published reviews?
- Do mistakes decrease with reviewer experience?
- How can we decrease the incidence of mistakes?

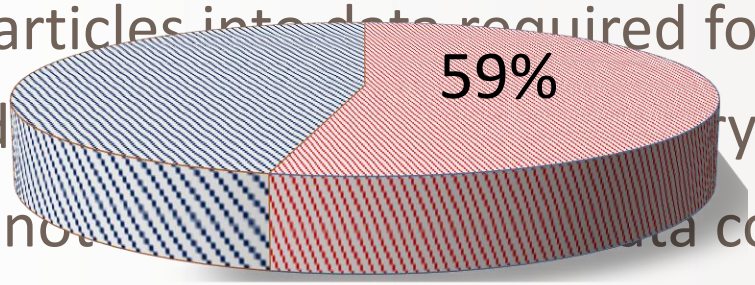
Do mistakes get past peer- & editorial-review?

- Type of discrepancies
  - Abstract-text discrepancies
  - Within-the-full-text discrepancies
  - Text-figure discrepancies
  - Text-table discrepancies
  - Multiple discrepancies

Puljaka, J Clin Epidemiol, 2020

Mistakes & review experience

- Retrospective study re-extracted data from 34 Cochrane reviews
- Found at least one error in 59% of reviews
- Types of errors:
  - misinterpretation of reported data
  - incorrect calculations made when converting data in primary articles into data required for the review
- All errors led [illegible]y effect estimates
- Conclusions not [illegible]ta corrections



Jones, J Clin Epidemiol, 2005

Mistakes & review experience

- Retrospective study compared results of 3 binary outcomes from published reviews compared to the authors own published review
- Found errors ranged from 8% to 42% (depending on outcome and review)
- Differences in pooled effect estimates were small (RR 0.01 to 0.05)

Carroll, BMC research notes, 2013

Mistakes & review experience

- Prospective, cross-sectional study on reviewer accuracy and efficiency of data extraction
- High, but similar, error rates across the various levels of reviewer experience (28% to 31%)
  - Errors of inaccuracy (14 to 18%)
  - Errors of omission (11 to 16%)
  - No significant differences in error rates or accuracy of meta-analysis results between groups
- However, time required for extraction tended to decrease with experience

Horton, J Clin Epidemiol, 2010

What can we do to decrease mistakes

- Randomized trial compared the frequency of errors from single vs. double data extraction
- Single data extraction resulted in more errors

(RD = 22%, P = 0.02)

- No substantial difference between methods in effect estimates for most outcomes
- Average time spent for single data extraction was less

(RD = 36%, P = 0.003)

Buscemi, J Clin Epidemiol, 2006

Example statement:

'The data extraction form will be pilot tested on a sample of three trial publications. Data from trial reports will be extracted independently by two reviewers with disagreements resolved through consensus, or by a third reviewer'

Components of Data Form

- Plan… Plan… Plan… Careful think and plan
- How much information to collect:

  overly vs. insufficiently detailed
- General rule of thumb:

'if you extract it then it has to have a place in the final review…

so where is that place?'
- Logical to entry into RevMan (e.g., copy/ paste)

Plan… Plan… Plan…

- Decide what you want to do before you start…

  before you start doing anything

- Sketch it out on paper or electronically
- Pilot test the form (e.g., 3 – 5 studies)
- Document changes between the protocol and conducting the review (e.g., new outcomes)

Generic Data Extraction Items

I. Coder Information
- Study ID

II. Publication
- First author, pub year
- Funding source

III. General Study Characteristics
- Inclusion/exclusion criteria
- Setting

IV. Baseline Participant Characteristics
- Number of participants randomized/ analyzed

V. Intervention
- drug name, dose

VI. Outcomes
- Primary outcome(s)
- Secondary outcome(s)
- Adverse outcome(s)

VII. Risk of Bias
- Sequence generation
- Allocation concealment
- Blinding
- Etc.

VIII. Notes

Major components of data extraction form

A. **PICO-based**
B. **Determine volume** of included studies and expected data available
C. **Determine data types** required for comparisons
D. **Determine data required** for subgroup analyses
E. **Determine data required** for risk of bias assessment

Major components of data extraction form

A. PICO-based:

1. Information on Participants (e.g., age, weight, height)
2. Information on Interventions (e.g., dosage, intervals)
3. Information on Comparisons (e.g., placebo, intervention)
4. Information on Outcomes (e.g., primary, secondary, AE)

Major components of data extraction form

B. Determine volume of included studies and expected data available

- **Scenario #1:**
  - ❖ 2 included trials
  - ❖ 3 outcomes per trial

- **Scenario #2:**
  - ❖ 100 included trials
  - ❖ 2 primary outcomes per trial
  - ❖ 5 secondary (surrogate) outcomes per trial
  - ❖ All adverse events described in all trials

Major components of data extraction form

C. Determine data types required for comparisons

- Most common:
  - Dichotomous (e.g. n/ N)
  - Continuous (e.g. mean ± SD)
- Special situations
  - Categorical
  - Ordinal
  - Counts and rates
  - Time to event (e.g. Survival analysis)
  - Computed effect sizes (e.g. MD, 95% CI)

Participants

Patient Baseline Demographics:

	Intervention 1	Intervention 2
Age (yr) *Mean ± SD*	±	±
Body weight (Kg) *Mean ± SD*	±	±
Height (cm) *Mean ± SD*	±	±
BMI (Kg/ m^2) *Mean ± SD*	±	±
Sex *n*		

Interventions/ comparators

	Intervention 1	Intervention 2	Intervention 3	Intervention 4
Classification	Starch	Crystalloid		
Type	6% HES (130/0.4)	0.9% NaCl		
Rate of administration	maximum dose 50ml/kg body wt/ day then 0.9% saline	maximum dose 50ml/kg body wt/ day then 0.9% saline		
Administration trigger	Hypovolemia	Hypovolemia		
Additional interventions				
Number randomized	3500	3500		

Outcomes

Primary outcome measures:

	Intervention (1)	*Intervention (2)*	*Intervention (3)*	*Intervention (4)*
Mortality	n = 597	n = 566	n =	n =
RIFLE categories				
Risk (R)	n = 1788	n = 1912	n =	n =
Injury (I)	n = 1130	n = 1253	n =	n =
Failure (F)	n = 336	n = 301	n =	n =
Loss (L)	n =	n =	n =	n =
End-stage (E)	n =	n =	n =	n =
Creatinine (post-treatment)				
Mean ± SD	112.11 ± (n = 3260)	102.11 ± (n = 3283)	± (n =)	± (n =)

Dichotomous (binary) outcomes

The outcome is one of two possibilities only
(e.g. alive vs. dead)

	Event	No event	
Intervention	a	b	$a+b = n_I$
Control	c	d	$c+d = n_C$

Dichotomous (binary) outcomes

- Make sure you know what outcome you are measuring
- Mortality…

  are we extracting how many died or

  how many are still alive???

- May experience difficulties with clearly identifying numbers:
  - poor reporting
  - provided in graph
  - percentage (%)
  - per protocol, not ITT

Dichotomous (binary) outcomes

- Per protocol analysis:
  - Number analyzed based on patient compliance and lack of protocol violations.
  - Leads to biased results not particularly reflective of the 'truth'
- Intention-to-treat analysis:
  - Number analyzed based on initial treatment assignment and not on the treatment eventually received
  - Everyone randomized is analyzed, even with missing data
  - ITT analysis avoids misleading artifacts and biases (e.g., non-random attrition)

Dichotomous (binary) outcomes

- Modified intention-to-treat analysis:
  - Same as intention-to-treat analysis but definition varies by study description
  - Commonly include all randomized participants that received at least one dose of intervention and had at least one post-treatment measure of efficacy or safety.

Continuous outcomes

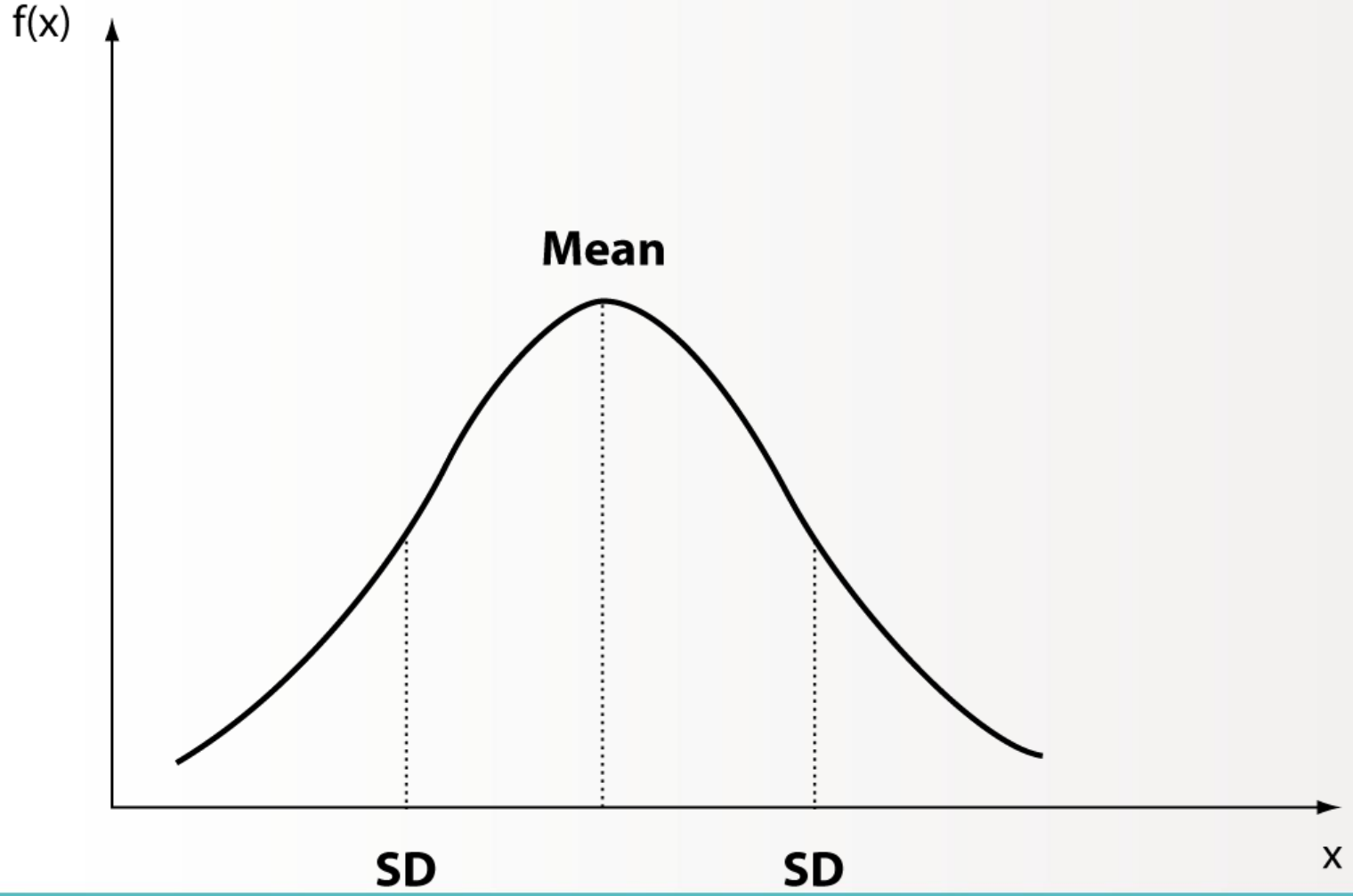
- Outcomes that can take any value in a specific range – numerical or ordered categories

  (e.g., weight, height, length of hospital stay)

- Is the scale validated (e.g., age, BP, VAS)?

- Is there a measure of variance (e.g., SD, SE, CI, P)?
- May experience difficulties with clearly identifying numbers:
  - poor reporting
  - provided in graph
  - per protocol, not ITT
  - measure of variance is not SD (convert first)

Parametric distribution



Categorical (nominal) outcomes

- Participants are classified into two or more categories with no intrinsic ordering to the categories (e.g., male, female)
- If clinically relevant, the data can be combined to form one group…
- Regardless of clinical relevance, data can be extracted separately to allow for subgroup analyses and in-between group comparisons

Ordinal outcomes

- Participants are classified into categories with a natural order (e.g., disease severity)
  - **Short**: small number of categories

    (disease severity: mild, moderate, severe)
  - **Long:** larger number of categories

    (e.g., risk assessment: low, low-moderate, moderate, moderate-high, high, very high)
- How to analyze:

  Short → often meta-analyzed as **binary data**

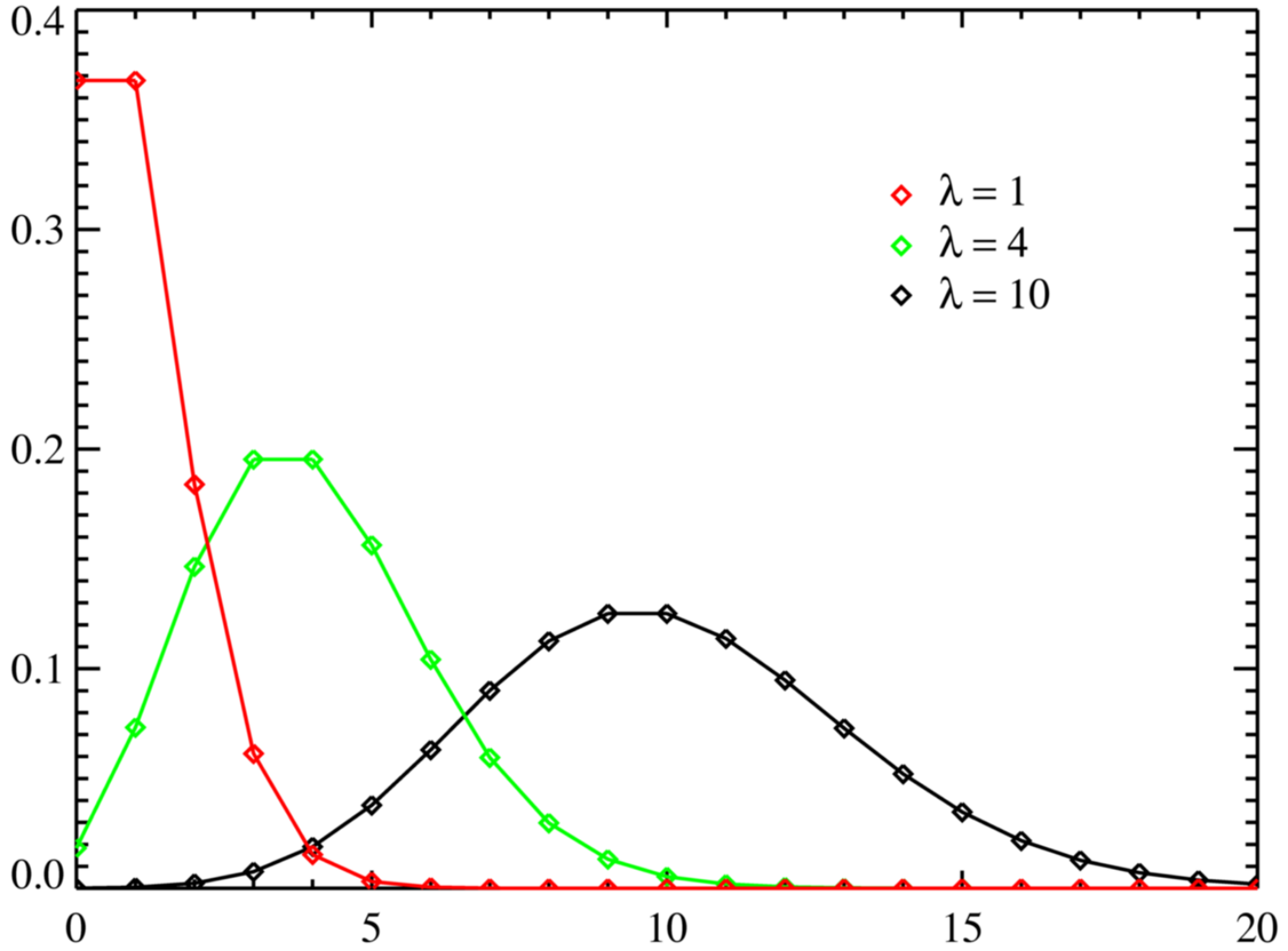
  Long → often meta-analyzed as **continuous data**

  All cases → if in doubt → consult a statistician

Count of events

- Events that can happen more than once to the same individual (e.g., MI, stroke, headache)
- Example: 100 reported cases of MI were reported during a study of 100 individuals… with 20 people each having 5 MI during the follow-up period
- If we assumed it was dichotomous data, then 100/100 (100%) of population had MI… while the truth is only 20/100 (20%) of population suffered from MI (unit-of-analysis error)
- Analyzing counts of events… statistician

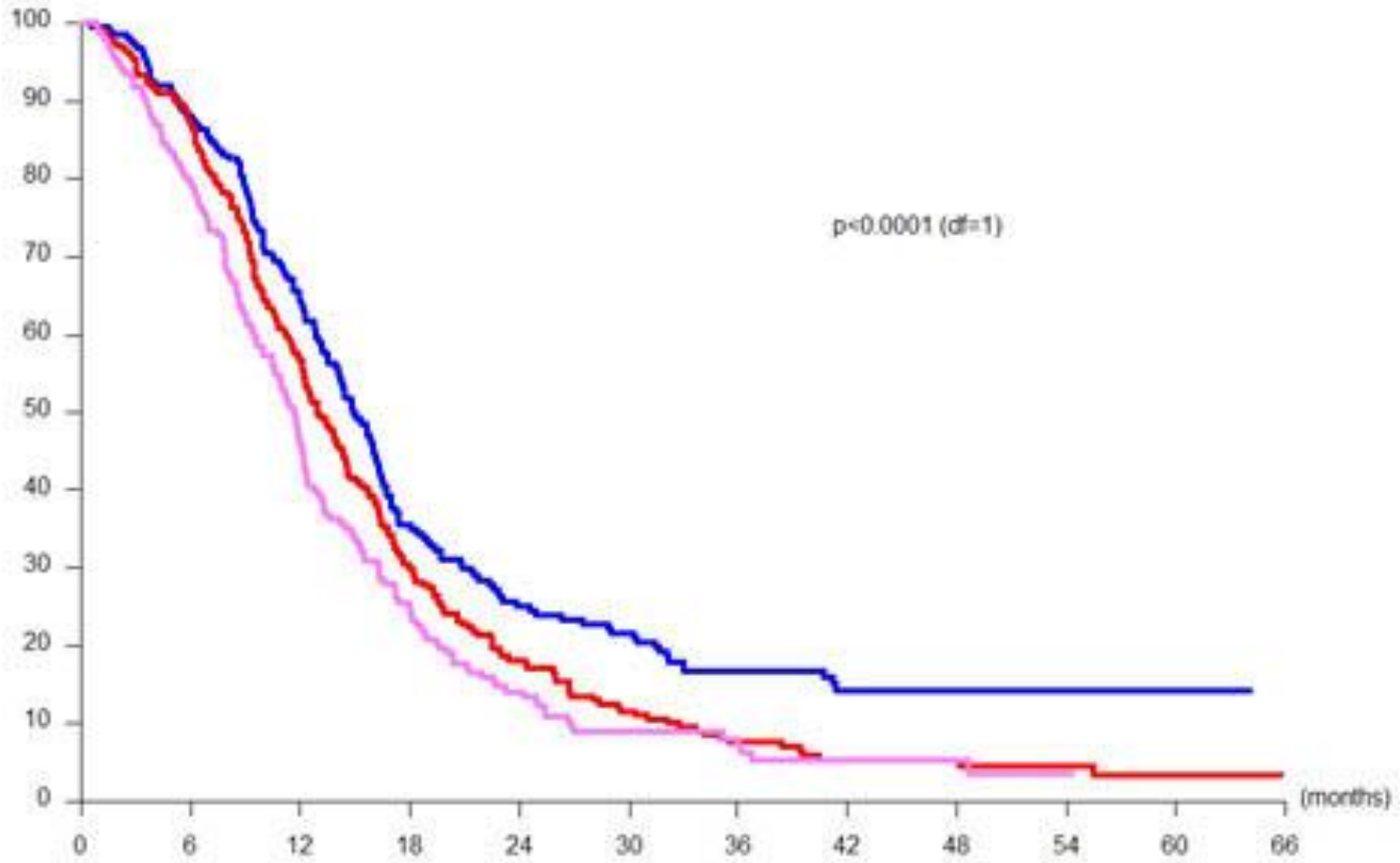
Poisson distribution



Time-to-Event Data

- Analysis of whether the event occurred and when
- 'Survival data' in stats (e.g., mortality, recurrence)
- Can sometimes be analyzed as dichotomous
- Hazard ratio analysis most appropriate

Time-to-Event Data (Kaplan Meier Curve)



O	N	Number of patients at risk									
154	183	161	118	65	45	36	25	17	15	8	3
207	220	191	124	66	39	24	15	10	6	5	2
157	170	134	77	42	23	14	8	4	3	1	0

Age
<=50 yrs
>50 & <=60 yrs
>60 yrs

Special situations

- Handling of missing data from included studies
  - Imputing data
  - Data from figures
  - Data from other sources
- How to reduce errors and bias associated with data extraction
  - Clear written guidelines (SOP) modified as changes occur
  - Data verification (double, independent data extraction vs. checking extracted data)

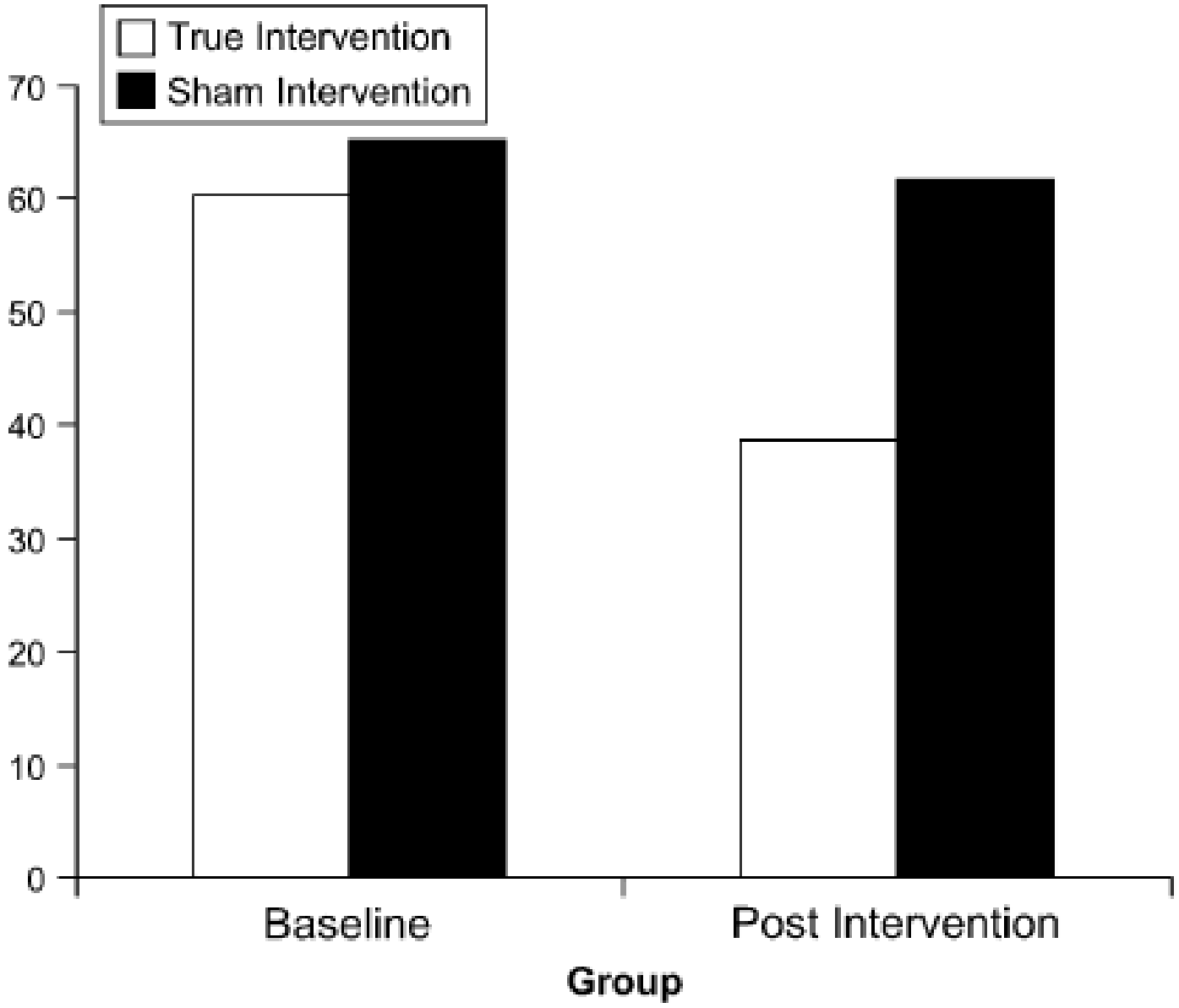


Figure 3. Visual analog scale values for anxiety. True intervention differs from sham intervention significantly ($p < 0.001$).

Handling of missing data

- Imputation of data
  - From the same study (e.g. P-value → SD)
  - From other studies in the same review
    (e.g. SD from another included study)
  - From other sources
    (e.g. Probability of event from other sources)
- Data from figures
- Data from other sources
  (e.g. Unpublished data presented in another review)

Questions

Introduction to RevMan 5

Structure in RevMan

- Comparison
  - Outcome
    - Sub-category (subgroup analyses, subdivision of outcome)
      - Study (data for each study entered in a standardized format, specific to each outcome

May not have subcategories…

then all studies fall directly under the outcome

EST. 1877

UNIVERSITY OF MANITOBA