Response

To the Editor:

We thank Drs Nathanson and Higgins for their interest in our work and recognize their contributions to the field through their development of the Mortality Probability Model (MPM) III. As they note, in our study models with a greater number of predictor variables provided better discriminatory capacity for the prediction of hospital mortality. When deciding whether a specific predictor variable should be included in a prognostic model, consideration must be given to the effort required to reliably obtain the data for that variable. In this regard, MPM III provides good discrimination using a small number of easily ascertained variables.

Drs Nathanson and Higgins write that “this study implies that DNR [do not resuscitate] status is not an important predictor of mortality in the ICU.” We believe this gives too much weight to the conclusions of our study. It has been previously documented that DNR status, by itself, is indeed a predictor of ICU and hospital mortality. Rather, DNR status did not—within the limits of our study as we discuss in the “Strengths and Limitations” section of our article—significantly improve the performance of recent versions of the Acute Physiology and Chronic Health Evaluation (APACHE) and Simplified Acute Physiology Score (SAPS) models.

By “performance,” we refer to the commonly used and reported measures of prognostic model assessment, namely discrimination (measured by the area under the receiver operating characteristic curve) and calibration (measured by the Hosmer-Lemeshow statistic [HLS]). We agree with Drs Nathanson and Higgins’ belief that both measures are imperfect—although they are widely reported and used. As we discuss, calibration is especially subject to a variety of influences, including sample size and case mix. We further agree that a significant HLS does not necessarily mean that a predictive model is suspect, although a nonsignificant HLS is desirable. Calibration plots for each model showed discrepancies between observed and predicted values especially at the highest and lowest deciles of risk, consistent with poor calibration.

Recognizing the problems inherent in the assessment of model performance and the trade-off between discrimination and calibration, we also reported Brier scores for the models studied (Tables 2 and 3 of our article). The Brier score, based on model prediction error, provides an overall estimate of performance. Our data demonstrated a progressive, albeit small, decrease in Brier score (ie, improved model performance) as model complexity increased. The addition of DNR status lowered the Brier scores for both APACHE models and for SAPS 3, suggesting its potential value as a predictor variable. As suggested, we also calculated both the Bayesian information criterion (BIC) and the corrected Akaile information criterion (AIC) for prognostic models with and without inclusion of DNR status. For each criterion, and similar to the Brier score analyses, the addition of DNR status was associated with a small improvement in model performance (approximately 3% decrease in AIC and BIC). The small improvements in Brier scores, BIC, and AIC associated with the addition of DNR status were not, however, reflected in statistically significant differences in area under the receiver operating characteristic curve or a consistent directional change in the value of the HLS.

Mark T. Keegan, MB
Ognjen Gajic, MD, FCCP
Bekele Afsessa, MD, FCCP
Rochester, MN

Affiliations: From the Division of Critical Care, Department of Anesthesiology, the Division of Pulmonary and Critical Care, Department of Medicine (Drs Gajic and Afsessa), and the Multidisciplinary Epidemiologic and Translational Research in Intensive Care (METRIC) group (Drs Keegan, Gajic, and Afsessa), Mayo Clinic.

Financial/nonfinancial disclosures: The authors have reported to CHEST that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

Address for correspondence: Mark T. Keegan, MB, Mayo Clinic, Department of Anesthesiology, Charlton 1145, 200 First St SW, Rochester, MN 55905; e-mail: keegan.mark@mayo.edu

© 2013 American College of Chest Physicians. Reproduction of this article is prohibited without written permission from the American College of Chest Physicians. See online for more details.

DOI: 10.1378/chest.12-2783

How Much Hypoxia Is Significant in Pulmonary Hypertension During Air Travel?

To the Editor:

In their interesting study published in CHEST (October 2012), Roubinian and colleagues reported a high incidence (26%) of hypoxemia in air passengers with pulmonary hypertension. This was based on the study’s definition of a “meaningful” oxygen desaturation as a mean arterial oxygen saturation ($\text{Sp}_2$) $< 85\%$. The study was predicated on the potential for in-flight hypoxemia to cause adverse effects through hypoxic pulmonary vasoconstriction and further elevation in pulmonary artery pressures. Since this work was conducted, our studies have been published in systemic pulmonary artery pressure (SPAP) during commercial flights using echocardiography, and our results suggest that a higher $\text{Sp}_2$ threshold may be more appropriate.

In a study of healthy passengers during a 9-h flight, we found that mean $\text{Sp}_2$ fell to 95% and SPAP increased by 6 mm Hg or about 20%. In a passenger with a genetic cause of increased pulmonary vasoreactivity (Chuvash polythemia) studied during a 6-h flight, $\text{Sp}_2$ only fell to 96%, yet SPAP increased by 15 mm Hg or about 50%. Interestingly, Roubinian and colleagues reported that 24% of the patients studied experienced symptoms without developing hypoxemia (defined as $\text{Sp}_2 < 85\%$). Although our echocardiographic findings have limitations, they nevertheless suggest that milder hypoxia could still have contributed to these