Commentary: Evaluation of primary graft dysfunction after lung transplantation--It is time to teach an old dog new tricks!

W. Hampton Gray, MD, and P. Michael McFadden, MD

Outcomes after lung transplantation have continued to improve during the last 20 years. Although the incidence of primary graft dysfunction (PGD) is decreasing, PGD remains a significant cause of morbidity and mortality in the posttransplant period. Grading of PGD is based on criteria proposed by the International Society for Heart and Lung Transplantation. These grades are determined by a scoring system, graded 0 to 3. The most recent revision by the International Society for Heart and Lung Transplantation was published in 2016. PGD is graded radiographically by evidence of infiltrates on chest radiography (Figure 1) and the ratio of partial pressure of oxygen to the fraction of inspired oxygen clinically. These parameters are evaluated at set time points after surgery (0-6 hours, 24 hours, 48 hours, and 72 hours). A patient with a poor ratio of partial pressure of oxygen to the fraction of inspired oxygen (<200) could potentially have grade 3 PGD versus grade 0 PGD, depending on interpretation of the observed chest radiograph. Often in the early postoperative period, chest radiographs can be difficult to interpret, which could affect grading for PGD and subsequently alter appropriate treatment. Evaluation of the accuracy and variability of specialized thoracic radiologists in determining PGD after lung transplantation is extremely important and has not been extensively studied.

In this issue of The Journal of Thoracic and Cardiovascular Surgery, Schwarz and colleagues report the results of a retrospective study of 497 lung transplants during a 6-year period. They examined whether interobserver variability exists in scoring of postoperative chest radiographs and the ultimate effect of such variability on the grade of PGD. In total, 1988 chest radiographs were evaluated in the study period. Full agreement of all 5 radiologists was present for only 43.0% of the entire cohort. There were 4 time points at which radiographs were examined (0, 24, 48, and 72 hours). Moderate agreement was found among the 5 radiologists at 0 and 24 hours (κ = 0.456 and κ = 0.456, respectively), and agreement continued to decrease at the later points (κ = 0.405 at 48 hours and κ = 0.409 at 72 hours). Best- and worst-case scenarios on further analysis led to PGD grade 3 rates of 8.4% versus 28.4% at 0 hours, 1.8% versus 4.8% at 24 hours, 2.0% versus 5.3% at 48 hours, and 0.2% versus 3.1% at 72 hours. Patients with an increased body mass index and those with size-reduced transplants were associated with higher rates of interobserver variability. Schwarz and colleagues conclude that substantial interobserver variability underlines the difficulty to grade posttransplant organ function adequately.

As we continue to strive to improve outcomes after lung transplantation, it is also important to validate the systems that we have in place for recognizing and treating complications in the perioperative period. PGD is an important example. Early recognition and treatment of severe PGD is essential to improving early outcomes. The findings of Schwarz and colleagues support the fact that the current grading system for PGD can vary significantly, depending on who interprets the chest radiograph after lung transplantation. Because the grading system for PGD includes 1 objective measure (ratio of partial pressure of oxygen to the fraction of inspired oxygen) and 1 subjective measure, it is not at all surprising that the findings of Schwarz and colleagues demonstrate substantial variability for such a
subjective interpretation. Fewer than half of the radiologists in the study (43%) were in full agreement about any of the chest radiographs after lung transplantation. This is no better than a mere “coin toss”; in fact, it is worse! This is concerning, given the fact that these were dedicated thoracic radiologists, and in most institutions radiographs are often read by a general radiologist. With such participants, full agreement likely would have been even lower.

Schwarz and colleagues3 are to be congratulated for such a large cohort of patients and a well-designed study that raises an interesting and important question of whether our current grading system for PGD is accurate, reliable, and reproducible. More objective measurements might allow for a more standardized and potentially reproducible grading system, but this remains unknown. To address this question adequately, further studies are needed at a multi-institutional level, with exponentially more radiologists truly to compare interobserver variability. Until then, it is important for us to understand the potential variability that exists in grading PGD after lung transplantation and what effect this may have on patient outcomes. It is clear that a close collaborative approach that includes both a strict radiographic interpretation and experienced medical and surgical clinical judgment have resulted in the best outcomes in PGD after lung transplantation.

References

FIGURE 1. Primary graft dysfunction 1 day after lung transplantation requiring ECMO support.