Figure E1: Effect of helmet NIV compared to facemask NIV on mortality. Observational data only. Df = degrees of freedom

Study or Subgroup	Helmet NIV Events	Total	Facemask NIV Events	Total	Weight	Risk Ratio M–H, Random, 95% CI
Antonelli 2002	4	33	18	66	14.9%	0.44 [0.16, 1.21]
Antonelli 2004	7	33	10	33	21.2%	0.70 [0.30, 1.62]
Conti 2007	6	25	10	25	20.7%	0.60 [0.26, 1.40]
Principi 2004	4	17	8	17	15.0%	0.50 [0.18, 1.35]
Rocco 2004	7	19	10	19	28.1%	0.70 [0.34, 1.45]
Total (95% CI)		**127**		**160**	**100.0%**	**0.60 [0.41, 0.89]**
Total events	28		56			

Heterogeneity. Tau² = 0.00; Chi² = 0.80, df = 4 (P = 0.94); I² = 0%
Test for overall effect: Z = 2.58 (P = 0.010)

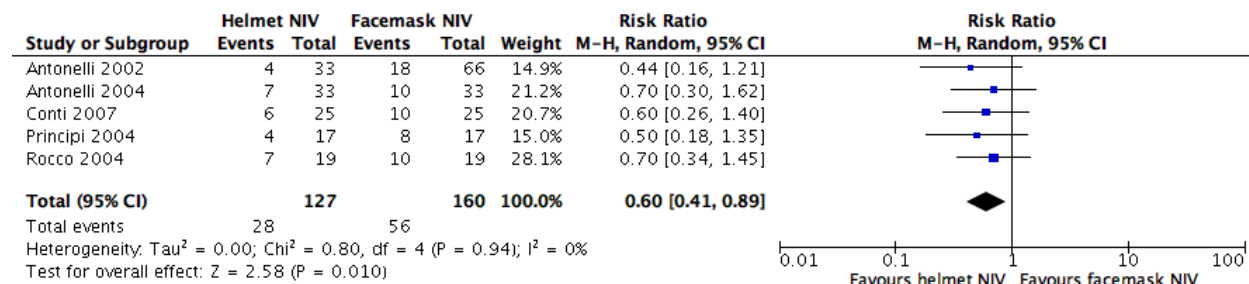Favours helmet NIV   Favours facemask NIV

Figure E2: Effect of helmet NIV compared to facemask NIV on intubation. Observational data only. Studies are grouped by type of respiratory failure. Df = degrees of freedom
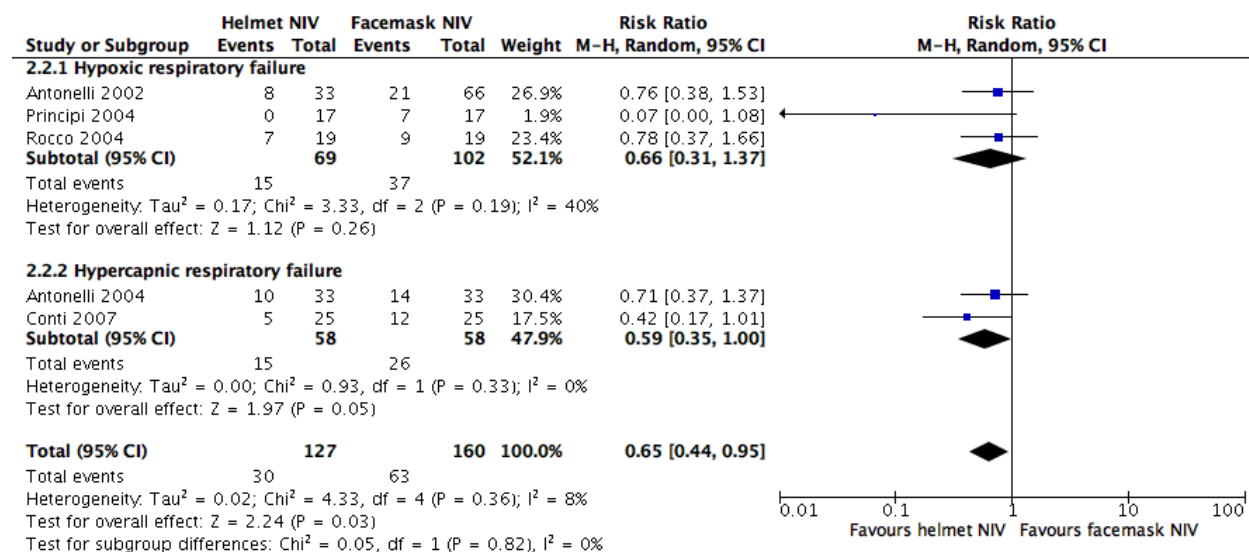
Study or Subgroup	Helmet NIV Events	Total	Facemask NIV Events	Total	Weight	Risk Ratio M–H, Random, 95% CI
2.2.1 Hypoxic respiratory failure						
Antonelli 2002	8	33	21	66	26.9%	0.76 [0.38, 1.53]
Principi 2004	0	17	7	17	1.9%	0.07 [0.00, 1.08]
Rocco 2004	7	19	9	19	23.4%	0.78 [0.37, 1.66]
Subtotal (95% CI)		**69**		**102**	**52.1%**	**0.66 [0.31, 1.37]**
Total events	15		37			
2.2.2 Hypercapnic respiratory failure						
Antonelli 2004	10	33	14	33	30.4%	0.71 [0.37, 1.37]
Conti 2007	5	25	12	25	17.5%	0.42 [0.17, 1.01]
Subtotal (95% CI)		**58**		**58**	**47.9%**	**0.59 [0.35, 1.00]**
Total events	15		26			
Total (95% CI)		**127**		**160**	**100.0%**	**0.65 [0.44, 0.95]**
Total events	30		63			

Heterogeneity. Tau² = 0.17; Chi² = 3.33, df = 2 (P = 0.19); I² = 40%
Test for overall effect: Z = 1.12 (P = 0.26)

Heterogeneity. Tau² = 0.00; Chi² = 0.93, df = 1 (P = 0.33); I² = 0%
Test for overall effect: Z = 1.97 (P = 0.05)

Heterogeneity. Tau² = 0.02; Chi² = 4.33, df = 4 (P = 0.36); I² = 8%
Test for overall effect: Z = 2.24 (P = 0.03)
Test for subgroup differences: Chi² = 0.05, df = 1 (P = 0.82), I² = 0%

Favours helmet NIV   Favours facemask NIV

Figure E3: Effect of helmet NIV compared to facemask NIV on ICU length of stay. Observational data only. Df = degrees of freedom

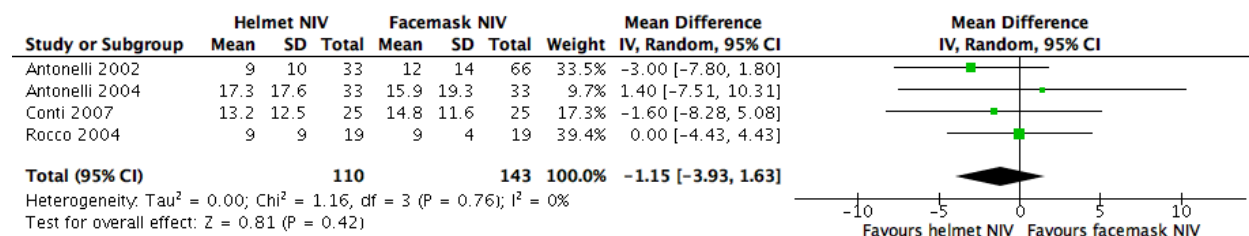| Study or Subgroup	Helmet NIV Mean	SD	Total	Facemask NIV Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Antonelli 2002	9	10	33	12	14	66	33.5%	–3.00 [–7.80, 1.80]
Antonelli 2004	17.3	17.6	33	15.9	19.3	33	9.7%	1.40 [–7.51, 10.31]
Conti 2007	13.2	12.5	25	14.8	11.6	25	17.3%	–1.60 [–8.28, 5.08]
Rocco 2004	9	9	19	9	4	19	39.4%	0.00 [–4.43, 4.43]
Total (95% CI)			**110**			**143**	**100.0%**	**–1.15 [–3.93, 1.63]**

Heterogeneity. Tau² = 0.00; Chi² = 1.16, df = 3 (P = 0.76); I² = 0%
Test for overall effect: Z = 0.81 (P = 0.42)

Favours helmet NIV   Favours facemask NIV

Figure E4: Effect of helmet NIV compared to facemask NIV on duration of NIV. Observational data only. Df = degrees of freedom

Study or Subgroup	Helmet NIV Mean	SD	Total	Facemask NIV Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Antonelli 2002	1.67	1.25	33	1.75	0.88	66	4.0%	-0.08 [-0.56, 0.40]
Antonelli 2004	1.28	1.52	33	1.3	0.81	33	2.7%	-0.02 [-0.61, 0.57]
Conti 2007	1.32	0.81	25	1.08	1.04	25	3.5%	0.24 [-0.28, 0.76]
Principi 2004	1.42	0.007	17	1.17	0.015	17	88.8%	0.25 [0.24, 0.26]
Rocco 2004	3.3	1	19	3.7	2	19	0.9%	-0.40 [-1.41, 0.61]
Total (95% CI)			127			160	100.0%	0.22 [0.12, 0.32]

Heterogeneity: Tau² = 0.00; Chi² = 4.26, df = 4 (P = 0.37); I² = 6%
Test for overall effect: Z = 4.45 (P < 0.00001)

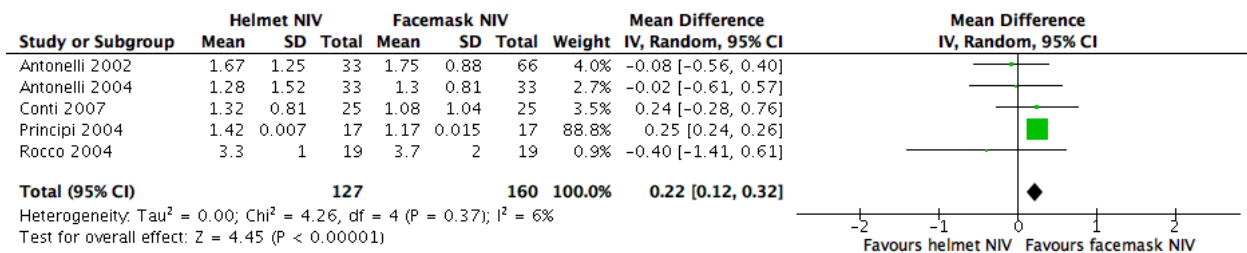Mean Difference IV, Random, 95% CI. Scale -2 to 2. Favours helmet NIV / Favours facemask NIV

Figure E5: Effect of helmet NIV compared to high flow nasal cannula on intubation. Observational data only. Df = degrees of freedom
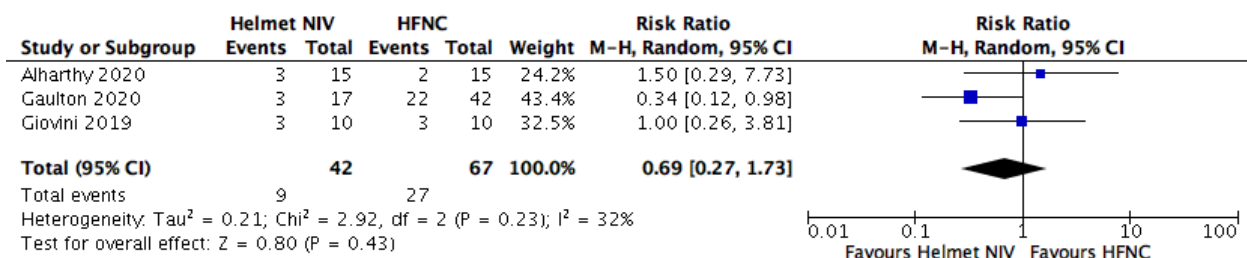
Study or Subgroup	Helmet NIV Events	Total	HFNC Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Alharthy 2020	3	15	2	15	24.2%	1.50 [0.29, 7.73]
Gaulton 2020	3	17	22	42	43.4%	0.34 [0.12, 0.98]
Giovini 2019	3	10	3	10	32.5%	1.00 [0.26, 3.81]
Total (95% CI)		42		67	100.0%	0.69 [0.27, 1.73]
Total events	9		27			

Heterogeneity: Tau² = 0.21; Chi² = 2.92, df = 2 (P = 0.23); I² = 32%
Test for overall effect: Z = 0.80 (P = 0.43)

Risk Ratio M-H, Random, 95% CI. Scale 0.01 to 100. Favours Helmet NIV / Favours HFNC

Figure E6: Effect of helmet NIV compared to high flow nasal cannula on mortality. Observational data only. Df = degrees of freedom

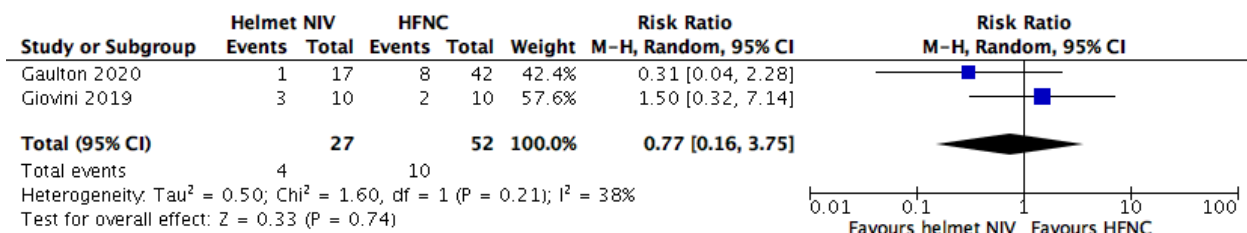| Study or Subgroup	Helmet NIV Events	Total	HFNC Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Gaulton 2020	1	17	8	42	42.4%	0.31 [0.04, 2.28]
Giovini 2019	3	10	2	10	57.6%	1.50 [0.32, 7.14]
Total (95% CI)		27		52	100.0%	0.77 [0.16, 3.75]
Total events	4		10			

Heterogeneity: Tau² = 0.50; Chi² = 1.60, df = 1 (P = 0.21); I² = 38%
Test for overall effect: Z = 0.33 (P = 0.74)

Risk Ratio M-H, Random, 95% CI. Scale 0.01 to 100. Favours helmet NIV / Favours HFNC

Figure E7: Effect of helmet NIV compared to facemask NIV on facial pressure sores. RCT data only. Df = degrees of freedom

Study or Subgroup	Helmet NIV Events	Total	Facemask NIV Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Ali 2011	0	15	1	15	9.1%	0.33 [0.01, 7.58]
Liu 2020	3	13	8	13	42.2%	0.38 [0.13, 1.11]
Liu 2020 2	2	29	0	30	9.8%	5.17 [0.26, 103.21]
Patel 2016	3	44	3	39	27.9%	0.89 [0.19, 4.14]
Yang 2015	0	20	7	20	11.0%	0.07 [0.00, 1.09]
Total (95% CI)		121		117	100.0%	0.50 [0.19, 1.37]
Total events	8		19			

Heterogeneity: Tau² = 0.31; Chi² = 5.22, df = 4 (P = 0.27); I² = 23%
Test for overall effect: Z = 1.35 (P = 0.18)

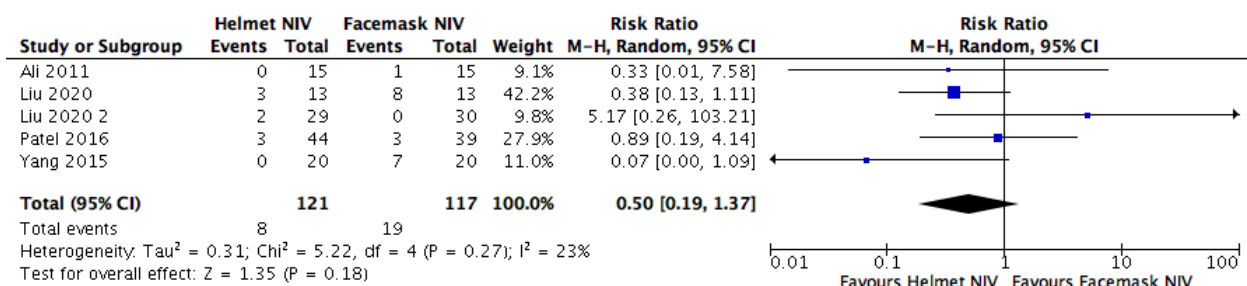Risk Ratio M-H, Random, 95% CI. Scale 0.01 to 100. Favours Helmet NIV / Favours Facemask NIV

Figure E8: Effect of helmet NIV compared to facemask NIV on intubation. Studies are group by risk of bias. RCT data only. Df = degrees of freedom
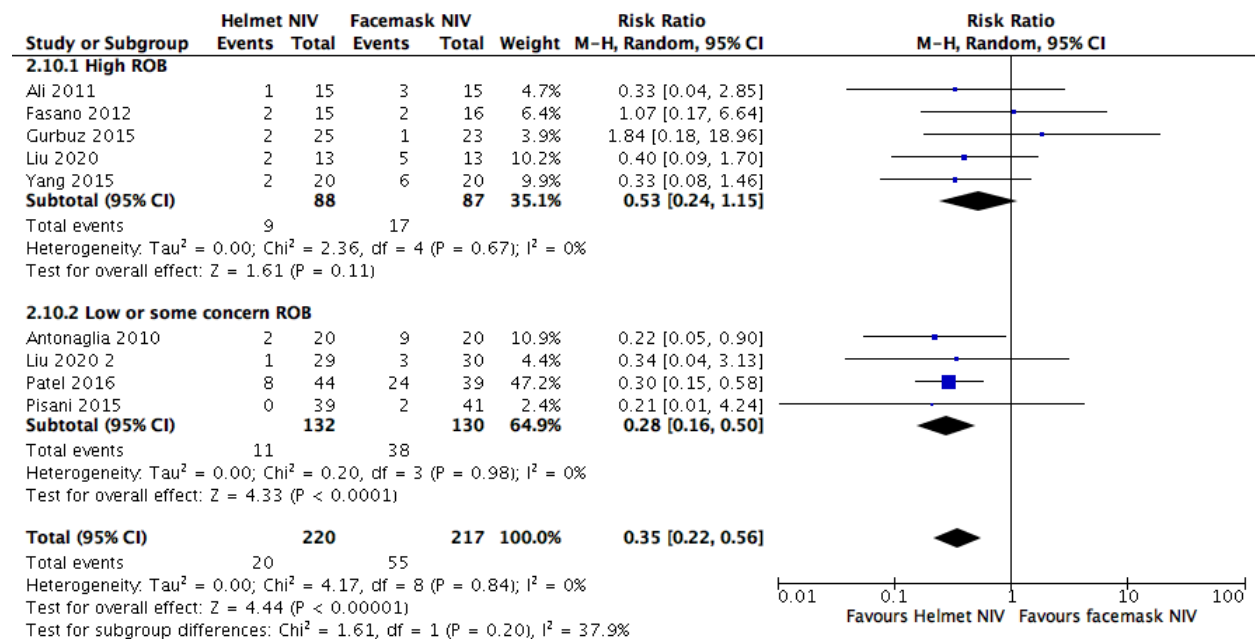
Study or Subgroup	Helmet NIV Events	Total	Facemask NIV Events	Total	Weight	Risk Ratio M–H, Random, 95% CI
2.10.1 High ROB						
Ali 2011	1	15	3	15	4.7%	0.33 [0.04, 2.85]
Fasano 2012	2	15	2	16	6.4%	1.07 [0.17, 6.64]
Gurbuz 2015	2	25	1	23	3.9%	1.84 [0.18, 18.96]
Liu 2020	2	13	5	13	10.2%	0.40 [0.09, 1.70]
Yang 2015	2	20	6	20	9.9%	0.33 [0.08, 1.46]
Subtotal (95% CI)		88		87	35.1%	0.53 [0.24, 1.15]
Total events	9		17			
Heterogeneity: Tau² = 0.00; Chi² = 2.36, df = 4 (P = 0.67); I² = 0%						
Test for overall effect: Z = 1.61 (P = 0.11)						
2.10.2 Low or some concern ROB						
Antonaglia 2010	2	20	9	20	10.9%	0.22 [0.05, 0.90]
Liu 2020 2	1	29	3	30	4.4%	0.34 [0.04, 3.13]
Patel 2016	8	44	24	39	47.2%	0.30 [0.15, 0.58]
Pisani 2015	0	39	2	41	2.4%	0.21 [0.01, 4.24]
Subtotal (95% CI)		132		130	64.9%	0.28 [0.16, 0.50]
Total events	11		38			
Heterogeneity: Tau² = 0.00; Chi² = 0.20, df = 3 (P = 0.98); I² = 0%						
Test for overall effect: Z = 4.33 (P < 0.0001)						
Total (95% CI)		220		217	100.0%	0.35 [0.22, 0.56]
Total events	20		55			
Heterogeneity: Tau² = 0.00; Chi² = 4.17, df = 8 (P = 0.84); I² = 0%						
Test for overall effect: Z = 4.44 (P < 0.00001)						
Test for subgroup differences: Chi² = 1.61, df = 1 (P = 0.20), I² = 37.9%						

Figure E9: Funnel plot of helmet NIV compared to facemask NIV for the outcome of mortality
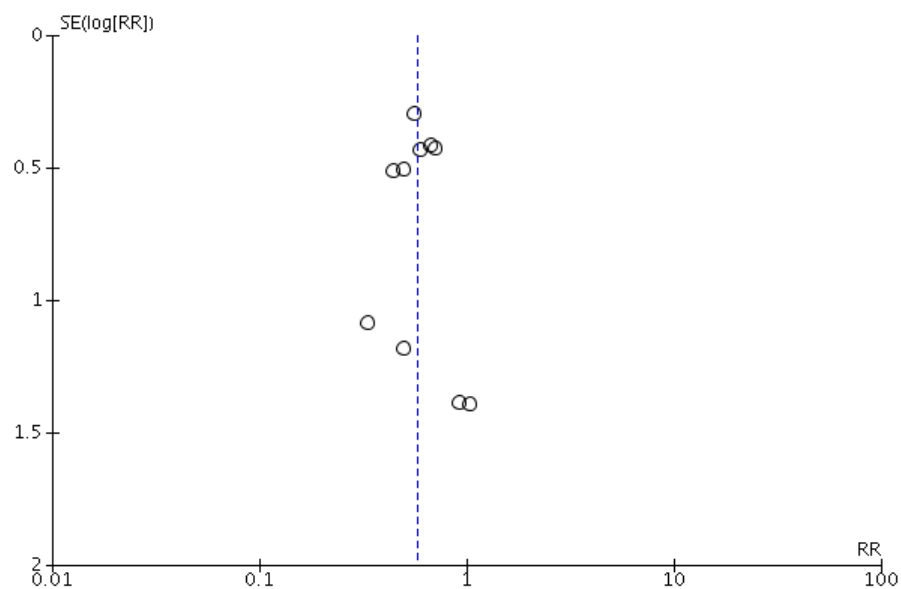
Figure E10: Funnel plot of helmet NIV compared to facemask NIV for the outcome of intubation
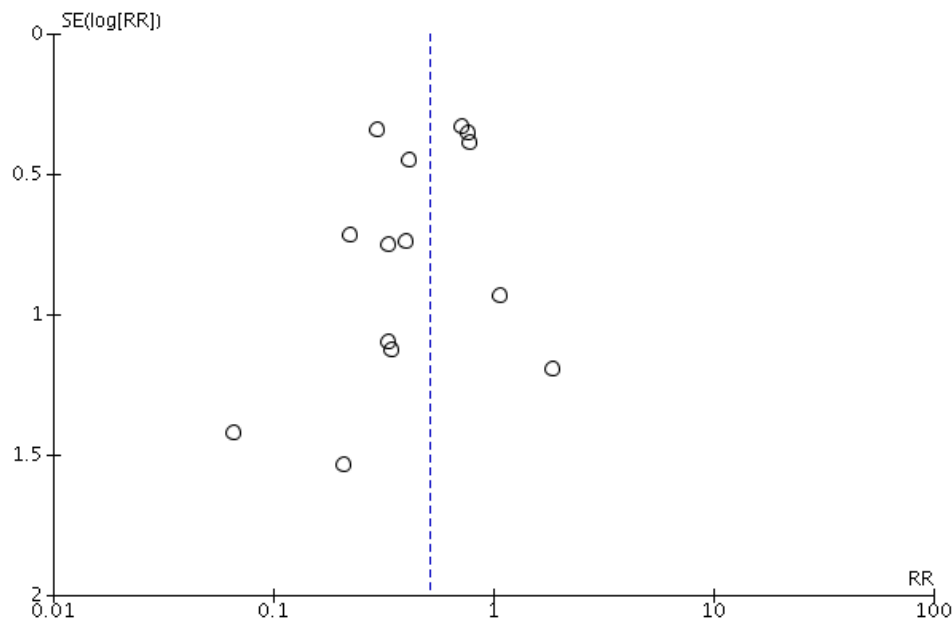
SE(log[RR])

RR

Figure E11: Effect of helmet NIV compared to facemask NIV on mortality. Studies are group by risk of bias. RCT data only. Df = degrees of freedom

Study or Subgroup	Helmet NIV Events	Total	Facemask NIV Events	Total	Weight	Risk Ratio M–H, Random, 95% CI
2.17.1 Low or some concerns						
Liu 2020 2	1	29	1	30	3.7%	1.03 [0.07, 15.77]
Patel 2016	12	44	19	39	81.4%	0.56 [0.31, 1.00]
Yang 2015	1	20	2	20	5.1%	0.50 [0.05, 5.08]
Subtotal (95% CI)		93		89	90.2%	0.57 [0.33, 0.99]
Total events	14		22			

Heterogeneity: Tau² = 0.00; Chi² = 0.20, df = 2 (P = 0.90); I² = 0%
Test for overall effect: Z = 2.00 (P = 0.05)

2.17.2 High ROB						
Gurbuz 2015	1	25	1	23	3.7%	0.92 [0.06, 13.87]
Liu 2020	1	13	3	13	6.0%	0.33 [0.04, 2.80]
Subtotal (95% CI)		38		36	9.8%	0.49 [0.09, 2.62]
Total events	2		4			

Heterogeneity: Tau² = 0.00; Chi² = 0.33, df = 1 (P = 0.56); I² = 0%
Test for overall effect: Z = 0.83 (P = 0.40)

Total (95% CI)		131		125	100.0%	0.56 [0.33, 0.95]
Total events	16		26			

Heterogeneity: Tau² = 0.00; Chi² = 0.56, df = 4 (P = 0.97); I² = 0%
Test for overall effect: Z = 2.16 (P = 0.03)
Test for subgroup differences: Chi² = 0.03, df = 1 (P = 0.87), I² = 0%

Favours helmet NIV Favours facemask NIV