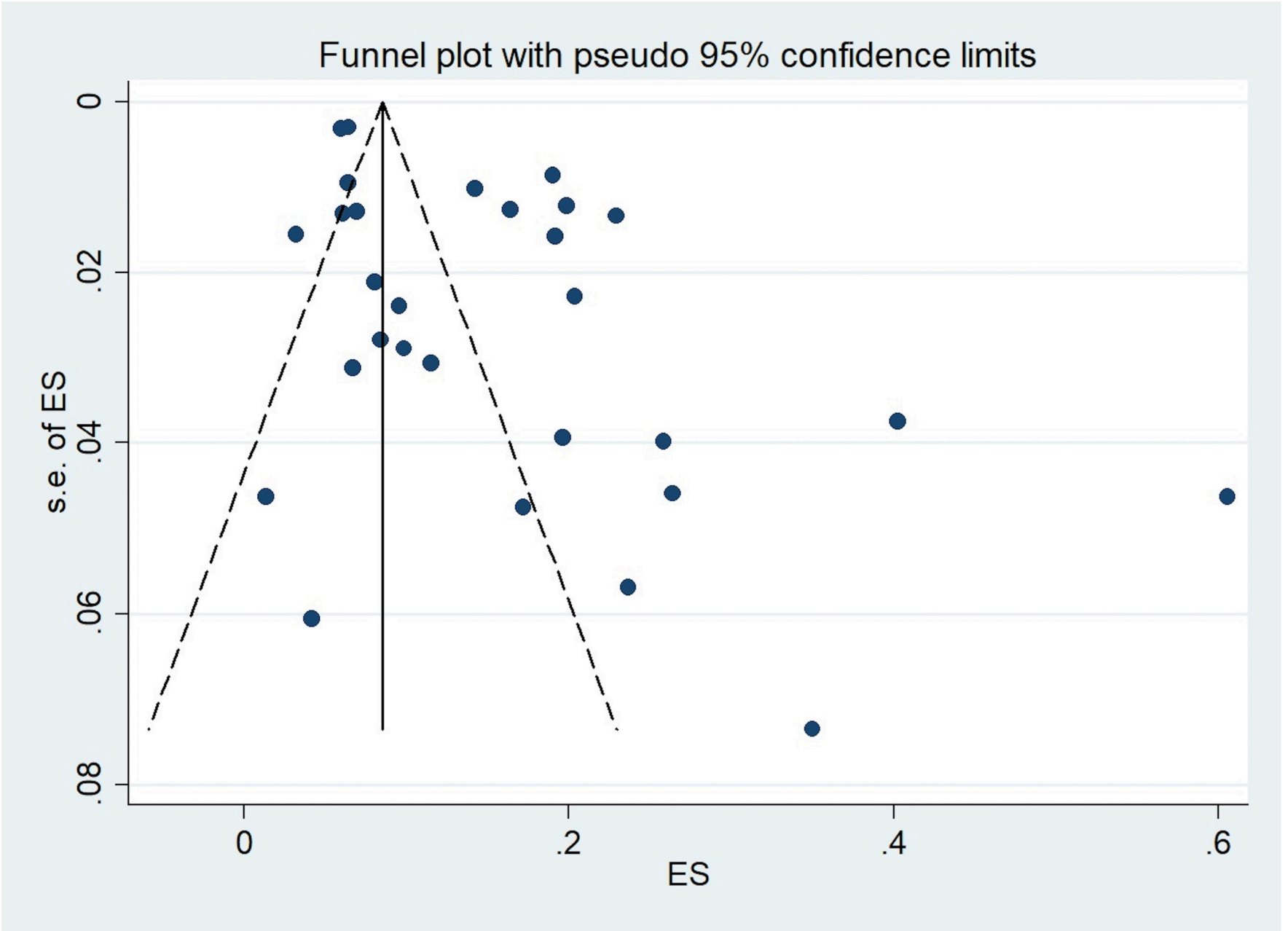
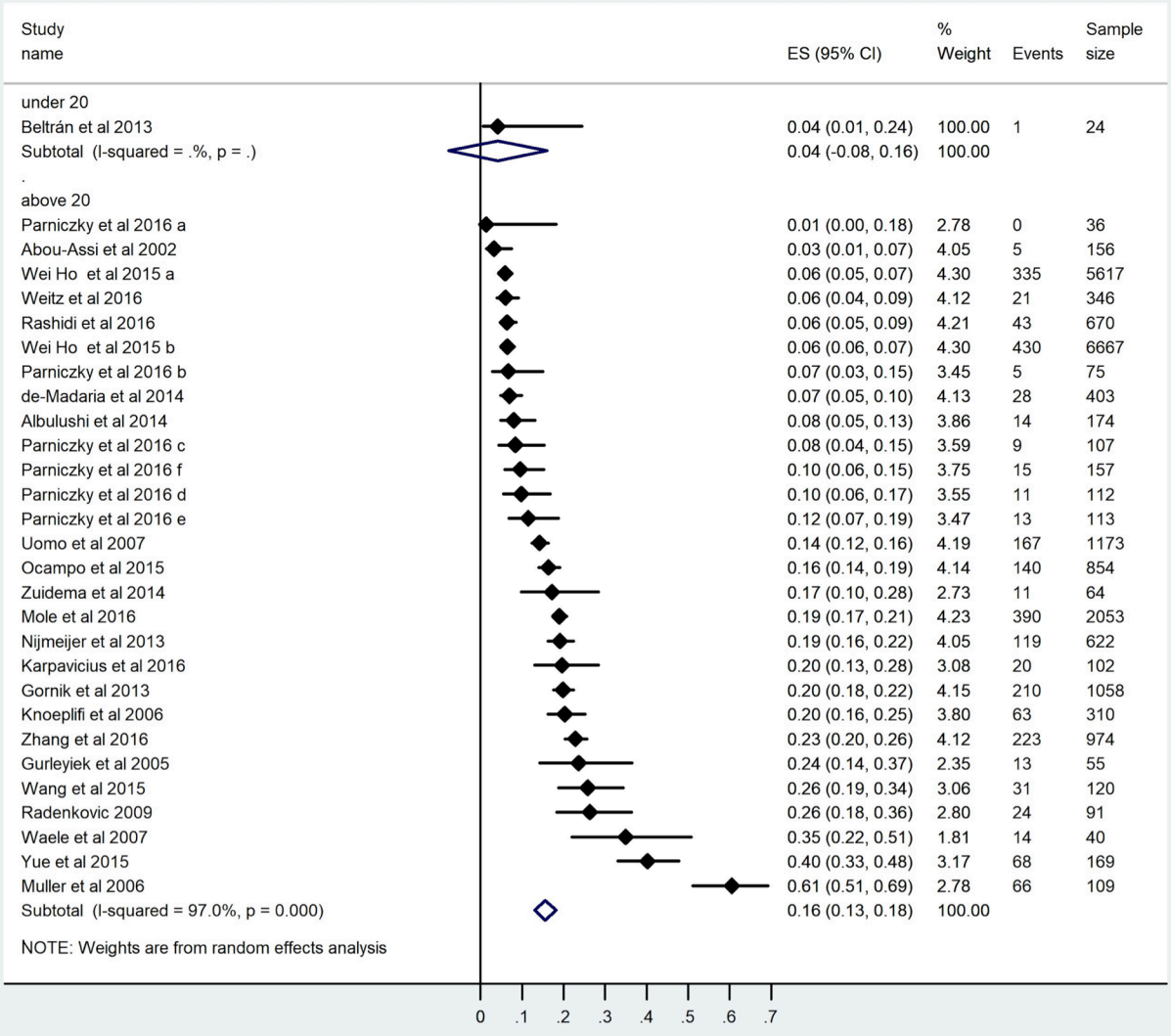
Assessment of the modified Newcastle-Ottawa scale in our study				
	(i) selection of study groups			
	S1	S2	S3	S4
low risk	non-selected etiology	confirmed AP diagnosis	AP diagnosis is confirmed using the latest guidelines	non-selected severity cases
high risk	selected etiology	unstable diagnosis concering the IAP/APA guideline	AP diagnosis is made using not the latest guidelines	selected severity cases
unclear risk	not mentioned	not mentioned	no information about diagnosis critera	no information
	(ii) comparability of the groups			
	C1			
low risk	exact age range is provided			
unlcear risk	age range is not provided			
	(iii) outcome of interest			
	O1.1	O1.2	O2	O3
low risk	severity assigned by the latest guidelines	described mortality (in-hospital and pancreas-related)	adequate follow-up for outcome occurrence (morality)	adequate follow-up for outcome occurrence (severity)
high risk	severity assigned by not the latest guidelines	in-hospital and pancreas-related	not adequate follow-up for outcome occurrence (morality)	not adequate follow-up for outcome occurrence (morality)
unclear risk	not mentioned	not described mortality	no information about follow-up	no information about follow-up

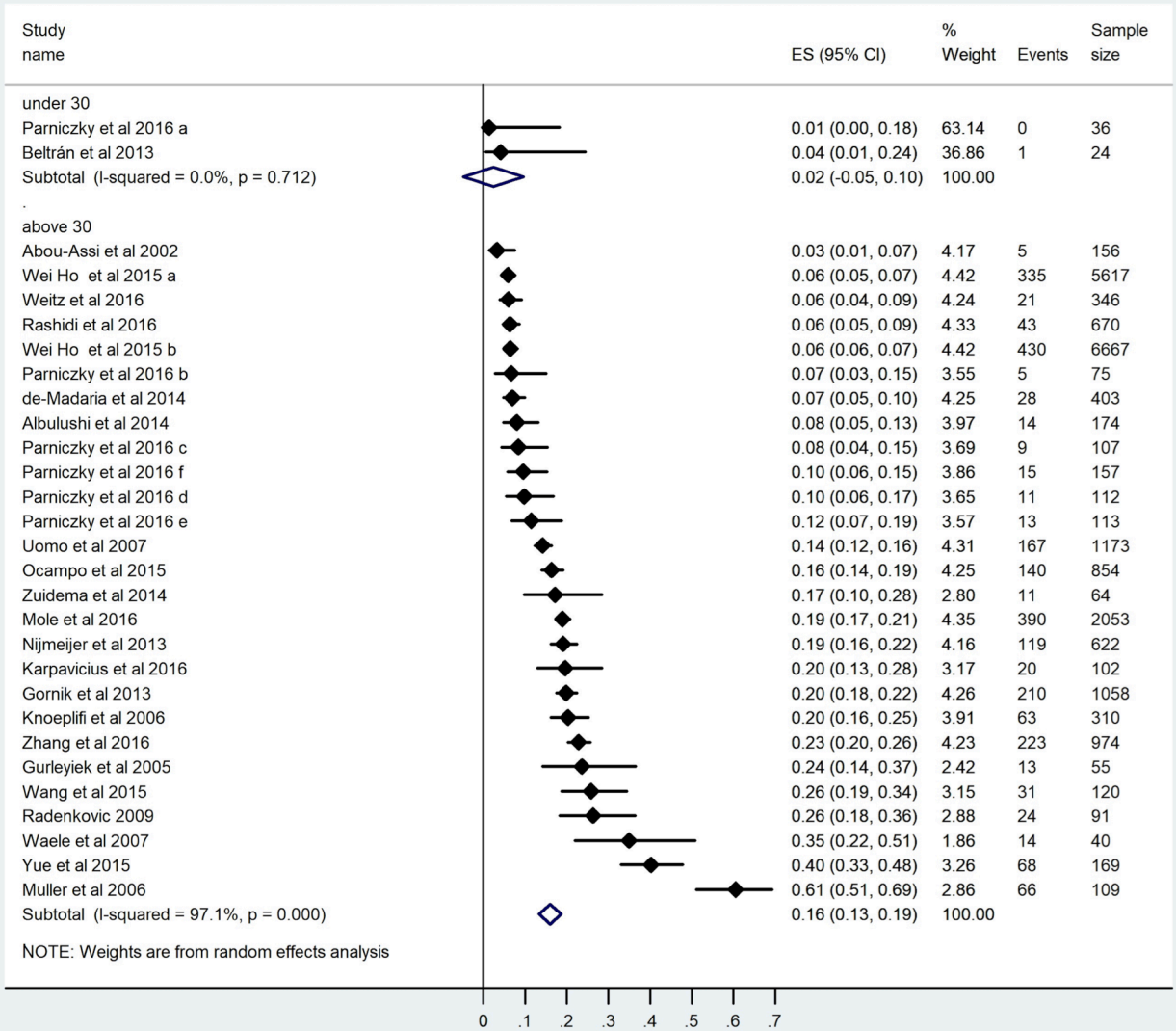
Suppl. Fig. 1. Assessment of the modified Newcastle-Ottawa scale. Low risk equals one point, low and unclear risk equals no point (see Table 2).



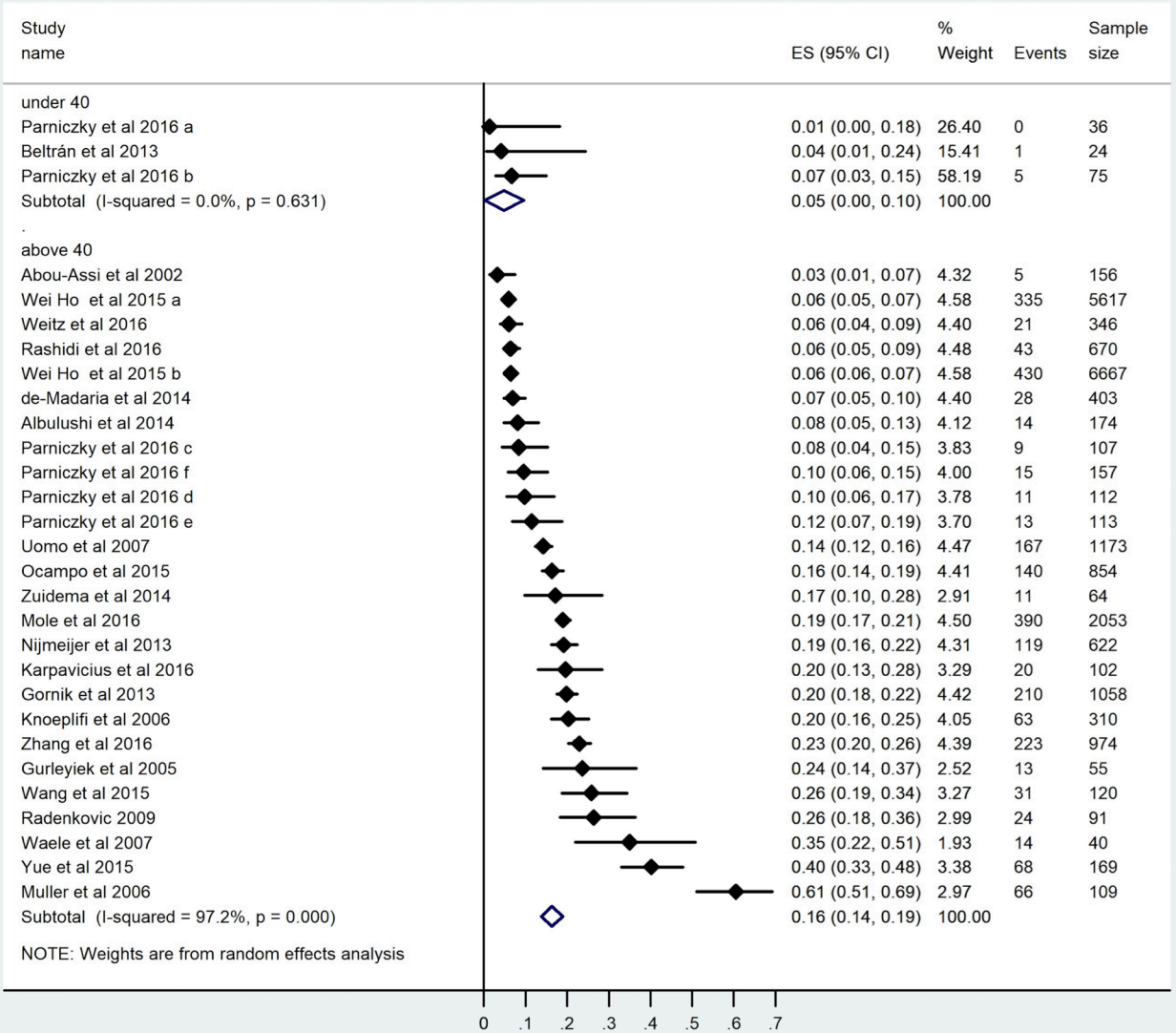
Suppl. Fig. 2. Funnel plot of severity in terms of publication bias. Funnel plots represent the standard error (SE) plotted against event rates (ES) for each study. The dotted line shows the 95% confidence limits. Plots are mostly on the right side showing that publication bias might present (CI: 1.961–6.728; p=0.001).



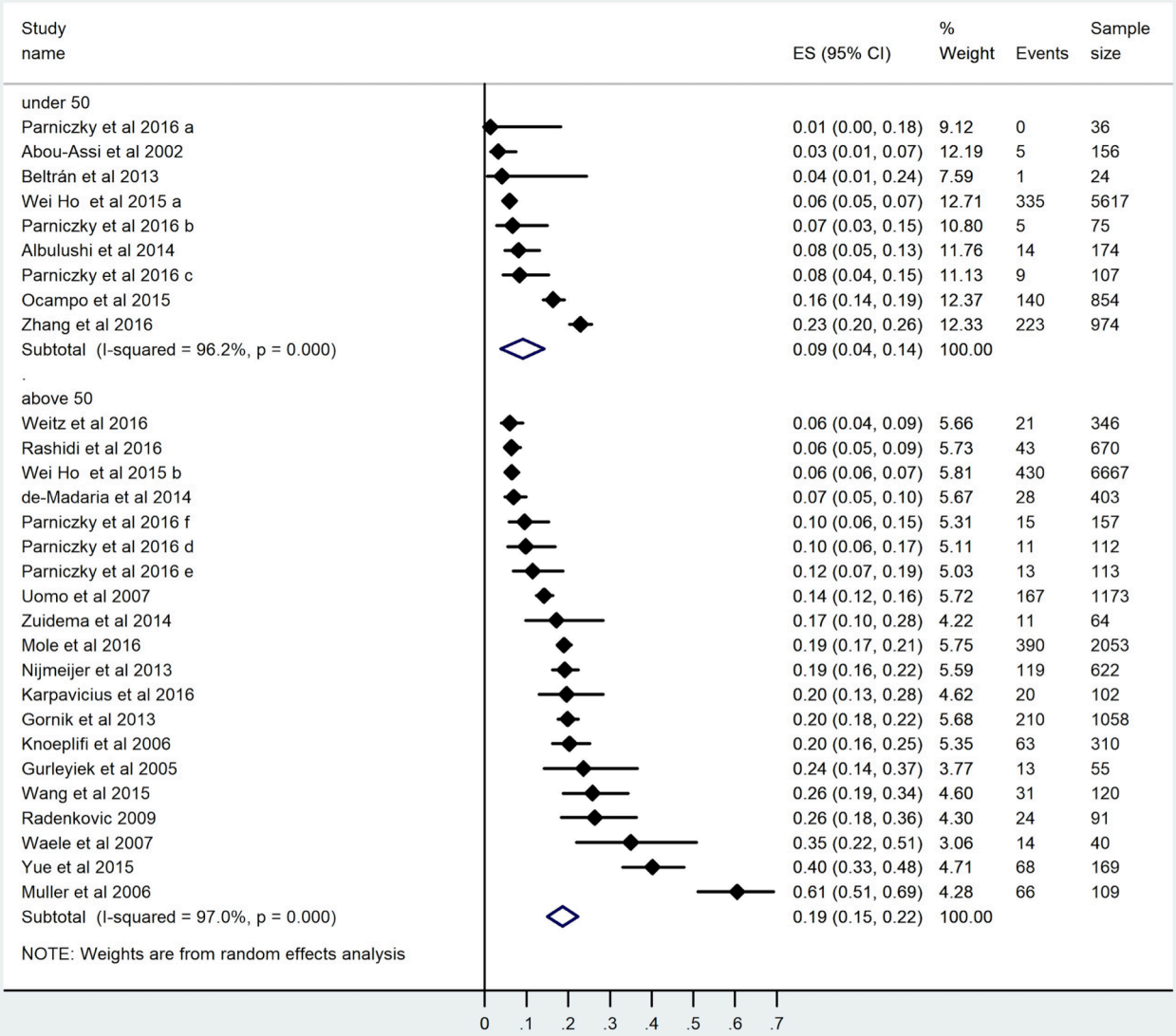
Suppl. Fig. 3. Forest plot of studies evaluating severity at age U20 compared to A20. Full diamonds show the weighted event rates for studies respectively, line represents the 95% confidence interval (CI), and empty diamonds show the pooled results of severe cases. A non-significant difference can be observed in severity under 20 and above 20 (p=0.188).



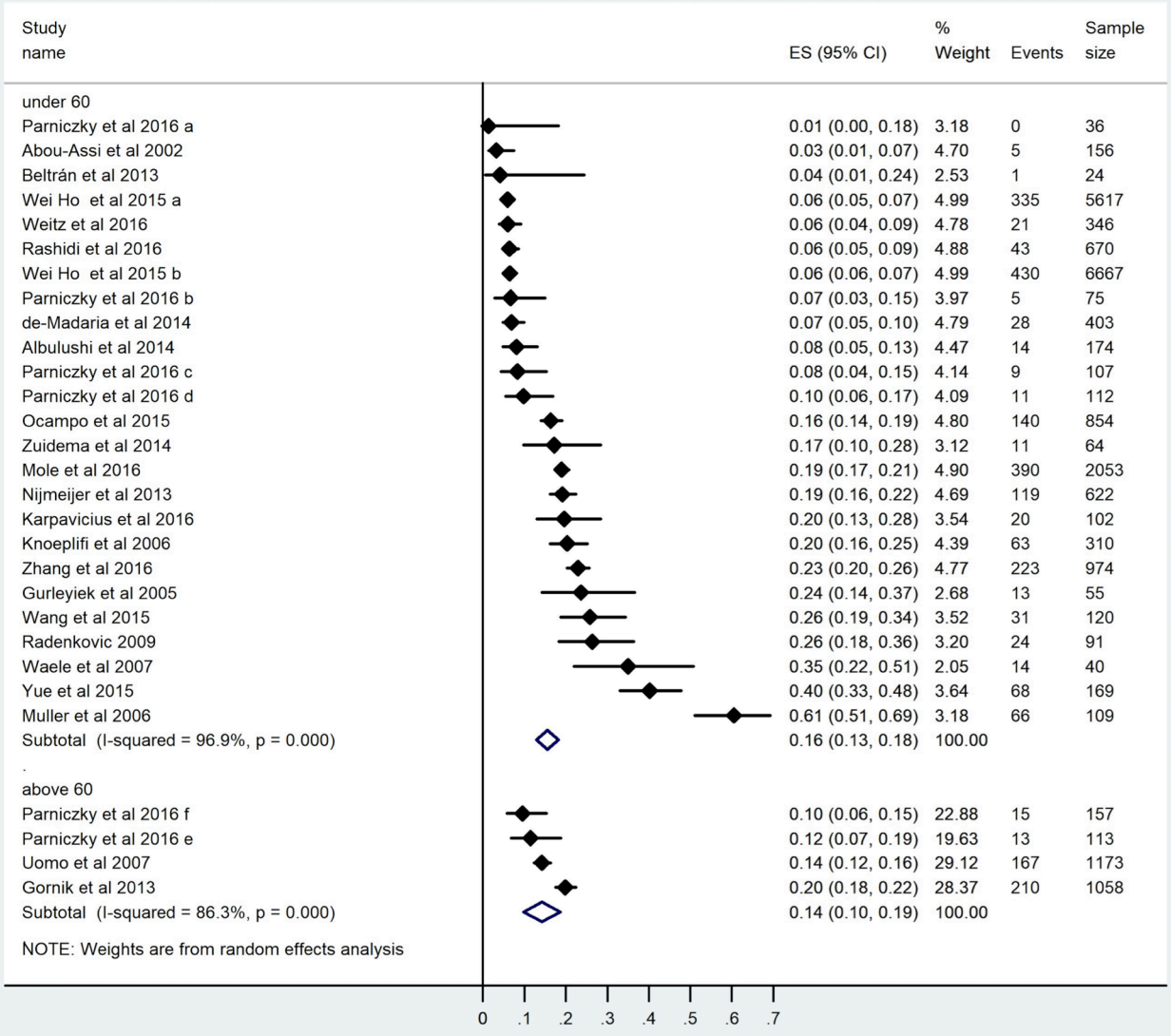
Suppl. Fig. 4. Forest plot of studies evaluating severity at age U30 compared to A30. Full diamonds show the weighted event rates for studies respectively, line represents the 95% confidence interval (CI), and empty diamonds show the pooled results of severe cases. A non-significant difference can be observed in severity under 30 and above 30 (p=0.036).



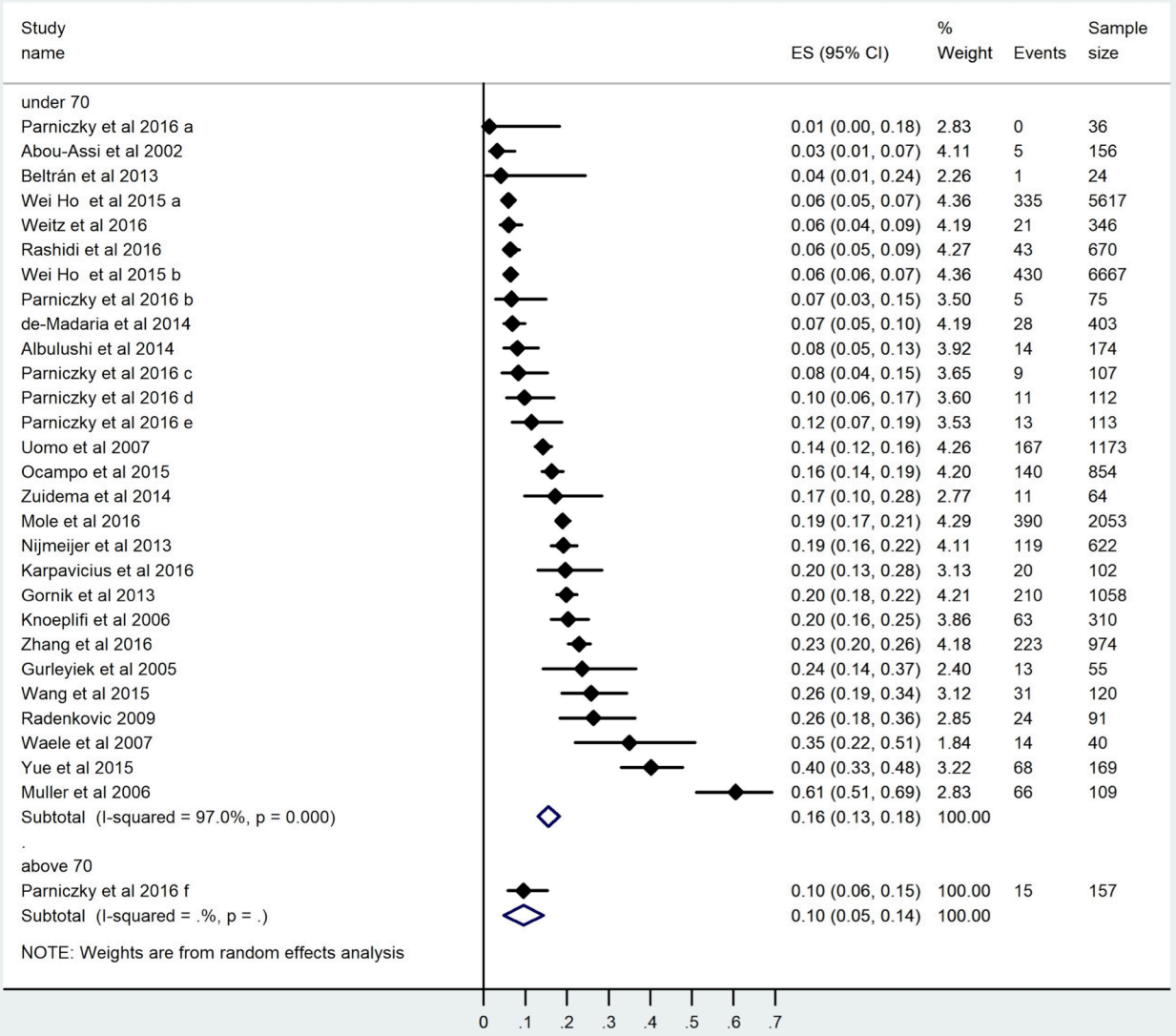
Suppl. Fig. 5. Forest plot of studies evaluating severity at age U40 compared to A40. Full diamonds show the weighted event rates for studies respectively, line represents the 95% confidence interval (CI), and empty diamonds show the pooled results of severe cases A significant difference can be observed in severity under 40 and above 40 (p=0.009).



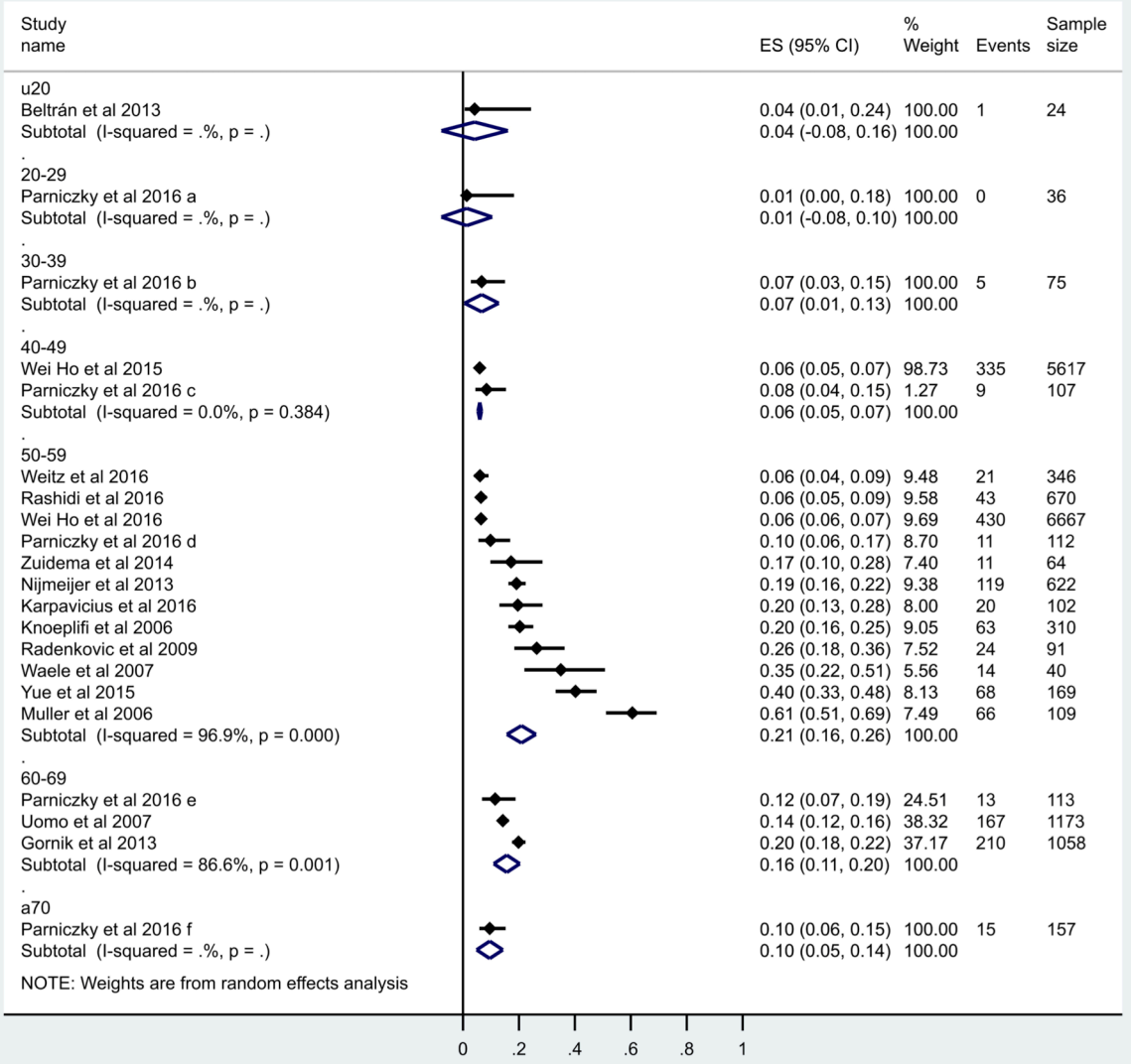
Suppl. Fig. 6. Forest plot of studies evaluating severity at age U50 compared to A50. Full diamonds show the weighted event rates for studies respectively, line represents the 95% confidence interval (CI), and empty diamonds show the pooled results of severe cases. A significant difference can be observed in severity under 50 and above 50 (p=0.021).



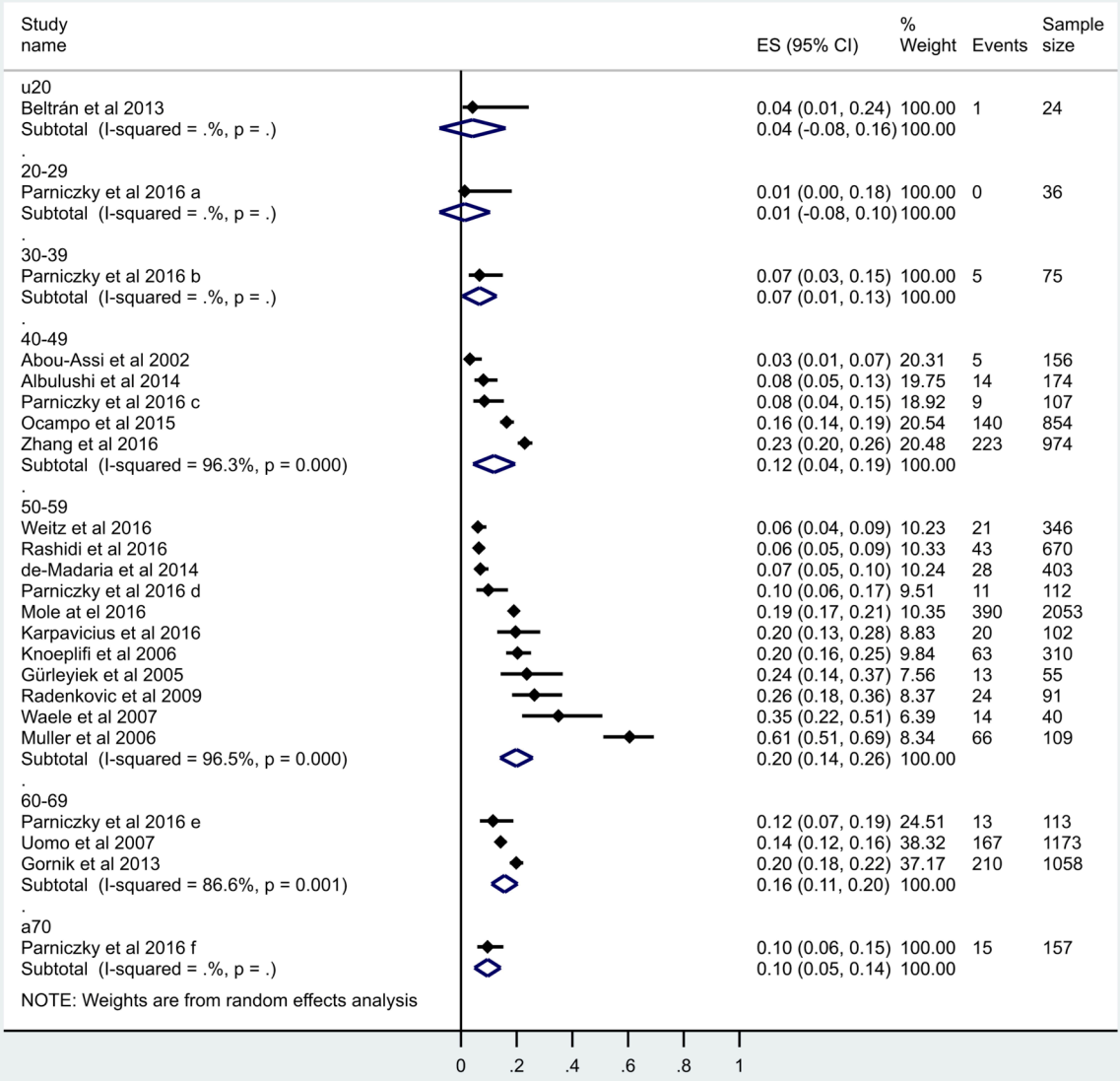
Suppl. Fig. 7. Forest plot of studies evaluating severity at age U60 compared to A60. Full diamonds show the weighted event rates for studies respectively, line represents the 95% confidence interval (CI), and empty diamonds show the pooled results of severe cases. A non-significant difference can be observed in severity under 60 and above 60 (p=0.994).



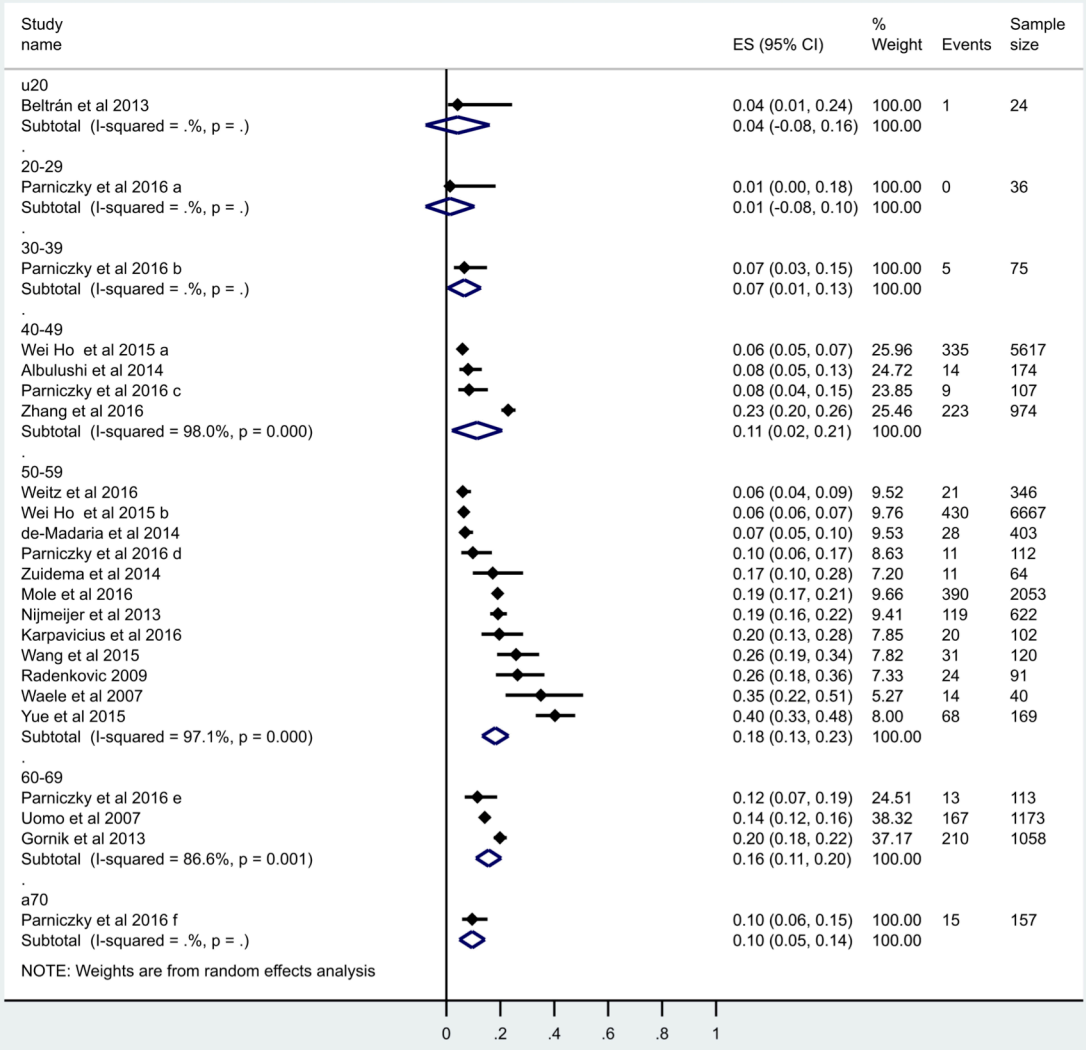
Suppl. Fig. 8. Forest plot of studies evaluating severity at age U70 compared to A70. Full diamonds show the weighted event rates for studies respectively, line represents the 95% confidence interval (CI), and empty diamonds show the pooled results of severe cases. A non-significant difference can be observed in severity under 70 and above 70 (p=0.133).



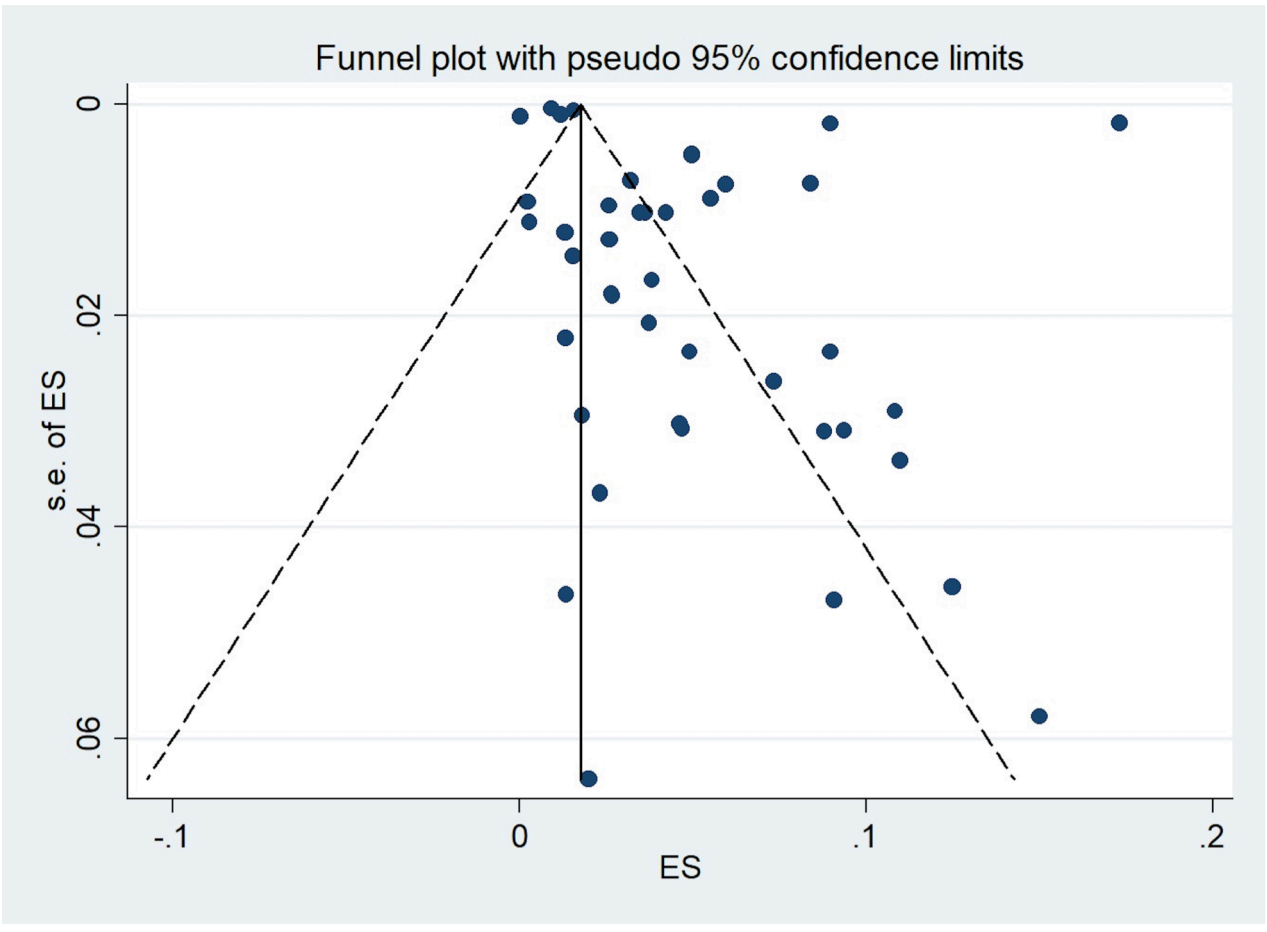
Suppl. Fig. 9. Forest plot of studies to decrease the heterogeneity. Articles where severity was assessed by the Atlanta or the revised Atlanta classification. Full diamonds show the weighted event rates for studies respectively, line represents the 95% confidence interval (CI), and empty diamonds show the pooled results of severe cases. Heterogeneity; I^2= 40-49: 0%, 50-59:96.9%, 60-69:86.6%.



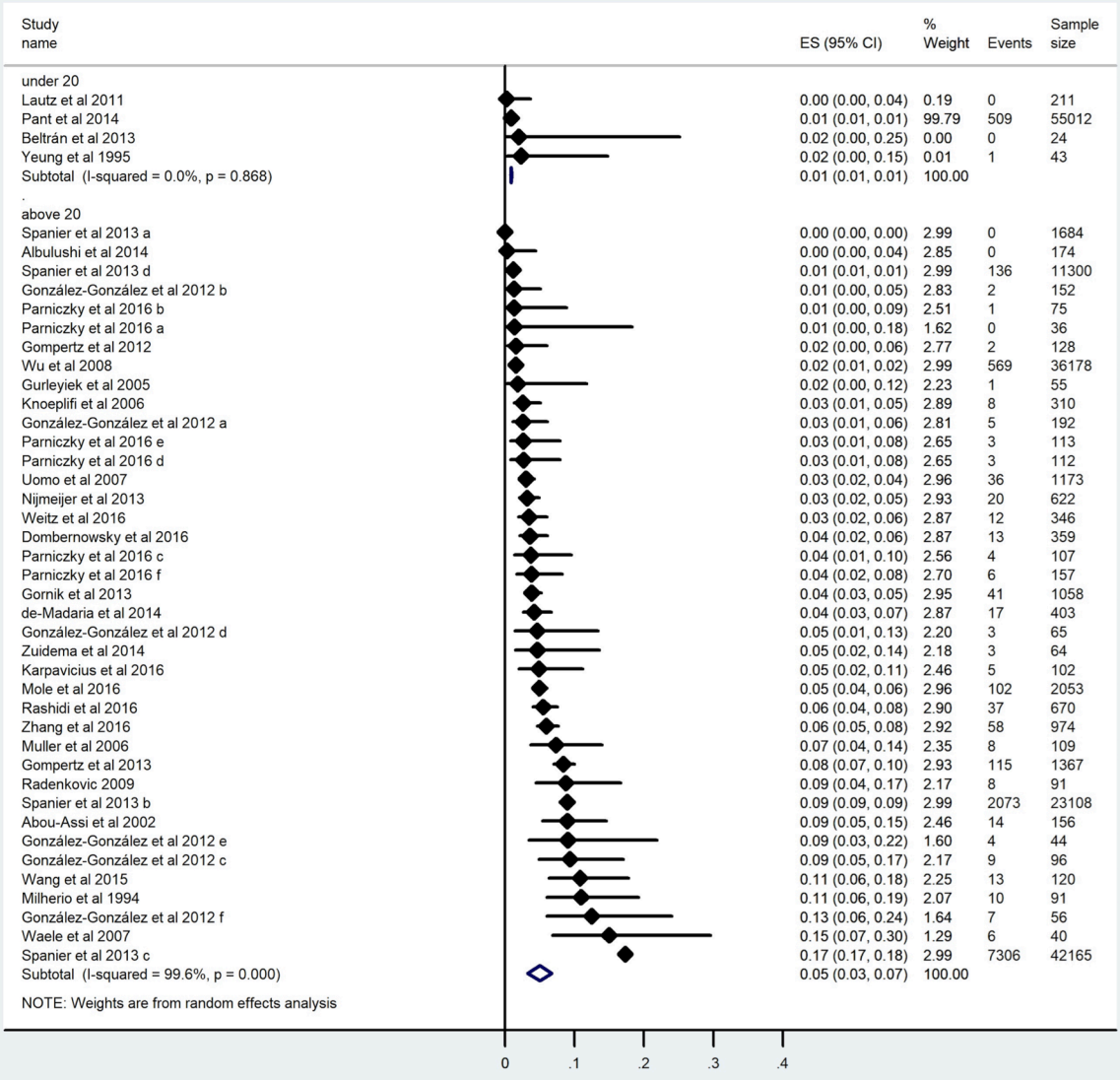
Suppl. Fig. 10. Forest plot of studies to decrease the heterogeneity. Analysis of high quality (NOS 4 and 5) studies. Full diamonds show the weighted event rates for studies respectively, line represents the 95% confidence interval (CI), and empty diamonds show the pooled results of severe cases. Heterogenity; I^2= 40-49: 96.3%, 50-59:96.5%, 60-69:86.6%.



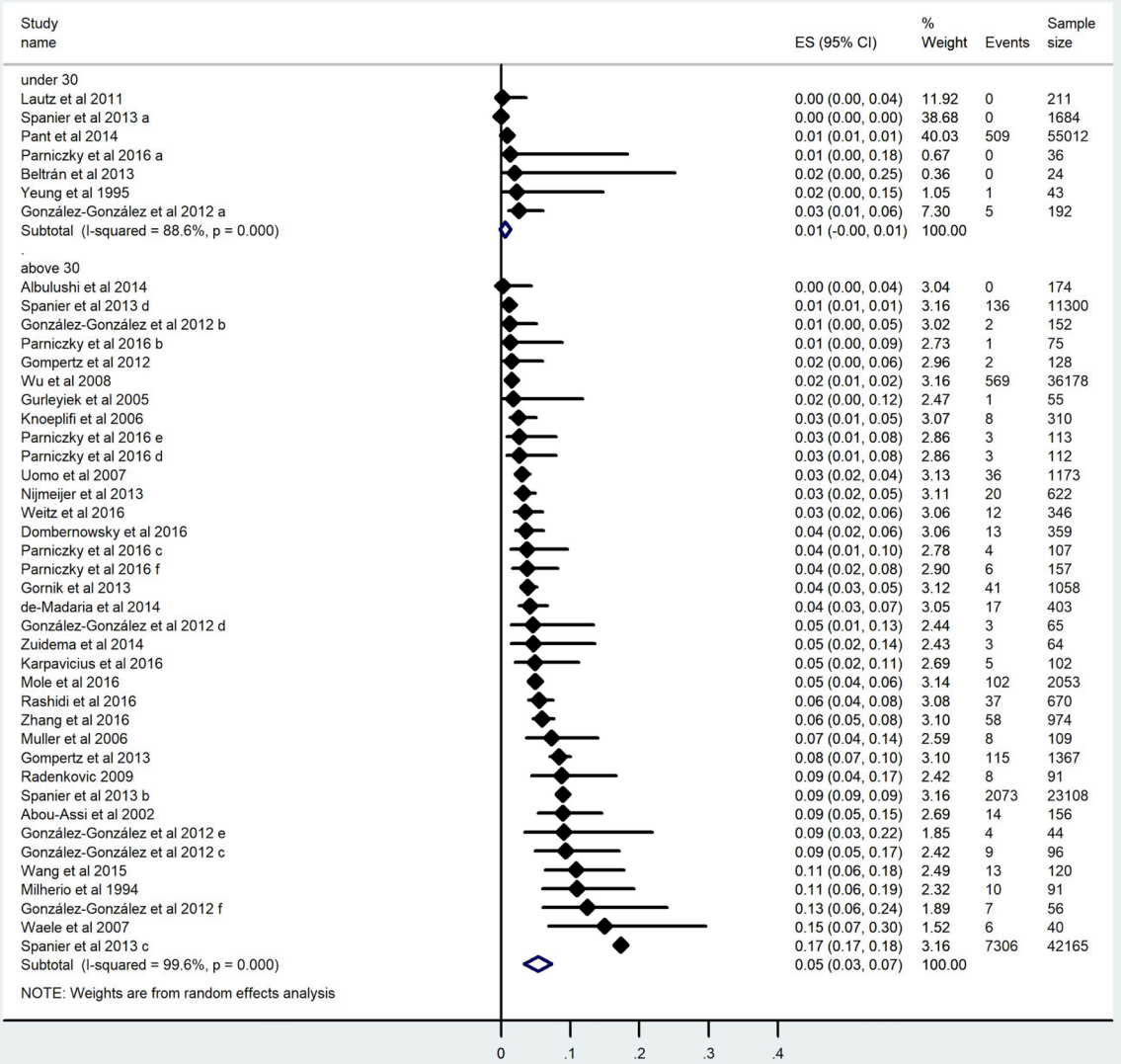
Suppl. Fig. 11. Forest plot of studies to decrease the heterogeneity. Analysis of where age ranges surly not overlap between the groups because of given age ranges concerning severity. Full diamonds show the weighted event rates for studies respectively, line represents the 95% confidence interval (CI), and empty diamonds show the pooled results of severe cases. Heterogenity; I^2= 40-49: 98%, 50-59:97.1%, 60-69:86.6%.



Suppl. Fig. 12. Funnel plot of mortality in terms of small study effect. Funnel plots represent the standard error (SE) plotted against event rates (ES) for each study. The dotted line shows the 95% confidence limits. Plots are slightly on the right side a mild asymmetry, but based on Egger's test publication bias was unlikely(CI: -0.901–9.234; p=0.104).



Suppl. Fig. 13. Forest plot of studies evaluating mortality at age U20 compared to A20. Full diamonds show the weighted event rates for studies respectively, line represents the 95% confidence interval (CI), and empty diamonds show the pooled results of severe cases. A significant difference can be observed in mortality under 20 and above 20 ($p<0.001$).



Suppl. Fig. 14. Forest plot of studies evaluating mortality at age U30 compared to A30. Full diamonds show the weighted event rates for studies respectively, line represents the 95% confidence interval (CI), and empty diamonds show the pooled results of severe cases. A significant difference can be observed in mortality under 30 and above 30 (p=0.001).