Shall we count the living or the dead?

Anders Huitfeldt

Coauthors

Background

- Epidemiology (and biostatistics) is often understood as data analysis aimed at helping doctors and patients make better health care decisions
- The dirty secret:
  - The implications of our research depends on the choice of effect measure
  - If using the relative risk, the implications also depend on the coding of the outcome variable
  - We have no consensus about how to choose between effect measures

Background

Background

- The switch relative risk has been independently rediscovered multiple times, and is *a natural scale for making effect homogeneity assumptions*

Outline of presentation

(1) What is the switch relative risk?

(2) Is stability what matters for choice of effect measure?

(3) Is the switch relative risk more stable than other effect measures?

(4) A practical example of clinical implications

(5) The switch relative risk in toxicology and entomology

Part 1:

Basic Concepts

The Switch Relative Risk

- In 1958, Mindel C. Sheps suggested that we should use
  - The normal risk ratio (which is based on the probability of the outcome event, and "counts the dead") for exposures that reduce the risk of the outcome event,
  - The *survival ratio* (which is based on the *probability of the complement of the outcome event,* i.e. "counts the living"), for exposures that reduce the incidence of the outcome.

$$RR = \frac{p_1}{p_0} \qquad\qquad SR = \frac{1 - p_1}{1 - p_0}$$

The Switch Relative Risk

- The "switch relative risk" (van der Laan 2007) automatically selects the version of the relative risk recommended by Sheps

The Switch Relative Risk

- Baker and Jackson proposed a notationally convenient representation of the switch relative risk, as a number θ in the range [-1,1], defined as follows:

$$\theta = \begin{cases} 1 - \frac{1-p^1}{1-p^0} & \text{if } p^1 > p^0 \\ 0 & \text{if } p^1 = p^0 \\ -1 + \frac{p^1}{p^0} & \text{if } p^1 < p^0 \end{cases}$$
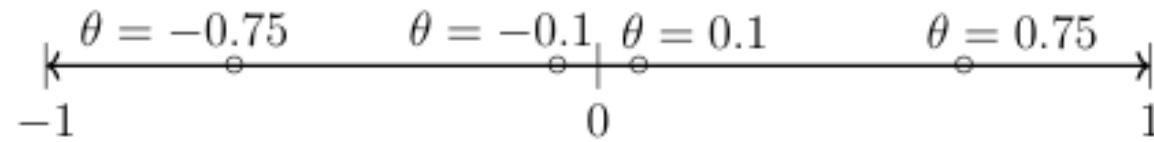
The Switch Relative Risk

Figure 3: Number line displaying the θ scale. $\theta = -0.75$ is a large protective effect corresponding to a risk ratio of 0.25. $\theta = -0.1$ is a moderate protective effect corresponding to a risk ratio of 0.9. $\theta = 0.1$ is a moderate harmful effect corresponding to a survival ratio of 0.9. $\theta = 0.75$ is a large harmful effect corresponding to a survival ratio of 0.25

The switch relative risk

- If we know the switch risk ratio and wish to combine it with an estimate of the baseline risk, in order to obtain the risk under treatment, this can be calculated using the effect function:

- $$g_\theta(p) = \begin{cases} 1 - (1 - p) \times (1 - \theta) & \text{if } \theta > 0 \\ p & \text{if } \theta = 0 \\ p \times (1 + \theta) & \text{if } \theta < 0 \end{cases}$$

Effect functions

- Sidenote: Every effect measure can be represented as an effect function g_Λ:

$$p_{v_i}^1 = g_\lambda(p_{v_i}^0)$$

This effect function governs many interesting features of the effect measures (collapsibility, closure)

Extension to time-to-event

- This mode of reasoning generalizes to time-to-event, if considering the *cumulative switch risk ratio* (at every time point/as a function of time)

Part 2:

How to think about the choice between effect measures

Choice between effect measures

- In order to set the stage, we must first discuss the considerations that guide this choice between effect measures

The standard procedure in evidence-based medicine

- In order to know whether the benefits of treatment outweigh the harms in a specific patient, a doctor needs to know the risk of the outcome that her patient will face if untreated (p_0) and the risk of the outcome he will face if treated (p_1)

- Ideally, we would like to tailor these risks to the specific patient's risk profile, and estimate the conditional risks $p^0_{v_i}$ and $p^1_{v_i}$ (where V is a set of baseline covariates that are relevant to how risk differs between patients)

- Unfortunately, it is usually not possible to power a study for estimating these quantities for every value of v.

The standard procedure in evidence-based medicine

- It is often assumed that the doctor has access to a reasonable estimate for the patient-specific baseline risk $p_{v_i}^0$, based on known diagnostic factors and other background information

- This can then combined with an estimate of the magnitude of the effect, in order to produce an estimate of $p_{v_i}^1$

The standard procedure in evidence-based medicine

- First, a measure of the magnitude of the effect is calculated in the study population:

$$RD = p_1 - p_0 \qquad\qquad RR = \frac{p_1}{p_0} \qquad\qquad OR = \frac{\dfrac{p_1}{1-p_1}}{\dfrac{p_0}{1-p_0}}$$

$$\theta = \begin{cases} 1 - \dfrac{1-p^1}{1-p^0} & \text{if } p^1 > p^0 \\[2mm] 0 & \text{if } p^1 = p^0 \\[2mm] -1 + \dfrac{p^1}{p^0} & \text{if } p^1 < p^0 \end{cases}$$

The standard procedure in evidence-based medicine

• Then, for any specific patient, we obtain an estimate of their risk of the outcome if untreated, using prognostic information,. This is combined with the magnitude of the effect in order to obtain an estimate of the risk of the outcome if treated:

$$p_{v_i}^1 = p_{v_i}^0 + RD \qquad\qquad\qquad p_{v_i}^1 = p_{v_i}^0 \times RR$$

$$p_{v_i}^1 = \frac{\dfrac{p_{v_i}^0}{1 - p_{v_i}^0} \times OR}{1 + \dfrac{p_{v_i}^0}{1 - p_{v_i}^0} \times OR} \qquad\qquad p_{v_i}^1 = \begin{cases} 1 - (1 - p) \times (1 - \theta) & \text{if } \theta > 0 \\ p & \text{if } \theta = 0 \\ p \times (1 + \theta) & \text{if } \theta < 0 \end{cases}$$

The standard procedure in evidence-based medicine

- This procedure will result in different predictions for the risk under treatment, depending on the choice of effect measure
- How do we know which one is correct?
  - Only if the chosen effect measure is *stable between patient groups* (homogeneity assumption) will we be justified in relying on the predictions

The standard procedure in evidence-based medicine

- The Cochrane Handbook explicitly tells us to use the risk ratio function:
  - ``The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI)''
- However, stability of the risk ratio is not a universally held belief, and has unclear theoretical support

The standard procedure in evidence-based medicine

- It is not possible to avoid relying on a homogeneity assumption, but it is possible to reason about the plausibility of each possible variant of the homogeneity assumption

Part 3:

A causal model for homogeneity of the switch relative risk

Switches

- Next, we provide a causal model for a *mechanism of action* that leads to stability of the switch relative risk.

- This mechanism is consistent with Patricia Cheng's Power-PC model for causal generative and preventive power, a model which has considerable support in the philosophy and psychology literature, where it has been argued that human reasoners use, and should use, these constructs to carry causal information from one setting to another

- The causal powers have been called "a brilliant piece of mathematical metaphysics" by Clark Glymour, a leading Philosopher of Science

Switches

- Suppose there is an attribute that determines, for every individual, whether they respond to treatment or not:
  - If the attribute is not present, treatment does not affect the outcome
  - If the attribute is present, treatment may change the outcome
- We refer to these attributes as "switches"

Switches

- We consider four types of switches:
  - If switch B is present, treatment is a sufficient cause of the outcome
  - If switch C is present, absence of treatment is a sufficient cause of the outcome
  - If switch D is present, treatment is sufficient cause of absence of outcome
  - If switch E is present, absence of treatment is sufficient cause of absence of outcome

Switches

- If there is more than one type of switch present, this can be generalized using sufficient-component cause models (causal pie models)

Switches

- Simple mathematics reveals that if treatment effect is determined by only one type of switch, then a characteristic effect measure will be determined by the prevalence of the switch
  - If switch B determines effect, SR=1-Pr(B) and treatment increases risk
  - If switch C determines effect, SR=1-Pr(C) and treatment decreases risk
  - If switch D determines effect, RR=1-Pr(D) and treatment decreases risk
  - If switch E determines effect, RR=1-Pr(E) and treatment increases risk

Switches

- Therefore, if treatment effect is determined by only one type of switch, the effect of treatment will be stable between groups that have the same distribution of switches

- If the prevalence of the switch differs between groups, this can be accounted for by conditioning on predictors of the prevalence of the switch as *effect modifiers*

Switches

- In reality, treatment effect is rarely determined by only one type of switch. This causes conceptual problems including monotonicity

- In those cases, there are significant advantages to identifying the *predominant switch*, using a model based on its characteristic effect measure, and considering heterogeneity in terms of deviation from that measure.

- We can then use sensitivity analysis and partial identification methods to bound the effect

If treatment increases risk

- If treatment increases risk, the effect of treatment must primarily be driven by switches of type B, or switches of type E
  - Do we have reason to believe that one type of switch is more prevalent?

If treatment increases risk

- If treatment increases risk, the effect of treatment must primarily be driven by switches of type B, or switches of type E
  - Do we have reason to believe that one type of switch is more prevalent?

  - YES

If treatment increases risk

- As an example, consider the risk of an anaphylactic response to penicillin treatment
  - A gene that functions as a switch of type B would say that the person who carries the gene gets an allergic reaction if they take Penicillin
  - A gene that functions as a switch of type E would say that the person who carries the gene does not get an allergic reaction if they don't take Penicillin

If treatment increases risk

- In other words, the gene that functions as a switch of type E would mean that the person is immune to any allergic reactions as long as they can avoid Penicillin
  - This is not consistent with any simple model of biology
  - We therefore conclude that the effect more likely depends on switches of type B, rather than switches of type E
  - Similar arguments can be made for most interventions that increase risk. Therefore, we adopt Sheps' conclusion that the effect of treatment will be more approximately stable on the survival ratio scale

If treatment increases risk

- This analysis assumes that the prevalence of the switch is independent of the baseline risk
  - This is a strong assumption that is often violated. For example, someone who has allergies to other things may be more likely to have a Penicillin allergy
  - This can be accounted for by controlling for predictors of general propensity to have allergic reactions, along with a sensitivity analysis

If treatment decreases risk

- As an example of a situation where treatment decreases risk, we will consider the effect of Penicillin at preventing rheumatic fever in patients with strep throat
  - Since treatment decreases risk, the effect is primarily driven either by switches of type C, or switches of type D

If treatment decreases risk

- Switches of Type D make treatment sufficient to prevent rheumatic fever. This could for example be a bacterial gene that determines antibiotic susceptibility

- Switches of Type C make the absence of treatment a sufficient cause of rheumatic fever. This means that the only possible way to avoid the outcome in a patient with such a switch, is to give Penicillin

If treatment decreases risk

- We would argue that switches of type C are much more prevalent than switches of type D

- Therefore, we again adopt Sheps' conclusion that the risk ratio is preferred for treatments that decrease risk

An evolutionary explanation

- If the evolutionary "baseline value" of an exposure was 0, switches of type B and D were inert in evolutionary history, whereas switches of type C and E were subjected to significant evolutionary pressure

- If the evolutionary baseline value of exposure was 1, switches of type C and E were inert, whereas switches of type B and D were subjected to significant evolutionary pressure
  - When a switch is subjected to evolutionary pressure, it quickly reaches fixation or elimination, and therefore becomes unviable as a factor in treatment effect variation

An evolutionary explanation

- For most treatments in pharmacoepidemiology, the "baseline" value of the exposure was 0

  - For example, very few people were exposed to Penicillin in our evolutionary past (perhaps excepting a small number of Shamans experimenting with fungi).

  - Whenever the evolutionary baseline value was 0, evolutionary arguments supports Sheps' argument

An evolutionary explanation

- Sometimes, epidemiologists consider exposures that have no evolutionary baseline value, for example sex
  - We all have ancestors who were subjected to evolutionary pressure as men, and ancestors who were subjected to evolutionary pressure as women
- In these situations, the analysis does not apply, and we again adopt Sheps' conclusion that in this situation, there is no principled reason to prefer one effect measure over another
  - This has important implications for research with active controls

An impossibility proof for the odds ratio

- We can prove mathematically that there can be no switch-based argument for stability of the odds ratio

  - This means that if we are guided by background beliefs about predictors of the covariates that turn the treatment effect "on" or "off", we will never conclude that the odds ratio (or any other non-collapsible effect measure) will be stable

Part 4:

A practical example

A practical example

- The Pfizer Covid vaccine has been estimated to be 96% effective at preventing cases of Covid-19
  - This effect size is presented on the correct scale according to Sheps' argument
  - Therefore, adopting her recommendations would not change anything about how the benefits of vaccination are measured

A practical example

- A nationwide observational study in Israel (Barda et al, NEJM 2021) has shown that the vaccine is associated with a small but possibly relevant elevated risk of Myocarditis. This was reported in terms of a risk ratio of 3.2.
  - Taking this result at face value, a clinician with a patient who has a baseline risk of Myocarditis of 1% would conclude that the patient has a 3.2% risk of myocarditis if given the vaccine.
  - Depending on the risk of infection if unvaccinated, and depending on the availability of other vaccines, this may well lead to a determination that the harms of vaccination outweigh the benefits for this particular patient.
- If the results from Barda et al had instead been presented in terms of the switch relative risk ($\theta = 0.000027$), as Sheps would have recommended, this would enable the clinician to conclude that the risk of myocarditis changes from 1% to approximately 1.0027% when the patient is vaccinated.

A practical example

- We would argue that this approach leads to a much more realistic estimate, consistent with a biologically interpretable hypothesis that approximately 0.0027% percent of the population carry some form of a "switch" that makes them susceptible to myocarditis if vaccinated.

- This hypothesis may not perfectly describe the underlying biology, it is for example possible that the presence of this switch is correlated with baseline risk of myocarditis, in which case a sensitivity analysis is needed to explore the potential consequences of such correlation.

- We maintain that even the upper bounds of this sensitivity analysis is unlikely to produce risk estimates as high as what one would obtain if the analysis relied on homogeneity of the risk ratio or the odds ratio. In our view, Sheps' approach provides a starting point for reasoning about what factors must be accounted for in the analysis, in order to meaningfully summarize the risk of adverse effect on a numerical scale.

A practical example

- In general, Sheps' recommendations are equivalent with the standard approach for measuring the benefits of medication
- It would however result in a change to how adverse effects are measured
  - Assuming Sheps' approach is correct, the following is true:
    - For patients whose risk profile is identical to the average in the study, the standard approach gives predictions about adverse affects that are correct on average
    - For patients whose risk profile is higher than the average in the study, the standard approach overestimates risks of adverse effects
    - For patients whose risk profile is lower than the average in the study, the standard approach underestimates risks of adverse effects

Part 5:

The switch relative risk in toxicology and entomology

A brief history of the switch relative risk

- In a very short paper from 1925, W.S. Abbott, an entomologist working at the US Department of Agriculture, proposed measuring the effect of insect sprays using what is now known as [Abbott's Formula](#).

- Abbott's formula, which is still used by entomologists today, is mathematically equivalent to Sheps' suggestion for the case where the intervention increases risk of the outcome.

- Abbott does not consider the situation where exposure reduces the risk of the outcome.

A brief history of the switch relative risk

- Consider a group of scientists testing the effect of an insecticide spray

- They first test the spray in a setting where 20% of mosquitoes die in the control condition, and observe that 60% die when exposed to the insecticide

  - What happened here? Did the spray triple the risk of death, or did it halve the probability of survival?

A brief history of the switch relative risk

- The scientists then proceed to test the same insecticide spray under conditions where 40% of the mosquitoes die under the control condition

- What do you expect to happen?

  - Will either effect parameter be stable between the two conditions?

A brief history of the switch relative risk

- In 1939, another entomologist, C.I. Bliss from the Institute for Plant Protection in Leningrad, extended Abbott's formula to the setting where exposure reduces incidence.

    - C.I. Bliss is also known as one of the creators of the probit model.

- To do so, he developed the [Joint Independent Action model](#), which has become central to how toxicologists think about interaction between poisons.

- Toxicologists have made attempts to convince epidemiologists about the utility of this framework for interaction, which according to [Howard and Webster has "firm biological foundations"](#) in contrast with epidemiological models that consider interaction in terms of departures from risk additivity.

A brief history of the switch relative risk

- Toxicologists have made multiple attempts to convince epidemiologists that the joint independent action model has advantages over the standard epidemiological approach
  - Clarice Weinberg has been writing about this since the 80s
  - Her suggestion is to use the log identity link in GLM models when exposure reduces risk, and the complimentary log function when exposure increases risk.
  - This leads to main effect estimates that are consistent with Sheps' suggestion

A brief history of the switch relative risk

- Along similar lines, Howard and Webster argued that the Joint Independent Action model has "firm biological foundations", in contrast with epidemiological models that consider interaction in terms of departures from risk additivity

Further reading

- "The choice of effect measure for binary outcomes: Introducing counterfactual outcome state transition parameters" Epidemiologic Methods 2016

- "Shall we count the living or the dead?" ArXiv 2021

- "Effect heterogeneity and variable selection for standardizing experimental findings to a target population" European Journal of Epidemiology, 2020

- "Regression by composition" forthcoming preprint