Title

Inconsistent results in meta-analyses for the prevention of falls are found between study-level data and patient-level data

Terry P Haines[1,2] ,Anne-Marie Hill[3,4]

1 – Director of Research, Southern Clinical School, Monash University, Physiotherapy Department, Frankston, Australia..

2 – Director of Allied Health Clinical Research Unit, Southern Health, Continuing Care, Cheltenham, Australia

3 – PhD candidate, The University of Queensland, Physiotherapy Department, St Lucia, Australia.

4 – Senior Lecturer, University of Notre Dame Australia, School of Health Sciences, Physiotherapy, Fremantle, Australia

Abstract

Introduction: Meta-analysis of study-level data is now a common tool for gathering insights within a research field. "Recurrent events" characterises the data for a range of outcomes that may be encountered in clinical practice. Methods to conduct meta-analyses using study-level data from investigations where recurrent events are of interest are yet to be specifically developed. This study seeks to examine whether existing study-level data meta-analysis approaches can be used to produce unbiased and precise effect estimates relative to meta-analyses conducted using patient level data where a recurrent event is the outcome of interest.

Method: Data from two studies focussing on the prevention of falls in the hospital setting (n=1838 total) was divided into the three hospital sites from which they were collected. Outcome data was considered as recurrent event survival data, single event survival data, count data, rate data and binary data. Analysis approaches including Andersen-Gill recurrent events survival analysis, Cox single event survival analysis, negative binomial regression, bootstrap resampling, relative risk and modified relative risk, and logistic regression were investigated.

Results: Andersen-Gill, negative binomial, bootstrap resampling and modified relative risk analysis approaches produced congruous point estimates of effect, though modified relative risk analysis produced 95% confidence intervals much narrower than other approaches. Pooling data using these analysis approaches resulted in biased and imprecise effect estimates when using study-level data as opposed to patient-level data.

Conclusion: Meta-analysis using study-level data may produce unbiased and precise estimates of effect where statistical heterogeneity is not present, where other forms of heterogeneity are not present, and where data has been presented using either Andersen-Gill or negative binomial regression approaches. When these conditions are not met it is recommended (?preferable) that researchers present results from individual studies rather than presenting a pooled effect estimate from study-level data.

Introduction

Meta-analysis is an increasingly popular tool that can be used to investigate a number of issues including the strength of association between a risk factor and a disease, the accuracy of a screening instrument, or the efficacy of a treatment protocol. The importance of this tool is highlighted by the eminence afforded to results generated using this approach. When evaluating the efficacy of treatment protocols, some rating scales of evidence for clinical practice place the results of meta-analysis above that gleaned from individual randomised controlled trials,[1] though other scales place results from high quality individual randomised trials on par with meta-analyses.[2, 3] Meta-analysis has not been without its critics and several strategies for improving the validity of findings derived from meta-analysis have previously been suggested.[4]

Two levels of data can be used to conduct a meta-analysis. More commonly, researchers use summative data from published literature (meta-analysis literature – MAL) to construct their pooled estimates of effect. Less frequent are investigations that utilise individual patient level-data (meta-analysis patient – MAP). Key reasons for MAL to be used preferentially to MAP include the difficulty in attaining individual patient level data with which to conduct MAP, and the considerably greater time and resource required to conduct MAP.[5] However, MAP is the gold standard approach to meta-analysis[6] making it the approach of choice where either a MAL or MAP could feasibly be conducted,

Previous authors have sought to compare the results of MAP and MAL within a particular field and in simulation studies. Several have found discordance between

the two approaches though this has frequently been attributable to different sets of patients and / or studies being included.[7] Another investigation of the risk of epithelial ovarian cancer from use of the oral contraceptive pill found these approaches to be concordant, though had to exclude studies identified as being sources of heterogeneity before being able to do so.[5] Other investigators demonstrated that MAP was mathematically identical to MAL for a model comparing multiple treatments to a control for a continuous outcome under the assumptions of fixed effects.[8] If MAL is able to produce equivalent results to MAP, then the time and resource savings associated with this approach would clearly make MAL the approach of preference. However, it also follows that if MAP is unable to be undertaken and MAL does not produce equivalent results to MAP, then it is questionable as to whether a meta-analysis should be pursued at all.

One area where it is unclear whether MAL produces equivalent results to MAP is the study of diseases that may recur (e.g. episodes of cancer, urinary tract infections, epileptic seizures), particularly where the rate of disease occurrence is of primary importance. One field where this is certainly the case is the prevention of accidental falls in hospitals. It has previously been argued that the rate of falls (number of falls by a patient divided by the length of time they are observed) is of greater importance in this field than the proportion of patients who experience one or more falls, or even the sum total number of falls because of the varied total length of the observation period for the participants in studies conducted in this field. The time period in falls studies is a natural variation due to variations in each patient's length of stay in hospital. An added contextual factor in this field is that this style of data can be handled in several different ways and that a range of analysis approaches could

potentially be used.. For example, in three recent randomised trials of targeted, multifactorial intervention programs, falls rate data has been considered as simple rate data (number of events divided by observation period) analysed using an adaptation of the relative risk calculation in one,[10] and by XXX in another, and recurrent events survival data in another.[11] This 'statistical analysis approach' heterogeneity may further threaten the validity of a MAL in this field.

In this study, we are concerned with determining whether a MAL, for evaluation of either risk factors or interventions, is able to be validly undertaken within a field where the primary outcome of interest is the rate of a disease that can recur. We first use a MAP approach to compare the findings of a range of analysis approaches to a gold standard for considering this style of data (aim 1). We argue that if an approach produces biased or imprecise effect size estimates using a MAP approach, then a MAL using this approach will also produce biased or imprecise effect size estimates and will be unsuitable for practical use. Second, we consider whether a MAL using a uniform approach to analysis of individual studies produces a precise and unbiased pooled effect estimate as that derived from a MAP using the same analysis approach (aim ii). Third we consider what the differential effect on a pooled effect estimate would be if a MAL was attempted using study-level data comprising heterogenous analysis approaches (aim iii).

Method

Design:

Meta-analysis (both MAL and MAP) utilising prospectively collected falls data from two large randomised controlled trials of falls prevention interventions conducted across three hospital sites.

Participants and setting:

The first trial, conducted between 2002 and 2003 took place on three subacute rehabilitation wards at the Peter James Centre, Melbourne, Australia. This study had n=626 participants and further details of this study have been published elsewhere.[11] The second, conducted between 2008 and 2009 took place on three geriatric assessment and rehabilitation wards, two acute orthopaedic wards, and one acute medicine / respiratory ward at the Princess Alexandra Hospital, Brisbane, Australia, and one medical and one surgical ward of the Swan Districts Hospital, Perth, Australia. This study had n=1206 participants in total, however data from n=1202 was used in this analysis (n=350 from Swan Districts Hospital, Perth, n=852 from Princess Alexandra Hospital, Brisbane) due to outliers still remaining in the study at the time of this analysis. Further information regarding this study has been published elsewhere.[12]

Measurements:

The primary outcome measure in both of these studies was in-hospital falls. In both studies, falls data was collated using local hospital incident reporting systems along with individual patient medical record review. In the first study staff were provided with a standardised definition of a fall.[11] In the second study, staff were provided with video-based training in how to apply a standardised definition of a fall,[13] and falls data

was additionally captured using weekly patient interview with a research assistant blinded to group allocation.

Three independent variables were selected for examination in this study; participant gender, age, and admission diagnosis of stroke versus 'not stroke'. The median age of participants was 78 years, so this variable was converted to a dummy variable of young (≤78 years) versus old (>78 years). These three dichotomous variables were specifically selected on the basis of one having variable relationships with the raw number of falls and length of stay in hospital. Participant gender had a strong relationship with falls but not patient length of stay in hospital. Participant admission diagnosis of stroke had a strong relationship with both falls and length of stay in hospital. Participant age did not have a strong relationship with either variable.

Analysis:

Aim i)

In addressing this aim we considered a range of analysis approaches that a researcher might encounter when attempting a meta-analysis in this field and analysis approaches that they could potentially encounter. A description of each approach and justification for selection is presented (table 1).

A MAP was undertaken for each approach with the combined dataset for the exposure variable to examine the relative bias and precision each approach had relative to the Andersen-Gill version of Cox semi-parametric survival analysis for recurrent events and negative binomial regression approaches. Previous authors have advocated use of the Andersen-Gill approach for recurrent events data,[14] while others have advocated

use of negative binomial regression approach ahead of the Andersen-Gill survival approach due to the ease of computation of the former and the risk of violating the proportional hazards assumption in the latter.[15] Thus we also checked the assumption of proportional hazards for each of the predictor variables (both visually and using the Schoenfeld residuals test[16]) and found no violations of the proportional hazards assumption.

Aim ii)

Analysis approaches identified from aim i) to produce relatively unbiased estimates of treatment effect from MAP were then used in MAL for each of the exposure variables. The combined dataset was broken down into three separate studies according to site of data collection (Melbourne: n=626, Brisbane: n=852, Perth, n=350). The Generic Inverse Variance Method was used in each case for undertaking the MAL. Application of the Generic Inverse Variance Method requires two numbers from each study: an effect estimate and its standard error. The effect estimate summarises the treatment effect in a clinical trial or epidemiologic investigation (for example, an odds ratio or hazard ratio). The standard error summarises the precision of the effect estimate. The Generic Inverse Variance Method is available in the statistical program RevMan (version 5). As the data analysis approaches considered in this study are ratios, using the Generic Inverse Variance Method required data to be entered as natural logarithms of the effect estimate (e.g. the natural logarithm of the hazard ratio) and standard errors of the natural logarithm of the effect estimate. Standard errors of the natural logarithm of the effect estimate were calculated using the difference between the natural logarithms of the upper and lower 95% confidence limits divided by 3.92 (a 95% confidence interval is 3.92 standard errors wide). Both

fixed and random effects analyses were undertaken, however results for the fixed effect analysis were only presented when $I^2=0$.

In addition to Generic Inverse Variance Method analyses, a "dichotomous data" MAL approach was employed entering the number of falls in the exposed and unexposed groups as the "events" for these groups, and the amount of participant observed time in the exposed and unexposed groups as the "total number of patients" data for these groups (Rate 5). Days were used as the unit of measure of time. A "Peto - observed minus expected" analysis was also considered for the pooling of survival analysis (first event) data. This approach can only be used for this type of analysis, and requires effect sizes to be close to 1.0, and even numbers in the groups being compared to produced unbiased results.[17] However this approach was not considered further as these three criteria were unlikely to be met in this field.

Aim iii) A strength of the Generic Inverse Variance Method is that it can be employed in fields where results have been analysed and presented using a variety of approaches. This approach weights the results of each study using the inverse of the variance of the effect estimate. Therefore, the standard errors derived from the three separate site analyses for each exposure variable were contrasted to determine which approach would most strongly influence the MAL results. Only the analysis approaches retained for examination in aim ii) were examined in aim iii).

Estimates of effect size were calculated using STATA I/C version 10.0. MAL were conducted using Review Manager (RevMan) version 5.0.

Results

The distribution of exposure, outcome and study observation period variables for the combined sample and each individual site are presented (table 2).

Aim i) -The results of the MAP for each of the analysis approaches examined are presented (figure 1). It is apparent across all three exposure variables that the "Survival 1" Andersen-Gill recurrent events survival analysis, the "Count" negative binomial regression, the "Rate 1 & 2" bootstrap simulation approaches to calculating a relative rate, and the "Rate 3" relative rate calculation using the modified relative risk formula (days as unit of time measure) produced congruous point estimates of effect. The 95% confidence intervals were relatively narrow using the "Rate 3" approach, followed by the "Count" approach. The "Survival 1" and "Rate 1 & 2" approaches produced 95% confidence intervals of similar width in each circumstance. The remaining analysis approaches did not produce unbiased estimates of effect across the three exposure variables.

Aim ii) The subset of analysis approaches that produced relatively unbiased effect estimates were then used in Generic Inverse Variance Measure MAL, the results of which (as opposed to their MAP results) are presented (figure 2). Included in this figure also are the results from the modified "dichotomous data" MAL (Rate 5). This figure demonstrates consistent incongruity between the MAP and MAL approaches for the gender and age exposure variables. It is possible that this may be due to the higher levels of statistical heterogeneity across the three sites for these variables, as demonstrated by the higher I^2 values. Noticeable also for these two exposure

variables were the considerably wider confidence intervals attained through the MAL analyses relative to their respective MAP. None of the MAL analysis approaches were able to produce unbiased effect estimates across all three exposure variables in comparison to the MAP effect estimates generated through "Survival 1", "Count", and "Rate 1, 2 &3" approaches, though the degree of bias was not excessive. The magnitude of effect was overstated in the MAL approaches for the gender and age exposure variables, and marginally understated for the stroke exposure variable.

Aim iii) The standard errors of the natural logarithm of the effect estimates generated through the analysis of individual site data for each exposure variable are presented (figure 3). It is evident that the "Rate 3" approach produced consistently smaller standard errors of the natural logarithm of the rate ratio, and the bootstrap simulation approach the largest. Hence, if combining data from any of these approaches into the one Generic Inverse Variance Method MAL, the results from the study employing the "Rate 3" approach will more heavily weighted, and those from the "Rate 1 & 2" bootstrap approach less heavily weighted, than if these studies had used the "Count" negative binomial regression or "Survival 1" Andersen-Gill survival analysis approach.

Discussion

Meta-analysis is a tool that has frequently been demonstrated to provide valuable insights across a range of fields and has been recommended as a means of reducing bias. However for the particular case of falls in hospitals, the utility of available meta-analysis instruments is less clear. The nature of the outcome being count data

and the inconsistency of follow-up duration in this setting as a natural consequence of patients' varied length's of stay in hospital complicates the pooling of data from separate studies. This study has demonstrated that several analysis approaches are available that can produce relatively unbiased estimates of effect within an individual study. Despite this, no currently available MAL approach was able to produce effect estimates that were congruent with those derived from MAP analyses. This was particularly the case with greater levels of statistical heterogeneity within the data.

Our initial research question confirmed the arguments put forward by previous authors that falls need to be modelled using approaches that treat them as events over time that may recur.[15, 18, 19] It was clear from our data that dichotomising data into groupings of fallers versus non-fallers or multiple fallers versus non multiple fallers were poor surrogates for examining the rate of falls, even when adjusting for length of observation.

Concerns previously raised that using a modified version of the relative risk formula produces variable results depending on the units that time is measured in were again confirmed.[20] However, it should be noted that using the modified relative risk formula with time measured in days consistently produced relatively unbiased point estimates of the relative rate of falls. This is notable as this is approach could be used in a MAL where individual studies have not presented an effect estimate but the relative number of falls and length of observation from exposed and unexposed groups can be gleaned. The confidence intervals from this approach were considerably narrower than those gleaned from the "Survival 1" Andersen-Gill recurrent events, "Count" negative binomial regression and "Rate 1 & 2" bootstrap

approaches indicating that use of these confidence intervals in a MAL could lead to spurious significant findings. This is not surprising as the application of the relative risk formula to this data in no way permits acknowledgement of the dependence of falls or time data within individual patients.

The bootstrap simulation approaches in contrast did produce conservative (possibly overly-conservative) 95% confidence intervals and unbiased effect point estimates, and were not affected by whether time data were measured in units of days or years. It was also notable that the bootstrap simulation approaches, which directly model the ratio of rate of falls between groups over the length of the study, produced point estimates of effect more consistent with the "Survival 1" Andersen-Gill approach than with the "Count" negative binomial regression approach, though the differences between these approaches was relatively small. The disadvantage of this approach relative to the "Rate 3" modified relative risk calculation is that it cannot be executed without full access to individual patient level data.

Contrasting the MAP and MAL findings where data has been analysed using a uniform approach demonstrated that pooled effect estimates using the MAL approach produced either biased or imprecise effect estimates relative to the MAP findings where statistical heterogeneity between study results was observed. This raises the spectre of misleading pooled results from a MAL in such circumstances. If different analysis approaches have been used, pooling of hazard ratios derived from Andersen-Gill analyses and Incidence Rate Ratios derived from negative binomial regression appears reasonable given their proximity of effect point estimates and comparable standard errors. Including data from bootstrap analysis or modified relative risk

calculations (where time is measured in days) will under or over influence the pooled effect estimate due to the relative size of their standard errors, though could always be acknowledged in the discussion and considered by readers if the number of studies contributing data in this way was very small.

For a researcher seeking to conduct a meta-analysis in the field of falls prevention in the hospital setting, several factors need to be considered. A MAL may produce unbiased and precise effect estimates if there is no heterogeneity in study findings and the effect estimates in individual studies have been presented using Andersen-Gill survival analysis (or other recurrent events survival analysis approaches demonstrated to produce equivalent results[18]), negative binomial regression, or bootstrap simulation approaches to calculation of a relative rate. Inclusion of data from a relative rate derived through modification of the relative risk formula (where time is measured in days) may also be considered, though is likely to overly influence the pooled effect point estimate and produce confidence intervals that are too narrow. In this field, heterogeneity in study results, even amongst the small number of large randomised trials to date, is likely.[10, 11, 21, 22] This study also demonstrated statistical heterogeneity in MAL of two out of three risk factors considered.

Given the likely difficulty in attaining individual patient level data in this field, a researcher seeking to conduct a meta-analysis has limited options. We would suggest systematic review and presentation of individual results rather than seeking to calculate a pooled effect estimate. This would also negate other serious concerns such as pooling data from interventions that by their content are heterogenous, from risk factors where exposure has been measured inconsistently, and studies that have used

varying patient populations. Meta-analysis can provide profound insight into a research field, however, in our opinion, should only be conducted when the results are expected to provide greater clarity than what is already afforded by individual study results. Heterogeneity of all forms, and difficulties in using a MAL approach that can produce a precise and unbiased pooled effect estimate are serious concerns that all authors, reviewers, and readers need to give greater attention to when considering results of a meta-analysis where the outcome of interest is a recurrent event.

This study had several limitations, principally, the number of factors and number of studies considered was small. Greater subtlety in strengths and limitations of the MAL approaches investigated could be revealed by examining a greater number and combinations of factors (~~eg.~~e.g. n~~N~~umber of studies, size of studies, strength of effect, statistical heterogeneity between studies) that may impact upon its accuracy. Other forms of recurrent events data will have data distributions different to those of the falls dataset examined in the present study warranting replication of this study with other clinical data distribution examples to determine consistency of results with the present study. There were also other approaches for analysis of recurrent events data (eg. Conditional survival analysis models) that were not investigated in the present study that could also be the focus of further investigation.

Table 1. Description and justification of analysis approaches used to address aim i).

Treatment of falls data	Description	Justification
Survival; time to event data (1)	Andersen-Gill survival analysis for recurrent events, using robust variance estimates to account for dependency of events within an individual. Effect estimate referred to as Hazard Ratio.	Recognised by previous authors as an appropriate approach for analysing recurrent event and specifically falls rate data.[14, 15, 18, 19]
Survival; time to event data (2)	Cox semi-parametric survival analysis, time until first event (ignores participant data after the first event). Effect estimate referred to as Hazard Ratio.	Time to first event analysis approach previously used to analyse fall risk factors amongst stroke patients.[31]
Count (quasi-rate) data	Negative binomial regression, a form of Poisson regression for overdispersed count data (where the outcome mean is exceeded by the standard deviation). Adjusting for individual participant period of observation in the study effectively makes this a comparison of rate data. Effect estimate referred to as Incidence	Recognised by previous authors as an appropriate standard for analysing falls rate data.[15, 18, 19]

Rate Ratio.

Rate data (1)	Bootstrap simulation for calculation of a rate ratio and normal-based 95% confidence intervals. Rate ratio calculated as (total falls in exposed group / total time in exposed group) / (total falls in unexposed group / total time in unexposed group). Time measured in days.	Has been used to calculate confidence intervals in evaluations of falls risk screening tools and falls risk factors in the hospital setting.[9, 23, 24]
Rate data (2)	As above however time measured in years.	The Cochrane Collaboration raises concern that changing the units that time is measured in changes the results.[17]
Rate data (3)	Calculation of a relative rate using the formula for relative risk whereby the number of falls in exposed / unexposed groups substitutes for the number of participants with the disease in these groups, and the amount of observed participant time in exposed / unexposed groups substitutes for the number of participants without the disease in these groups. Time measured in days.	Previously used to evaluate the effectiveness of a multi-factorial intervention program to reduce falls in this setting.[10]

Rate data (4)	As above however time measured in years.	It has previously been demonstrated that changing the units that time is measured in changes the results when relative rate is calculated in this way.[20]
Binary data (1)	Logistic regression classifying all patients experiencing one or more falls as a faller as opposed to being a non-faller. Effect estimate referred to as Odds Ratio.	Frequently presented data in epidemiologic and experimental studies in this setting. An analysis option forwarded by the Cochrane Collaboration for considering count data where a minority of participants have the event and the counts for those who do are mostly low.[17]
Binary data (2)	As above however including adjustment for participant period of observation in the study.	Adjusting for participant exposure may compensate for insensitivity of the analysis approach above to variation in patient length of stay in hospital.
Binary data (3)	Logistic regression classifying all patients experiencing two or	Frequently presented data in falls prevention

	more falls as a multiple faller as opposed to being a non-multiple faller.	studies more broadly.
Binary data (4)	As above however including adjustment for participant period of observation in the study.	Adjusting for participant exposure may compensate for insensitivity of the analysis approach above to variation in patient length of stay in hospital.
Binary data (5)	Relative risk analysis classifying all patients experiencing one or more falls as a faller as opposed to being a non-faller. Effect presented as Relative Risk.[32]	Previously used to evaluate the effectiveness of a multi-factorial intervention program to reduce falls in this setting.[11]

Table 2. Distribution of exposure, outcome and participant observation variables for the combined dataset and each site.

	Complete dataset	Melbourne	Brisbane	Perth
n	1828	626	852	350
Gender; male – n (%)	767 (42%)	206 (33%)	426 (50%)	135 (39%)
Age – mean (sd)	76.7 (10.7)	79.8 (9.3)	73.2 (11.4)	79.5 (8.5)
Admission diagnosis stroke – n (%)	170 (9%)	72 (12%)	65 (8%)	33 (9%)
Falls; total	497	254	177	66
Fallers; 1 or more falls – n (%)	275 (15%)	125 (20%)	105 (12%)	45 (13%)
Multi-fallers; 2 or more falls – n (%)	89 (5%)	41 (7%)	34 (4%)	14 (4%)
Length of observation; mean (sd) per patient	25.6 (26.2)	29.7 (22.2)	26.0 (30.7)	17.2 (18.0)

Figure 1. MAP pooled effect estimates with 95% CIs utilising each analysis approach described for each exposure variable. Effect estimates >1 indicate males, people with stroke and older patients at higher risk for falls.

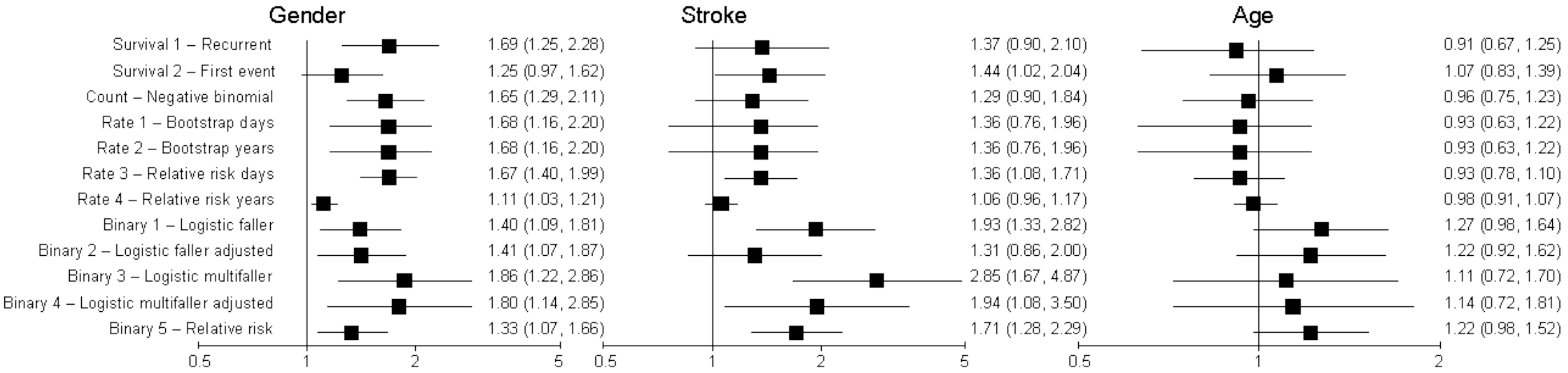


Figure 2. MAP versus MAL pooled effect estimates (95% CIs), and I^2 values (for MAL analyses) for each exposure variable. Effect estimates >1 indicate males, people with stroke and older patients at higher risk for falls.

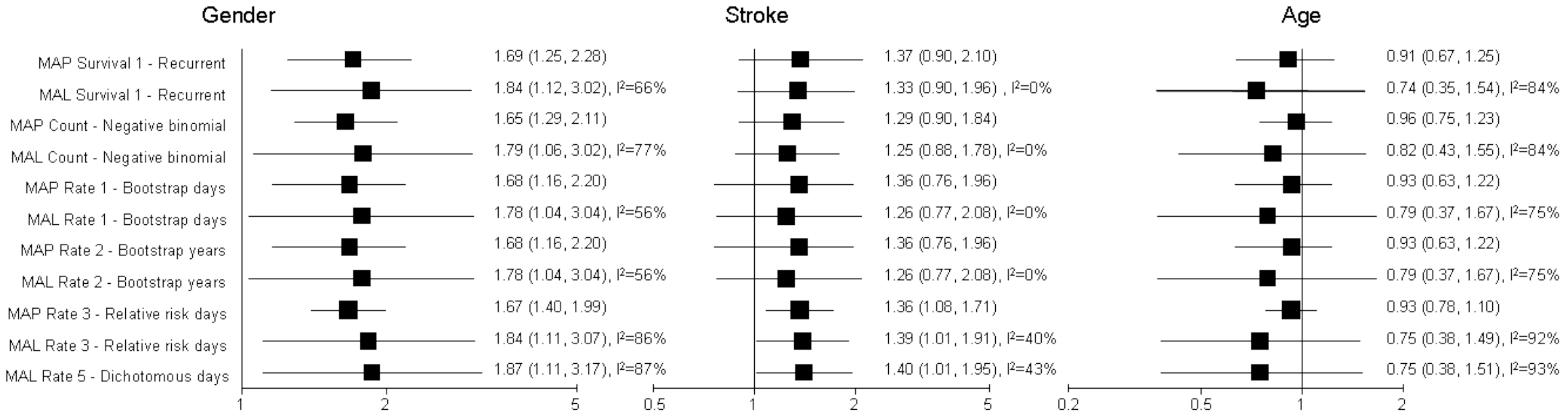
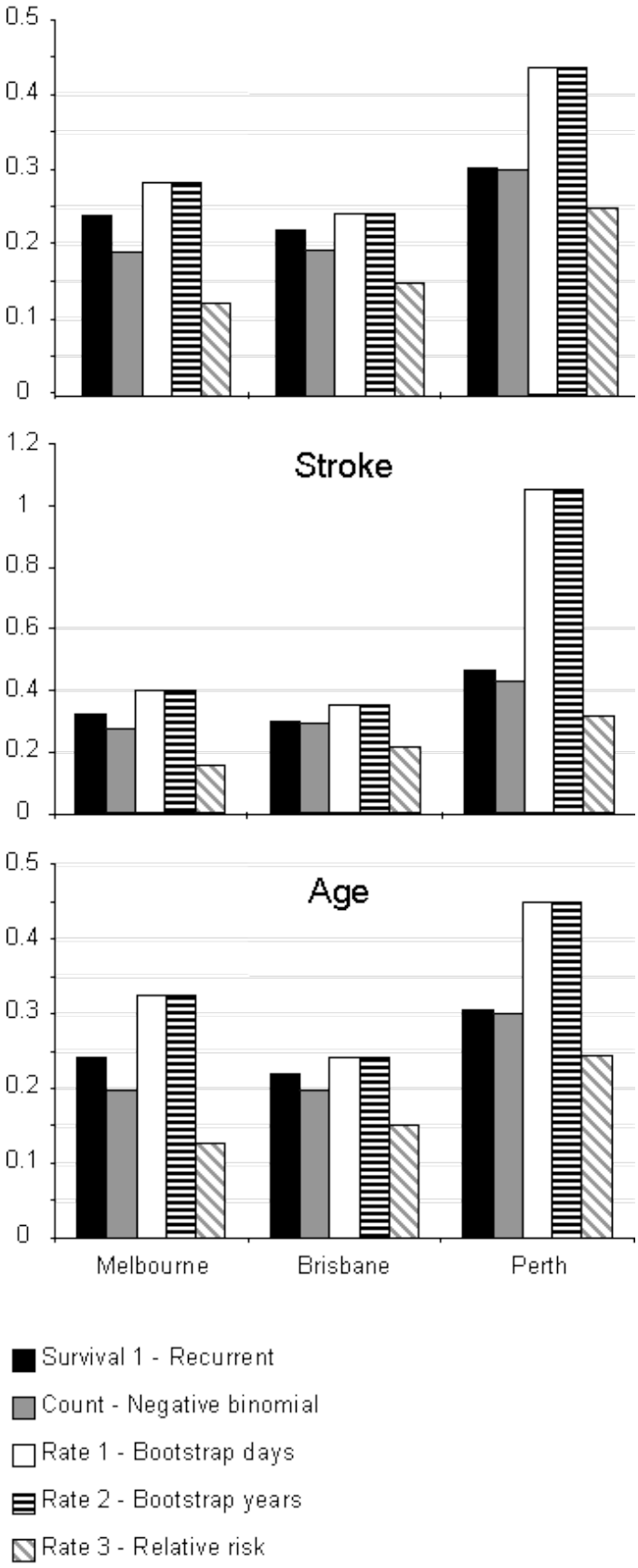


Figure 3. Standard errors of the natural logarithm of the effect estimates by site and exposure variable.



References

1. Melnyk B (2004) Integrating levels of evidence into clinical decision making. *Pediatric Nursing*;30:323-325.
2. Kroke A, Boeing H, Rossnagel K, Willich S (2003) History of the concept of 'levels of evidnece' and their current status in relation to primary prevention through lifestyle interventions. *Public Health Nutrition*;7:279-284.
3. American Family Physician. *Levels of evidence in AFP*. 2009 [cited Accessed August 12th, 2009]; Available from: www.aafp.org/online/en/home/publications/journals/afp/afplevels.html.
4. Egger M, Smith G (1995) Misleading meta-analysis. *British Medical Journal*;310:752-754.
5. Steinberg K, Smith S, Stroup D, Olkin I, lee N, Williamson G, Thacker S (1997) Comparison of effect estimates from a meta-analysis of summary data from published studies and from a meta-analysis using individual patient data for ovarian cancer studies. *Am J Epidemiol*;14:917-925.
6. Oxman AD, Clarke M, Stewart L (1995) From science to practice. Meta-analyses using individual patient data are needed. *J Am Med Assoc*;274:845-846.
7. Lambert P, Sutton A, Abrams K, Jones D (2002) A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *Journal of Clinical Epidemiology*;55:86-94.
8. Olkin I, Sampson A (1998) Comparison of meta-analysis versus analysis of variance of individual patient data. *Biometrics*;54:317-322.

9. Haines TP, Hill KD, Bennell KL, Osborne RH (2006) Recurrent events counted in evaluations of predictive accuracy. *J Clin Epidemiol*;59:1155-61.

10. Healey F, Monro A, Cockram A, Adams V, Heseltine D (2004) Using targeted risk factor reduction to prevent falls in older in-patients: a randomised controlled trial. *Age and Ageing*;33:390-395.

11. Haines T, Bennell K, Osborne R, Hill K (2004) Effectiveness of targeted falls prevention programme in subacute hospital setting: randomised controlled trial. *British Medical Journal*;328:676-679.

12. Hill A, Hill K, Brauer S, Oliver D, Hoffmann T, Beer C, McPhail S, Haines T (2009) Evaluation of the effect of patient education on rates of falls in older hospital patients: Description of a randomised controlled trial. *BMC Geriatrics*;9:14.

13. Haines T, Massey B, Varghese P, Fleming J, Gray L (2009) Inconsistency in classification and reporting of in-hospital falls. *Journal of the American Geriatrics Society*;57:517-523.

14. Therneau T, Hamilton S (1997) rhDNase as an example of recurrent event analysis. *Stat Med*;16:2029-2047.

15. Robertson MC, Campbell AJ, Herbison P (2005) Statistical analysis of efficacy in falls prevention trials. *Journals of Gerontology Medical Sciences*;60A:530-534.

16. Grambsch P, Thernea T (1994) Proportional hazards tests and diagnositics based on weighted residuals. *Biometrica*;81:515-526.

17. The Cochrane Collaboration (2008) Cochrane handbook for systematic reviews of interventions version 5.0.1, ed. J. Higgins, et al.

18. Kuramoto L, Sobolev B, Donaldson M (2008) On reporting results from randomized controlled trials with recurrent events. *BMC Medical Research Methodology*;8:35.

19. Donaldson M, Sobolev B, Cook W, Janssen P, Khan KM (2009) Analysis of recurrent events: a systematic review of randomised controlled trials of interventions to prevent falls. *Age Ageing*;38:151-155.

20. Haines T, Hill K (2005) Difficulties encountered in hospital falls prevention research. *Age & Ageing.*;34:311.

21. Cumming RG, Sherrington C, Lord SR, Simpson JM, Vogler C, Cameron ID, Naganathan V (2008) Cluster randomised trial of a targeted multifactorial intervention to prevent falls among older people in hospital. *BMJ*;336:758-60.

22. Stenvall M, Olofsson B, Lundstrom M, Englund U, Borssen B, Svensson O, Nyberg L, Gustafson Y (2007) A multidisciplinary, multifactorial intervention program reduces postoperative falls and injuries after femoral neck fracture. *Osteoporos Int*;18:167-75.

23. Haines T, Kuys S, Morrison G, Clarke J, Bew P (2008) Balance impairment not predictive of falls in geriatric rehabilitation wards. *Journals of Gerontology Medical Sciences*;63:523-28.

24. Haines T, Hill K, Bennell K, Osborne R (2006) New instrument for targeting falls prevention interventions accurate and clinically applicable in hospital setting. *Journal of Clinical Epidemiology*;59:168-175.

25. Oliver D, Britton M, Seed P, Martin F, Hopper A (1997) Development and evaluation of evidence based risk assessment tool (STRATIFY) to predict which elderly inpatients will fall: case-control and cohort studies. *British Medical Journal*;315:1049-1053.

26. Haines TP, Hill K, Walsh W, Osborne R (2007) Design-related bias in hospital fall risk screening tool predictive accuracy evaluations: systematic review and meta-analysis. *J Gerontol A Biol Sci Med Sci*;62:664-72.

27. Evans D, Hodgkinson B, Lambert L, Wood J (2001) Falls risk factors in the hospital setting: a systematic review. *International Journal of Nursing Practice*;7:38-45.

28. Oliver D, Connelly J, Victor C, Shaw F, Whitehead A, Genc Y, Vanoli A, Martin F, Gosney M (2007) Strategies to prevent falls and fracture in hospitals and care homes and effect of cognitive impairment: Systematic review and meta-analyses. *British Medical Journal*;334:82.

29. Oliver D, Daly F, Martin FC, McMurdo MET (2004) Risk factors and risk assessment tools for falls in hospital in-patients: a systematic review. *Age Ageing*;33:122-130.

30. Coussement J, De Paepe L, Schwendimann R, Denhaerynck K, Dejaeger E, Milisen K (2008) Interventions for preventing falls in acute and chronic care hospitals: A systematic review and meta-analysis. *Journal of the American Geriatrics Society*;56:29-37.

31. Nyberg L, Gustafson Y (1997) Fall prediction index for patients in stroke rehabilitation. *Stroke*;28:716-21.

32. Altman D (1991) Practical statistics for medical research. 1st ed., London: Chapman and Hall.