FLINDERS CLINICAL CHANGE

Filling critical gaps in our knowledge on chronic disease management, and later and end-of-life health care

Consistently estimating risk difference in a jurisdiction of interest

By Simon Eckermann, Michael Coory, Andrew Willan

Flinders Centre for Clinical Change and Health Care Research

Consistently estimating risk difference in a jurisdiction of interest

Simon Eckermann, PhD*, *Flinders University*

Michael Coory, PhD, *University of Queensland*

Andrew R. Willan, PhD, *SickKids Research Institute and*

University of Toronto

Running Title: Consistently estimating risk difference

Keywords: absolute risk difference; relative risk; odds ratios; direct comparison; indirect comparison

Word Count: 3,408

*Simon Eckermann, Professor in Health Economics

Flinders University Centre for Clinical Change and Health Care Research

Room 52, 1st Floor, A Block, Repatriation General Hospital, Daws Road, Daw Park, South Australia, 5041

E-mail: simon.eckermann@flinders.edu.au;

Telephone: +618 8275 2858

Facsimile: +618 8275 2854

Consistently estimating risk difference in a jurisdiction of interest

Abstract   (Word count 245)

Objective: Economic analysis and assessment of net clinical benefit often require estimation of absolute risk difference (ARD) for binary outcomes such as survival or progression, given baseline epidemiological risk in a jurisdiction of interest and trial evidence of treatment effects. Typically, the assumption is made that relative treatment effects are constant across baseline risk, in which case relative risk (RR) or odds ratios (OR) could be applied to estimate ARD. Our objective is to establish whether use of RR or OR allows consistent estimates of ARD.

Methods: ARD is calculated from alternate framing of effects (e.g., mortality vs. survival) using relative risk (RR) and odds ratios (OR) in direct and indirect comparisons.

Results: ARD is shown to be consistently estimated with OR but changes with framing of effects using RR wherever epidemiological risk differs from trial risk. Additionally, in indirect comparisons ARD is shown to be consistently estimated with OR, while the direction let alone extent of ARD is shown to not necessarily be consistent with RR where risk in common comparator arms differ.

Conclusion: Odds ratios allow consistent calculation of absolute risk difference in translating evidence from trial settings and across trials in direct and indirect comparisons, avoiding selection biases from framing of effects with relative risk. These findings are critical for translating evidence to inform economic analysis and assessment of net clinical benefit, given that use of RR in estimating ARD is commonly suggested precisely where risk differs in practice or across arms.

Introduction

Summary statistics derived from binary measures of treatment outcome can be divided into two categories. The first category comprises arithmetic measures such as absolute risk difference (ARD) and, its reciprocal, the number-needed-to-treat (NNT); the second category comprises relative (ratio) measures such as relative risk (RR) and odds ratio (OR). Economic analysis and assessment of net clinical benefit typically require estimates of ARD; however, ARD is seldom taken directly from trials due to concerns it might vary according to patients' baseline risk of the outcome [1-3]. Instead, the convention is to adopt a two-step procedure. First, the baseline risk for the particular circumstances of the economic analysis is estimated. Then, this risk is multiplied by an appropriate algebraic form of the average relative treatment effect (e.g., RR, OR) from a trial (or meta-analysis of trials) to estimate an arithmetic measure of treatment effect (e.g., ARD, NNT) for the particular circumstances of economic analysis or assessment of net clinical benefit.

There is continuing debate about whether it is better to use OR or RR as the measure of relative treatment effect. Several clinical epidemiologists have stated their preference for RR on the basis that clinicians may be inclined to think on the RR scale and may mistakenly interpret OR as RR. If OR is mistaken as RR, then the treatment effect can be overestimated as OR tends to be further away from the null value (1.0) where baseline risk is not small, as is usually the case in clinical trials [4-6].

However, there are problems associated with using RR in modeling binary data. The core issue is that, unlike OR, RR is not symmetric [7]. Specifically, if the event of interest is switched to its complement (e.g., if *survival* is used as the outcome instead

of *mortality*), then RR is not symmetrical around 1, in contrast to OR and ARD, which are symmetrical around 1 and 0 respectively. Walter describes this property of the RR as 'very troubling' [8] and Fleiss argued that this property of RR effectively rules it out as a useful metric to use in meta-analysis [9].

For indirect comparisons (i.e., comparison of two drugs to each other using separate trials against a common comparator), RR has a more severe problem. Not only can the size of the treatment effect change, but an inferential fallacy can occur where the direction of the treatment effect can reverse depending on whether effects were framed from a utility bearing perspective (e.g., survival) or a disutility bearing perspective (e.g., mortality) [10]. This paper extends the findings of previous research to show that, for both direct (i.e. A vs. B) and indirect comparisons (i.e. A vs. B via common comparator C), use of OR has advantages over RR in consistently (i.e. free of framing effects) estimating ARD, such as that required for use in economic analysis.

We first consider standard methods for estimating ARD in a jurisdiction of interest for direct and then indirect comparisons calculated with relative risk and odds ratios. We illustrate, by algebraic proof and with 'real world' examples, that OR, unlike RR, allow consistent estimation of ARD in both direct and indirect comparisons. That is, the size and direction of risk differences are the same when estimated using odds ratios, regardless of whether risks are framed from a positive (e.g., survival) or negative (e.g., mortality) direction. We conclude by critically discussing these findings in the context of previously identified properties of RR and OR in estimating treatment risk and modeling ARD.

Estimating absolute risk difference to inform economic analysis

The estimation of ARD from relative treatment effects in general requires that the risk of the outcome for the therapy being compared (B in either A vs. B with direct, or A vs. B via C with indirect, comparison) has been estimated in the jurisdiction of interest. This estimate should reflect the epidemiological expected risk of the outcome with B and, hence, in many cases, will not necessarily be that observed in trial evidence [11]. For example, consider where the jurisdiction of interest is the UK, Australia or Canada and evidence comes from a direct trial in the US of A vs. B. The expected risk with B in the jurisdiction of interest could differ from that observed in the US trial due to factors including population risk, health care access or health system practice. Differences in health system practice between that in trials and the jurisdiction of interest may be driven by change in diagnostic build up or different use of complementary therapies across time and between jurisdictions, which in turn may be related to differences in factors including relative prices, patient preferences or clinical training and practice. Hence, even where current practice for use of technology B in the jurisdiction of interest reflects that in trial settings, the best estimate of the risk of B will usually be based on current epidemiological evidence in the jurisdiction of interest rather than necessarily trial data.

The problem – inconsistent estimation of risk difference with relative risk

For the jurisdiction of interest, let the estimate of risk framed from a utility bearing perspective (e.g., survival rate) be given by $R_{UB}*$ and from a disutility perspective by $R_{DB}* = 1 - R_{UB}*$. In a direct comparison, relative risk from a trial setting

4

$RR_{UAB} = \dfrac{R_{UA}}{R_{UB}}$ can be directly applied to $R_{UB}*$ to estimate $R_{UA}*$ and, hence, the ARD

as difference in effects with

$$\Delta E_u = R_{UB}*(\frac{R_{UA}}{R_{UB}}) - R_{UB}* = (R_{UB}*)(\frac{R_{UA}}{R_{UB}} - 1) \qquad (1)$$

under the assumption that RR is constant regardless of baseline risk.

Alternatively, from a disutility bearing perspective relative risk $RR_{DAB} = \dfrac{(1 - R_{UA})}{(1 - R_{UB})}$ can

be directly applied to $R_{DB}* = 1 - R_{UB}*$ to estimate $R_{DA}*$ (e.g., mortality rate) and,

hence, difference in effects with

$$\Delta E_D = (1 - R_{UB}*) - (1 - R_{UB}*)[\frac{(1 - R_{UA})}{(1 - R_{UB})}] = (1 - R_{UB}*)(1 - \frac{(1 - R_{UA})}{(1 - R_{UB})}) \qquad (2)$$

The difference between these two estimates of risk difference, $\Delta E_U - \Delta E_D$, calculated

with RR from a utility (1) and disutility (2) bearing perspective, simplifies, as shown

in Appendix 1, to

$$\Delta E_U - \Delta E_D = \frac{(R_{UA} - R_{UB})}{(1 - R_{UB})}(\frac{R_{UB}*}{R_{UB}} - 1) \qquad (3)$$

Now the ARD will be the same with alternative framing only if the first or second

term in equation (3) is 0. The first term in equation (3) is 0 only if there is no

treatment effect, while the second term represents the proportion by which

epidemiological risk in the jurisdiction of interest differs from that observed in trial

evidence. Where the base risk in the jurisdiction of interest exactly matches that from

direct trials (i.e., where $R_{UB}* = R_{UB}$), then $\Delta E_U - \Delta E_D = 0$, with

$\Delta E_U = \Delta E_D = R_{UA} - R_{UB}$.

However, where the risk in the jurisdiction of interest is not the same as the trial evidence, that is $r = R_{UB}^* / R_{UB} \neq 1$, and there is any treatment effect, ARD calculated with RR will depend on whether effects have been framed from a utility or disutility bearing perspective. Hence, estimation of the extent of ARD using RR differs with framing of effects where the epidemiological (base) risk differs from that observed in the trial setting and there is any treatment effect.

In the usual case of interest where there is an RR in favour of A, equation (3) indicates that ARD calculated from a utility bearing perspective (ΔE_U) will be greater (less) than that from a disutility perspective (ΔE_D) where the epidemiological risk in the jurisdiction of interest is greater (less) than that in the trial. Conversely, if there is an RR in favour of B, ARD calculated from a utility bearing perspective (ΔE_U) will be greater (less) than that from a disutility perspective (ΔE_D) where the epidemiological risk in the jurisdiction of interest is less (greater) than that in the trial. Consequently, where ARD is calculated based on RR, there is clearly scope for selection bias in framing of effects whenever base risk for treatment B differs in the jurisdiction of interest to that in the trial setting.

To illustrate the potential for selection bias when estimating ARD with RR, take the simple example of a direct comparison between A and B where risk of mortality in the trial setting is 0.2 with A and 0.4 with B, while in two separate jurisdictions of interest (X and Y) the epidemiologically determined risk of mortality with current treatment B is 0.4 (the same as the trial setting) and 0.3, respectively. Table 1 shows the calculation of ARD in these two jurisdictions, given trial evidence and their

epidemiological risk. For jurisdiction X, which has the same risk in arm B as that of the trial, risk difference is estimated as 0.2 in favour of A, regardless of whether effects are framed as mortality or survival. However, for jurisdiction Y, where epidemiological risk differs from that in the trial setting, risk difference in favour of A is 0.15 with mortality, while 0.233 with survival. Having established that ARD is not consistently estimated using RR with alternative framing of binary outcomes, we now consider whether the OR provides a consistent estimate.

Consistent estimation of absolute risk difference with odds ratios

To estimate ARD using OR with effects framed from a utility bearing perspective (e.g., survival) involves:

(1)   that the baseline risk in a jurisdiction of interest $R_{UB}*$ is converted to an odds ratio with $O_{UB}* = R_{UB}*/(1 - R_{UB}*)$

(2)   the odds ratio from trial evidence O_{UAB} is then applied to $O_{UB}*$ to estimate O_{UA} with $O_{UA} = O_{UAB}O_{UB}* = O_{UAB}R_{UB}*/(1 - R_{UB}*)$

(3)   the odds ratio O_{UA} is converted to a risk R_{UA} using

$$R_{UA} = O_{UA}/(1 + O_{UA}) = (O_{UAB} \times \frac{R_{UB}*}{1 - R_{UB}*})/(1 + \frac{O_{UAB}R_{UB}*}{1 - R_{UB}*})),$$

$$R_{UA} = \frac{O_{UAB} \times R_{UB}*}{(1 - R_{UB}* + O_{UAB}R_{UB}*)}$$

(4)   risk difference is then calculated as

$$\Delta E_U = R_{UA} - R_{UB}* = \frac{O_{UAB} \times R_{UB}*}{(1 - R_{UB}* + O_{UAB}R_{UB}*)} - R_{UB}*$$, which can

alternatively be expressed as

$$\Delta E_U = R_{UB}{}^* \left(\frac{O_{UAB}}{(1 - R_{UB}{}^* + O_{UAB}R_{UB}{}^*)} - 1 \right)$$

$$= R_{UB} \left(\frac{O_{UAB} - 1 + R_{UB}{}^* - O_{UAB}R_{UB}{}^*}{(1 - R_{UB}{}^* + O_{UAB}R_{UB}{}^*)} \right)$$

$$= R_{UB}{}^* (1 - R_{UB}{}^*)(O_{UAB} - 1)/(1 - R_{UB}{}^* + O_{UAB}R_{UB}{}^*)$$

Now we show that the same ARD is estimated with use of odds ratios where comparison is undertaken with outcomes framed from a disutility bearing perspective.

(1) Recalling that base risk, framed from a disutility perspective (e.g., mortality rate), is given by $R_{DB}{}^* = 1 - R_{UB}{}^*$, $R_{DB}{}^*$ is converted to an odds ratio with $O_{DB}{}^* = R_{DB}{}^* /(1 - R_{DB}{}^*) = (1 - R_{UB}{}^*)/R_{UB}{}^*$

(2) The odds ratio O_{DAB} is applied to O_{DB} to estimate O_{DA} with

$$O_{DA} = O_{DAB} \times O_{DB}{}^* = \frac{1}{O_{UAB}} \times \frac{(1 - R_{UB}{}^*)}{R_{UB}{}^*}$$

(3) the odds ratio O_{DA} is then converted to a risk R_{DA} using

$$R_{DA} = \frac{O_{DA}}{1 + O_{DA}} = \frac{(1 - R_{UB}{}^*)}{O_{UAB}R_{UB}{}^*} \Big/ \left(1 + \frac{(1 - R_{UB}{}^*)}{O_{UAB}R_{UB}{}^*} \right)$$

$$= \frac{(1 - R_{UB}{}^*)}{O_{UAB}R_{UB}{}^* + (1 - R_{UB}{}^*)}$$

(4) risk difference is calculated as

$$\Delta E_D = R_{DB} - R_{DA} = 1 - R_{UB}{}^* - \frac{(1 - R_{UB}{}^*)}{(1 - R_{UB}{}^* + O_{UAB}R_{UB}{}^*)}$$

$$= (1 - R_{UB}{}^*)\left(1 - \frac{1}{(1 - R_{UB}{}^* + O_{UAB}R_{UB}{}^*)} \right)$$

$$= (1 - R_{UB}{}^*)\left(\frac{O_{UAB}R_{UB}{}^* - R_{UB}{}^*}{(1 - R_{UB}{}^* + O_{UAB}R_{UB}{}^*)} \right)$$

$$= R_{UB}{}^* (1 - R_{UB}{}^*)(O_{UAB} - 1)/(1 - R_{UB}{}^* + O_{UAB}R_{UB}{}^*)$$

This is the same equation for ARD as that estimated from a utility bearing perspective. Hence, use of OR ensures ARD is consistently estimated with alternative framing of effects for any given baseline risk in the jurisdiction of interest. Consequently, use of OR prevents the selection bias with respect to framing of effects that occurs with RR in translating evidence to a jurisdiction of interest. Table 2 takes the same example used in Table 1 to illustrate consistency in estimating ARD with OR regardless of baseline risk in the jurisdiction of interest. For jurisdiction Y there is an ARD of 21/130=0.16 in favour of A with mortality or survival.

Additional problems estimating ARD with RR in indirect comparisons

We have shown that in direct comparisons the extent of ARD for binary outcomes is not consistently estimated with RR where baseline risk in the jurisdiction of interest differs from that in trials. In indirect comparisons, RR faces a further problem in interpreting trial evidence before considering translation of such evidence to a jurisdiction of interest. Eckermann, Coory and Willan [10] demonstrate that, in indirect comparisons RR does not guarantee a consistent direction of treatment effect where base risk differs across common comparator arm, while OR does. For example, in indirect comparison of Natalizumab vs. Interferon via placebo for multiple sclerosis [12-14], Natalizumab was suggested to be 30% more effective than Interferon for progression (RR=0.70) but 16% less effective than Interferon for no progression (RR=0.84) [10]. The OR was 0.83 for progression and its reciprocal (1/0.83=1.21) for no progression in favour of Natalizumab in each case. Here, we extend this result to show that ARD calculated with RR also differs in direction with alternate framing of outcomes, while ARD is consistently estimated with OR.

Standard use of RR to estimate ARD in a jurisdiction with, for example, a 50% base risk of progression with Interferon, suggests a 15.0% advantage for progression in favour of Natalizumab ($0.50 \times 0.30 = 0.15$), but an 8% advantage in favour of Interferon with no progression (0.50×0.16), as shown in Table 3. Such reversal of the direction of ARD with alternate framing of effects using RR will occur in indirect comparisons whenever the direction of treatment effect reverses with RR. Eckermann, Coory and Willan [10] show that such reversals do not necessarily require large differences in base risk to occur. More generally, even if differences in common comparator risk in indirect comparison are not large enough to cause direction of treatment effect to be reversed, the extent of ARD will be influenced if there is any difference in common comparator risk across arms. This is especially problematic given that using RR in indirect comparisons has been suggested as valuable where base risks differ across arms [11], which is precisely where problems with inconsistency occur. The positive and more substantive question of interest to which we now turn is whether OR provides a consistent estimate of ARD for indirect as well as direct comparisons.

ARD is estimated with OR under the same conditions as RR in table 3 as 4.6% in favour of Natalizumab with progression or no progression. That is, given

$R_{DB} = 0.5$ and $O_{UAB} = 1.21$ or $O_{DAB} = 0.83$

$\qquad$ (1) $\qquad O_{UB} = 0.5/0.5 = 1$, or $O_{DB} = 0.5/0.5 = 1$

$\qquad$ (2) $\qquad O_{UA} = O_{UB} \times O_{UAB} = 1 \times 1.21 = 1.21$, or $O_{DA} = O_{DB} \times O_{DAB} = 1 \times 0.83 = 0.83$

$\qquad$ (3) $\qquad R_{UA} = \dfrac{1.21}{1+1.21} = 0.546$, or $R_{DA} = \dfrac{0.83}{1+0.83} = 0.454$

$\qquad$ (4) $\qquad \Delta E = 0.546 - 0.50 = 0.5 - 0.454 = 0.046$

More generally, the finding that OR consistently estimate ARD for binary outcomes in direct comparisons will extend to indirect comparisons given treatment effect is consistently estimated with OR, following Eckermann, Coory and Willan [10]. Hence, in indirect comparisons, advantages of OR over RR in consistently estimating treatment effect [10] extends to consistently estimating ARD.

Consequently, OR has distinct advantages over RR in consistently estimating ARD in direct and indirect comparisons. The only exceptions to this are:

(1) in direct comparisons where base risk in the jurisdiction of interest is the same as that of the trial; and

(2) in indirect comparisons where base risk in the jurisdiction of interest is the same as that of the trial and, additionally, there is no difference in risk across common comparator arms.

However, in neither of these special cases is RR useful, as ARD can be consistently estimated from trial evidence of risk differences directly. Consequently, use of RR in estimating ARD is inconsistent and creates scope for selection biases in framing of effects in all cases where applying a treatment effect from a trial to a base risk in a jurisdiction of interest may be valuable. These inconsistencies are overcome with OR, which allow consistent estimation of ARD regardless of epidemiological risk in translating evidence or risk differences across common comparator arms in indirect comparisons.

Further advantages of odds ratios in modeling binary outcomes
Odds ratios have further advantages over RR and direct use ARD when modeling binary events, such as mortality and progression, in ensuring treatment risk remains

bounded between 0 and 1 when applying trial data to base risk in a jurisdiction of interest. For example, consider a case where the event rate from a positive (utility bearing) perspective is progression free survival (PFS) at two years. The appropriate epidemiological estimate in the jurisdiction of interest (e.g., the UK, Canada or Australia) is 90%, while evidence (predominantly from the US in an, on average, more complex population for the same indication) suggests a relative risk for PFS of 1.2 (based on an increase in PFS from 70% to 84%). The estimate of PFS with treatment in the jurisdiction of interest in this case will be 95.3% using odds ratios, while clearly out of appropriate bounds at 108% (0.9×1.2) using relative risk or 104% (0.90+0.14) using ARD. Greenland suggests these range limitations on RR and ARD indicate "purely logical reasons for disbelieving constancy of the difference or relatives of proportions."[15]

The advantage of odds ratios in such situations arises as odds will always convert back to a risk bounded between 0 and 1. Hence, an OR from the trial (2.25 =63/28=(0.84/0.16)/(0.7/0.3)) multiplied by the appropriate epidemiological base odds (9 =0.9/0.1) leads to odds of 20.25, which converts back to a risk of 95.3% (=20.25/21.25). This advantage of OR over RR in appropriately bounding treatment risk for binary outcomes between 0 and 1 also arises in modeling risk under uncertainty, where unbounded values can always arise by chance using RR with methods such as Monte-Carlo simulation. The bounding of treatment risk between 0 and 1 with OR more generally enables feasible constancy of odds across the full range of potential baseline risks, unlike ARD and RR. There is, therefore, no logical reason for disbelieving constancy of OR with binary outcomes.

Further, it is also noted that odds ratios are naturally modeled as lognormal distributions[16], while modeling relative risk as a lognormal distribution becomes increasingly problematic as risk increases [17]. This can be seen in the divergence of RR from OR, when considering application of treatment effect to baseline risk in direct comparisons where $RR = \dfrac{OR}{1 - p_0 + p_0 \times OR}$. Hence, for small probabilities of binary events ($p_0 \approx 0$), relative risk approximates odds ratios and a lognormal distribution, but the relationship increasingly diverges from the OR and a lognormal distribution as base risk increases and a treatment effect is evident (OR diverges from 1).

In summary, OR have distinct advantages over RR with binary outcomes in allowing consistent estimation of ARD, ensures treatment risk remains bounded between 0 and 1 and naturally fit the lognormal distribution. Consequently, OR are suggested as the preferred metric for estimating or modeling net clinical benefit and economic modeling for health technology assessment.

Conclusion

Risk differences estimated using relative risk have been shown to differ with positive and negative framing of binary outcomes (e.g., mortality or survival) in direct and indirect comparison whenever the epidemiological risk to be modified differs from that in the trial setting. In indirect comparisons, the direction as well as the size of the risk difference can change where risks differ between common comparator arms. Use of symmetric odds ratios has been shown to overcome these problems by consistently

estimating absolute risk difference in direct or indirect comparisons with alternate framing of binary outcomes, regardless of baseline risk.

These findings point to distinct advantages of OR over RR whenever application of a relative treatment effect to baseline risk has value. It is precisely where the baseline risk differs from that in trials when translating evidence across jurisdictions or interpreting evidence in indirect comparisons that use of relative treatment effects is suggested. Hence, in translating evidence with an assumption of constant treatment effect in direct or indirect comparisons the OR should be used to allow consistent estimation of ARD for binary outcomes.

Further, as previously shown, risks of binary events calculated with odds ratios remain bounded between 0 and 1, while relative risk from a trial applied to the appropriate epidemiological risk can easily result in values above 1. This can be particularly problematic in modeling risk of binary outcomes under uncertainty, where, more generally, OR naturally fit a lognormal distribution from which RR increasingly diverges as probabilities and treatment effects increase. Consequently, odds ratios are shown to be the preferred metric for calculating or modeling risk differences in informing process of health technology assessment and decision making in direct or indirect comparisons.

The cumulative impact of these findings casts significant doubt on the 'common wisdom' that relative risk is more interpretable by clinicians than odds ratios in interpreting and translating evidence. Clinicians would be unlikely to admit to such a 'common wisdom' if they were made aware that estimating risk differences for

assessment of net clinical benefit or economic analysis with RR, rather than OR, is inconsistent with alternative framing of effects, does not appropriately bound treatment risk between 0 and 1 and increasingly diverges from the lognormal distribution as base risk and treatment effect increase. The last vestige of credibility for interpretability of RR in translating evidence to estimate ARD falls in indirect comparisons. For cases of interest with differences across arms in common comparator risk, the direction of ARD calculated with RR can be inconsistent and the extent will be inconsistent with alternative framing of effects. In contrast, when calculated with OR, the direction and extent of ARD remain consistent with alternative framing of effects, regardless of differences in baseline risk between trials and jurisdiction and across common comparator arms in indirect comparisons.

These findings also have implications for consistently and appropriately standardising effects in practice, which can be framed from a positive (utility bearing) perspective or negative (disutility bearing) perspective. For example, in application of the net benefit correspondence theorem [18, 19], use of odds ratios, unlike relative risk, would enable consistent differences between providers to be estimated with standardised event rates. That is, analogous to the findings presented here, risk differences for each provider, relative to a standard (e.g., industry average) population risk, will be the same (and appropriately bounded) for mortality and survival rates when calculated by applying odds ratios to the standard populations odds and converting back to risk.

References:

1.   Engels, E., et al., *Heterogeneity and statistical significance in meta-analysis: an empirical study of 125 meta-analyses.* Statistics in Medicine, 2000. **19**(13): p. 1707-1728.

2.   Deeks, J., *Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes.* Statistics in Medicine., 2002. **21**: p. 1575-1600.

3.   Furukawa, T., G. Guyatt, and L. Griffith, *Can we individualize the 'number needed to treat'? An empirical study of summary effect measures in meta-analyses.* Int J Epidemiol, 2002. **31**: p. 72-76.

4.   Sinclair, J. and M. Bracken, *Clinically useful measures of effect in binary analyses of randomised trials.* J Clin Epidemiol, 1994. **47**: p. 881-889.

5.   Davies, H., I. Crombie, and M. Travokoli, *When can odds ratios mislead?* BMJ: British Medical Journal, 1998. **316**: p. 989-991.

6.   Sackett, D., J. Deeks, and A. DG, *Down with odds ratios!* Evidence-Based Medicine, 1996. **1**: p. 164-167.

7.   Sheps, M., *Shall we count the living or the dead?* New Engl J Med, 1958. **259**: p. 1210-1214.

8.   Walter, S., *Choice of effect measure for epidemiological data.* J Clin
        Epidemiol, 2000. **53**: p. 931-939.

9.   Fleiss, J., *Measures of effect size for categorical data*, in *The Handbook of
        Research Synthesis*, H. Cooper and L. Hedges, Editors. 1994, Russell Sage:
        New York.

10.   Eckermann, S., M. Coory, and A.R. Willan, *Indirect comparison: relative risk
        fallacies and odds solution.* Journal of Clinical Epidemiology, 2008. **in
        submission**.

11.   Australian Government Department of Health & Ageing. *Guidelines for
        preparing submissions to the Pharmaceutical Benefits Advisory Committee
        (PBAC).* 2006 [cited 19 December 2008]; Available from:
        http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pbs-
        general-pubs-pharmpac-gusubpac.htm.

12.   Australian Government Department of Health & Ageing. *Public Summary
        Documents by Product: Natalizumab, concentrated solution for IV infusion,
        300 mg per 15 mL, Tysabri® November 2006.* 2007 [cited 14 October 2008];
        Available from:
        http://www.health.gov.au/internet/main/publishing.nsf/Content/pbac-psd-
        natalizumab-nov06.

13.   Polman, C.H., et al., *A Randomized, Placebo-Controlled Trial of Natalizumab for Relapsing Multiple Sclerosis.* The New England Journal of Medicine, 2006. **354**(9): p. 899-910.

14.   Duquette, P. and e. al, *Interferon beta-1b in the treatment of multiple sclerosis: Final outcome of the randomized controlled trial.* Neurology, 1995. **45**(7): p. 1277-1285.

15.   Greenland, S., *Interpretation and choice of effect measures in epidemiologic analyses.* Am. J. Epidemiol., 1987. **125**(5): p. 761-768.

16.   Altman, D.G., *Practical statistics for medical research.* 1991, London: CRC Press. 611.

17.   Localio, A.R., D.J. Margolis, and J.A. Berlin, *Relative risks and confidence intervals were easily computed indirectly from multivariable logistic regression.* Journal of clinical epidemiology, 2007. **60**(9): p. 874-882.

18.   Eckermann, S., *Hospital Performance Including Quality: Creating Economic Incentives Consistent with Evidence-Based Medicine. PhD Dissertation.* 2004, University of New South Wales, available at http://www.library.unsw.edu.au/~thesis/adt-NUN/public/adt-NUN20051018.135506/, accessed 29 May 2007: Sydney.

19.   Eckermann, S., A. Briggs, and A.R. Willan, *Health Technology Assessment in the Cost-Disutility Plane.* Medical Decision Making, 2008. **28**(2): p. 172-181.

Table 1: Inconsistency of ARD with alternative framing of effects when calculated with RR and baseline risk differs from trial evidence (Jurisdiction Y)

A mortality	B mortality	A survival	B survival
20%	40%	80%	60%
RR mortality A vs. B in trial		RR survival A vs. B in trial	
0.2/0.4=0.50		8/6=1.33	
Jurisdiction X: BR mortality 40%		Jurisdiction X: BR survival 60%	
ARD = 0.4×0.5 = 0.20 in favour A		ARD = 0.6×0.33 = 0.20 in favour A	
Jurisdiction Y: BR mortality 30%		Jurisdiction Y: BR survival 70%	
ARD = 0.3×0.5 = 0.15 in favour A		ARD = 0.7×0.33 = 0.233 in favour A	

Table 2: Consistency of ARD with alternative framing of effects when calculated with

OR and baseline risk (BR) in jurisdictions of interest (X, Y)

A mortality	B mortality	A survival	B survival
20%	40%	80%	60%
OR= (0.2/0.8)/(0.4/0.6)=3/8		(0.8/0.2)/(0.4/0.6)=8/3	
Jurisdiction X: BR mortality 40%		Jurisdiction X: BR survival 60%	
B Odds mortality= 0.4/0.6=2/3		B Odds survival = 0.6/0.4= 3/2	
A Odds mortality= 2/3×3/8=1/4		A Odds survival = 3/2×8/3=4	
A Risk mortality = 1/5		A Risk survival = 4/5	
ARD mortality = 4/10-1/5 =1/5 in favour of A		ARD survival =4/5-6/10 =1/5 in favour of A	
Jurisdiction Y: prob. mortality 30%		Jurisdiction Y: prob. survival 70%	
B Odds mortality= 0.3/0.7=3/7		B Odds survival = 0.7/0.3= 7/3	
A Odds mortality= 3/8×3/7=9/56		A Odds survival = 8/3×7/3=56/9	
A Risk mortality = 9/65		A Risk survival = 56/65	
ARD mortality =3/10-9/65 =21/130 = 0.16 in favour A		ARD survival =56/65-7/10 =21/130 = 0.16 in favour A	

Table 3: Inconsistency of ARD with alternative framing of effects calculated with RR in indirect comparison of Natalizumab and Interferon

Natalizumab Progression	Placebo Progression	Natalizumab No progression	Placebo No progression
33.3%	59.0%	66.7%	41.0%
Natalizumab vs. placebo Progression RR = 0.57		Natalizumab vs. placebo No progression RR = 1.63	
Interferon Progression	Placebo Progression	Interferon No progression	Placebo No progression
68.7%	83.9%	31.3%	16.1%
Interferon vs. placebo [6] Progression RR = 0.82		Interferon vs. placebo No progression RR = 1.95	
RR for progression Nat. vs. Int. 0.70 (favours Natalizumab)		RR for no progression Nat vs. Int 0.84 (favours Interferon)	
BR progression 50%		BR no progression 50%	
ARD =0.5×(1-0.7) = 0.15 in favour of Natalizumab		ARD = 0.5×(1-0.84) = 0.08 in favour of Interferon	

Appendix 1: Selection biases in risk difference under alternative conditions

From Section 1, consider risk differences (ARD, ΔE) estimated with relative risk (RR) in direct comparisons with effects framed from a utility bearing perspective as

$$\Delta E_u = (R_{UB}*)(\frac{R_{UA}}{R_{UB}} - 1) \tag{1}$$

and with effects framed from a disutility bearing perspective as

$$\Delta E_D = (1 - R_{UB}*)(1 - \frac{(1 - R_{UA})}{(1 - R_{UB})}) . \tag{2}$$

Where $R_{UB}*$ is the epidemiological risk in the jurisdiction of interest for treatment B with effects framed from a utility bearing perspective, R_{UB} and R_{UA} are the same risk for treatment B and treatment A from trial evidence. Hence,

$$\begin{aligned}
\Delta E_U - \Delta E_D &= R_{UB}*(\frac{R_{UA}}{R_{UB}} - 1) - (1 - R_{UB}*)(1 - \frac{(1 - R_{UA})}{(1 - R_{UB})}) \\
&= \frac{R_{UB}*}{R_{UB}}(R_{UA} - R_{UB}) - \frac{1 - R_{UB}*}{(1 - R_{UB})}(R_{UA} - R_{UB}) \\
&= (R_{UA} - R_{UB})(\frac{R_{UB}*}{R_{UB}} - \frac{1 - R_{UB}*}{(1 - R_{UB})}) \\
&= \frac{(R_{UA} - R_{UB})}{(1 - R_{UB})}(\frac{R_{UB}*}{R_{UB}} - 1)
\end{aligned} \tag{3}$$