Appendix

This appendix was part of the submitted manuscript and has been peer reviewed. It is posted as supplied by the authors.

e-Supplement Box 1

The low target range studied in the SUPPORT trial (2) was not usual care in the U.S. at the time the trial was conducted (6,7). The Tin study (5) used to justify the low range selected as usual care was an outlier for the following reasons: 1) it was the only study reporting a center using an upper limit of a target range as low as 90%; 2) it involved data collected from 1990 to 1994, which was a decade earlier than SUPPORT (2003 to 2006); 3) it had a mortality rate in infants (~50%) that was substantially higher than the rates observed in more recent reports (15%-20%); and 4) usual care in northern England may well have differed from usual care in the U.S., even at the earlier time. The investigators relied on the American Association of Pediatrics (AAP) guidelines (4) recommending a range of 85% to 95% to justify the low target oxygen saturation range (85%-89%) studied in SUPPORT. However, based on published data for infants treated in more than 100 centers close to the time of the SUPPORT study, (6) target ranges sometimes included a lower limit that dipped below 85%, but those limits were always paired with an upper limit for the range ≥ 92%. Usual care was almost never confined to the lower half of the AAP recommended range (i.e., 85%-89%). Moreover, health care providers kept infants in the top half and above the target ranges most of the time, so the bottom half of the AAP range was rarely used (7). The high range studied (91%-95%) in SUPPORT overlapped with the upper half of the AAP recommended range and was consistent with usual care. But infants randomized to the arm with a narrow low oxygen range did not receive usual care; that range was experimental.

e-Supplement Box 2

In the ARMA trial, a 70-kg subject needing mechanical ventilation support would be randomized to a low volume (~ 400 ml machine breath) as opposed to a "traditional" high volume (~800 ml machine breath) (12). In usual care at the time of the trial, individual breaths from the ventilator could be set for this 70-kg person anywhere from 1000 to 400 ml, but healthcare providers would rapidly titrate the breaths downward after the initial setting, as needed, to keep the airway pressures below or near what was believed to be a safe level (~ 30.0 cm H2O) (11,20). At the enrolling centers, a breath size of 400 ml as studied in ARMA would place subjects in the 3rd percentile of usual care, and one of 800 ml as studied in the 80th percentile. In the high "traditional" breath arm, the mean airway pressure went from a believed safe pressure level (~ 30 cm H2O) before randomization to directly after randomization mean airway pressures of what was believed to be unsafe (11,16,20). The mean airway pressures reached in the high traditional arm were 33, 34 and 37 cm H2O at 24, 72 and 96 hours after enrollment, respectively (16). In fact, unless the plateau pressure reached > 50 cm H2O in an individual subject no action was allowed to lower airway pressures according to the protocol (12).

The ARMA trial enrolled 861 subjects. At its conclusion, the mortality rate was statistically significantly greater in the "traditional" control arm than in the low-size breath group (40% versus 31%) (12). 2,587 patients were screened for the trial and met inclusion criteria but were ineligible for various reasons (e.g., no consent, physician refusal) (17,18,19). Those patients continued to receive usual care by the same physicians and hospitals caring for patients enrolled in the trial. The mortality rate for the unenrolled group was only 31.7%, substantially lower than that of the "traditional" control group in the study (12,17,18,19). The study did not include a control arm using titrated, usual care to monitor safety and enable the trial to be stopped early if mortality rates were unexpectedly high in the "traditional" control arm.

e-Supplement Box 3

The TRICC trial investigated whether a hemoglobin level of 10 or 7.0 g/dL, as a single fixed trigger to transfuse RBCs, is associated with comparable mortality rates in critically ill patients (21). Before the trial, a survey of physicians by the TRICC investigators showed that age, Acute Physiology and Chronic Health Evaluation (APACHE) II scores, preoperative risk status, the existence of shock, presence of coronary ischemia, and development of anemia all significantly modified transfusion thresholds (22). In that survey, a heathy young patient would not be transfused RBCs until hemoglobin levels fell to approximately 8.3 g/dL, and this was adjusted significantly upward to 9.5 g/dL in a patient with cardiovascular disease who was actively bleeding.

A sub-group analysis of the TRICC trial demonstrated that patients with severe ischemic heart disease had a different response to the two fixed transfusion thresholds compared to patients without severe ischemic heart disease ($P = 0.03$ for an interaction) (17,24,25). In patients with ischemic heart disease, a restrictive strategy increased 30-day mortality compared to patients without ischemic heart disease, whereas the opposite pattern was seen in patients without ischemic heart disease. In the liberal arm, the younger, healthier patients were randomized to a strategy that led to more aggressive RBC transfusions than would likely have been the routine practice. Furthermore, subgroup analyses within the original report of the TRICC trial (21) indicated that the increase in mortality in the liberal transfusion group occurred in the younger (< 55 years old) or healthier patients (APACHE II scores < 20). The increased risks incurred by different subgroups in each arm makes the comparison of mortality rates between the two treatment arms uninformative (17,24,25).