**Supplementary figure legends**

Supplementary figure 1: Risk of bias summary – RCTs

Supplementary figure 2: Risk of bias summary – cohort studies

Supplementary Figure 3: Forest plots for secondary outcomes – mortality domain: hospital mortality

Supplementary Figure 4: Forest plots for secondary outcomes – physiological domain: a) time to extubation; b) ventilator free days to day 28; c) delirium

Supplementary Figure 5: Forest plots – secondary outcomes – resource use domain: a) ICU LOS; b) hospital LOS; c) tracheostomies

Supplementary Figure 6: Forest plots – secondary outcomes – adverse events domain: a) self-extubation; b) re-intubation; c) Ventilator associated pneumonia

| | Random sequence generation (selection bias) | Allocation concealment (selection bias) | Blinding of participants and personnel (performance bias) | Blinding of outcome assessment (detection bias) | Incomplete outcome data (attrition bias) | Selective reporting (reporting bias) |
|---|---|---|---|---|---|---|
| Mehta 2012 | + | + | − | − | + | + |
| Nassar Junior 2014 | + | + | − | − | + | + |
| Olsen 2020 | + | + | − | + | + | + |
| Samuelson 2008 | + | + | − | − | + | + |
| Shehabi 2013 | ? | + | − | − | + | + |
| Strom 2010 | + | + | − | − | + | + |
| Strom 2011 | + | + | − | + | − | + |
| Treggiari 2009 | + | ? | − | + | + | + |

| Study | Method of recruitment (selection bias) | Measurement of depth of sedation bias | Outcome assessment (detection bias) | Important confounding factors (bias) | Length & Completeness of follow-up (attrition bias) | Other bias |
|---|---|---|---|---|---|---|
| Arabi 2007 | + | + | − | − | ? | − |
| Balzer 2015 | + | + | − | ? | + | ? |
| Bugedo 2013 | ? | + | − | − | + | + |
| Burry 2015 | ? | + | + | + | − | + |
| Capuzzo 2001 | + | − | − | ? | + | + |
| Costa 2014 | + | − | − | − | − | + |
| Dale 2014 | + | + | − | + | + | + |
| Faust 2016 | + | + | − | ? | + | + |
| Guttormson 2011 | − | − | − | ? | − | ? |
| Khan 2014 | + | + | − | ? | + | + |
| Mendes 2008 | ? | + | − | − | + | + |
| Quenot 2007 | + | − | − | ? | + | + |
| Ren 2017 | ? | − | − | − | + | + |
| Samuelson 2006 | + | − | + | ? | + | + |
| Samuelson 2007 | + | − | + | ? | + | + |
| Sen 2017 | + | + | − | ? | + | − |
| Shehabi 2012 | ? | + | − | + | + | ? |
| Shehabi 2018 | ? | ? | − | ? | + | + |

| RCTs | Light Deaths | Light Number | Deep Deaths | Deep Number | Risk Ratio | RR | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|
| Mehta et al. 2012 | 63 | 209 | 63 | 214 | | 1.02 | [0.76; 1.37] | 55.5% |
| Nassar et al. 2014 | 9 | 30 | 13 | 30 | | 0.69 | [0.35; 1.37] | 10.2% |
| Shehabi et al. 2013 | 3 | 21 | 2 | 16 | | 1.14 | [0.22; 6.05] | 1.7% |
| Strom et al. 2010 | 20 | 55 | 27 | 58 | | 0.78 | [0.50; 1.22] | 24.0% |
| Treggiari et al. 2009 | 12 | 65 | 11 | 64 | | 1.07 | [0.51; 2.26] | 8.6% |
| **Random effects model** | | 380 | | 382 | | **0.93** | **[0.75; 1.15]** | **100.0%** |

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.75$
Test for overall effect: $z = -0.68$ ($p = 0.50$)

0.2 0.5 1 2 5
Favours light    Favours deep

| Cohort studies | Light Deaths | Light Number | Deep Deaths | Deep Number | Risk Ratio | RR | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|
| Balzer et al. 2015 | 131 | 1371 | 175 | 513 | | 0.28 | [0.23; 0.34] | 20.7% |
| Dale et al. 2014 | 107 | 780 | 96 | 703 | | 1.00 | [0.78; 1.30] | 20.3% |
| Faust et al. 2016 | 24 | 79 | 22 | 65 | | 0.90 | [0.56; 1.45] | 18.5% |
| Khan et al. 2014 | 51 | 262 | 86 | 440 | | 1.00 | [0.73; 1.36] | 20.0% |
| Quenot et al. 2007 | 75 | 197 | 101 | 226 | | 0.85 | [0.68; 1.07] | 20.5% |
| **Random effects model** | | 2689 | | 1947 | | **0.73** | **[0.41; 1.30]** | **100.0%** |

Heterogeneity: $I^2 = 96\%$, $\tau^2 = 0.4103$, $p < 0.01$
Test for overall effect: $z = -1.07$ ($p = 0.28$)

0.2 0.5 1 2 5
Favours light    Favours deep

**RCTs**

| RCTs | Light Number | Light Days | Light SD | Deep Number | Deep Days | Deep SD | Mean Difference | MD | 95%-CI |
|---|---|---|---|---|---|---|---|---|---|
| Mehta et al. 2012 | 209 | 7.33 | 6.72 | 214 | 8.00 | 6.72 | | -0.67 | [-1.95; 0.61] |

Test for overall effect: $z = -1.02$ ($p = 0.31$)

-6 -4 -2 0 2
Favours light    Favours deep

**Cohort studies**

| Cohort studies | Light Number | Light Days | Light SD | Deep Number | Deep Days | Deep SD | Mean Difference | MD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|---|
| Balzer et al. 2015 | 1371 | 0.80 | 0.78 | 513 | 3.72 | 3.68 | | -2.92 | [-3.24; -2.60] | 51.1% |
| Shehabi et al. 2012 | 35 | 2.77 | 1.62 | 213 | 7.43 | 1.94 | | -4.67 | [-5.26; -4.07] | 48.9% |
| Random effects model | 1406 | | | 726 | | | | -3.77 | [-5.49; -2.06] | 100.0% |

Heterogeneity: $I^2 = 96\%$, $\tau^2 = 1.4702$, $p < 0.01$
Test for overall effect: $z = -4.31$ ($p < 0.01$)

-6 -4 -2 0 2
Favours light    Favours deep

a)

**RCTs**

| RCTs | Light Number | Light Days | Light SD | Deep Number | Deep Days | Deep SD | Mean Difference | MD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|---|
| Nassar et al. 2014 | 30 | 24.33 | 4.7 | 30 | 16.67 | 20.2 | | 7.67 | [0.23; 15.10] | 11.4% |
| Olsen et al. 2020 | 349 | 15.33 | 19.4 | 351 | 14.67 | 18.6 | | 0.67 | [-2.15; 3.48] | 43.2% |
| Shehabi et al. 2013 | 21 | 21.30 | 9.2 | 16 | 20.10 | 10.1 | | 1.20 | [-5.12; 7.52] | 15.0% |
| Strom et al. 2010 | 55 | 13.80 | 11.0 | 58 | 9.60 | 10.0 | | 4.20 | [0.32; 8.08] | 30.4% |
| Random effects model | 455 | | | 455 | | | | 2.62 | [-0.09; 5.34] | 100.0% |

Heterogeneity: $I^2 = 31\%$, $\tau^2 = 2.3849$, $p = 0.23$
Test for overall effect: $z = 1.89$ ($p = 0.06$)

-5 0 5 10 15 20
Favours deep    Favours light

**Cohort studies**

| Cohort studies | Light Number | Light Days | Light SD | Deep Number | Deep Days | Deep SD | Mean Difference | MD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|---|
| Bugedo et al. 2013 | 132 | 12.33 | 18.0 | 155 | 10.33 | 17.2 | | 2.00 | [-2.09; 6.09] | 10.1% |
| Faust et al. 2016 | 79 | 24.10 | 3.1 | 65 | 23.60 | 4.9 | | 0.50 | [-0.87; 1.87] | 89.9% |
| Random effects model | 211 | | | 220 | | | | 0.65 | [-0.65; 1.95] | 100.0% |

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.50$
Test for overall effect: $z = 0.98$ ($p = 0.33$)

-5 0 5 10 15 20
Favours deep    Favours light

b)

**RCTs**

| RCTs | Light Delirium | Light Number | Deep Delirium | Deep Number | Risk Ratio | RR | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|
| Mehta et al. 2012 | 113 | 209 | 113 | 214 | | 1.02 | [0.86; 1.22] | 90.1% |
| Nassar et al. 2014 | 12 | 30 | 9 | 30 | | 1.33 | [0.66; 2.69] | 5.8% |
| Samuelson 2008 | 0 | 18 | 0 | 18 | | | | 0.0% |
| Shehabi et al. 2013 | 8 | 21 | 6 | 16 | | 1.02 | [0.44; 2.34] | 4.1% |
| Random effects model | | 278 | | 278 | | 1.04 | [0.88; 1.23] | 100.0% |

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.77$
Test for overall effect: $z = 0.45$ ($p = 0.65$)

0.2 0.5 1 2 5
Favours light    Favours deep

**Cohort studies**

| Cohort studies | Light Delirium | Light Number | Deep Delirium | Deep Number | Risk Ratio | RR | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|
| Balzer et al. 2015 | 445 | 1371 | 213 | 513 | | 0.78 | [0.69; 0.89] | 27.3% |
| Dale et al. 2014 | 176 | 780 | 75 | 703 | | 2.12 | [1.65; 2.72] | 25.8% |
| Khan et al. 2014 | 94 | 141 | 167 | 302 | | 1.21 | [1.03; 1.41] | 27.0% |
| Ren et al. 2017 | 13 | 73 | 29 | 70 | | 0.43 | [0.24; 0.76] | 19.9% |
| Random effects model | | 2365 | | 1588 | | 1.01 | [0.63; 1.62] | 100.0% |

Heterogeneity: $I^2 = 95\%$, $\tau^2 = 0.2088$, $p < 0.01$
Test for overall effect: $z = 0.04$ ($p = 0.97$)

0.2 0.5 1 2 5
Favours light    Favours deep

c)

Supplementary Figure 4: Forest plots for secondary outcomes – physiological domain: a) time to extubation; b) ventilator free days to day 28 (VFD28); c) delirium

Note: data converted from median/IRQ to mean/SD[12] – time to extubation in the following studies: Balzer et al 2015; Mehta et al 2012; Shehabi et al 2012; VFD28 in the following studies: Bugedo et al 2013; Nassar Jr et al 2014.

### a)

**RCTs** — Light / Deep — Mean Difference

| RCTs | Number | Days | SD | Number | Days | SD | MD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|
| Mehta et al. 2012 | 209 | 12.00 | 10.5 | 214 | 10.67 | 9.0 | 1.33 | [-0.52; 3.19] | 34.0% |
| Nassar et al. 2014 | 30 | 11.00 | 7.8 | 30 | 10.67 | 10.9 | 0.33 | [-4.46; 5.13] | 10.7% |
| Olsen et al 2020 | 349 | 12.00 | 17.1 | 351 | 12.33 | 17.1 | -0.33 | [-2.87; 2.20] | 25.5% |
| Shehabi et al. 2013 | 21 | 6.67 | 5.0 | 16 | 6.33 | 5.6 | 0.33 | [-3.15; 3.82] | 17.2% |
| Strom et al. 2010 | 55 | 15.60 | 17.0 | 58 | 20.83 | 12.4 | -5.23 | [-10.74; 0.27] | 8.5% |
| Treggiari et al. 2009 | 65 | 34.50 | 27.3 | 64 | 28.00 | 20.8 | 6.50 | [-1.87; 14.87] | 4.0% |
| **Random effects model** | **729** | | | **733** | | | **0.28** | **[-1.46; 2.02]** | **100.0%** |

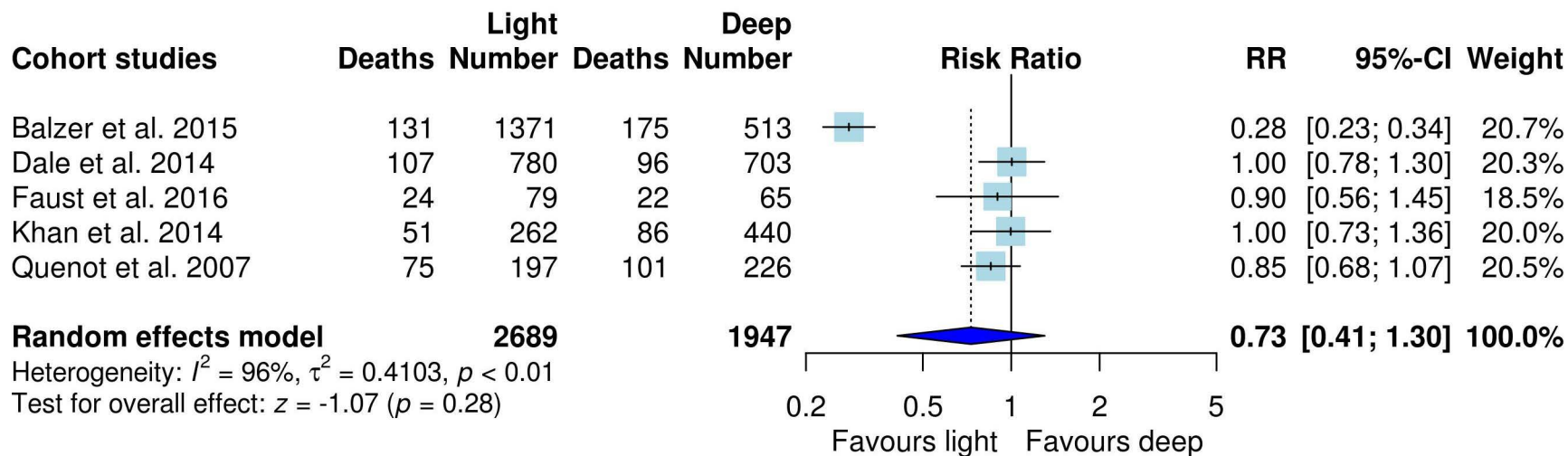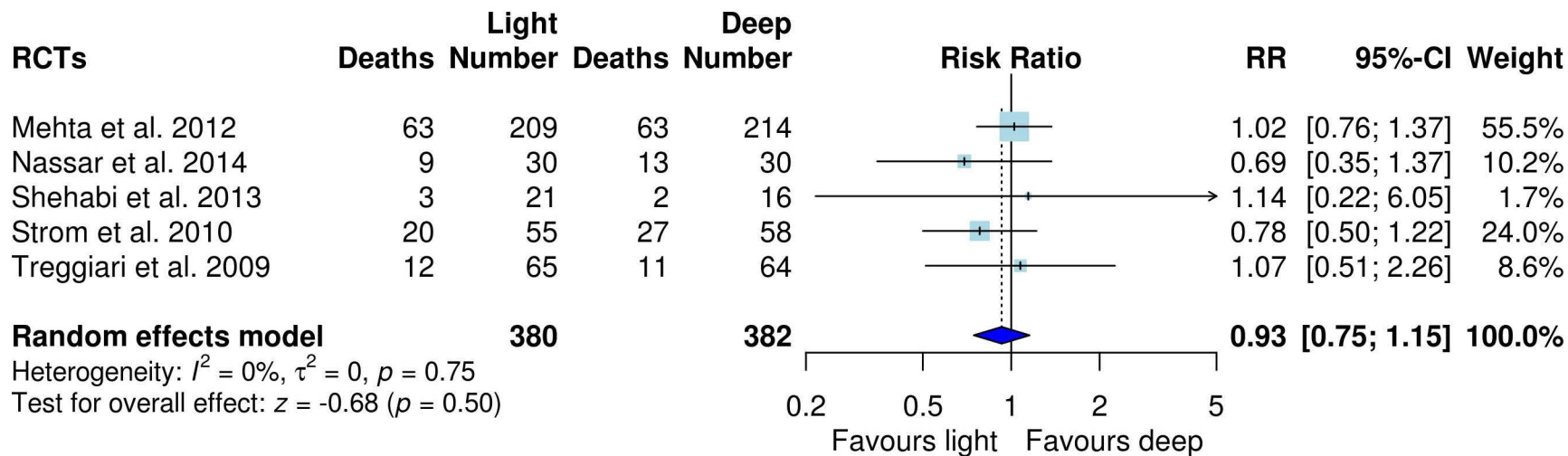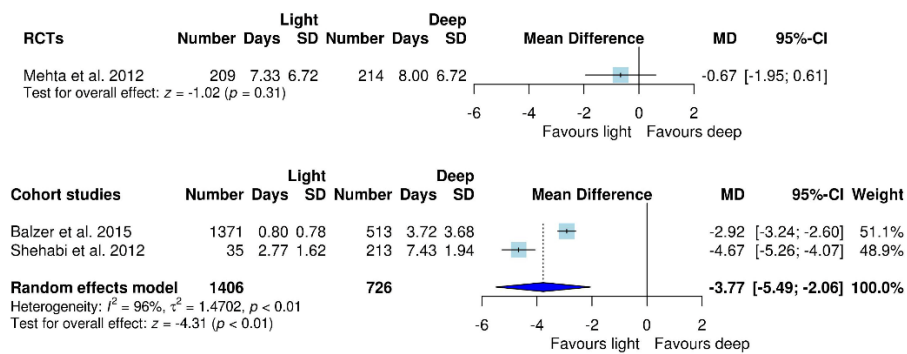Heterogeneity: $I^2 = 32\%$, $\tau^2 = 1.4277$, $p = 0.20$
Test for overall effect: $z = 0.31$ ($p = 0.75$)

-25 -20 -15 -10 -5 0 5 10 — Favours light / Favours deep

**Cohort studies** — Light / Deep — Mean Difference

| Cohort studies | Number | Days | SD | Number | Days | SD | MD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|
| Balzer et al. 2015 | 1371 | 9.67 | 8.2 | 513 | 23.67 | 19.3 | -14.00 | [-15.73; -12.27] | 13.5% |
| Bugedo et al. 2013 | 132 | 11.67 | 9.0 | 155 | 10.33 | 6.7 | 1.33 | [-0.53; 3.20] | 13.4% |
| Dale et al. 2014 | 780 | 3.33 | 3.7 | 703 | 3.67 | 4.5 | -0.33 | [-0.75; 0.09] | 14.0% |
| Faust et al. 2016 | 79 | 6.70 | 5.2 | 65 | 8.81 | 6.8 | -2.11 | [-4.14; -0.08] | 13.3% |
| Guttormson 2011 | 25 | 5.97 | 4.2 | 10 | 18.23 | 16.0 | -12.27 | [-22.32; -2.21] | 5.7% |
| Quenot et al. 2007 | 197 | 6.83 | 7.8 | 226 | 13.50 | 18.3 | -6.67 | [-9.29; -4.04] | 12.8% |
| Ren et al. 2017 | 73 | 7.47 | 2.5 | 70 | 9.76 | 3.8 | -2.29 | [-3.34; -1.24] | 13.8% |
| Sen et al. 2017 | 32 | 4.33 | 3.9 | 135 | 7.33 | 5.2 | -3.00 | [-4.61; -1.39] | 13.6% |
| **Random effects model** | **2689** | | | **1877** | | | **-4.30** | **[-7.39; -1.21]** | **100.0%** |

Heterogeneity: $I^2 = 97\%$, $\tau^2 = 17.6542$, $p < 0.01$
Test for overall effect: $z = -2.73$ ($p < 0.01$)

-25 -20 -15 -10 -5 0 5 10 — Favours light / Favours deep

### b)

**RCTs** — Light / Deep — Mean Difference

| RCTs | Number | Days | SD | Number | Days | SD | MD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|
| Mehta et al 2012 | 209 | 26.00 | 28.37 | 214 | 22.00 | 19.41 | 4.00 | [-0.64; 8.64] | 26.0% |
| Nassar et al. 2014 | 30 | 24.33 | 19.46 | 30 | 17.33 | 14.79 | 7.00 | [-1.75; 15.75] | 19.1% |
| Shehabi et al. 2013 | 21 | 19.57 | 19.08 | 16 | 16.67 | 20.32 | 2.90 | [-9.97; 15.77] | 13.3% |
| Strom et al. 2010 | 55 | 38.67 | 36.54 | 58 | 58.67 | 39.53 | -20.00 | [-34.03; -5.97] | 12.0% |
| Treggiari et al. 2009 | 65 | 19.25 | 4.27 | 64 | 22.75 | 5.35 | -3.50 | [-5.17; -1.83] | 29.6% |
| **Random effects model** | **380** | | | **382** | | | **-0.69** | **[-6.96; 5.58]** | **100.0%** |

Heterogeneity: $I^2 = 80\%$, $\tau^2 = 33.7865$, $p < 0.01$
Test for overall effect: $z = -0.21$ ($p = 0.83$)

-25 -20 -15 -10 -5 0 5 10 — Favours light / Favours deep

**Cohort studies** — Light / Deep — Mean Difference

| Cohort studies | Number | Days | SD | Number | Days | SD | MD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|
| Balzer et al. 2015 | 1371 | 21.00 | 15.58 | 513 | 30.67 | 23.79 | -9.67 | [-11.88; -7.45] | 18.6% |
| Bugedo et al. 2013 | 132 | 19.67 | 15.74 | 155 | 20.33 | 17.21 | -0.67 | [-4.48; 3.15] | 15.5% |
| Dale et al. 2014 | 780 | 10.67 | 10.40 | 703 | 12.33 | 11.89 | -1.67 | [-2.81; -0.52] | 20.1% |
| Khan et al. 2014 | 440 | 19.23 | 20.46 | 262 | 22.95 | 52.12 | -3.72 | [-10.31; 2.87] | 10.4% |
| Quenot et al. 2007 | 197 | 14.67 | 12.70 | 226 | 19.67 | 20.89 | -5.00 | [-8.25; -1.75] | 16.7% |
| Sen et al. 2017 | 32 | 9.33 | 5.43 | 135 | 13.33 | 6.74 | -4.00 | [-6.20; -1.80] | 18.6% |
| **Random effects model** | **2952** | | | **1994** | | | **-4.21** | **[-7.22; -1.19]** | **100.0%** |

Heterogeneity: $I^2 = 88\%$, $\tau^2 = 11.4184$, $p < 0.01$
Test for overall effect: $z = -2.74$ ($p < 0.01$)

-25 -20 -15 -10 -5 0 5 10 — Favours light / Favours deep

### c)

**RCTs** — Light / Deep — Risk Ratio

| RCTs | Tracheostomy | Number | Tracheostomy | Number | RR | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|
| Mehta et al. 2012 | 54 | 209 | 49 | 214 | 1.13 | [0.81; 1.58] | 70.8% |
| Nassar et al. 2014 | 1 | 30 | 1 | 30 | 1.00 | [0.07; 15.26] | 1.1% |
| Strom et al. 2010 | 16 | 55 | 17 | 58 | 0.99 | [0.56; 1.76] | 24.3% |
| Treggiari et al. 2009 | 3 | 65 | 4 | 64 | 0.74 | [0.17; 3.17] | 3.8% |
| **Random effects model** | | **359** | | **366** | **1.07** | **[0.81; 1.43]** | **100.0%** |

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.94$
Test for overall effect: $z = 0.50$ ($p = 0.62$)

0.1 0.2 0.5 1 2 5 10 — Favours light / Favours deep

**Cohort studies** — Light / Deep — Risk Ratio

| Cohort studies | Tracheostomy | Number | Tracheostomy | Number | RR | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|
| Bugedo et al. 2013 | 6 | 132 | 12 | 155 | 0.59 | [0.23; 1.52] | 44.3% |
| Faust et al. 2016 | 8 | 79 | 11 | 65 | 0.60 | [0.26; 1.40] | 55.7% |
| **Random effects model** | | **211** | | **220** | **0.59** | **[0.31; 1.12]** | **100.0%** |

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.98$
Test for overall effect: $z = -1.61$ ($p = 0.11$)

0.1 0.2 0.5 1 2 5 10 — Favours light / Favours deep

Supplementary Figure 5: Forest plots – secondary outcomes – resource use domain: a) ICU LOS; b) hospital LOS; c) tracheostomies

Note: data converted from median/IRQ to mean/SD[12] – ICU LOS& hospital LOS in the following studies: Balzer et al 2015; Bugedo et al 2013; Dale et al 2014; Guttormson et al 2011 (ICU LOS only); Mehta et al 2012; Nassar Jr et al 2014; Quenot et al 2007; Sen et al 2017; Shehabi et al 2013; Strom et al 2010; Treggiari et al 2009.

a)



b)



c)

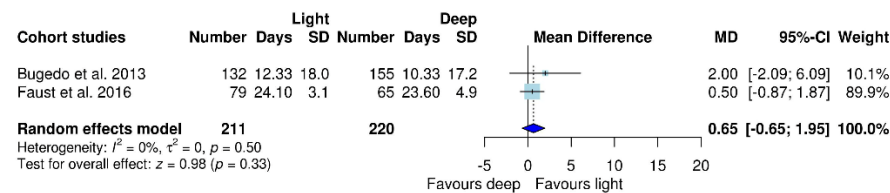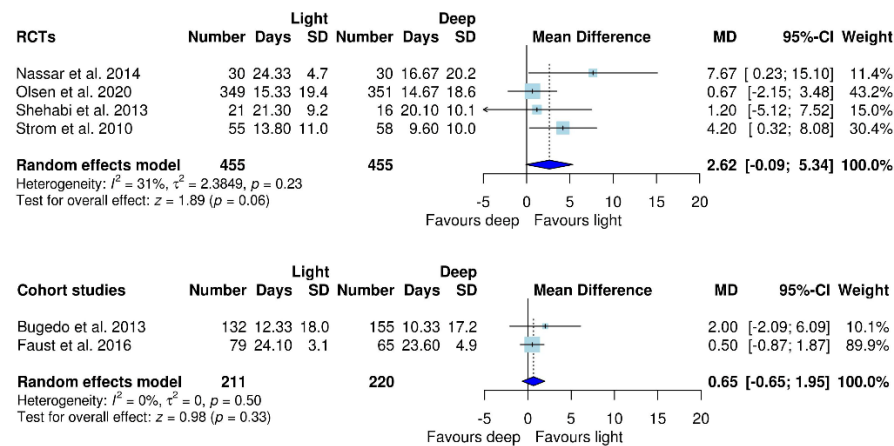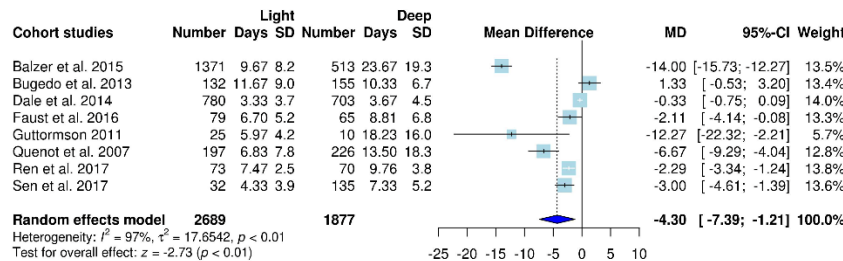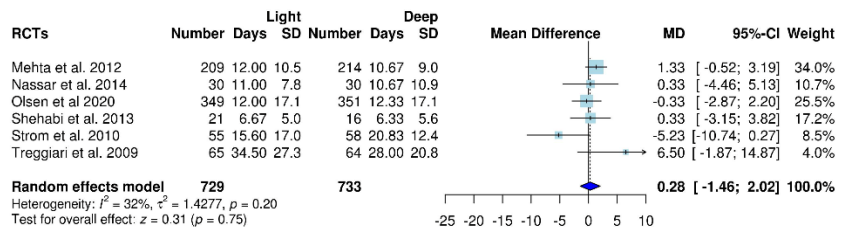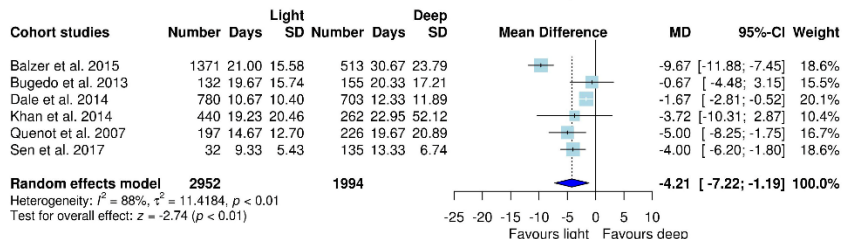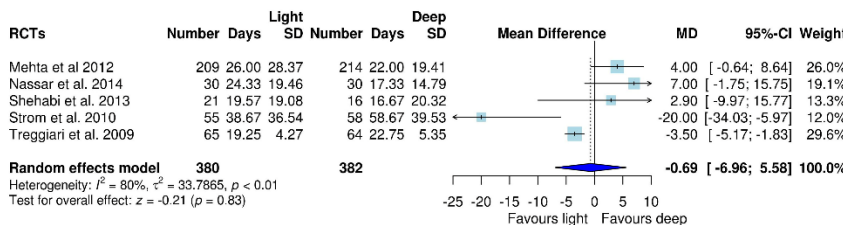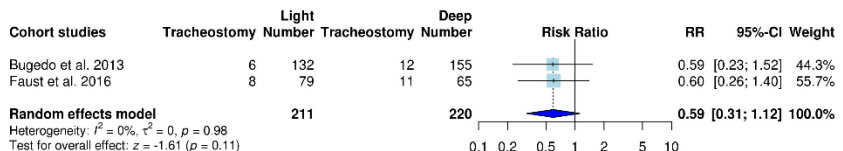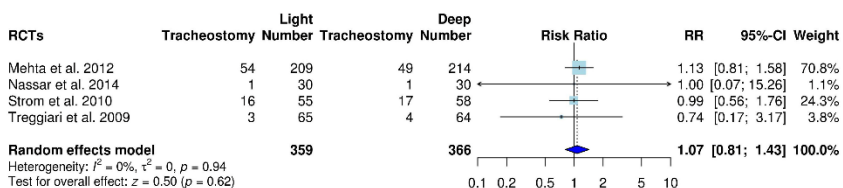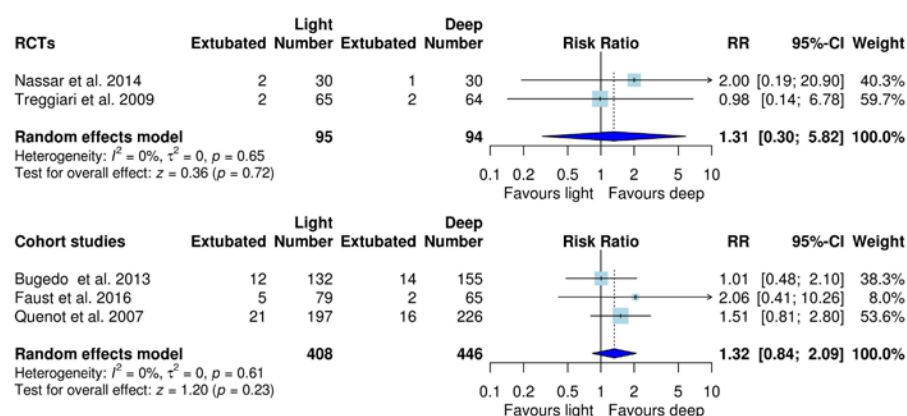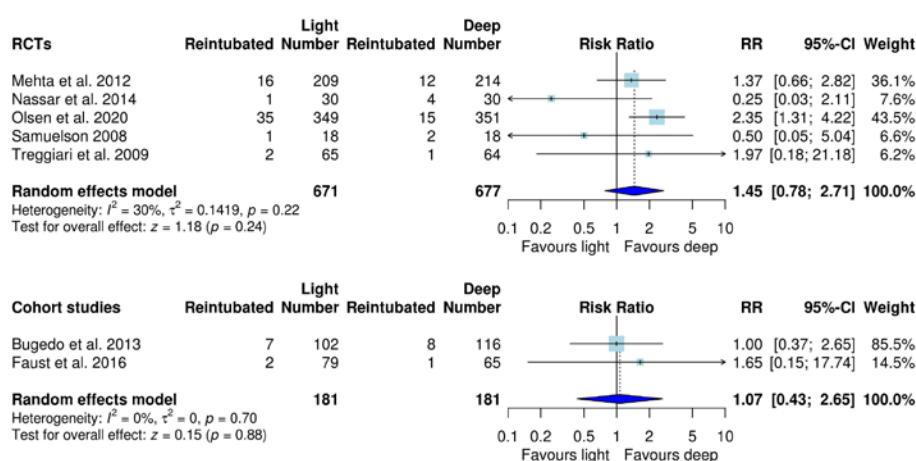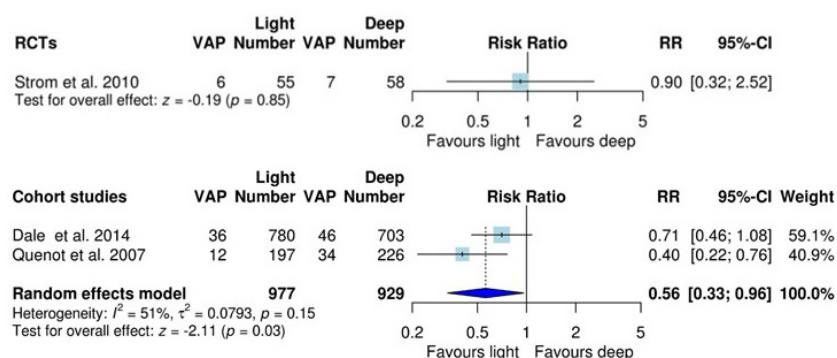Supplementary Figure 6: Forest plots – secondary outcomes – adverse events domain: a) self-extubation; b) re-intubation; c) Ventilator associated pneumonia