Evidence of Harm

https://pixabay.com/vectors/warning-danger-signs-symbols-41310/

Objectives

* This presentation aims to increase the participants' knowledge to appraise scientific evidence on harm.
* At the end of this session, the participants are expected to;
  * Discuss the appraisal of the validity of a harm evidence
  * Discuss the importance of a valid evidence about harm for clinical practice
  * Describe hazard ratio.
  * Explain the significance of confounding factors in harm articles.
  * Discuss the application possibility of a valid and important evidence into clinical practice

Critical Appraisal

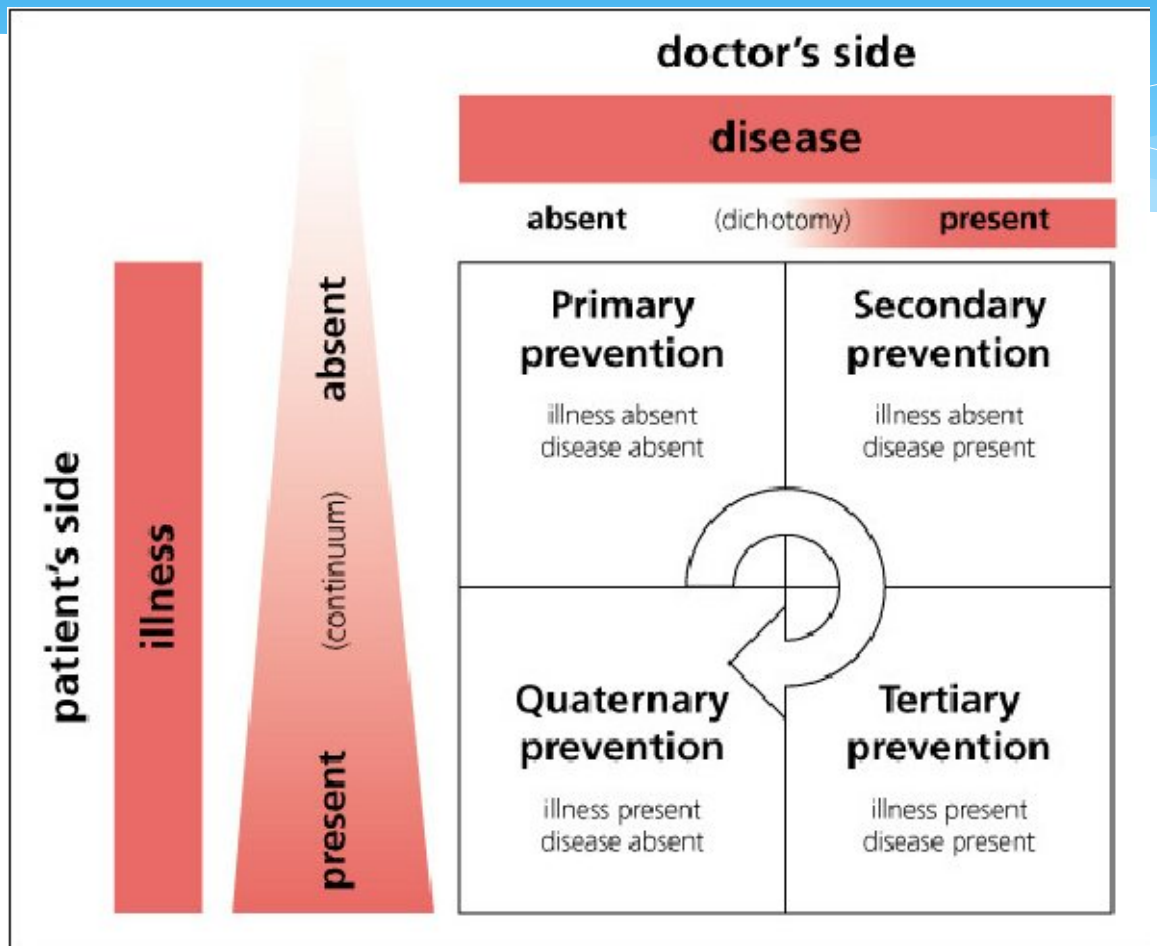
* Validity
  * Bias in the study?
  * Fatal flaw or weakness?
* Results
  * Clinical significance?
  * Statistical significance?
* Applicability

* Basic study types
  * Therapy
  * Harm
  * Diagnosis
  * Prognosis
* Advanced study types
  * Systematic review
  * Practice guideline
  * Disease prevalence
  * Economic analysis
  * Decision analysis

Why study harm?

* Harm represents the dark half of therapy
* The ideal therapy has great benefit and minimal risk
* The reality is that many therapies have associated harms that only come to light after years of observations
* Waiting until we have a proof of 100% safety?

Levels of Prevention



https://www.researchgate.net/profile/Marc_Jamoulle

* **+ primordial prevention**
* In 1978, Strasser suggested that prevention of cardiovascular disease (CVD) should go beyond the concept of primary prevention. He coined the term ‘primordial prevention’ to denote activities that prevented the penetration of risk factors into the population[1] by intervening to stop the appearance of the risk factors.

https://www.ecrjournal.com/articles/primordial-prevention-cardiovascular

How to appraise harm studies?

* 1. Is the evidence valid?
* 2. Is the valid evidence important?
* 3. Is the evidence applicable?

How should we study harm?

* Cohort or case-control studies are often the designs of choice for harm, especially for relatively uncommon events.
* A well-designed non-RCT can provide excellent evidence, but the susceptibility to bias is greater for these study designs.
* No solution to this problem other than to critically appraise harm articles for how well they minimize bias.

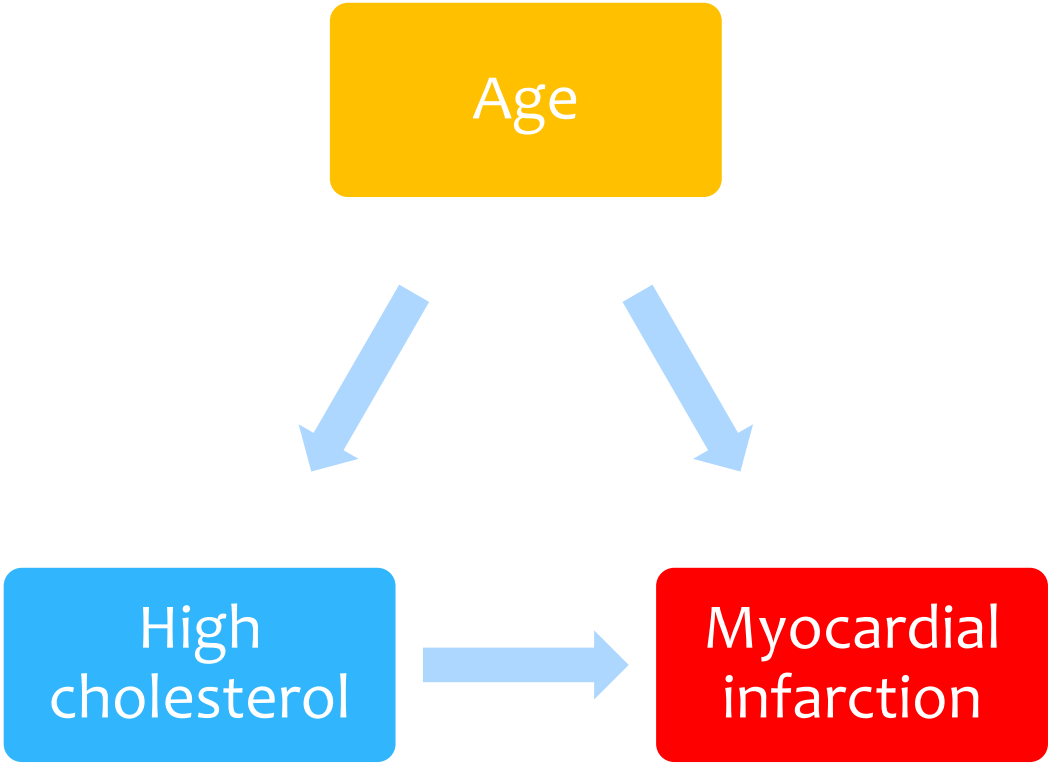
1. Are the results valid?

* Did the investigators demonstrate similarity in all known determinants of outcome?

* Did they adjust for any differences in their analyses?

Confounders

* Confounders are factors affecting both input and outcome variables.



Are the results valid?

* For RCT's, the randomization process helps us balance groups by both known and unknown risk factors.
* Therefore, RCT's remain a powerful design for harm studies
* The problem is that RCT's are very often not feasible for studies of risk.

Are the results valid?

* For cohort and case-control studies, our options are more limited.
* Cohort design
  * Need to measure known potential confounders very carefully.
  * Can adjust for these statistically
  * CANNOT adjust for unknown differences between the cohorts, however.

Are the results valid?

* Case-control design
  * Can match cases and controls by important factors such as age, sex, etc.
  * Can adjust statistically
  * Still vulnerable to unrecognized confounders, as well as the biases we have discussed previously for case-control designs

Are the results valid?

* **Were exposed patients equally likely to be identified in the two groups?**
* A particular issue for case-control studies (recall bias)
* Birth defects and prenatal alcohol exposure
* Who is more likely to remember that one glass of wine during the first trimester?
  * Mom with normal, healthy baby?
  * Mom with baby with cleft palate?

Are the results valid?

* Look for attempts to control these biases:
  * Blinding interviewers and participants to study hypothesis
  * Scripted interviewing
  * Confirmation of exposures using clinical databases and medical records

Are the results valid?

* **Were the outcomes measured in the same way in the groups being compared?**
* For RCT's and cohort studies, patients in the exposure group might be followed more closely (surveillance bias).
* Look for clear description of outcome assessment
  * Blinded, groups treated equally in all ways

Are the results valid?

* **Was follow-up sufficiently complete?**
* If the exposed and nonexposed groups are not followed equally, bias may be introduced.
* Especially if those who are lost to follow-up DIFFER from those not lost.

Are the results valid?

* Example: a cohort study examines the association of coffee with upper GI bleeding.
* The conclusion is that coffee is protective.
* However, in the coffee-exposed group, a large number of participants are lost to follow-up.
* These participants turn out to be heavy drinkers who forget to return for follow-up – the true rate of UGI bleeding in the coffee group is likely higher than found in the study, because these subjects were excluded and are at higher risk of bleeding.

Are the results valid?

* What is an acceptable loss to follow-up?
* Difficult to answer:
  * If the results would not change under all possible values for the lost data, we can be confident in the results
  * This is rare – usually we have to trust that those lost to follow-up do not differ from those included in the study.
* Is there a percentage of missing data that is always acceptable?

Are the results valid?

* Example:
  * Exposed group risk = 2/100
  * Non-exposed group risk = 1/100
  * RR = 2
* What if we lost only 1% of our participants?
* What if the 1 lost in the exposed group was an event, and the 1 lost in the non-exposed group was not:
  * Exposed group risk = 1/99
  * Non-exposed group risk = 1/99
  * RR = 1

2. What are the results?

* **How strong is the association between exposure and outcome?**
* We discussed RR, ARR, NNT, OR, etc. during the previous sessions (and will see these again as we review papers in the small groups).
* In general, larger values of RR or OR garner greater attention.
* Dose-response also is consistent with causation.

What are the results?

* **How precise is the estimate of risk?**
* What is our measure of precision?
* Confidence intervals, NOT p-values!
* A wider confidence interval leads us to place less faith in the reported results.

What are the results?

* How would you interpret each of the following 95% CI's for a relative risk?
* (11, 15)
* (1.1, 212)
* (0.1, 14)
* (0.97, 1.02)

Hazard Ratio (HR)

* Regression models are used to obtain HR and their CI.
* HR is a measure of an effect of an intervention on an outcome of interest over time. Hazard ratio is reported most commonly in time-to-event analysis or survival analysis.
* The outcome could be an adverse/negative outcome (e.g. time from treatment/surgery until death/relapse) or a positive outcome (e.g. time to cure/discharge/conceive/heal or disease-free survival).
* Hazard Ratio (i.e. the ratio of hazards) = Hazard in the intervention group ÷ Hazard in the control group

https://s4be.cochrane.org/blog/2016/04/05/tutorial-hazard-ratios/

3. How can I apply the results to patient care?

* **Were the study patients similar to the patient in my practice?**
* If the study examined only patients over the age of 65 and your patient is 21, the results may not be relevant (depending on the harm/disease in question).
* One way to ask this: is there any compelling reason why I could NOT apply this study's results to my patient?

How can I apply the results to patient care?

* **Was the duration of follow-up adequate?**
* Why would this matter?
* Need enough time for a potential harm to occur
  * For example, many cancers take years or decades to become detectable

How can I apply the results to patient care?

* **What was the magnitude of the risk?**
* Or, even if the risk is real, does it matter?
* NNT and NNH help us with this.

How can I apply the results to patient care?

* **Should I attempt to stop the exposure?**
* How strong is the evidence?
* How great is the risk?
* How much benefit would be lost by withholding the treatment?
* EXAMPLES: Vioxx, HRT, PPA, …

* Your 25 years-old female patient reports using echinacea for common cold and wants to inquire about the safety of her practice.
* You make a literature search and come across the following article:
  * Barrett BP, Brown RL, Locken K, Maberry R, Bobula JA, D'Alessio D. Treatment of the common cold with unrefined echinacea. A randomized, double-blind, placebo-controlled trial. Ann Intern Med. 2002 Dec 17;137(12):939-46. doi: 10.7326/0003-4819-137-12-200212170-00006. PMID: 12484708.

https://www.acpjournals.org/doi/full/10.7326/0003-4819-137-12-200212170-00006?journalCode=aim

* 1. Is the evidence valid?
* 2. Is the valid evidence important?
* 3. Is the evidence applicable?

1. Is this evidence about harm valid?

* 1. Were there clearly defined groups of patients, similar in all important ways other than exposure to the treatment or other cause?

Table. **Characteristics of Study Participants**

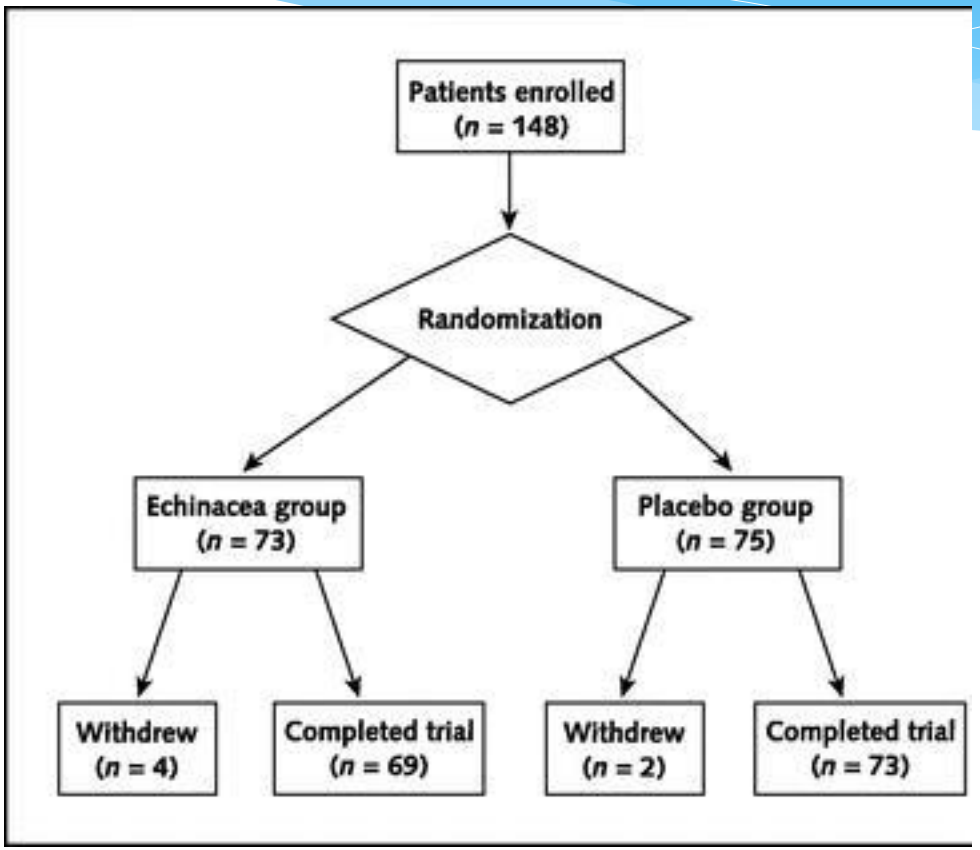
Characteristic	Echinacea Group	Placebo Group
Participants entering study, *n*	73	75
Participants completing protocol, *n*	69	73
Mean age ± SD, *y*	20.8 ± 2.4	21.0 ± 3.4
Women, *n (%)*	50 (72)	48 (66)
Current use of tobacco, *n (%)*	15 (22)	15 (21)
Nonwhite ethnicity, *n (%)*	4 (5.8)	4 (5.5)
Participants who had taken echinacea before, *n (%)*	30 (43)	28 (38)

* *“The placebo and echinacea capsules were indistinguishable to study personnel and to the participants. Allocation to echinacea or to placebo was concealed from participants and from the investigational team until all data had been collected, entered, and cleaned."*

* 2. Were treatments/exposures and clinical outcomes measured in the same ways in both groups? (Was the assessment of outcomes either objective or blinded to exposure?)

* *"Primary outcomes were defined prospectively as severity and duration of self-reported symptoms. Duration was defined as the number of days from study enrollment to the last day before the participant answered "No" to the question, "Do you think that you are still sick today?"*
* *Symptom severity was measured daily on nine-point Likert scales. The 15 symptoms assessed were dry cough, productive cough, cough interfering with sleep, sore throat, scratchy throat, hoarseness, runny nose, plugged or stuffy nose, sneezing, headache, fever, sweats, muscle aches, feeling "run down," and loss of appetite.*

* 3. Was the follow-up of the study patients sufficiently long (for the outcome to occur in a large enough number of patients) and complete (so that no or very few patients were lost to follow-up)?


Patients enrolled
(n = 148)
Randomization
Echinacea group
(n = 73)
Placebo group
(n = 75)
Withdrew
(n = 4)
Completed trial
(n = 69)
Withdrew
(n = 2)
Completed trial
(n = 73)

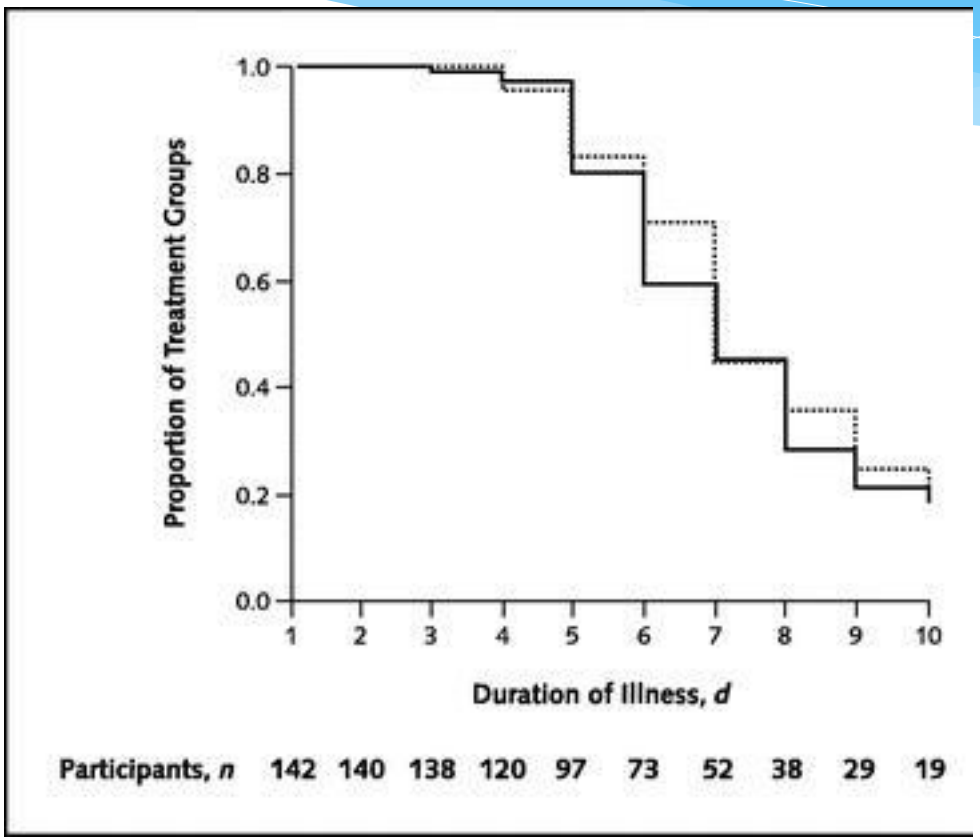
What about confounders?

* Five covariates were investigated as possible confounders: 1) duration of symptoms before study entry, 2) severity of illness at enrollment, 3) date of enrollment [seasonal or etiologic agent effect], 4) use of nonprotocol medications, and 5) sex.

* 4. Do the results of the harm study fulfill some of the diagnostic tests for causation? (Hill‘s criteria)

2. Are the valid results of this harm study important?

* After controlling for severity and duration of symptoms before study entry, sex, date of enrollment, and use of nonprotocol medications, researchers found no statistically significant treatment effect (adjusted hazard ratio, 1.24 [CI, 0.86 to 1.78]).
* Multivariable regression models assessing severity scores over time failed to detect statistically significant differences between the echinacea and placebo groups.


Proportion of Treatment Groups
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
4
5
6
7
8
9
10
Duration of Illness, d
Participants, n
142
140
138
120
97
73
52
38
29
19

Summary

* How do we appraise the validity of a harm evidence?
* What is a “hazard ratio” and how is it interpreted?
* Explain the significance of confounding factors.
* What is the significance of applicability of harm evidence in clinical practice?