The final publication is available in: Accident Analysis and Prevention, 60, 2013, 245-253. 10.1016/j.aap.2012.12.003

Corrigendum to:

Publication bias and time-trend bias in meta-analysis of bicycle helmet efficacy: A re-analysis of Attewell, Glase and McFadden 2001

Rune Elvik
Institute of Transport Economics
Gaustadalléen 21
NO-0349 Oslo, Norway
(E-mail: re@toi.no)

ABSTRACT

This paper is a corrigendum to the previously published paper: “Publication bias and time-trend bias in meta-analysis of bicycle helmet efficacy: A re-analysis of Attewell, Glase and McFadden 2001” (Accident Analysis and Prevention 2011, 1245-1251). This corrigendum was prepared to correct errors in data and analysis in the previously published paper. Like the previously published paper, this paper confirms that the meta-analysis of bicycle helmet efficacy reported by Attewell, Glase and McFadden (Accident Analysis and Prevention, 2001, 345-352) was influenced by publication bias and time-trend bias that was not controlled for. As a result, the analysis reported inflated estimates of the effects of bicycle helmets. This paper presents a re-analysis of the study. The re-analysis included: (1) Ensuring the inclusion of all published studies by means of continuity corrections of estimates of effect relying on zero counts; (2) Detecting and adjusting for publication bias by means of the trim-and-fill method; (3) Detecting and trying to account for a time-trend bias in estimates of the effects of bicycle helmets; (4) Updating the study by including recently published studies evaluating the effects of bicycle helmets. The re-analysis shows smaller safety benefits associated with the use of bicycle helmets than the original study.

Key words: bicycle helmets, meta-analysis, publication bias, time-trend bias, re-analysis, trim-and-fill

1 INTRODUCTION

Numerous studies have found that bicycle helmets are effective in reducing head injury to bicyclists. A meta-analysis based on 13 estimates of the effect on head injury of wearing a bicycle helmet concluded that the risk of head injury is reduced by 60% (Attewell, Glase and McFadden 2001). The same study concluded that the risk of brain injury is reduced by 58% and the risk of facial injury reduced by 47%. All these reductions in risk were statistically significant at the 5% level. These results were confirmed in a meta-analysis performed for the Cochrane collaboration by Thompson, Rivara and Thompson (2009), who reported even more impressive reductions in the risk of head injury, brain injury and facial injury. A meta-analysis by Elvik, Høye, Vaa and Sørensen (2009) reported a 64% reduction in the risk of head injury when a hard helmet is worn and a 41% reduction in risk when a soft helmet is worn. According to this meta-analysis, the risk of facial injury is reduced by 34% if a hard helmet is worn; wearing a soft helmet was associated with a statistically non-significant (5% level of significance) increase of 14% in the risk of facial injury.

While these meta-analyses are broadly in agreement with respect to the effects of wearing a bicycle helmet, they differ in important respects. The most important difference between them concerns the set of studies included. The Cochrane review (Thompson, Rivara and Thompson 2009) is the most restrictive, omitting several studies because they were judged not to employ an appropriate study design. The review of Elvik et al. (2009), on the other hand, included all studies that were retrieved.

In meta-analysis, an ideal of including all studies that deal with a topic has wide support. Assessing study quality is also widely supported, but there are many ways of doing so, none of them without a large element of subjectivity. Rather than omitting studies classified as poorly designed, most meta-analysts would prefer to include these studies and assess how excluding them would influence summary estimates of effect as part of a sensitivity analysis.

The objective of this paper is to critically assess the meta-analysis reported by Attewell et al. (2001). The authors of that study discussed the possibility of publication bias, admitting that it could not be ruled out, but concluding that it was unlikely to greatly influence summary estimates of effect. Since publication of the paper, new techniques for detecting and adjusting for publication bias have been developed (Rothstein, Sutton and Borenstein 2005). It is now possible to test and adjust for the possible presence of publication bias more rigorously than at the time when Attewell et al. prepared their paper. Moreover, analysts have become increasingly aware of other potential biases that may influence meta-analyses. A case can therefore be made for re-analysing the study of Attewell et al. in order to test for the possible presence of various sources of bias in the study.

2 BIASES IN META-ANALYSIS

There are many sources of bias in meta-analyses. Briefly, the following are the most important (Rothstein, Sutton and Borenstein 2005, Sweeting, Sutton and Lambert 2004, Borenstein et al. 2009):

1. Publication bias, which denotes a tendency not to publish studies if findings are not statistically significant or contradict prior expectations or the vested interests of sponsors of the research,
2. Time trend bias, which refers to a tendency for study findings to change over time; if all findings are pooled independently of when they were published, the trend will be pasted over and the summary estimate of effect will be misleading,
3. Zero count bias, which is bias arising if studies with zero counts are omitted or if inefficient continuity corrections are applied to such studies.

It is possible to detect and adjust for all these sources of bias. The techniques for doing so are not perfect and some of them rely on assumptions whose validity cannot be tested in each study. It is nevertheless of interest to examine the extent to which summary estimates of effect could be influenced by the various sources of bias.

3 BIASES IN META-ANALYSIS OF ATTEWELL ET AL.

3.1 Continuity correction

The first potential source of bias to be discussed is the method adopted for continuity correction. Attewell et al. (2001) included studies in which one of the four numbers used to calculate the odds ratio was zero by adding 0.25 to each cell of the 2 x 2 table. Various approaches that can be taken to continuity correction in meta-analysis are discussed by Sweeting et al. (2004). They argue that adding a constant to each cell of a 2 x 2 table is arbitrary and propose two other techniques that can be used for continuity correction. One of these techniques, empirical continuity correction, has been applied in the re-analysis of the Attewell et al. study.

Empirical continuity correction develops correction factors that sum to one by using the summary estimate of effect based on studies with non-zero counts. Denote this summary estimate by $\widehat{\Omega}$. Then, in a 2 x 2 table, such as those used by Attewell et al. to estimate the effect of bicycle helmets, four numbers can be identified (see table 1). Adopting the notation in the leftmost 2 x 2 part of Table 1, Sweeting et al. (2004) define the empirical continuity correction k_T and k_C as follows:

$$\frac{k_T(n_T R + k_C)}{k_C(n_T + k_T)} = \widehat{\Omega} \qquad (1)$$

The letters are defined as in the leftmost 2 x 2 part of Table 1; R denotes the ratio of the number of observations in the control group, no event cell to the number of observations in the treatment, no event cell. It is initially assumed (counterfactually) that both case counts (i.e. those with head injury) are zero. The following approximate continuity correction factors are then derived (given the condition that the correction factors should sum to one):

$$k_C \approx \frac{R}{R + \widehat{\Omega}} \qquad (2)$$

$$k_T \approx \frac{\widehat{\Omega}}{R + \widehat{\Omega}} \qquad (3)$$

To give an example: In the study reported by Attewell et al. (2001) the summary estimate of effect on head injury based on studies with no zero counts was 0.419 (i.e. 58% reduction in the risk of injury; fixed-effects model). In one study, the following counts were observed:

Head injury, helmet worn: 0

Head injury, no helmet: 7

Other injury, helmet worn: 8

Other injury, no helmet: 6

This results in the following correction factors:

$K_C = (6/8)/[(6/8) + 0.419] = 0.642$ and $K_T = 0.419/[(6/8) + 0.419] = 0.358$

The adjusted estimate then becomes:

$(0.358/7.642)/(8.358/6.642) = 0.037$, which suggests that the zero count of head injuries among helmeted cyclists in this sample is consistent with a larger effect of bicycle helmets than observed in other studies, but not a 100% protective effect.

Table 1 about here

Other effects based on zero counts were similarly adjusted and included in the meta-analysis in adjusted form.

3.2 Publication bias

The possible presence of publication bias in the meta-analysis reported by Attewell et al. (2001) was tested for by means of the trim-and-fill technique (Duval and Tweedie 2000A, 2000B, Duval 2005). This is a non-parametric method based on funnel plots. A funnel plot is a diagram that shows estimates of effect on the abscissa and the statistical precision of each estimate of effect on the ordinate. Data points in a funnel plot should ideally speaking distribute like a funnel turned upside down, since precise estimates, located near the top of the diagram, should display less dispersion than imprecise estimates, located near the bottom of the diagram. The trim-and-fill technique is based on the assumption that a funnel plot should be symmetric if there is no publication bias. If one of the tails of the distribution is missing, or is markedly thinner populated by data points than the other, this is taken to indicate publication bias. To illustrate the technique, studies of the effect of bicycle helmets on facial injury will be used as example.

The trim-and-fill method is based on the assumption that publication bias leads to the suppression of data points in the left part of the funnel plot. When data points refer to a protective treatment, for which most odds ratios are smaller than 1, analysis therefore starts by inverting the diagram. This is done by multiplying all log-odds ratios by –1. Log-odds ratios are then sorted in order of ascending values. This is shown in the first column of Table 2.

Table 2 about here

The next two columns show the statistical weights associated with each estimate of effect, and the product of the log-odds ratios and statistical weights. At the bottom the sum of the statistical weights and the weights multiplied by the log-odds ratios are shown. The ratio of these numbers, i.e.:

$$\frac{\sum_{i=0}^{n} LnOR_i \cdot W_i}{\sum_{i=0}^{n} W_i} = \text{summary log-odds ratio} = \frac{126.878}{275.142} = 0.461$$

Is the summary log-odds ratio, or summary estimate of effect, in this case inverted. The differences between the summary log-odds ratio and each of the log-odds ratios are taken. For the first row of Table 2, the difference is 0.174 – 0.461 = –0.287. For the last row, it is 2.522 – 0.461 = 2.061. These differences are then ranked according to the absolute magnitude. The smallest difference, –0.039, is given the rank –1. The second smallest difference –0.247, is given the rank –2, and so on until all the differences have been ranked. The two estimators of publication bias, L and R, are based on these ranks. The estimator L is defined as follows (Duval 2005):

$$L_0 = \frac{4T_n - n(n+1)}{2n-1}$$

T_n denotes the sum of the positive ranks. With respect to Table 2, the sum is 3 + 5 + 6 + 7 + 8 = 29. Thus L becomes:

$$L = \frac{(4 \cdot 29) - (8 \cdot 9)}{((2 \cdot 8) - 1)} = 2.93 = 3$$

When the estimate for L has decimals, it is rounded to the nearest whole number. The estimator R is based on the length of the rightmost number of ranks associated with positive effects, i.e. the number of positive ranks larger than the absolute value of any of the negative ranks. Denoting this length with γ, the estimator is defined by $R_0 = \gamma - 1$. In Table 2, the length of the positive ranks is five (3, 5, 6, 7, 8). The length of ranks larger than the absolute value of any of the negative ranks is four (5, 6, 7, 8; since the largest negative rank is –4). Hence R becomes 4 – 1 = 3. The first iteration of the trim-and-fill analysis thus resulted in trimming away the three rightmost estimates of effect. When the summary log-odds ratio is re-estimated deleting the three rightmost estimates of effect, it becomes:

Summary log-odds ratio = $\frac{60.904}{200.593} = 0.304$

The differences and signed ranks are re-estimated. The new values are shown in columns six and seven of Table 2. The estimate for R is now 4, and the estimate for L 3.73, which is rounded to 4. The summary log-odds ratio is re-estimated omitting the four rightmost estimates of effect. The new summary estimate is 47.154/182.448 = 0.258. The differences and signed ranks are re-estimated. The new values are shown in columns eight and nine of Table 2. The estimate for R is 5; the estimate for L is 4. In general R and L tend to produce similar, but not necessarily identical estimates.

A new round of re-estimation is done by omitting the five rightmost studies. The summary estimate of the log odds ratio now becomes 46.325/181.279 = 0.256. Re-estimating the differences and signed ranks using this value produces estimates of R and L that are both identical to those obtained at the previous iteration. Iteration stops when the estimates for R and L no longer change. Thus, the final estimates in this case are 5 for R and 4 for L.

It is not uncommon that the estimates for R and L differ. In such cases, the analyst must choose which estimator to use in adjusting for publication bias. A conservative choice is to rely on the estimator indicating the highest number of missing studies. In this example, R suggests that 5 studies are missing L suggests that 4 studies are missing. Using R, the five missing studies have been added in Figure 1. The five missing studies are inserted symmetrically around the trimmed mean and are the mirror image of the five studies that were trimmed in the trim part of the analysis. For an easily accessible technical introduction to the trim-and-fill technique, see Duval (2005).

Figure 1 about here

The scales of the axes have been chosen as recommended by Sterne and Egger (2001). The abscissa shows the logarithm of the odds ratio (log odds ratio). Negative values indicate a reduction in risk; positive values indicate an increase in risk. The ordinate shows the fixed-effects value of the standard error of each estimate of effect, with the scale inverted so that the most precise estimates are located at the top of the funnel plot.

The meta-analysis reported by Attewell et al. (2001) presents summary estimates of the effect of bicycle helmets in five categories: (1) Head injury, (2) Brain injury, (3) Facial injury, (4) Neck injury, and (5) Fatal injury. The trim-and-fill technique has been applied to four of these categories. There were only three estimates of effect with respect to neck injury; too few for meaningfully testing for publication bias. Evidence suggesting publication bias was found in all the four categories of results for which it was tested. The trim-and-fill analyses were applied both to the fixed-effects and random-effects estimates of effect, although Duval (2005:134) points out that the method appears to be more tenable when applied to a set of fixed-effects estimates of effect than when applied to a set of random-effects estimates of effect.

Table 3 shows the original and adjusted summary estimates of effect according to a fixed-effects model of analysis and a random-effects model of analysis. A fixed-effects model is based on the assumption that the variation of estimates of effect between studies is random only. A random-effects model allows for systematic between-study variation in estimates of effects. A random-effects model was applied in all cases except for fatal injury, where the heterogeneity test did not indicate that there was any systematic between-study variation in estimates of effect.

Table 3 about here

Summary estimates of effect adjusted for publication bias invariably indicate smaller effects than unadjusted summary estimates. In some cases, in particular effects on facial injury according to a fixed-effects model, the effects of adjusting for publication bias are large. In most cases, however, adjusting for publication bias has a very small effect on summary estimates of effect. Attewell et al. (2001) were therefore correct in their conjecture that the possible presence of publication bias did not greatly influence summary estimates of effect.

3.3 Time trend bias

The effects of a safety measure, like bicycle helmets, may change over time. Often, one would expect that technological innovation made a safety measure more effective over time. As far as bicycle helmets are concerned, however, the opposite appears to be the case. Figure 2 shows changes in the summary effect of bicycle helmets on head injury as a function of cumulative statistical weights.

Figure 2 about here

The first studies, based on small samples, indicated a reduction in the risk of head injury of about 75% (odds ratio 0.25). As new studies were added, the summary estimate of effect became smaller, reducing to about 55% (odds ratio 0.45). This means that, on the average, recently published studies show considerably smaller benefits of bicycle helmets than older studies. It is important to keep in mind that the data points in Figure 2 are cumulative; thus the rightmost data point summarises the contributions of all studies, not just the most recent studies. A summary estimate of effect not recognising this trend may be misleading and give a too optimistic impression of the effects of new bicycle helmets.

One may test if the time trend bias is stable by using it to predict the summary estimate of effect when new studies are added. In section 4 of the paper, the analysis is updated by adding new studies to test the persistence of the time trend.

4 UPDATING THE META-ANALYSIS

Since the analysis presented by Attewell et al. was published, new studies evaluating the effects of bicycle helmets have been published. The analysis has been updated by including studies reported by Hausotter (2000), Hansen et al. (2003), Heng et al. (2006) and Amoros et al. (2009). Inclusion of these studies produced five new estimates for head injury, three new estimates for facial injury and one new estimate for neck injury. A new-meta-analysis was performed, adding the new estimates to those included in the study by Attewell et al. (2001).

As far as effects of bicycle helmets on head injury are concerned, the addition of five new estimates confirmed the time trend observed for the original thirteen estimates. This is shown in Figure 3. Figure 3 reproduces the data points in Figure 2 as well as the function fitted to those data points. When the new estimates are added to the original, the new data point to right in Figure 3 is created. As can be seen, extrapolating the curve fitted to the original data points predicts this new data point remarkably well.

Figure 3 about here

Table 4 shows summary estimates of effect based on the original estimates, the new estimates and all estimates. Estimates based on recently published studies show much smaller effects of bicycle helmets on head injury and facial injury than the original study. In fact, in the random-effects model, there is a statistically non-significant tendency for the wearing of bicycle helmets to be associated with an increase of the risk of facial injury. As far as neck injury is concerned, the tendency found in the original study for the risk of injury to increase when a helmet is worn is confirmed when a new estimate is added.

Table 4 about here

The head, the face and the neck can be viewed as three distinct regions of the body. Hence, it makes sense to develop summary estimates of effect of bicycle helmets for the head, the face and the neck. These estimates are shown at the bottom of Table 4. In general, the estimates suggest a moderate overall effect of bicycle helmets. In the random-effects analysis, based on the new estimates only, the effect is not statistically significant at the five percent level. For all studies, based on a random-effects model adjusted for publication bias, the best estimate is a 33 % reduction of the risk of injury to the head, the face or the neck if a bicycle helmet is worn. This summary estimate is statistically significant at the 5 % level.

The addition of new estimates did not remove publication bias. Figure 4 shows that six new data points were added in a trim-and-fill analysis of estimates of the effect of bicycle helmets on head injury.

Figure 4 about here

The trim-and-fill analysis was applied to all data points (thirteen original plus five new). It was judged that the five new data points were too few to apply the trim-and-fill technique. A trim-and-fill analysis was nevertheless attempted for these five data points and it converged at the value of three, suggesting that three new data points should be added to adjust for publication bias. No test of publication bias was performed for the four estimates of effect bicycle helmets on neck injury.

5 DISCUSSION

Do bicycle helmets reduce the risk of injury to the head, face or neck? With respect to head injury, the answer is clearly yes, and the re-analysis of the meta-analysis reported by Attewell et al. (2001) in this paper has not changed this answer. As far as facial injury is concerned, evidence suggests that the protective effect is smaller, but on balance there does seem to be a slight protective effect. The risk of neck injury does not seem to be reduced by bicycle helmets. There are only four estimates of effect, but they all indicate an increased risk of injury. When the risk of injury to head, face or neck is viewed as a whole, bicycle helmets do provide a protective effect. This effect is statistically significant in older studies. New studies, summarised by a random-effects model of analysis, indicate only a statistically non-significant protective effect.

These findings raise a number of issues. In the first place, why do recent studies show a smaller protective effect of bicycle helmets than older studies? In the second place, should a meta-analysis include all studies or just studies that satisfy certain selection criteria, like those applied in the Cochrane review of bicycle helmets (Thompson et al. 2009)? In the third place, why are the findings of some studies that have evaluated the effects of laws mandating the use of bicycle helmets apparently inconsistent with the findings of studies of the protective effect of bicycle helmets for each user?

There are two main reasons why the findings of studies that have evaluated the effects of bicycle helmets can vary: substantive and methodological. One reason for varying findings is that different types of helmets do not have the same protective effect. The first studies of bicycle helmets included mostly hard shell helmets. These have been found to offer better protection against head injury than soft shell helmets, which have become more popular over time. Thus, in the study of Hansen et al. (2003) more than one third of the helmets were soft shell helmets and these helmets were found to protect less well against head injury than hard shell helmets. Thompson et al. (2009) dismiss this argument, claiming that (page 7): “Bicycle helmets of all types ... provide substantial protection for cyclists of all ages.” However, the same authors (Rivara, Thompson and Thompson 1997) reported a study which found that only hard shell helmets protected against neck injury; use of a soft shell helmet was associated with an increased incidence of neck injury. On balance, the evidence suggests that: (1) Soft shell helmets offer less protection than hard shell helmets, and (2) Soft shell helmets appear to have become more common over time.

Thompson et al. (2009) were concerned about the quality of studies that have evaluated the effects of bicycle helmets and defined a set of selection criteria for including studies in a meta-analysis. Applying these criteria, seven studies were included and eight studies were omitted. Two studies were awaiting assessment at the time of publication of the review. Four of the seven studies that were included were performed by the same researchers as the Cochrane review. In short, Thompson et al. classified four of their own studies as good enough to be included in the meta-analysis, but excluded eight studies, none of which they were involved in. Littell, Corcoran and Pillai (2008) regard involvement in the conduct of one or more studies included in a review, or publication of a previous review on the same topic as a case of conflict of interest. This conflict of interest is relevant for the Cochrane review reported by Thompson et al. (2009). They were themselves authors of four of the seven studies included and had performed a similar Cochrane review twice before (in 2003 and 2006). To their credit, however, Thompson et al. included a very comprehensive section discussing criticisms of their review.

Study quality assessment is not an exact science. The Department for Transport in Great Britain issued a report in 2002 entitled: “Bicycle helmets – a review of their effectiveness: A critical review if the literature.” The report includes an assessment of the quality of 16 studies that have evaluated the effects of bicycle helmets. Studies were rated as good, reasonable or weak. Of the seven studies Thompson et al. (2009) included in the Cochrane review, one was rated as good, two as good/reasonable, three as reasonable and one as reasonable/weak. Three of the eight studies Thompson et al. omitted were also rated by the Department for Transport (2002). One was rated as good, one as good/reasonable and one as reasonable. Thus, if the rating developed by the Department for Transport is applied, it is by no means obvious that all the seven studies that were included by Thomson et al. ought to have been included. Nor is it clear that all the omitted studies were of lower quality than the studies included.

An alternative to omitting studies classified as bad would be to develop a quality score for each study and use that score in a sensitivity analysis, as illustrated by Elvik (2005). Although it is clear that any numerical quality score will contain an element of arbitrariness, including all studies and performing a sensitivity analysis allows readers to judge how study quality influences study findings. This opportunity does not exist if studies rated as “bad” are simply omitted.

Several researchers have been puzzled by the fact that, on the one hand, studies have reported large protective effects of bicycle helmets; on the other hand, studies of the effect of legislation that has been associated with large increases in the rate of helmet wearing have not always shown a clear decline in the number of head injuries among cyclists. There are at least two reasons why even a large increase in the rate of helmet wearing will not necessarily lead to a major reduction of the number of cyclists sustaining head injury. One reason could be selective recruitment, which means that it is the most cautious and safety-minded cyclists, with a lower rate of accident involvement than other cyclists, who first start wearing helmets. If, for example, in a population of cyclists 60 % have a 20 % lower rate of accident involvement than an average cyclist (i.e. a relative risk of 0.8), and these cyclists start wearing helmets that reduce their risk of head injury by 40 %, the total number of head injuries would be reduced by 19% (0.8 · 0.60 = 0.48; i.e. the safe cyclists are involved in 48 % of all accidents before starting to wear helmets; this reduces by 0.8 · 0.6 · 0.40 = 0.19; ceteris paribus the number of injuries is reduced by 19 %). This is less than one would expect if aggregate effects were strictly proportional to individual effects. In the latter case, one would expect the number of head injuries to be reduced by 0.40 · 0.60 = 0.24 = 24 %. If there is very selective recruitment, aggregate effects could be substantially smaller than implied by the individual protective effects of bicycle helmets.

Another possible reason why the aggregate effects of bicycle helmets could be smaller than expected on the basis of individual effects is behavioural adaptation. Once helmeted, cyclists might feel better protected and adopt more risky riding behaviour. While this cannot be ruled out, there is no direct evidence for it and performing a convincing study of such behavioural adaptation would be very difficult. The issue remains unresolved (Robinson 2007).

6 CONCLUSIONS

Based on the studies reviewed in this paper, the following conclusions can be drawn:

1. A re-analysis has been performed of a meta-analysis of the protective effects of bicycle helmets reported in Accident Analysis and Prevention (Attewell et al. 2001). The original analysis was found to be influenced by publication bias and time-trend bias that were not controlled for.
2. When these sources of bias are controlled for, the protective effects attributed to bicycle helmets become smaller than originally estimated.
3. When the analysis is updated by adding four new studies, the protective effects attributed to bicycle helmets are further reduced. According to the new studies, no statistically significant overall effect of bicycle helmets could be found when injuries to head, face or neck are considered as a whole.
4. The findings of this study are inconsistent with other meta-analyses, in particular a Cochrane review published in 2009. However, the study inclusion criteria applied in the Cochrane review are debatable.

ACKNOWLEDGEMENT

This study was funded by the Research Council of Norway, grant number 208437.

REFERENCES

Amoros, E., Chiron, M., Ndiaye, A., Laumon, B. 2009. Cyclistes Victimes d’Accidents (CVA) Partie 2: Études cas-témoins. Effet du casque sur les blessures à la tête, à la face et au cou. Convention InVS J06-24. INRETS, Lyon.

Attewell, R. G., Glase, K., McFadden, M. 2001. Bicycle helmet efficacy: a meta-analysis. Accident Analysis and Prevention, 33, 345-352.

Borenstein, M., Hedges, L. V., Higgins, J. T., Rothstein, H. 2009. Introduction to meta-analysis. John Wiley and Sons, New York.

Department for Transport. 2002. Bicycle helmets – a review of their effectiveness: a critical review of the literature. Road Safety Research Report No. 30. Department for Transport, London.

Duval, S., Tweedie, R. 2000A. Trim and fill: a simple funnel plot based method of testing and adjusting for publication bias in meta-analysis. Journal of the American Statistical Association, 95, 89-98.

Duval, S., Tweedie, R. 2000B. A non-parametric trim and fill method of assessing publication bias in meta-analysis. Biometrics, 56, 455-463.

Duval, S. 2005. The trim and fill method. In Rothstein, H., Sutton, A. J., Borenstein, M. (Eds): Publication bias in meta-analysis: Prevention, assessment and adjustments, 127-144. John Wiley and Sons, Chichester.

Elvik, R. 2005. Can we trust the results of meta-analyses? A systematic approach to sensitivity analysis in meta-analyses. Transportation Research Record, 1980, 221-229.

Elvik, R., Høye, A., Vaa, T., Sørensen, M. The Handbook of Road Safety Measures. Second edition. Emerald Publishing, Bingley.

Hansen, K. S., Engesæter, L. B., Viste, A. 2003. Protective effect of different types of bicycle helmets. Traffic Injury Prevention, 4, 285-290.

Hausotter, W. 2000. Fahrradunfälle mit und ohne Fahrradhelm. Versicherungsmedizin, 52, 28-32.

Heng, K. W. J., Lee, A. H., Zhu, S., Tham, K. Y., Seow, E. 2006. Helmet use and bicycle-related trauma in patients presenting to an acute hospital in Singapore. Singapore Medical Journal, 47, 367-372.

Littell, J. H., Corcoran, J., Pillai, V. 2008. Systematic Reviews and Meta-Analysis. Oxford University Press, Oxford.

Rivara, F. P., Thompson, D. C., Thompson, R. S. 1997. Epidemiology of bicycle injuries and risk factors for serious injury. Injury Prevention, 3, 110-114.

Robinson, D. J. 2007. Bicycle helmet legislation: can we reach a consensus? Accident Analysis and Prevention, 39, 86-93.

Rothstein, H., Sutton, A. J., Borenstein, M. 2005. Publication bias in meta-analysis. Prevention, assessment and adjustments. John Wiley and Sons, Chichester.

Sterne, J. A., Egger, M. 2001. Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. Journal of Clinical Epidemiology, 54, 1046-1055.

Sweeting, M. J., Sutton, A. J., Lambert, P. C. 2004. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. Statistics in Medicine, 23, 1351-1375.

Thompson, D. C., Rivara, F., Thompson, R. 2009. Helmets for preventing head and facial injuries in bicyclists. Cochrane Review. The Cochrane Library, Issue 1, 2009.

LIST OF FIGURES AND TABLES

Figure 1:

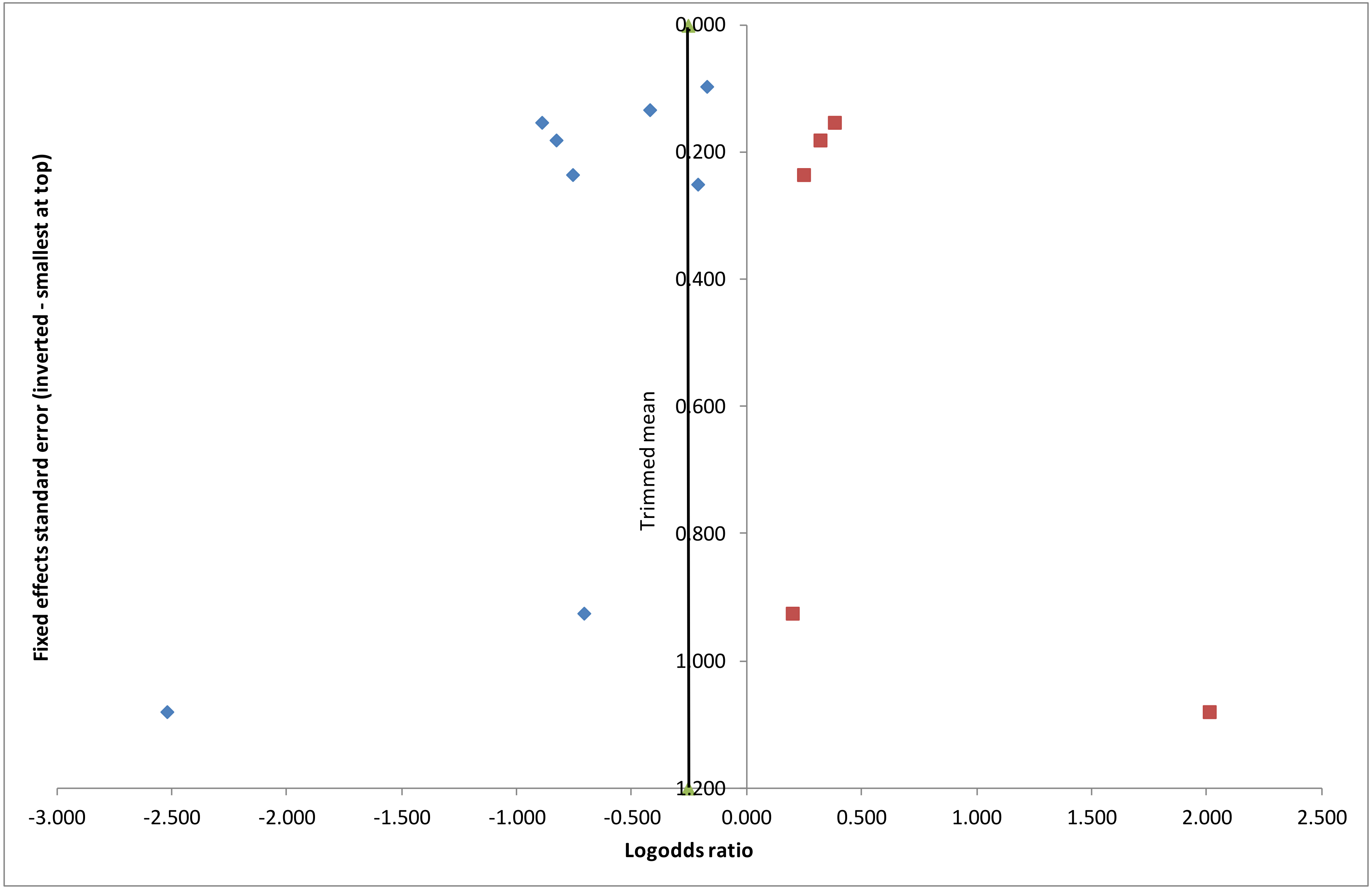
0.000
0.200
0.400
0.600
0.800
1.000
1.200
Trimmed mean
Fixed effects standard error (inverted - smallest at top)
-3.000
-2.500
-2.000
-1.500
-1.000
-0.500
0.000
0.500
1.000
1.500
2.000
2.500
Logodds ratio

Figure 2:

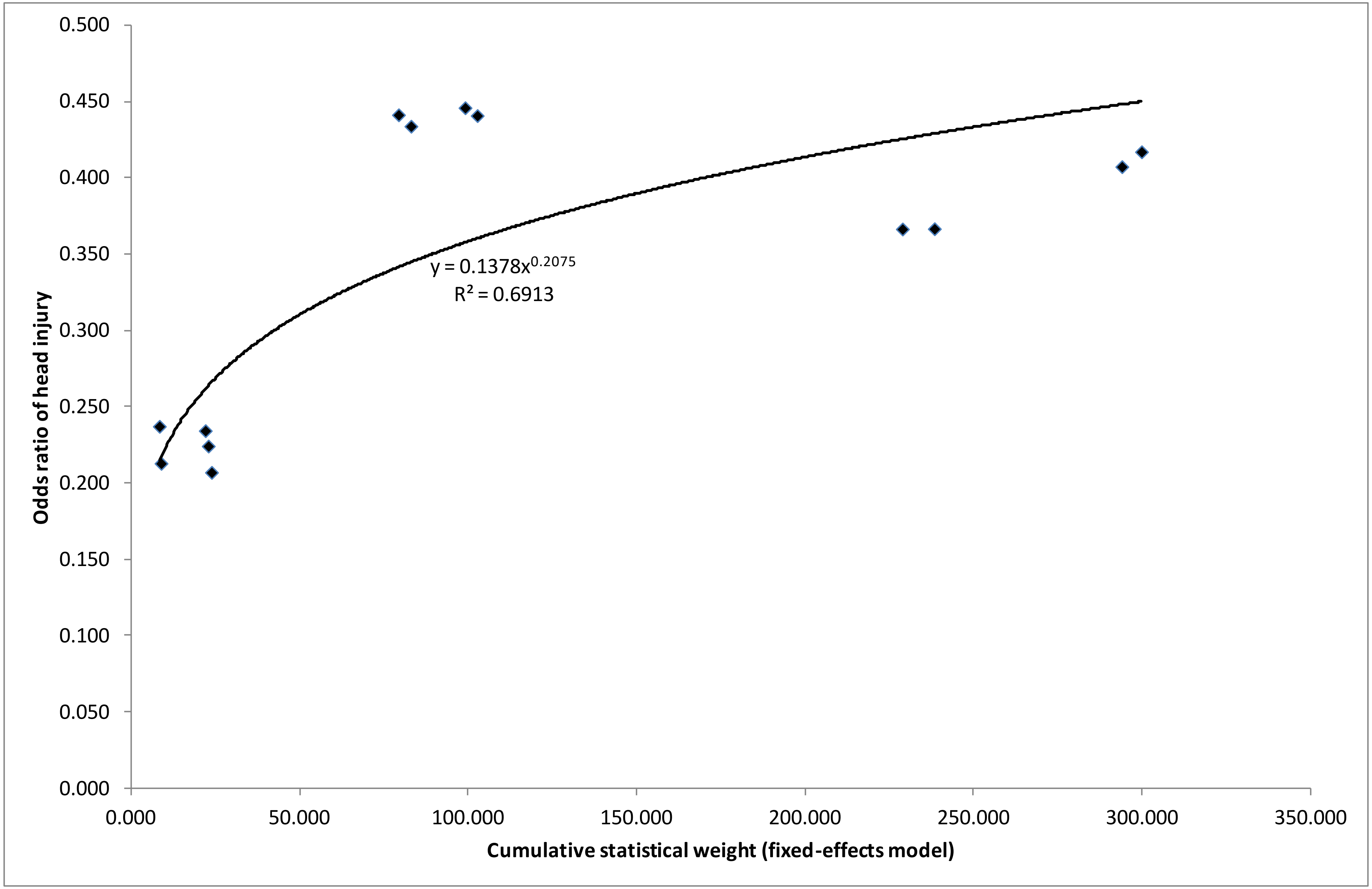
0.500
0.450
0.400
0.350
0.300
0.250
0.200
0.150
0.100
0.050
0.000
Odds ratio of head injury
$y = 0.1378x^{0.2075}$
$R^2 = 0.6913$
0.000
50.000
100.000
150.000
200.000
250.000
300.000
350.000
Cumulative statistical weight (fixed-effects model)

Figure 3:

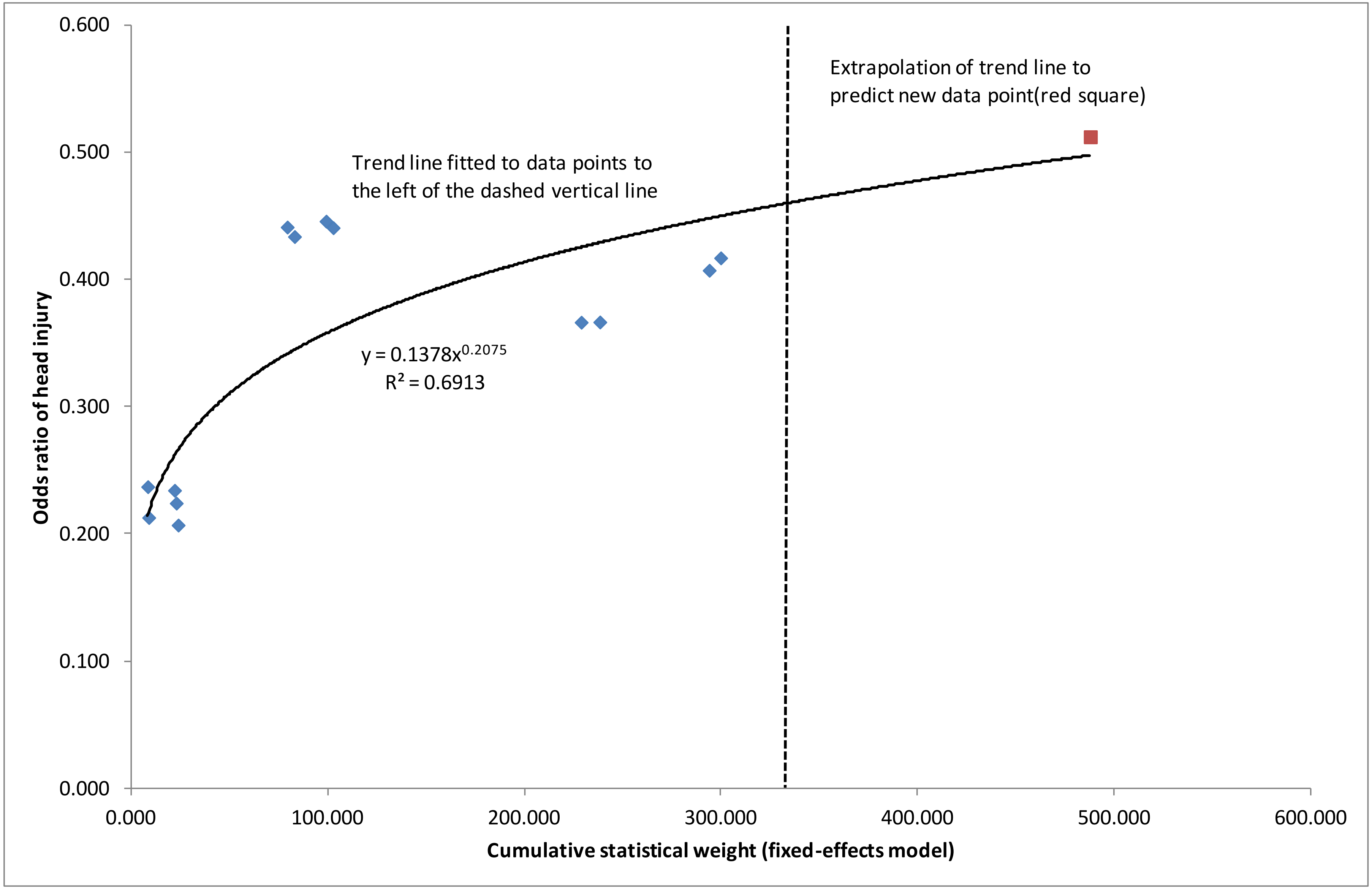
0.600
0.500
0.400
0.300
0.200
0.100
0.000
Odds ratio of head injury
0.000
100.000
200.000
300.000
400.000
500.000
600.000
Cumulative statistical weight (fixed-effects model)
Trend line fitted to data points to the left of the dashed vertical line
Extrapolation of trend line to predict new data point(red square)
y = 0.1378x^0.2075
R² = 0.6913

Figure 4:

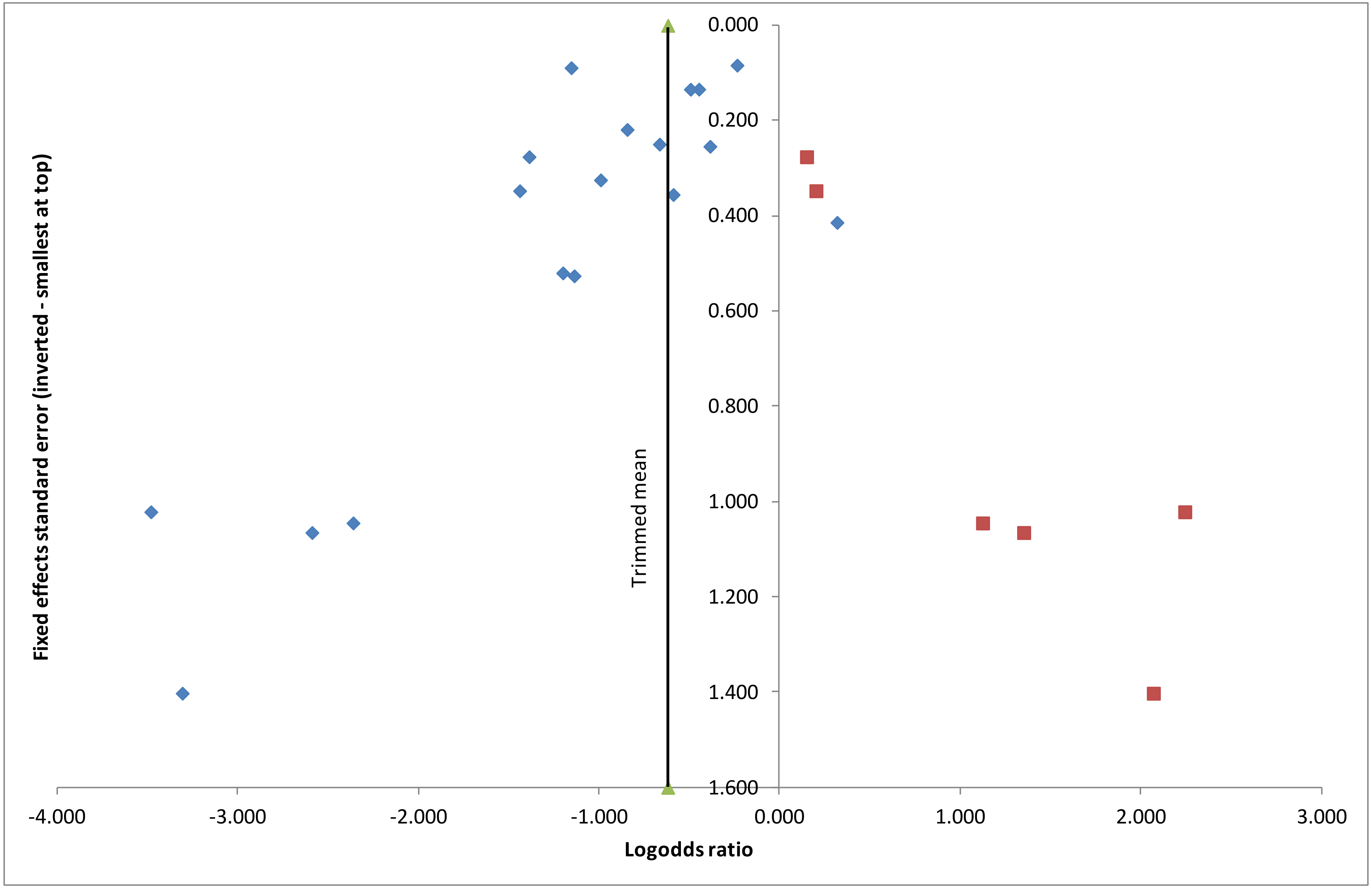
Fixed effects standard error (inverted - smallest at top)
Trimmed mean
Logodds ratio
0.000
0.200
0.400
0.600
0.800
1.000
1.200
1.400
1.600
-4.000
-3.000
-2.000
-1.000
0.000
1.000
2.000
3.000

Table 1:

	Notation for derivation		Empirical continuity correction (prior)		Data (observed cell counts)		Data and empirical continuity correction	
Group	**Event (head injury)**	**No event (no head injury)**	**Event (head injury)**	**No event (no head injury)**	**Event (head injury)**	**No event (no head injury)**	**Event (head injury)**	**No event (no head injury)**
Treatment (helmet)	$0 + k_T$	$n_T + k_T$	0.358	8.358	0	8	0.358	8.358
Control (no helmet)	$0 + k_C$	n_T x $R + k_C$	0.642	6.642	7	6	7.642	6.642

Table 2:

Ln OR (sorted)	Fixed effects weight	Ln OR * weight	Centered difference	Rank (iteration 1)	Centered difference	Rank (iteration 2)	Centered difference	Rank (iteration 3)	Centered difference	Rank (iteration 4)
0.174	108.465	18.908	-0.287	-4	-0.129	-3	-0.084	-2	-0.081	-2
0.214	16.016	3.422	-0.247	-2	-0.090	-1	-0.045	-1	-0.042	-1
0.422	56.798	23.995	-0.039	-1	0.119	2	0.164	3	0.167	3
0.709	1.169	0.829	0.248	3	0.405	4	0.450	4	0.453	4
0.758	18.144	13.750	0.297	5	0.454	5	0.499	5	0.502	5
0.829	30.740	25.498	0.368	6	0.526	6	0.571	6	0.574	6
0.892	42.952	38.313	0.431	7	0.588	7	0.634	7	0.636	7
2.522	0.858	2.163	2.061	8	2.219	8	2.264	8	2.267	8
Sum	275.142	126.878								
Log-odds ratio			0.461		0.304		0.258		0.256	
L (#)				2.93		3.73		4		4
R (§)				3		4		5		5

(#) L = estimator of publication bias based on $L_0 = \frac{4T_n - n(n+1)}{2n-1}$, where n = number of estimates (8) and T_n = sum of positive ranks (e.g. 3+5+6+7+8)

(§) R = estimator of publication bias based on $R_0 = \gamma - 1$, where γ equals the number of positive ranks greater than the largest absolute value of the negative ranks, e.g. (5, 6, 7, 8 > 4), 4 – 1 = 3

Table 3:

	Summary odds ratios; 95% confidence intervals ( ); number of estimates [ ]	Summary odds ratios; fixed-effects model; 95% confidence interval in parentheses ( ); number of estimates in brackets [ ]		Summary odds ratios; random-effects model; 95% confidence interval in parentheses ( ); number of estimates in brackets [ ]	
Type of injury	**Original study**	**Re-analysis, not adjusting for publication bias**	**Re-analysis, adjusting for publication bias**	**Re-analysis, not adjusting for publication bias**	**Re-analysis, adjusting for publication bias**
Head injury	0.40 (0,29, 0.55) [12]	0.42 (0.37, 0.47) [13]	0.43 (0.38, 0.48) [16]	0.38 (0.28, 0.53) [13]	0.43 (0.33, 0.59) [16]
Brain injury	0.42 (0.26, 0.67) [8]	0.40 (0.32, 0.50) [8]	0.40 (0.33, 0.50) [9]	0.42 (0.26, 0.67) [8]	0.42 (0.26, 0.67) [8]
Facial injury	0.53 (0.39, 0.73) [7]	0.63 (0.56, 0.71) [8]	0.77 (0.70, 0.86) [13]	0.56 (0.42, 0.74) [8]	0.58 (0.44, 0.77) [9]
Neck injury	1.36 (1.00, 1.86) [3]	1.36 (1.00, 1.86) [3]	Not adjusted	1.40 (0.97, 2.02) [3]	Not adjusted
Fatal injury	0.27 (0.10, 0.72) [6]	0.23 (0.08, 0.64) [6]	0.37 (0.15, 0.90) [10]	0.23 (0.08, 0.64) [6]	0.37 (0.15, 0.90) [10]

Table 4:

		Summary odds ratios; fixed-effects model; 95% confidence interval in parentheses ( ); number of estimates in brackets [ ]		Summary odds ratios; random-effects model; 95% confidence interval in parentheses ( ); number of estimates in brackets [ ]	
Type of injury	**Studies included**	**Re-analysis, not adjusting for publication bias**	**Re-analysis, adjusting for publication bias**	**Re-analysis, not adjusting for publication bias**	**Re-analysis, adjusting for publication bias**
Head injury	As in Attewell et al.	0.42 (0.37, 0.47) [13]	0.43 (0.38, 0,48) [16]	0.38 (0.28, 0.53) [13]	0.43 (0.31, 0.59) [16]
	New studies	0.71 (0.62, 0.82) [5]	0.78 (0.68, 0.89) [8]	0.58 (0.41, 0.84) [5]	0.70 (0.51, 0.97) [7]
	All studies	0.51 (0.47, 0.56) [18]	0.54 (0.49, 0.59) [24]	0.43 (0.33, 0.56) [18]	0.50 (0.39, 0.65) [23]
Facial injury	As in Attewell et al.	0.63 (0.56, 0.71) [8]	0.77 (0.70, 0.86) [13]	0.56 (0.42, 0.74) [8]	0.58 (0.44, 0.77) [9]
	New studies	0.94 (0.81, 1.09) [3]	Too few to adjust	1.20 (0.72, 2.00) [3]	Too few to adjust
	All studies	0.74 (0.67, 0.81) [11]	0.79 (0.72, 0.86) [13]	0.71 (0.55, 0.92) [11]	0.79 (0.62, 1.01) [13]
Neck injury	As in Attewell et al.	1.36 (1.00, 1.86) [3]	Too few to adjust	1.40 (0.97, 2.02) [3]	Too few to adjust
	New studies	1.24 (0.98, 1.57) [1]	Too few to adjust	1.24 (0.98, 1.57) [1]	Too few to adjust
	All studies	1.28 (1.06, 1.55) [4]	Too few to adjust	1.28 (1.06, 1.55) [4]	Too few to adjust
Head, face or neck injury	As in Attewell et al.	0.54 (0.50, 0.59) [24]	0.55 (0.51, 0.59) [28]	0.52 (0.41, 0.66) [24]	0.57 (0.45, 0.72) [28]
	New studies	0.87 (0.79, 0.95) [9]	0.87 (0.79, 0.96) [10]	0.85 (0.66, 1.11) [9]	0.88 (0.68, 1.14) [10]
	All studies	0.66 (0.62, 0.70) [33]	0.67 (0.63, 0.71) [39]	0.60 (0.49, 0.73) [33]	0.67 (0.56, 0.82) [39]