**Adjusting for disease severity across Intensive Care Units in multicenter studies**

**Supplementary Information**

Supplementary Table 1. Center level characteristics for the selected study centers

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Center | Teaching/non-teaching hospital (yes/no) | High volume  (>1500 admissions in 2011-2015, yes/no) | High APACHE IV predicted mortality (mean >25%, yes/no) | High proportion of medical admissions (>70%, yes/no) |
| 1 | no | no | no | no |
| 2 | no | no | no | no |
| 3 | no | no | no | no |
| 4 | no | no | no | no |
| 5 | no | no | no | yes |
| 6 | no | no | no | yes |
| 7 | no | yes | no | no |
| 8 | no | yes | no | no |
| 9 | no | yes | no | no |
| 10 | no | yes | no | yes |
| 11 | yes | yes | no | no |
| 12 | yes | yes | no | no |
| 13 | yes | yes | no | no |
| 14 | yes | yes | no | no |
| 15 | yes | yes | yes | no |
| 16 | yes | yes | yes | no |
| 17 | yes | yes | yes | no |
| 18 | yes | yes | yes | yes |
| 19 | yes | yes | yes | yes |
| 20 | yes | yes | yes | yes |

Supplementary Table 2. Relative bias, for the different methods (columns)*a* per scenario (rows), in comparison with scenario 0 (correction for APACHE IV score in all centers)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | **1. Naive** | **2. Restriction** | **3. Two-stage** | **4. Standardization (groups) *b*** | **5. Standardization (centers) *c*** | **6. Multiple imputation** | **7. Regression calibration** | **8. Propensity score modelling** |
|  | **APACHE IV score** | **APACHE II score** |  |  |  |  |  |  |  |  |
| 1A | teaching (10) | non-teaching (10) | 60.6% | 6.2% | 18.4% | 18.4% | 18.4% | 0.7% | -17.0% | 27.9% |
| 2A | non-teaching (10) | teaching (10) | 168.1% | -15.6% | 70.1% | 67.4% | 68.8% | 11.6% | -3.4% | 70.1% |
| 3A | >1500 adm. (14) | < 1500 adm (6) | 26.6% | 2.1% | 8.9% | 7.5% | 8.9% | 0.7% | -14.3% | 21.1% |
| 4A | < 1500 adm (6) | >1500 adm. (14) | 138.2% | -3.4% | 79.6% | 76.9% | 78.3% | 17.0% | -11.5% | 78.3% |
| 5A | pred. mort >25% (6) | pred. mort <25% (14) | 161.3% | 25.2% | 55.1% | 53.8% | 55.1% | 7.5% | -14.3% | 57.9% |
| 6A | pred. mort <25% (14) | pred. mort >25% (6) | 115.0% | -14.3% | 36.1% | 33.4% | 33.4% | 0.7% | -11.5% | 41.5% |
| 7A | >70% med. adm (6) | <70% med. adm (14) | 173.6% | -21.1% | 63.3% | 62.0% | 63.3% | -6.1% | -15.6% | 64.7% |
| 8A | <70% med. adm (14) | >70% med. adm (6) | 71.5% | 8.9% | 25.2% | 23.8% | 25.2% | 6.2% | -12.9% | 33.4% |
| 9A | random (12) | random (8) | 110.9% | 11.6% | 33.4% | 30.7% | 32.0% | 7.5% | -11.5% | 42.9% |
| **Mean relative bias over APACHE II scenarios** | | | **114.0%** | **0.0%** | **43.4%** | **41.5%** | **42.6%** | **5.1%** | **-12.4%** | **48.6%** |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **APACHE IV score** | **SAPS II score** |  |  |  |  |  |  |  |  |
| 1B | teaching (10) | non-teaching (10) | 19.8% | 6.2% | 8.9% | 7.5% | 7.5% | -0.7% | -63.3% | 21.1% |
| 2B | non-teaching (10) | teaching (10) | 48.3% | -15.6% | 30.7% | 29.3% | 29.3% | 7.5% | -40.1% | 42.9% |
| 3B | >1500 adm. (14) | < 1500 adm. (6) | 10.2% | 2.1% | 4.8% | 3.4% | 3.4% | 0.7% | -59.2% | 18.4% |
| 4B | < 1500 adm. (6) | >1500 adm. (14) | 45.6% | -3.4% | 36.1% | 33.4% | 34.7% | 17.0% | -46.9% | 48.3% |
| 5B | pred. mort. >25% (6) | pred. mort. <25% (14) | 53.8% | 25.2% | 27.9% | 26.6% | 27.9% | 13.0% | -60.5% | 42.9% |
| 6B | pred. mort. <25% (14) | pred. mort. >25% (6) | 32.0% | -14.3% | 11.6% | 10.2% | 10.2% | -7.5% | -55.1% | 23.8% |
| 7B | >70% med. adm. (6) | <70% med. adm. (14) | 51.1% | -21.1% | 27.9% | 26.6% | 26.6% | -0.7% | -52.4% | 41.5% |
| 8B | <70% med. adm. (14) | >70% med. adm. (6) | 23.8% | 8.9% | 11.6% | 10.2% | 10.2% | 2.1% | -59.2% | 23.8% |
| 9B | random (12) | random (8) | 41.5% | 11.6% | 19.8% | 18.4% | 18.4% | 10.2% | -56.4% | 36.1% |
| **Mean relative bias over SAPS II scenarios** | | | **36.2%** | **0.0%** | **19.9%** | **18.4%** | **18.7%** | **4.6%** | **-54.8%** | **33.2%** |

Legend: adm., admissions; APACHE, acute physiology and chronic health evaluation; med., medical; pred. mort., predicted in-hospital mortality; SAPS, simplified acute physiology score.

*a* For a description of methods, see main text (“Methods to adjust for confounding by disease severity”).

*b* APACHE II / SAPS II scores were standardized across all centers that had the respective score available.

*c* APACHE II / SAPS II scores were standardized within each center that had the respective score available.

|  |  |  |
| --- | --- | --- |
| Colour code | |0%| - |25%| | |25%| - |50%| |
| |50%| - |75%| | |75%| - |100%| | >|100%| |

Supplementary Table 3. Absolute bias, expressed as odds ratios, for the different methods (columns)*a* per scenario (rows), in comparison with scenario 0 (correction for APACHE IV score in all centers) *,*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | **1. Naive** | **2. Restriction** | **3. Two-stage** | **4. Standardization (groups) *b*** | **5. Standardization (centers)*c*** | **6. Multiple imputation** | **7. Regression calibration** | **8. Propensity score modelling** |
|  | **APACHE IV score** | **APACHE II score** |  |  |  |  |  |  |  |  |
| 1A | teaching (10) | non-teaching (10) | 1.04 | 1.00 | 1.01 | 1.01 | 1.01 | 1.00 | 0.99 | 1.02 |
| 2A | non-teaching (10) | teaching (10) | 1.13 | 0.99 | 1.05 | 1.05 | 1.05 | 1.01 | 1.00 | 1.05 |
| 3A | >1500 adm. (14) | < 1500 adm.(6) | 1.02 | 1.00 | 1.01 | 1.00 | 1.01 | 1.00 | 0.99 | 1.02 |
| 4A | < 1500 adm. (6) | >1500 adm. (14) | 1.11 | 1.00 | 1.06 | 1.06 | 1.06 | 1.01 | 0.99 | 1.06 |
| 5A | pred. mort. >25% (6) | pred. mort. <25% (14) | 1.12 | 1.02 | 1.04 | 1.04 | 1.04 | 1.00 | 0.99 | 1.04 |
| 6A | pred. mort. <25% (14) | pred. mort. >25% (6) | 1.09 | 0.99 | 1.03 | 1.02 | 1.02 | 1.00 | 0.99 | 1.03 |
| 7A | >70% med. adm. (6) | <70% med. adm. (14) | 1.13 | 0.98 | 1.05 | 1.05 | 1.05 | 1.00 | 0.99 | 1.05 |
| 8A | <70% med. adm. (14) | >70% med. adm. (6) | 1.05 | 1.01 | 1.02 | 1.02 | 1.02 | 1.00 | 0.99 | 1.02 |
| 9A | random (12) | random (8) | 1.08 | 1.01 | 1.02 | 1.02 | 1.02 | 1.01 | 0.99 | 1.03 |
| **Mean absolute bias over APACHE II scenarios** | | | **1.09** | **1.00** | **1.03** | **1.03** | **1.03** | **1.00** | **0.99** | **1.04** |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **APACHE IV score** | **SAPS II score** |  |  |  |  |  |  |  |  |
| 1B | teaching (10) | non-teaching (10) | 1.01 | 1.00 | 1.01 | 1.00 | 1.01 | 1.00 | 0.95 | 1.01 |
| 2B | non-teaching (10) | teaching (10) | 1.03 | 0.99 | 1.02 | 1.02 | 1.02 | 1.01 | 0.97 | 1.03 |
| 3B | >1500 adm. (14) | < 1500 adm. (6) | 1.01 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.96 | 1.01 |
| 4B | < 1500 adm. (6) | >1500 adm. (14) | 1.03 | 1.00 | 1.03 | 1.02 | 1.02 | 1.01 | 0.97 | 1.03 |
| 5B | pred. mort. >25% (6) | pred. mort. <25% (14) | 1.04 | 1.02 | 1.02 | 1.02 | 1.02 | 1.01 | 0.96 | 1.03 |
| 6B | pred. mort. <25% (14) | pred. mort. >25% (6) | 1.02 | 0.99 | 1.01 | 1.01 | 1.01 | 0.99 | 0.96 | 1.02 |
| 7B | >70% med. adm. (6) | <70% med. adm. (14) | 1.04 | 0.98 | 1.02 | 1.02 | 1.02 | 1.00 | 0.96 | 1.03 |
| 8B | <70% med. adm. (14) | >70% med. adm. (6) | 1.02 | 1.01 | 1.01 | 1.01 | 1.01 | 1.00 | 0.96 | 1.02 |
| 9B | random (12) | random (8) | 1.03 | 1.01 | 1.01 | 1.01 | 1.01 | 1.01 | 0.96 | 1.03 |
| **Mean absolute bias over SAPS II scenarios** | | | **1.03** | **1.00** | **1.01** | **1.01** | **1.01** | **1.00** | **0.96** | **1.02** |

Legend: adm., admissions; APACHE, acute physiology and chronic health evaluation; med., medical; pred. mort., predicted in-hospital mortality; SAPS, simplified acute physiology score.

*a* For a description of methods, see main text (“Methods to adjust for confounding by disease severity”).

*b* APACHE II / SAPS II scores were standardized across all centers that had the respective score available.

c APACHE II / SAPS II scores were standardized within each center that had the respective score available

|  |  |  |
| --- | --- | --- |
| Colour code | <= δ0.05 | > δ0.05 |

Supplementary Table 4. Coverage of 95% confidence interval of estimated treatment effect when using the different methods (columns)*a* per scenario (rows), relative to scenario 0 (correction for APACHE IV score in all centers)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | **1. Naive** | **2. Restriction** | **3. Two-stage** | **4. Standardization (groups) *b*** | **5. Standardization (centers) *c*** | **6. Multiple imputation** | **7. Regression calibration** | **8. Propensity score modelling** |
|  | **APACHE IV score** | **APACHE II score** |  |  |  |  |  |  |  |  |
| 1A | teaching (10) | non-teaching (10) | 62.9% | 93.6% | 92.0% | 92.8% | 92.6% | 94.2% | 92.9% | 89.7% |
| 2A | non-teaching (10) | teaching (10) | 0.2% | 94.3% | 50.6% | 53.4% | 52.3% | 95.2% | 95.1% | 42.1% |
| 3A | >1500 adm. (14) | < 1500 adm. (6) | 88.4% | 94.3% | 94.3% | 94.7% | 94.7% | 94.3% | 92.9% | 92.9% |
| 4A | < 1500 adm. (6) | >1500 adm. (14) | 2.5% | 96.3% | 40.8% | 43.4% | 42.8% | 93.1% | 92.4% | 31.4% |
| 5A | pred. mort. >25% (6) | pred. mort. <25% (14) | 0.4% | 93.2% | 69.0% | 69.3% | 68.2% | 94.5% | 93.5% | 57.2% |
| 6A | pred. mort. <25% (14) | pred. mort. >25% (6) | 10.6% | 93.7% | 83.6% | 84.5% | 84.5% | 94.4% | 93.1% | 77.6% |
| 7A | >70% med. adm. (6) | <70% med. adm. (14) | 0.1% | 94.5% | 58.2% | 59.7% | 58.9% | 93.6% | 92.6% | 48.6% |
| 8A | <70% med. adm. (14) | >70% med. adm. (6) | 48.0% | 93.9% | 89.8% | 90.6% | 90.1% | 94.3% | 93.4% | 84.9% |
| 9A | random (12) | random (8) | 13.5% | 94.9% | 84.2% | 87.1% | 86.2% | 94.7% | 93.1% | 75.1% |
| **Mean coverage over APACHE II scenarios** | | | **25.2%** | **94.3%** | **73.6%** | **75.1%** | **74.5%** | **94.3%** | **93.2%** | **66.6%** |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **APACHE IV score** | **SAPS II score** |  |  |  |  |  |  |  |  |
| 1B | teaching (10) | non-teaching (10) | 92.1% | 93.6% | 94.5% | 94.7% | 94.8% | 94.1% | 58.9% | 93.4% |
| 2B | non-teaching (10) | teaching (10) | 73.6% | 94.3% | 86.7% | 87.3% | 87.2% | 94.0% | 80.4% | 75.4% |
| 3B | >1500 adm. (14) | < 1500 adm. (6) | 94.4% | 94.3% | 94.4% | 94.8% | 94.9% | 93.9% | 63.0% | 93.9% |
| 4B | < 1500 adm. (6) | >1500 adm. (14) | 76.5% | 96.3% | 83.9% | 84.7% | 84.8% | 92.7% | 79.9% | 70.2% |
| 5B | pred. mort. >25% (6) | pred. mort. <25% (14) | 68.7% | 93.2% | 88.6% | 90.1% | 89.5% | 95.4% | 64.6% | 76.7% |
| 6B | pred. mort. <25% (14) | pred. mort. >25% (6) | 86.0% | 93.7% | 93.1% | 94.1% | 94.2% | 93.0% | 67.3% | 90.9% |
| 7B | >70% med. adm. (6) | <70% med. adm. (14) | 70.7% | 94.5% | 88.7% | 89.3% | 88.9% | 95.1% | 73.2% | 76.9% |
| 8B | <70% med. adm. (14) | >70% med. adm. (6) | 90.4% | 93.9% | 94.0% | 93.9% | 93.5% | 93.8% | 62.0% | 91.5% |
| 9B | random (12) | random (8) | 79.9% | 94.9% | 92.5% | 92.7% | 92.6% | 94.3% | 66.8% | 83.5% |
| **Mean coverage over SAPS II scenarios** | | | **81.4%** | **94.3%** | **90.7%** | **91.3%** | **91.2%** | **94.0%** | **68.5%** | **83.6%** |

Legend: adm., admissions; APACHE, acute physiology and chronic health evaluation; med., medical; pred. mort., predicted in-hospital mortality; SAPS, simplified acute physiology score.

*a* For a description of methods, see main text (“Methods to adjust for confounding by disease severity”).

*b* APACHE II / SAPS II scores were standardized across all centers that had the respective score available.

*c* APACHE II / SAPS II scores were standardized within each center that had the respective score available.

|  |  |  |
| --- | --- | --- |
| Colour code | 0% - 25% | 25% - 50% |
| 50% - 75% | 75% - 100% |

Supplementary Table 5. Treatment effects (odds ratio [OR]), when all data were analyzed using confounding adjustment based on the different disease severity scores and predicted in-hospital mortality

|  |  |  |
| --- | --- | --- |
| **Disease severity measure** | **βtreatment** | **ORtreatment** |
| no confounding adjustment *a* | 0.396 | 1.49 |
| APACHE IV score (scenario 0) *b* | 0.073 | 1.08 |
| APACHE II score | 0.135 | 1.14 |
| SAPS II score | 0.100 | 1.11 |
| APACHE IV predicted in-hospital mortality *c* | 0.001 | 1.00 |
| APACHE II predicted in-hospital mortality | 0.122 | 1.13 |
| SAPS II predicted in-hospital mortality | 0.103 | 1.11 |

Legend APACHE, Acute Physiology and Chronic Health Evaluation; OR, odds ratio; SAPS, Simplified Acute Physiology Scores

*a* Conform expected; treatment was simulated such that the unadjusted treatment effect was OR 1.5

*b* Relative bias in the main article was measured with this scenario as a reference

*c* Conform expected; treatment was simulated such that – corrected for APACHE IV predicted in-hospital mortality – the adjusted treatment effect was OR 1.0

Supplementary Table 6. Treatment effects for in-hospital mortality expressed as odds ratio’s with 95% percentile confidence intervals*a,b*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | **1. Naive** | **2. Restriction** | **3. Two-stage** | **4. Standardization (groups) *c*** | **5. Standardization (centers)*d*** | **6. Multiple imputation** | **7. Regression calibration** | **8. Propensity score modelling** |
|  | **APACHE IV score** | **APACHE II score** |  |  |  |  |  |  |  |  |
| 1A | teaching (10) | non-teaching (10) | 1.12  (1.07 - 1.18) | 1.08  (1.02 - 1.15) | 1.09  (1.03 - 1.15) | 1.09  (1.03 - 1.15) | 1.09  (1.04 - 1.15) | 1.08  (1.02 - 1.14) | 1.06  (1.01 - 1.12) | 1.10  (1.05 - 1.15) |
| 2A | non-teaching (10) | teaching (10) | 1.22  (1.16 - 1.28) | 1.06  (0.95 -1.19) | 1.13  (1.07 - 1.19) | 1.13  (1.07 - 1.19) | 1.13  (1.07 - 1.19) | 1.09  (1.02 -1.14) | 1.07  (1.01 - 1.13) | 1.13  (1.08 - 1.18) |
| 3A | >1500 adm. (14) | < 1500 adm.(6) | 1.10  (1.04 -1.16) | 1.08  (1.02 -1.14) | 1.08  (1.03 -1.14) | 1.08  (1.03 - 1.14) | 1.08  (1.03 -1.14) | 1.08  (1.02 -1.14) | 1.07  (1.01 - 1.12) | 1.09  (1.05 -1.14) |
| 4A | < 1500 adm. (6) | >1500 adm. (14) | 1.19  (1.13 - 1.25) | 1.07  (0.91 - 1.27) | 1.14  (1.08 - 1.20) | 1.14  (1.08 - 1.20) | 1.14  (1.08 - 1.20) | 1.09  (1.02 - 1.16) | 1.07  (1.00 - 1.13) | 1.14  (1.09 - 1.19) |
| 5A | pred. mort. >25% (6) | pred. mort. <25% (14) | 1.21  (1.15 - 1.27) | 1.10  (1.01 - 1.19) | 1.12  (1.06 - 1.18) | 1.12  (1.06 - 1.18) | 1.12  (1.06 - 1.18) | 1.08  (1.02 - 1.14) | 1.07  (1.01 -1.12) | 1.12  (1.07 - 1.17) |
| 6A | pred. mort. <25% (14) | pred. mort. >25% (6) | 1.17  (1.11 - 1.23) | 1.07  (1.00 - 1.14) | 1.10  (1.05 - 1.16) | 1.10  (1.04 - 1.16) | 1.10  (1.04 - 1.16) | 1.08  (1.02 - 1.14) | 1.07  (1.01 - 1.12) | 1.11  (1.06 - 1.16) |
| 7A | >70% med. adm. (6) | <70% med. adm. (14) | 1.22  (1.16 - 1.28) | 1.06  (0.94 - 1.17) | 1.13  (1.07 - 1.19) | 1.13  (1.07 - 1.19) | 1.13  (1.07 - 1.19) | 1.07  (1.01 - 1.13) | 1.06  (1.01 - 1.12) | 1.13  (1.08 - 1.18) |
| 8A | <70% med. adm. (14) | >70% med. adm. (6) | 1.13  (1.08 - 1.19) | 1.08  (1.02 - 1.15) | 1.10  (1.04 - 1.15) | 1.10  (1.04 - 1.15) | 1.10  (1.04 - 1.15) | 1.08  (1.02 - 1.14) | 1.07  (1.01 - 1.12) | 1.10  (1.05 - 1.15) |
| 9A | random (12) | random (8) | 1.17  (1.11 - 1.23) | 1.09  (1.02 - 1.16) | 1.10  (1.05 - 1.16) | 1.10  (1.04 - 1.16) | 1.10  (1.04 - 1.16) | 1.08  (1.03 - 1.14) | 1.07  (1.01 - 1.12) | 1.11  (1.06 - 1.16) |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **APACHE IV score** | **SAPS II score** |  |  |  |  |  |  |  |  |
| 1B | teaching (10) | non-teaching (10) | 1.09  (1.03 - 1.15) | 1.08  (1.02 - 1.15) | 1.08  (1.03 - 1.14) | 1.08  (1.03 - 1.14) | 1.08  (1.03 - 1.14) | 1.08  (1.02 - 1.14) | 1.03  (0.97 - 1.08) | 1.09  (1.05 - 1.14) |
| 2B | non-teaching (10) | teaching (10) | 1.12  (1.06 - 1.17) | 1.06  (0.95 - 1.19) | 1.10  (1.04 - 1.16) | 1.10  (1.04 - 1.16) | 1.10  (1.04 - 1.16) | 1.08  (1.02 - 1.15) | 1.04 (0.99 - 1.11) | 1.11  (1.06 - 1.16) |
| 3B | >1500 adm. (14) | < 1500 adm. (6) | 1.08  (1.03 - 1.14) | 1.08  (1.02 - 1.14) | 1.08  (1.02 - 1.14) | 1.08  (1.02 - 1.14) | 1.08  (1.02 - 1.14) | 1.08  (1.02 - 1.13) | 1.03  (0.97 - 1.08) | 1.09  (1.04 - 1.14) |
| 4B | < 1500 adm. (6) | >1500 adm. (14) | 1.11  (1.05 - 1.17) | 1.07  (0.91 - 1.27) | 1.10  (1.05 - 1.17) | 1.10  (1.05 - 1.16) | 1.10  (1.05 - 1.16) | 1.09  (1.02 - 1.16) | 1.04  (0.97 - 1.11) | 1.12  (1.07 - 1.16) |
| 5B | pred. mort. >25% (6) | pred. mort. <25% (14) | 1.12  (1.06 - 1.18) | 1.10  (1.01 - 1.19) | 1.10  (1.04 - 1.16) | 1.10  (1.04 - 1.16) | 1.10  (1.04 - 1.16) | 1.09  (1.03 - 1.14) | 1.03  (0.97 - 1.09) | 1.11  (1.06 - 1.16) |
| 6B | pred. mort. <25% (14) | pred. mort. >25% (6) | 1.10  (1.04 - 1.16) | 1.07  (1.00 - 1.14) | 1.09  (1.03 - 1.14) | 1.08  (1.03 - 1.14) | 1.08  (1.03 - 1.14) | 1.07  (1.01 - 1.13) | 1.03  (0.98 - 1.09) | 1.10  (1.05 - 1.14) |
| 7B | >70% med. adm. (6) | <70% med. adm. (14) | 1.12  (1.06 - 1.18) | 1.06  (0.94 - 1.17) | 1.10  (1.04 - 1.16) | 1.10  (1.04 - 1.16) | 1.10  (1.04 - 1.16) | 1.08  (1.02 - 1.14) | 1.04  (0.98 - 1.09) | 1.11  (1.06 - 1.16) |
| 8B | <70% med. adm. (14) | >70% med. adm. (6) | 1.09  (1.04 - 1.15) | 1.08  (1.02 - 1.15) | 1.09  (1.03 - 1.14) | 1.08  (1.03 - 1.14) | 1.08  (1.03 - 1.14) | 1.08  (1.02 - 1.14) | 1.03  (0.97 - 1.09) | 1.10  (1.05 - 1.14) |
| 9B | random (12) | random (8) | 1.11  (1.05 - 1.17) | 1.09  (1.02 - 1.16) | 1.09  (1.04 - 1.15) | 1.09  (1.03 - 1.15) | 1.09  (1.03 - 1.15) | 1.08  (1.03 - 1.14) | 1.03  (0.98 - 1.09) | 1.11  (1.06 - 1.15) |

Legend: adm., admissions; APACHE, acute physiology and chronic health evaluation; med., medical; pred. mort., predicted in-hospital mortality; SAPS, simplified acute physiology score.

*a* Reference is data-generating model, where the treatment effect was OR 1.00, corrected for APACHE IV predicted in-hospital mortality

*b* For a description of methods, see main text (“Methods to adjust for confounding by disease severity”).

*c* APACHE II / SAPS II scores were standardized across all centers that had the respective score available.

*d* APACHE II / SAPS II scores were standardized within each center that had the respective score available