

Thank You to Jacob Iasha Sznajder, M.D., Editor in Chief of the *American Journal of Respiratory and Critical Care Medicine*



As we look into the new year, the ATS leadership would like to take this opportunity to thank Jacob I. Sznajder, M.D., for his outstanding stewardship as Editor in Chief of the *American Journal of Respiratory and Critical Care Medicine* over the past 4 years. We celebrate all he and his team have accomplished, including maintaining the premier status of the *Journal* in the fields of pulmonary disease, critical care medicine, and sleep-disordered breathing.

Dr. Sznajder has overseen numerous improvements to the *Journal*, including the addition of new, timely content such as concise clinical reviews, pulmonary perspectives, recommended readings by fellows, the image section, podcasts, videos, and downloadable PowerPoint figures. All of his efforts have strengthened the reputation of the *Journal*, increased readability, and maintained the impact factor of 11.04, the highest impact factor of all journals in the fields of critical care medicine and the respiratory system.

Dr. Sznajder's term will end on December 31, 2014. As we begin our search for a new Editor, per the search

announcement at atsjournals.org, the Search Committee is seeking candidates who are active investigators in the field and have substantial editorial experience. We look forward to receiving applications from candidates who possess all the virtues of Dr. Sznajder with a vision to lead the *Journal* into the future. ■

Patricia Finn, M.D.
President, the American Thoracic Society

Roy Brower, M.D.
Chair, Publications Policy Committee

Diane Gern
Publisher

Copyright © 2014 by the American Thoracic Society

Challenges in Interpreting Wheeze Phenotypes: The Clinical Implications of Statistical Learning Techniques



"We are drowning in information but starving for knowledge" (1). This is especially relevant to biomedicine, where our ability to generate new data has increased exponentially, resulting in a deluge of information. We aspire to use this information to predict disease outcomes and understand their causes, so that we can design stratified prevention strategies and targeted treatments. Along with the exponential increase in the amount of data being generated, there have been major advances in computationally intensive statistical machine learning techniques to identify patterns within such large data sets (2). One particular area of interest is using data reduction techniques (such as latent class analysis) on longitudinal profiles of patients' symptoms to identify meaningful patterns and help disentangle subtypes of diseases. Latent class analysis is a statistical method that allows identification of heterogeneous groups of patients who have different symptom characteristics. These latent classes are hypothetical constructs that may represent distinct diseases with common observable characteristics but different profiles of manifestations and comorbidities over time. Different external manifestations of symptom profiles may indicate distinct causes and underlying biological mechanisms, and associations with different genetic and environmental factors (3). The hope is that in the future, we may be able to better understand the nature

of complex diseases and therefore develop more personalized, targeted treatment and management strategies, thus transforming clinical care. But can we also identify meaningful patterns in these data that may be useful to practicing physicians now?

One of the areas where statistical machine learning and pattern recognition have been increasingly applied is in asthma and allergic disease, specifically in identifying distinct subtypes of wheezing phenotypes during childhood (4–9) and of atopy based on skin test and IgE responses (10, 11). The basic hypothesis is that asthma is not a single disease, but rather a diagnostic label applied to patients with wheeze, breathlessness, cough, and chest tightness, which are variable across time. Underneath this umbrella term may lie a collection of heterogeneous diseases. This asthma heterogeneity may be more prominent in adults than in children, but begins with different patterns of wheezing in childhood. In this issue of the *Journal*, Depner and colleagues (pp. 129–138) of the Protection against Allergy—Study in Rural Environments (PASTURE) study make an important contribution in systematically comparing statistically inferred latent classes based on the natural history of wheeze during childhood and clinically identified phenotypes based on patient histories, wheeze triggers, and treatment response (12). As in other studies using similar techniques, the authors identify

four phenotypes of childhood wheezing illness using latent class analysis: transient early wheeze, persistent wheeze, late-onset wheeze, and an intermediate-onset wheeze. Of note, these distinct phenotypes are associated with different biological and environmental characteristics. One of the important features of this work is that although there was strong correlation between latent classes and clinically assigned diagnoses based on symptom severity and triggers, the dichotomous clinical phenotypes (such as asthma diagnosis, multitrigger wheeze, unremitting wheeze, and frequent wheeze) seem to be equally correlated with distinct latent classes of late-onset, persistent, and intermittent wheeze. This points toward the divergence between currently used aggregated dichotomous disease diagnoses and epidemiologically derived data-driven distinct disease subtypes with possibly different underlying biological mechanisms.

It is worth noting that latent class analysis is just one of many different dimensionality reduction techniques that can be used to make sense of the current proliferation of data. Expanding the current practice of looking only at profiles of wheeze to joint modeling of groups of symptoms (such as cough or congestion [13, 14]) may help better understanding of distinct asthma subgroups.

The clinical community often perceives that one of the problems of data-driven approaches is that different techniques often provide different classifications of asthma subgroups, and that the results may differ from a clinician-based approach (e.g., in ascertaining disease severity) (15). We wish to emphasize that cautious interpretation of the results from unsupervised learning analyses with respect to their direct application to clinical practice is warranted. One of the major challenges facing unsupervised statistical learning techniques is to identify an optimally meaningful group of symptoms and other phenotypic markers on which to base the inference, so that the outcomes are consistent and clinically meaningful. Prosperi and colleagues opened up the discussion surrounding the challenges in identifying asthma subgroups using statistical machine learning, suggesting that if we are to fully capitalize on these techniques, we need urgent conversation and agreement on how best to select the phenotypic and/or other markers and how to use these consistently across different study populations (15).

The time has come for a more reflective process to optimally describe and understand the problem space we are trying to define. This requires an iterative and interdisciplinary dialogue between the clinician, the statistician, and the basic scientist. Such a cross-discipline iterative process is essential if we are to refine our current tools for disaggregating complex phenotypes and find optimal models to describe the underlying latent structure in the data. Understanding the underlying disease heterogeneity and the associated causal mechanisms will enable the discovery of clinically meaningful subgroups of asthma. The identification of such subgroups should lead to the novel therapeutic target discovery and enable a genuinely personalized approach to disease management. This process, rather than being hypothesis driven, is an iterative hypothesis-generating approach that is informed by clinical knowledge.

It is encouraging to see an article like this with a cross-disciplinary team contributing to the generation of new knowledge (12). It sends the message that speculation on the clinical implications of complex statistical techniques is important. Although we have not yet reached a point at which we fully understand the clinical relevance of disease phenotypes that have been identified using different statistical learning methodologies,

with the recognition that asthma is not a single disease, we may be at the end of the beginning of this journey. It is worth pursuing this journey with the more holistic approach offered by systems biology, integrating different levels of information to combine high-dimensional datasets consisting of genetic, environmental, and biological data to fully understand asthma and allergic disease. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Danielle Belgrave, Ph.D.
Institute of Inflammation and Repair
University of Manchester and University Hospital of South Manchester
Manchester, United Kingdom
and
Institute of Population Health
University of Manchester
Manchester, United Kingdom

Angela Simpson, M.D., Ph.D.
Adnan Custovic, M.D., Ph.D.
Institute of Inflammation and Repair
University of Manchester and University Hospital of South Manchester
Manchester, United Kingdom

References

1. Naisbitt J. Megatrends. New York: Warner; 1982. pp. 222–229.
2. Bishop CM. Model-based machine learning. *Philos Trans A Math Phys Eng Sci* 2013;371:20120222.
3. Belgrave DC, Custovic A, Simpson A. Characterizing wheeze phenotypes to identify endotypes of childhood asthma, and the implications for future management. *Expert Rev Clin Immunol* 2013;9:921–936.
4. Belgrave DC, Simpson A, Semic-Jusufagic A, Murray CS, Buchan I, Pickles A, Custovic A. Joint modeling of parentally reported and physician-confirmed wheeze identifies children with persistent troublesome wheezing. *J Allergy Clin Immunol* 2013;132:575–583.e12.
5. Chen Q, Just AC, Miller RL, Perzanowski MS, Goldstein IF, Perera FP, Whyatt RM. Using latent class growth analysis to identify childhood wheeze phenotypes in an urban birth cohort. *Ann Allergy Asthma Immunol* 2012;108:311–315.e1.
6. Granell R, Henderson AJ, Timpson N, St Pourcain B, Kemp JP, Ring SM, Ho K, Montgomery SB, Dermitzakis ET, Evans DM, et al. Examination of the relationship between variation at 17q21 and childhood wheeze phenotypes. *J Allergy Clin Immunol* 2013;131:685–694.
7. Granell R, Sterne JA, Henderson J. Associations of different phenotypes of wheezing illness in early childhood with environmental variables implicated in the aetiology of asthma. *PLoS ONE* 2012;7:e48359.
8. Henderson J, Granell R, Heron J, Sherriff A, Simpson A, Woodcock A, Strachan DP, Shaheen SO, Sterne JA. Associations of wheezing phenotypes in the first 6 years of life with atopy, lung function and airway responsiveness in mid-childhood. *Thorax* 2008;63:974–980.
9. Savenije OE, Granell R, Caudri D, Koppelman GH, Smit HA, Wijga A, de Jongste JC, Brunekreef B, Sterne JA, Postma DS, Henderson J, Kerkhof M. Comparison of childhood wheezing phenotypes in 2 birth cohorts: ALSPAC and PIAMA. *J Allergy Clin Immunol* 2011;127:1505–1512.e14.
10. Lazić N, Roberts G, Custovic A, Belgrave D, Bishop CM, Winn J, Curtin JA, Hasan Arshad S, Simpson A. Multiple atopy phenotypes and their associations with asthma: similar findings from two birth cohorts. *Allergy* 2013;68:764–770.
11. Simpson A, Tan VY, Winn J, Svensén M, Bishop CM, Heckerman DE, Buchan I, Custovic A. Beyond atopy: multiple patterns of sensitization in relation to asthma in a birth cohort study. *Am J Respir Crit Care Med* 2010;181:1200–1206.
12. Depner M, Fuchs O, Genuneit J, Karvonen AM, Hyvärinen A, Kaulek V, Roduit C, Weber J, Schaub B, Lauener R, et al.; The PASTURE Study Group. Clinical and epidemiologic phenotypes of childhood asthma. *Am J Respir Crit Care Med* 2014;189:129–138.

13. Smith JA, Drake R, Simpson A, Woodcock A, Pickles A, Custovic A. Dimensions of respiratory symptoms in preschool children: population-based birth cohort study. *Am J Respir Crit Care Med* 2008;177:1358–1363.
14. Spycher BD, Silverman M, Pescatore AM, Beardsmore CS, Kuehni CE. Comparison of phenotypes of childhood wheeze and cough in 2 independent cohorts. *J Allergy Clin Immunol* 2013;132:1058–1067.
15. Prosperi MC, Sahiner UM, Belgrave D, Sackesen C, Buchan IE, Simpson A, Yavuz TS, Kalayci O, Custovic A. Challenges in identifying asthma subgroups using unsupervised statistical learning techniques. *Am J Respir Crit Care Med* 2013;188:1303–1312.

Copyright © 2014 by the American Thoracic Society

Inhomogeneous Computed Tomographic Densities in Lungs in Acute Respiratory Distress Syndrome: Stress Multipliers Leading to Ventilator-induced Injury?



In this issue of the *Journal*, Cressoni and colleagues (pp. 149–158) extend their investigations into the mechanisms of ventilator-induced lung injury (VILI) in patients with acute respiratory distress syndrome (ARDS) (1). In approaching the question of how relatively small tidal volumes in ARDS can cause lung injury, they first acknowledge the concepts of “the second hit” and “baby lung.” They then go on to ask if stress amplification at the boundary between regions of lung of differing density might further contribute to the development of VILI. They base their hypothesis on the theoretical analysis of Mead and colleagues (2). Mead reasoned that the stress needed to expand a region of homogeneous lung would be the transpulmonary pressure, and that if a small region of the lung were to decrease in volume, or become atelectatic, it would be subjected to a greater expanding stress, thus amplifying the inflating pressure needed to restore its original volume.

To explore this question, Cressoni and colleagues used computed tomographic (CT) scans from 148 patients with ARDS and 100 control subjects to estimate normal and pathological levels of lung inhomogeneity that could lead to stress amplification and localized lung injury. Because the relative expansion of different parts of the lung could not be directly measured, they used the spatial inhomogeneity of lung CT density as a measure of underinflation in the denser regions. Inhomogeneous density was assessed over the entire lung by comparing the density of each individual voxel with the average density of a surrounding spherical shell with a diameter of 2.4 to 3.6 mm. Because the density of 95% of these voxels in normal lungs differed from that of their surrounding shells by a factor of less than 1.61, this value was chosen as the normal limit for inhomogeneity, and higher ratios were defined as abnormal inhomogeneity. In normal lungs, inhomogeneous regions accounted for about 5% of the voxels and occurred principally at the periphery of the lung and near the large vessels and hilar structures. In patients with ARDS, inhomogeneous voxels were 18% of the total and occurred largely at the interface between “healthy” and “diseased” regions (i.e., at the border between regions with high and normal CT density).

Patients were categorized according to the severity of ARDS, which was significantly related to the proportion of abnormally inhomogeneous voxels, being 14% in mild ARDS and 23% in severe ARDS. As would be expected from this, the proportion of abnormally inhomogeneous voxels was also significantly correlated

with $\text{PaO}_2/\text{FiO}_2$, physiological shunt fraction, plateau pressure, and dead-space fraction. Finally, the lung inhomogeneity was significantly greater in patients who died than in survivors. The extent of lung inhomogeneity appears to be an important indicator of ARDS severity and predicts outcome, and the authors are to be congratulated for this novel finding.

However, whether inhomogeneity in CT density implies stress amplification of the type Mead and colleagues described remains an open question. Their analysis predicted that a transpulmonary pressure of 30 cm H_2O throughout a homogeneous lung would increase to approximately 140 cm H_2O at the surface of a spherical region that had decreased to one-tenth of its original volume. This assumes that the constant tensile force on alveolar walls would be applied to a smaller spherical surface of the contracted region. Their analysis, which was illustrated with a lung model constructed from condoms, required that the shape of the lung region undergoing contraction be preserved as its volume decreased, so that a spherical region would contract to a smaller sphere, etc. To the extent that alveolar contraction occurs principally in one dimension, as in plate-like atelectasis, stress amplification could be far less than suggested by Mead and colleagues.

More importantly, as acknowledged by Cressoni and colleagues, there is uncertainty concerning the interpretation of CT density, specifically, whether it can actually be interpreted as indicating alveolar collapse in diseased lungs. As articulated by Hubmayr (3), the dorsoventral gradient of CT density seen in ARDS may be due largely to differences in the amount of alveolar fluid content, and may not simply reflect a gradient of alveolar size. Hubmayr suggests that the denser regions of diseased lungs could have alveolar size similar to surrounding regions, whose alveoli are not flooded. If this were true, it would preclude increases in stress associated with alveolar size heterogeneity.

If this alternative hypothesis is correct and stress amplification is not important, how would we explain the correlations of extent of inhomogeneity with assessments of disease severity, physiological indices of disease, and mortality? To the extent that more severely diseased lungs are likely to have more regions with either flooding or consolidation, it would not be surprising that the extent of inhomogeneity reflects the extent of pathology and correlates with respiratory dysfunction.

A recent imaging study may bear on the question of whether stress amplification occurs at the margins of collapsed regions in