Clinical Appraisal

RC ROUNDS

DR A JEWETT

Edits/Updates - L Jia 2021

Study Designs

Descriptive = describe status of a population (case reports, case series, cross-sectional surveys)

Analytical = analyze the relation of two or more factors (cohort, case control, cross sectional)

Experimental = manipulate one or more and observe the other (RCT)

Prevalence

Number of cases of a disease present in a population at a specific time over the number of persons in that population at that time

Point prevalence = prevalence of disease at a certain point in time (common)

- ie. Do you have depression currently

Period prevalence = how many people have disease at any point during a certain time period

- ie. Have you had depression in the last 10 years

Lifetime prevalence = proportion of patients who experienced an illness up until that point in their life

- prone to age effect, recall bias, fatality effect

Incidence

Number of new cases of a disease occurring in the population during a specified period of time over number of persons at risk of developing the disease during that period of time

Case Control Studies

Cases = have the disease

- divide into were exposed vs were not exposed

Controls = do not have the disease

- divide into were exposed and were not exposed

Cohort Studies

Defined population → non-randomized into exposed and non-exposed → each into disease and no disease

Prospective – looking forward

Retrospective – looking back

Pros and cons

Case control

- cheaper
- retrospective
- both exposure and disease has already taken place
- more bias prone, esp to recall
- quicker
- useful in rare diseases ie scz

Cohort

- costly
- prospective
- exposure has happened, not outcome
- less bias prone
- prolonged
- useful in rare exposures (ie. Chernobyl)

Secondary Reviews

Primary research = recruit individuals

Secondary research = recruit individual research studies

- narrative review
- systematic review – with or without meta-analysis

Quality of meta analysis

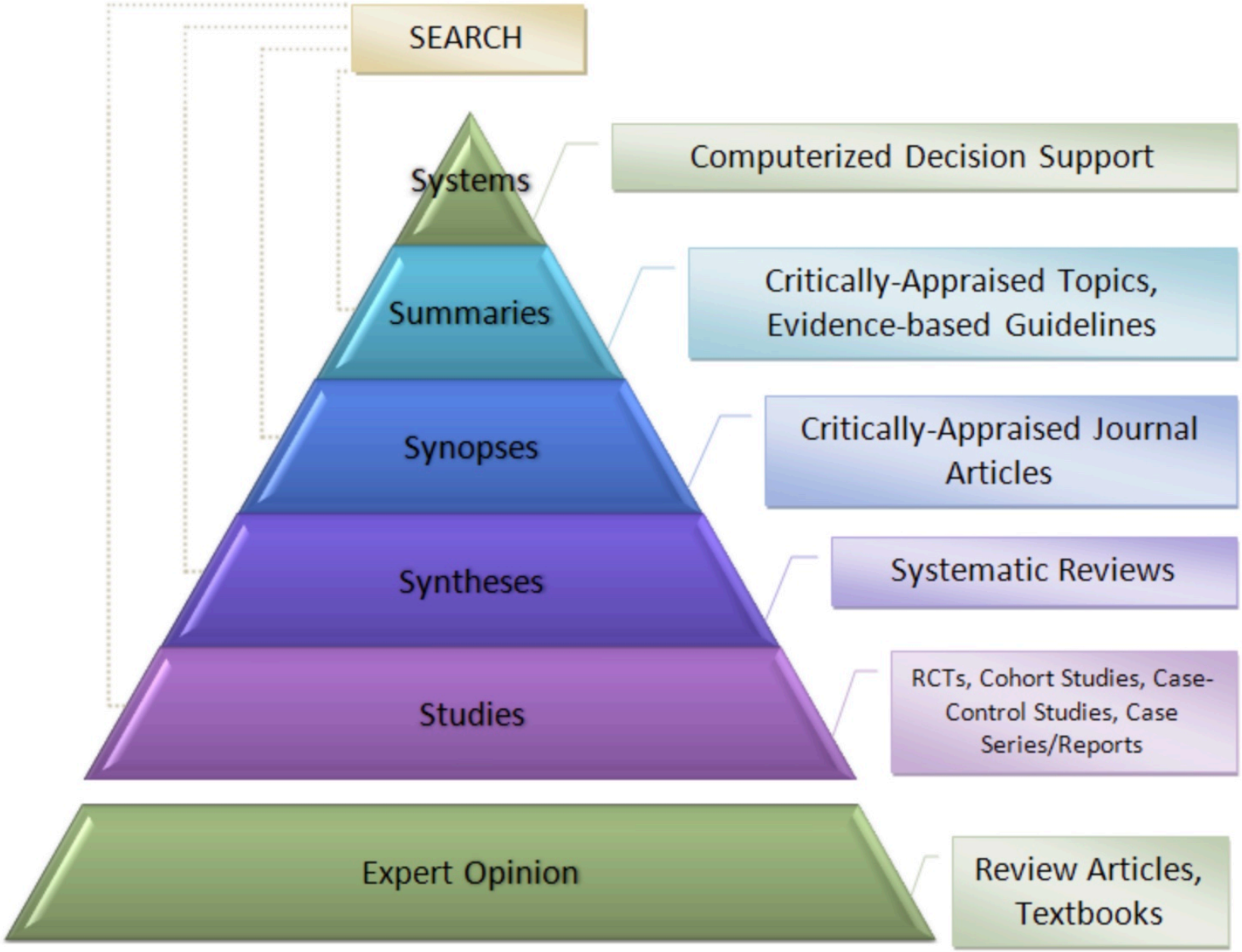
- higher the number of included studies
- larger the sample size of individual studies
- better the methodological quality of studies (weighting)
- more similar the process and conduct of studies = better quality meta-analysis
- garbage in garbage out – if you use bad studies, meta-analysis will not be good

Systematic review = SELECT, SYNTHESIZE and SUMMARIZE high quality primary RESEARCH relevant to a particular RESEARCH Q —> basis for EBM

Meta-analysis = statistical APPROACH to combining the data derived from a systematic-review to generate statistical OUTCOMES

Hierarchy of evidence: Systematic Review/Meta Analysis → RCT → Cohort → Case Control → Cross sectional survey → Case reports

The 5S Pyramid



Systems: decision support technology that matches individual patient characteristics with best available evidence that applies
- no currently available examples for mental health (more of a theoretical construct than a reality at present)

Summaries: an integration of evidence related to a particular clinical problem; i.e. clinical practice guidelines, Uptodate
- pro: easy access to a range of evidence relevant to the management of a particular condition
- con: only as good as the process by which this information is selected, interpreted and kept up to date

Synopses: filtered and pre-appraised evidence, i.e. CJP
- pro: succinct appraisals of studies that are selected by peers for high quality and relevance to clinical practice
- con: these journals doe not include all important studies, and the synopsis is subject to the interpretation of its author

Syntheses: systematic reviews (and meta-analyses); i.e. Cochrane
- pro: evidence at this level is particularly helpful when many single (small) studies have been published and a single pooled estimate of effect is desired
- con: in areas where there is scant high quality evidence, systematic reviews are usually not helpful (“garbage in, garbage out”); systematic reviews and meta-analyses can also be difficult to interpret and apply to individual patients

Studies: single studies, searched using Medline or PubMed
- pro: searching the evidence at this level nets the most complete and up to date list of studies
- con: it also yields mostly misses (studies that cannot be readily applied to practice or that are of low quality) and is inefficient and time-consuming

Ref: Haynes, B. ACP Journal Club. 145(3): 8-9, 2006.

Terms RE Causation

- temporal relationship = exposure precedes disease
- strength of association = large relative risk or odds ratio
- dose-response relationship = increasing exposure increases risk
- reversibility = reducing exposure decreases risk
- consistency = similar results from other studies
- biological plausibility = consistent with pharmacological or toxicological data
- analogy = relationship established for similar cause and disease
- elimination of other explanations = not the result of confounding or bias
- specificity = one cause, one effect

Odds Ratio

Make table – Exposed, Non exposed, total; case, controls

Odds ratio = how likely you are to have disease if exposed

Case control and cohort studies

AD/BC = OR = A/C (odds that case was exposed = cases exposed/cases NOT exposed)
B/D (odds that control was exposed = controls exposed/controls NOT exposed)

	Case/Disease	Controls/No disease
Exposed	A	B
Not Exposed	C	D

Risk Rates

Risk is a probability of an event – incidence of outcome

- need prospective observation (cohort or RCT design)

Risk difference – absolute risk of exposed – absolute risk of non-exposed

- also called attributable risk (epidem) or absolute risk reduction (ARR in RCTs)

Relative risk = risk in exposed/risk in non-exposed

RR = A/A+B divided by C/C+D

	Case/Disease	Controls/No disease
Exposed	A	B
Not Exposed	C	D

Probability of an event (good or bad) occurring in exposed people compared to the probability of the event in non-exposed people

RR = 1 = risk in exposed equals risk in non-exposed = no association

RR > 1 = risk in exposed greater than risk in non exposed = positive association, possibly causal

RR < 1 = risk in exposed less than in non-exposed = negative association, possibly protective

Other Terms

NNT = number of patients need to be treated for one to show response

- NNT = 1/ARR, round to next highest whole number

ARR = AbR (treatment) – AbR (placebo) – risk difference

Association

A research study finds that X is associated with Y

- causal link = X may be causing Y
- chance finding = spurious, no real link (random error)
- confounding = through a third variable
- bias

Bias = any systematic error in the design, conduct or analysis of a study that results in a mistaken estimate of an exposure's effect on the outcome (ie. Selection, information, measurement, analysis, confounder)

- Selection bias: selection of individuals, groups or data for analysis in such a way that proper randomization is not achieved, thereby ensuring that the sample obtained is not representative of the population intended to be analyzed.

- Publication bias or reporting bias: the distortion produced by NOT publishing negative results or results which go against the experimenter's prejudices, a sponsor's interests, or community expectations.

- Confirmation bias: the distortion produced by experiments that are designed to seek confirmatory evidence instead of trying to disprove the hypothesis.

- Exclusion bias: results from applying different criteria to cases and controls in regards to participation eligibility for a study/ different variables serving as basis for exclusion.

RCTs

Subjects do well

- volunteerism
- elite group – stringent eligibility criteria filters out poor prognostic subjects
- placebo effect
- Hawthorne effect – people who know they are being studied modify their behavior and do better than the average patient

For trials to be GENERALIZED

Need external validity = generalizability → caused by sampling and experimental set up

Internal validity = chance, confounding and bias affect internal validity

Pragmatic design

MBP = Measurement based practice

Internal validity is the INTEGRITY of the experiment - degree of confidence that the causal relationship you are testing is not influenced by other factors or variables.

External validity is the EXTENT which your results can be generalized to other contexts.

The validity of your experiment depends on your experimental design.

Test Properties

	Disease	No disease
Test +	A TP	B FP
Test -	C FN	D TN

- sensitivity (true positive rate) → how good is the test at picking up people who have the condition
  - true positive/total diseased
  - TP / disease (A / A+C)
  - SnOUT = rule disease OUT
  - i.e. testing negative, likely DON'T have disease; testing positive might lead to unnecessary distress due to false +
- character of the instrument, not the patient's diagnosis
- a very sensitive instrument makes more false positive, okay if diagnosis should not be missed, but diagnosis not traumatic
- if highly sensitive test is negative – helps rule out disease
- specificity (true negative rate) → how good is the test at correctly excluding people without the condition
  - true negative/total non-diseased
  - TN / no disease (D / B+D)
  - SpIN = rule disease IN
  - i.e. test positive, likely HAVE the disease; testing negative might miss some who actually have disease (false -)
- character of the instrument not the patient's diagnosis
- very specific instrument will have more false negatives → good if diagnosis can be missed but false diagnosis is bad or costly, helps to rule in a disease

Predictive Value

- positive predictive value = PPV
- true positive/total test positive
- more useful measure, informs the chance of having the disease if positive
- negative predictive value = NPV
- true negative/total test negative
- informs chance of not having the disease if tested negative

Qualitative:
Nominal —> variable with categories that do not have a natural order or ranking. Calculations i.e. mean, median, or standard deviation, would be meaningless.
Examples: genotype, blood type, zip code, gender, race, eye color, political party
Ordinal —> order matters but NOT the difference between values.
Examples: education level ("high school","BS","MS","PhD"), satisfaction rating ("dislike", "neutral", "like", "extremely like").

Statistics

Quantitative:
Interval —> there is order AND the difference between two values is meaningful.
Examples: temperature, pH, SAT score (200-800), credit score (300-850).
Ratio —> all the properties of an interval variable, and a clear definition of 0. When the variable equals 0, there is none of that variable. Examples: enzyme activity, dose amount, reaction rate, flow rate, concentration, pulse, weight, length, survival time.

Qualitative data = nominal or categorical, gender, ethnicity, etc.

Ordinal = cancer staging, pain rating, order

Quantitative data = discrete - counts only, no units - not measured

- continuous = interval (added up scores of rating scales) vs. ratio scales (duration of a problem, age, weight, etc.)

Independent variable = controlled by experimenter

Dependent variable = depends on the value of independent variable

Mean = average

Median central value in a rank order, 50th %ile

Mode = most commonly occurring value

Variance = average of the squared differences of individual observations from the mean

SD = square root of variance

Range = upper limit - lower limit

IQR = 25th percentile to 75th %ile

Statistics

- proving is more difficult than disproving
- null hypothesis – what the researcher sets out to disprove
- alternate hypothesis – what the researcher believes and wants to prove (research question)
- type I error = error of commission – chance that a positive result is spurious (false positive)
- type II error = error of omission (false negative)

Pharma companies:
WANT to be #1 (F+), want to be ALPHA (0.05)
DONT WANT #2 (F-), dont want to be beta (0.2)

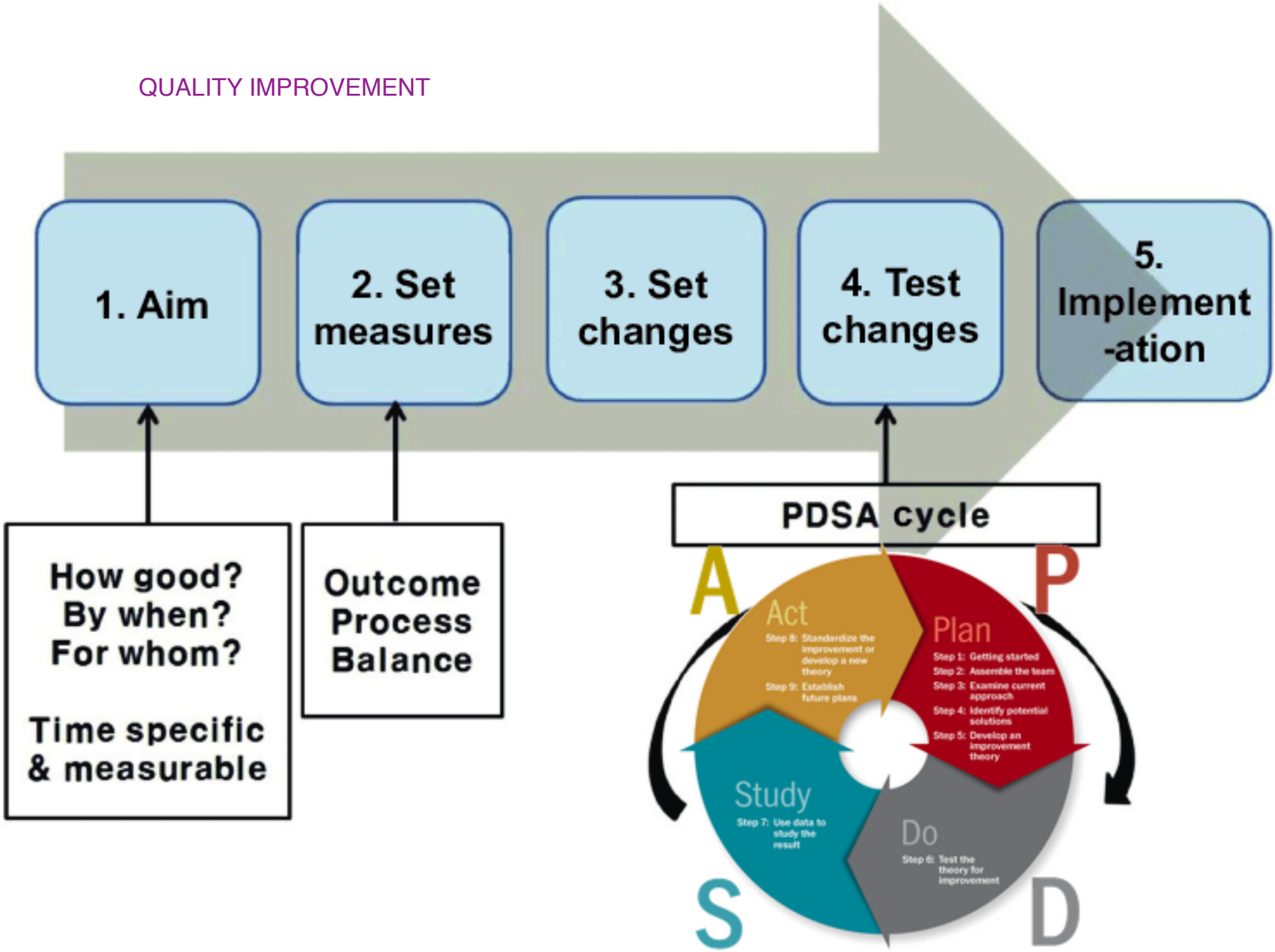
- power = ability to detect true effect (want at least 80%) POWER = 1 - beta
- confidence intervals = support point estimate with degree of confidence

CI = the range/interval of values to estimate the true value of a parameter = 1 - alpha = 95%
if the confidence level is 95% —> then the likelihood that the true value lies within the interval is 95%
Confidence level = 1 - alpha = 1 - 0.05 (usually) —> alpha is the probability of rejecting a TRUE null hyp
p-value = probability that observed value occurred by chance - if p < alpha, then result is statistically significant

Statistics

- reliability = replicability
- test-retest correlation
- inter-rater reliability
- validity = applicability
- face (subjective), construct (deeper)
- construct validity criterion = concurrent (external, subgroups), predictive (future), incremental validity

QUALITY IMPROVEMENT



Q bank pearls

- specificity = known cases plus positive control cases over cases without disease
- false positive rate 1-specificity (ie. 85% specificity = 15% false positive rate)
- incidence rate = number of new cases per person-year of observation
- quality improvement 80-20 rule of management = pareto chart
- cross sectional study can measure point prevalence
- cohort study can measure incidence - need to observe population long enough to pick all new cases
- prevalence = incidence x duration
- lifetime morbid risk = probability a person developing a disorder during an entire period
- lifetime prevalence = proportion of individuals in the population who have ever manifested a disorder alive on a given day ifetime prevalence is the proportion of a population who, at some point in life has ever had the characteristic.

Q bank pearls

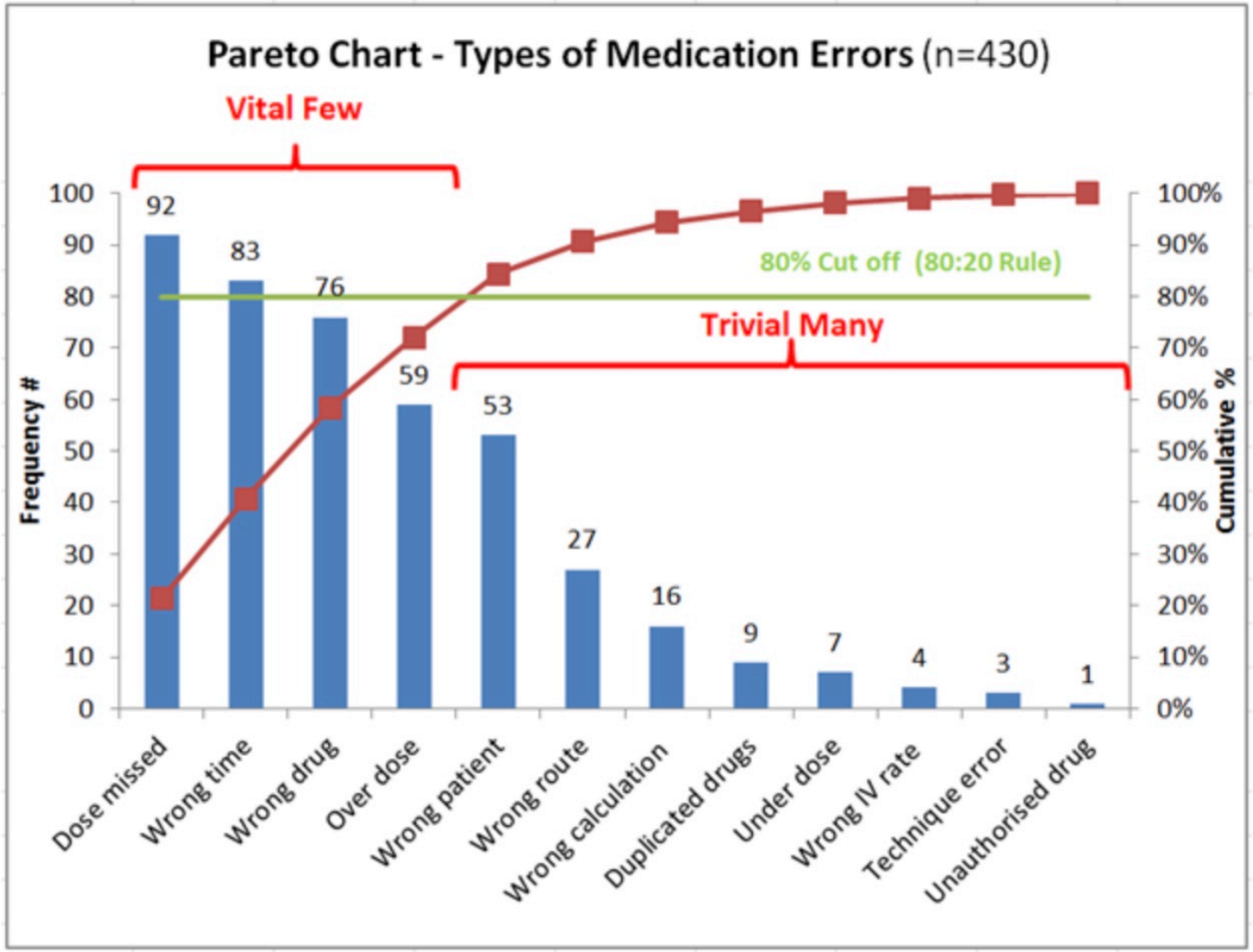
Low N = chance of a type 2 error may increase (if you don't examine enough rabbits, you may conclude blue rabbit does not exist – false negative – type II)

Use a pareto chart for quality improvement – if you focus on the 20% of high impact circumstances, 80% of problems can be managed, bar chart with most frequent events on the left and least frequent to the right

A study may lack internal validity if the differing outcomes are due to factors other than the difference in the interventions – ie. Unequal distributions of prognostically important subject characteristics, differences in intervention delivery, or inconsistent measurement of significant elements such as comorbidities or outcomes → randomization reduces unequal distributions and boosts internal validity

QI – can be achieved through case conferences, no need for special committee, define problem so efforts focused, often informs utilization management

Figure 1: Pareto Chart – Audit of types of medication errors



80/20 Rule – The Pareto Principle

The 80/20 Rule (also known as the **Pareto principle** or the law of the vital few & trivial many) states that, for many events, roughly 80% of the effects come from 20% of the

Q Bank Pearls

Face validity = does this make sense, commonsense appearance, conducted by laypersons and experts – ie. OCD scale

Distortion of the apparent lifetime prevalence could be caused by the distribution of age in a sample (ie. Too many young people can lead to lower estimate if illness has late onset)

Field trial = early stage testing of a tool to establish practical utility

Two major disadvantages in interpreting meta-analysis or secondary research = publication bias and heterogeneity in terms of clinical and methodological variables

Paired T test – two scores from same individuals (ie. Pre and post)

Q bank pearls

Higher prevalence will increase PPV and decrease NPV

Likelihood ratio = sensitivity/1-specificity

In a positively skewed distribution, most data will fall to left of mean, while tail will be on right; mean is to the right of the median and mode to the left; neg skewed opposite

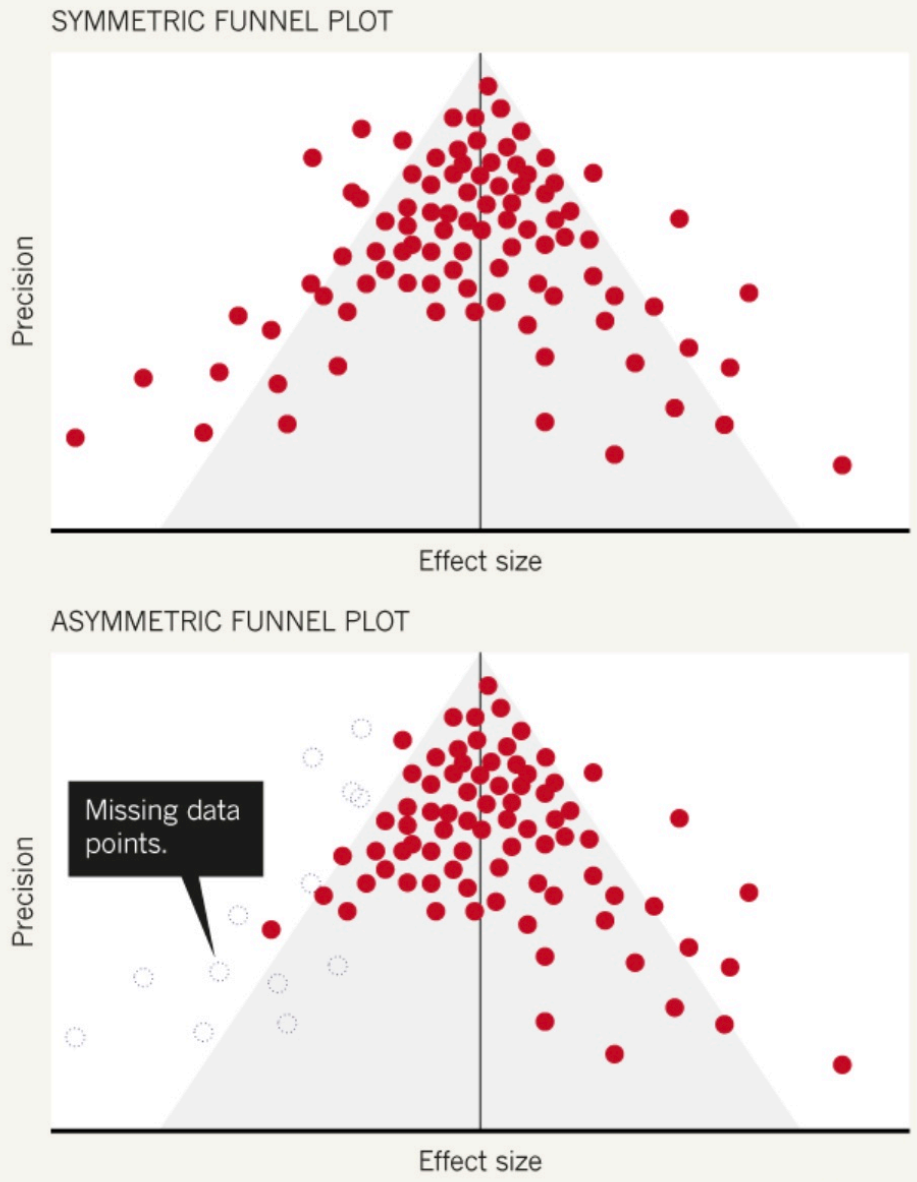
2 standard deviations = 2.5%

Chi square tests = counted numbers = compares tallies of categorical responses between two or more groups = best for comparing observed vs. expected proportions

Most common method to detect publication bias is to investigate asymmetry in inverted funnel plots – shows relation between study effect size and precision (small studies likely to remain unpublished if results nonsignificant or unfavorable, larger studies get published regardless)

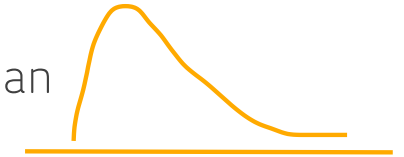
PLOTTING FOR PUBLICATION BIAS

Funnel plots show the data from multiple experiments. In some – but not all – cases, a wildly asymmetric shape can indicate that some negative results are missing from the literature.



Toronto Review Course

Right skewed distributions common – mean to the right of the mode and median



Rt skewed = + skewed

68-95-99

If categorical/percentages/proportions – use chi square, if continuous (means), use T test if 2, ANOVA if more than 2

"Continuous TAN"

Linear regression if two continuous variables like depression scores with anxiety scores, logistic regression if categorical outcome (ie. Continuous independent, categorical dependent like anxiety score w greater odd of meeting criteria for MDD)

Correlation strength over 0.7 is strong

NNT = 1/ARR

Resources

Thank you to Dr. Palaniyappan from the London Review Course, from which this is heavily derived!