

Issues of Applicability in Prevention Studies

Russell Harris, MD, MPH
University of North Carolina,
USA



My Agenda

- (1) Some general definitions and comments about applicability
- (2) Example: breast cancer screening (looking back)
- (3) Example: prostate cancer screening (looking forward)



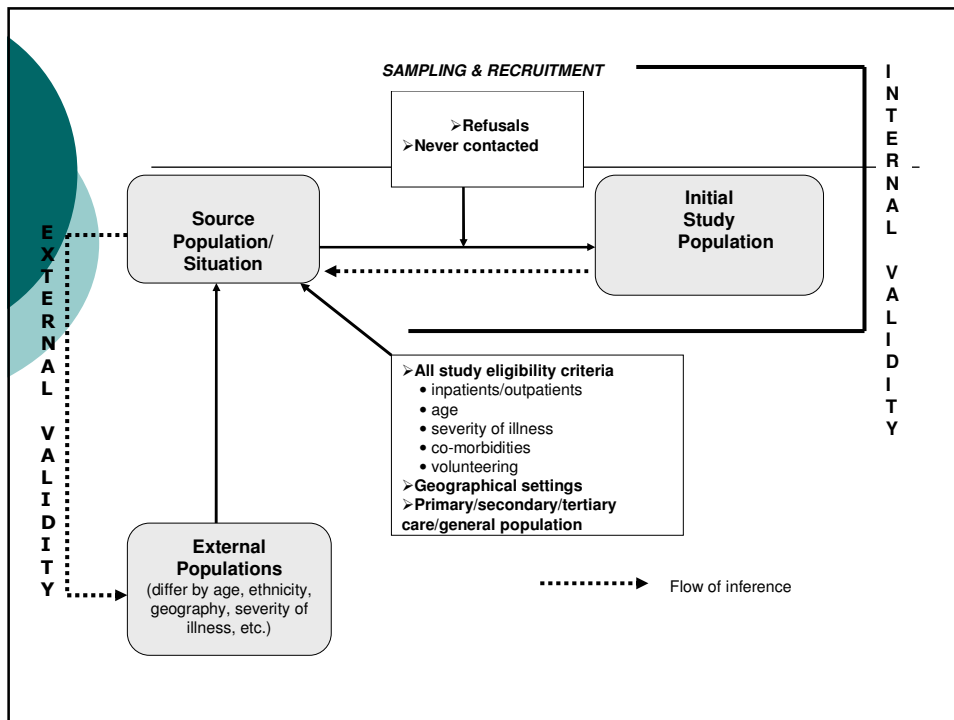
Terminology

- Applicability = External Validity = Generalizability
- Efficacy and Effectiveness are descriptions of studies. Efficacy studies are less and effectiveness studies are more applicable.
- Internal Validity is different. It is the degree to which the study was unbiased for the initial study population.



Terminology: Applicability/External Validity

- The extent to which the results (both the direction and the magnitude) of the study can be generalized (or applied) to “external” populations, external to the source/study population.
- External is defined by not eligible for the study population.



Issues that Affect Applicability/External Validity

- (1) Participant factors: susceptibility for the outcome; comorbidity; demographics; preferences; motivation; run-ins; volunteers
- (2) Health system factors: health care centers and clinicians; training; technology; environment
- (3) Research factors: asking different question; assessing different outcome; time of follow-up; duration of study
- (4) Time: have populations, health systems, treatments changed over time?



General Comments About Applicability/External Validity

- (1) A study may be applicable for one external population/situation but not another.
- (2) Issue is how external and source populations differ in:
 - the factors that determine the outcome
 - not general statistical representativeness
- (3) A continuum rather than a dichotomy



Two Common Decision Rules about Applicability

- 1. Less restrictive: Consider the evidence applicable unless there is reason not to.
- 2. More restrictive: Only consider the evidence applicable to people who would have been eligible for the study.
- My own position is somewhere between in these two. There is room for judgment.



Example: Breast Cancer Screening

- (1) 8 RCTs done in 1980s and before: pooled RR of 0.78 (0.70-0.87) for mammography vs no mammography for women ages 50-74 for outcome of breast cancer mortality.
- (2) Mammography now widespread
- (3) The applicability question: Looking back, has mammography reduced breast cancer mortality to the extent expected in the community?



Example: Breast Cancer Screening in US

- Breast cancer mortality 1975-2000 declined by 24%. Best modeling study shows that about half of this (about 12%) due to mammography.
- Are the RCTs still applicable to the general population today? Does applicability change over time?



Reasons that Breast Cancer Screening Studies May Be Less Applicable Today

- Non-use of mammography
- Failure of technology/interpretation
- Improved background detection
- Lack of benefit from more sensitive mammography
- Improved treatment



Example: Breast Cancer Screening

- Bottom line: Older breast cancer studies are likely less applicable today than they were 25 years ago.
- They may be even less applicable in the future.



Example: Prostate Cancer Screening

- The newly published initial results from the ERSPC trial show a 20% reduction in prostate cancer mortality over 9 years.
- NNS = ~ 1400 ; NNT = 48
- $\sim 45\%$ of cancers diagnosed were overdiagnosis



Example: Prostate Cancer Screening

- Is the ERSPC trial applicable to the general population?
- In benefits?
 - subgroups by age, ethnicity
 - would longer follow-up show a larger (or a smaller) benefit?
 - treatment in the community vs trial



Example: Prostate Cancer Screening

- In harms?
 - harms not well reported in RCTs
 - other observational studies tell us more about harms in the community
- Bottom line: There are major uncertainties about the applicability of the ERSPC trial results to the general population.



Conclusions

- (1) Assessing applicability is complex. We should continue to evaluate the programs we implement.
- (2) Even established prevention programs may become less effective with time.
- (3) The long-term effects of prevention programs are often unknown. We may still judge it is a good idea to implement a prevention program, but we should still remember what we don't know.



Mark Twain

- It ain't what we don't know that gets us into trouble. It's what we know for sure that just ain't so.
- (Donald Rumsfeld: It is the things we don't know that we don't know.)



Thank you



General Comments About Applicability/External Validity

- Often a major problem for RCTs
- Highly selective eligibility criteria may improve internal validity, but may worsen external validity.
- Often less of a problem for large cohort studies (where internal validity is a problem)



General Approach to Critically Appraising Studies

- (1) Assess systematic error (internal validity: issue of bias; clinical confidence)
- (2) Assess random error (precision; power; statistical confidence)
- (3) When internal validity and precision are adequate, then assess applicability