Dominique Laroche, M.D., Ph.D., Philippe Gomis, M.D., Jean-Marc Malinovsky, M.D., Ph.D., Paul-Michel Mertes, M.D., Ph.D. Hôpitaux Universitaires de Strasbourg, Service d'Anesthésie-Réanimation, Nouvel Hôpital Civil, Strasbourg, Cedex, France (P.-M.M.). mertes.paul@wanadoo.fr; paul-michel.mertes@chru-strasbourg.fr

References

1. Laroche D, Gomis P, Gallimidi E, Malinovsky JM, Mertes PM: Diagnostic value of histamine and tryptase concentrations in severe anaphylaxis with shock or cardiac arrest during anesthesia. Anesthesiology 2014; 121:272–9
2. Sirvent AE, González C, Enríquez R, Fernández J, Millán I, Barber X, Amorós F: Serum tryptase levels and markers of renal dysfunction in a population with chronic kidney disease. J Nephrol 2010; 23:282–90
3. Volcheck GW, Mertes PM: Local and general anesthetics immediate hypersensitivity reactions. Immunol Allergy Clin North Am 2014; 34:525–46, viii
4. Mertes PM, Malinovsky JM, Jouffroy L, Aberer W, Terreehorst I, Brockow K, Demoly P; Working Group of the SFAR and SFA; ENDA; EAACI Interest Group on Drug Allergy: Reducing the risk of anaphylaxis during anesthesia: 2011 updated guidelines for clinical practice. J Investig Allergol Clin Immunol 2011; 21:442–53
5. Fisher M, Baldo BA: Anaphylaxis during anaesthesia: Current aspects of diagnosis and prevention. Eur J Anaesthesiol 1994; 11:263–84
6. Laxenaire MC, Moneret-Vautrin DA: Allergy and anaesthesia. Cur Opinion Anaesth 1992; 5:436–41

(Accepted for publication November 25, 2014.)

More Attention to Respiration: A Simple but Effective Approach to Reduce Postoperative Mortality?

To the Editor:

Congratulations to Mazo *et al.*[1] for their elaborate and extensive work. They designed the Prospective Evaluation of a RIsk Score for postoperative pulmonary COmPlications in Europe (PERISCOPE) study to improve the external validity of the Assess Respiratory Risk in Surgical Patients in Catalonia (ARISCAT) score, which they describe as the only prospective internal validated score to predict postoperative pulmonary complications.[2] Therefore, they tested the generalizability of this score in a large European cohort and three subsamples: Spain, Western Europe, and Eastern Europe. They conclude that their risk score predicts three levels of postoperative pulmonary complications in an area outside the development setting.

Yet, to us the generalizability of the score seems doubtful, because the postoperative mortality reported in this study is inconsistent with the mortality reported in the European Surgical Outcomes Study (EUSOS).[3]

This inconsistency is conspicuous because the design, especially sampling strategies, of PERISCOPE resembles in many details the EUSOS study:

- Both groups performed a multicenter design including numerous European hospitals.
- Both groups defined continuous 7-day cohort periods to collect data of patients undergoing an in-hospital surgical procedure.
- Both studies excluded patients undergoing obstetric procedures.
- Both studies observed in-hospital mortality as an important outcome variable.

PERISCOPE (n = 5.099 patients) reports an overall in-hospital mortality of 0.9%, Spain (n = 2.000): 1.0%, Western Europe (n = 1.538): 0.8%, Eastern Europe (n = 1561): 0.9%. The crude mortality in the EUSOS study (n = 46.539 patients) was 4% ranging from 1.2% in the participating hospitals in Iceland (n = 162) to 21.5% in Latvia (n = 302). The United Kingdom provided the biggest sample of n = 10.630 patients, the mortality rate was 3.6%. In Spain (n = 5.433), 3.8% of surgical patients died.

This significant difference between both studies is especially remarkable due to the high-risk surgical procedures like cardiac or neurosurgery, which are included in the PERISCOPE but not in the EUSOS study. With respect to the aim of the PERISCOPE study, which is to improve generalizability, we consider it therefore indispensable to include this observation in the validation of the predictive score.

Two of the authors were involved in both publications. We wonder, why they did not discuss this important possible restriction of their validation study. Possibly the analysis of the ARISCAT study, in which the score to predict pulmonary complications was developed, gives an important clue to interpret the data.

Mazo *et al.*[1] refer to the excellent internal validity of the ARISCAT study. Internal validity means optimal control of study conditions to ensure that the covariation of predictive score and outcome is not biased (nonspuriousness). Nonetheless, the better the internal validity is, the more limited is the external validity, *i.e.*, the more elaborated the strategies are to control confounding influences, the more limited is the generalizability of a study.

The PERISCOPE study increases the external validity of the predictive score by a large degree of replication of the ARISCAT design in a new sample of patients. This strategy limits this generalizability to the special conditions as reported in the ARISCAT study. These conditions differ from the EUSOS investigation with high external and less controlled internal validity. Thus, the differences in mortality reported in the optimal controlled ARISCAT and PERISCAT studies compared with the EUSOS study may be explained simply by the effects induced into the participating hospitals by the studies itself. It is possible that the fact that respiration was studied directed the attention of hospital staff toward more careful observation of the respiration of postsurgical patients.

From our point of view, is it useful to consider this following aspect: if the authors conclude that increased attention to respiration (*e.g.*, simply measuring oxygen saturation) may have contributed to reduce mortality, we will obtain a very easy to handle but highly effective approach to significantly reduce mortality in our hospitals.

Competing Interests

The authors declare no competing interests.

Gerhard Brodner, M.D., Ph.D., Hugo Van Aken, M.D., Ph.D. Fachklinik Hornheide, Muenster, Germany (G.B.). gerhard.brodner@fachklinik-hornheide.de

References

1. Mazo V, Sabaté S, Canet J, Gallart L, Gama de Abreu M, Belda J, Langeron O, Hoeft A, Pelosi P: Prospective external validation of a predictive score for postoperative pulmonary complications. Anesthesiology 2014; 121:219–31
2. Canet J, Gallart L, Gomar C, Paluzie G, Vallès J, Castillo J, Sabaté S, Mazo V, Briones Z, Sanchis J; ARISCAT Group: Prediction of postoperative pulmonary complications in a population-based surgical cohort. Anesthesiology 2010; 113:1338–50
3. Pearse RM, Moreno RP, Bauer P, Pelosi P, Metnitz P, Spies C, Vallet B, Vincent JL, Hoeft A, Rhodes A; European Surgical Outcomes Study (EuSOS) group for the Trials groups of the European Society of Intensive Care Medicine and the European Society of Anaesthesiology: Mortality after surgery in Europe: A 7 day cohort study. Lancet 2012; 380:1059–65

(Accepted for publication November 25, 2014.)

In Reply:

We thank Drs. Brodner and Van Aken for their interest in our study of the Prospective Evaluation of a RIsk Score for postoperative pulmonary COmPlications in Europe (PERISCOPE)[1] and for giving us the opportunity to extend the discussion of perioperative risk assessment through this correspondence. They have questioned the evidence for the generalizability of the score developed in the Assess Respiratory Risk in Surgical Patients in Catalonia (ARISCAT) study[2] based on two issues. One is the different mortality rates in the PERISCOPE cohort and the larger cohort of the European Surgical Outcomes Study (EUSOS).[3] The second issue is the external validation process used to explore the utility of the ARISCAT score in wider European settings, by applying it to the PERISCOPE sample and subsamples. We will comment on these two issues separately.

First, we point out that the question of postoperative mortality as reported after EUSOS[3] has been discussed in correspondence between Drs. Brodner and Van Aken[4] and the EUSOS authors.[5] Thus, the need for caution before assuming that 4% is the true incidence of postoperative mortality in Europe has already been covered, and it has been pointed out that the heterogeneity of countries and hospitals and the differences in the sample sizes contributed by each of the EUSOS centers account for the mortality observed and the dispersion of rates.

Neither the PERISCOPE[1] nor EUSOS[3] mortality rates of 0.9% and 4%, respectively, should be interpreted as representative of any particular population because both cohorts were convenience samples rather than random population-based ones. This aspect of design, however, does not mean that either study is biased or inaccurate. We note here that other large series[6,7] have recorded postoperative mortality rates very close to ours. In any case, we emphasize that neither mortality (as an outcome) nor other variables that were not predictors in the ARISCAT score[2] can be discussed as central concerns in the context of the PERISCOPE validation study. We only contribute these observations to reflect on the profiles of the two European samples. Even if the EUSOS mortality rate, found in a larger population, could somehow be considered a definitive standard, or reference figure, it would still be entirely valid to perform an external validation of a predictive model for complications in a population with a different mortality profile from the EUSOS cohort's. External validation is a dynamic process in which an understanding of performance in different settings progressively increases confidence in a score's generalizability or clinical reliability. If we find a score is unhelpful, we will know we need to learn more.

Our last comment regarding the issue of comparison to EUSOS[3] is that other than similar reliance on convenience cohorts in that study and ours, we cannot agree that the two designs were similar. However, as mentioned above, we do not wish to go into extensive detail in comparing the studies, so suffice it to say that the primary outcome of our study was the presentation of a postoperative pulmonary complication not mortality.

Next, Drs. Brodner and Van Aken make certain affirmations about internal and external validity that do merit discussion here because we would not wish readers to be misled. It is wrong to argue that a finding of excellent internal validity, such as the ARISCAT[2] score showed, represents a limitation to external validity. It is precise when a predictive model has shown internal validity that its generalizability is worth exploring externally.[8] It is true that the discrimination and calibration of models are usually optimistic in their development sample, but this is the very reason why they should then undergo external validation, which might support or rule out transportability. Following recommendations from specialists in the field,[8–10] we used rigorous collection and analysis methods in the PERISCOPE study, whose design was praised in an editorial in this journal.[11] We take this opportunity to express our thanks for that praise, but to have done otherwise than control the design carefully would certainly have led to confusing results.

The concluding hypothesis of Drs. Brodner and Van Aken, that increased attention to measuring oxygen saturation may have helped to reduce mortality in the PERISCOPE[1] cohort, is attractive but we cannot, of course, confirm it based on our data. We think it might be a strategy worth studying in an appropriate clinical trial, however.

Finally, we want to emphasize that, in our opinion, the greatest strength of our study lies in the replication itself, which is an essential and often overlooked procedure to verify the validity of a predictive model. We agree with Eisenach and Houle[11] that reproducibility, replication, and generalizability