Critical Care Research Using "Big Data": A Reality in the Near Future

Kwangha Lee

Division of Pulmonary, Allergy and Critical Care Medicine, Department of Internal Medicine, Pusan National University School of Medicine, Busan, Korea

Data science can be defined as "the set of fundamental principles that support and guide the principled extraction of information and knowledge from data" [1]. Current advances in data science have resulted in a new era of "big data." Digitalization of the health-care system and the use of big data have changed the way that we can conduct clinical research [2], because size, complexity, and dynamic nature are beyond the scope of traditional data collection and analysis methods [3]. Also, widespread implementation of electrical medical records is bringing the world of data science to the patient's bedside; the potential benefits of big data to health care will be increasingly significant. Recently, big data and data science have been gradually infiltrating many aspects of clinical research in critical care, which could be familiar to many intensivists.

Regarding data science applications in critical care [4], the most common applications of data science are predictive and prognostic models using supervised learning algorithms. One of the oldest and best-known prognostic models to estimate mortality risk in intensive care unit (ICU) patients is the Acute Physiology and Chronic Health Evaluation (APACHE) score [5]. In other examples of data science use in critical care, much research with clustering and phenotypes, physiologic waveform analysis, image analysis, and natural language processing has been recently performed [4]. In Korea, many studies in critical care using big data have been based on data from the National Health Insurance Service-National Sample Cohort (NHIS-NSC), which can be considered representative of the general population [6].

Currently, some studies of the *Acute and Critical Care* have been performed using NHIS-NSC. Oh et al. [7] reported the influence of immunosuppressants on in-hospital mortality. Recently, Heo et al. [8] reported the characteristics and mortality changes of all ICU patients over the previous 8 years. They concluded that older patients admitted to the ICU were at risk for higher Charlson Comorbidity Index score [9]. Also, a tendency toward a decrease in the hazard ratio of long-term mortality was reported. In this issue of the *Acute and Critical Care*, Oh et al. [10] reported association between socioeconomic status of ICU patients and mortality. They reported that 30-day mortality after ICU admission was not associated with socioeconomic status in the National Health Insurance coverage setting, and occupation was associated with 1-year mortality after ICU admission. However, epidemiologic concepts suggest moving to critical care studies using big data on contemporary issues.

In the future, many critical care studies using big data will be analyzed by computational systems and made available to physicians to augment the clinician's decision-making capac-

ity. To maximize the benefit of big data, critical care physicians need to gain familiarity with the analytical techniques used by data scientists. Also, clinicians need further understanding on related fields of study, such as molecular characteristics, physiological analysis, and image analysis, through a collaborative and scientifically rigorous approach.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

ORCID

Kwangha Lee   https://orcid.org/0000-0001-9878-201X

REFERENCES

1. Provost F, Fawcett T. Data science and its relationship to big data and data-driven decision making. Big Data 2013;1:51-9.
2. Bates DW, Saria S, Ohno-Machado L, Shah A, Escobar G. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. Health Aff (Millwood) 2014; 33:1123-31.
3. Docherty AB, Lone NI. Exploiting big data for critical care research. Curr Opin Crit Care 2015;21:467-72.
4. Sanchez-Pinto LN, Luo Y, Churpek MM. Big data and data science in critical care. Chest 2018;154:1239-48.
5. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985;13:818-29.
6. Lee J, Lee JS, Park SH, Shin SA, Kim K. Cohort profile: the National Health Insurance Service-National Sample Cohort (NHIS-NSC), South Korea. Int J Epidemiol 2017;46:e15.
7. Oh SY, Cho S, Lee H, Chang EJ, Min SH, Ryu HG. Sepsis in patients receiving immunosuppressive drugs in Korea: analysis of the national insurance database from 2009 to 2013. Korean J Crit Care Med 2015;30:249-57.
8. Heo J, Hong Y, Han SS, Kim WJ, Kwon JW, Moon KW, et al. Changes in the characteristics and long-term mortality rates of intensive care unit patients from 2003 to 2010: a nationwide population-based cohort study performed in the Republic of Korea. Acute Crit Care 2018;33:135-45.
9. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373-83.
10. Oh TK, Jo J, Jeon YT, Song IA. Impact of socioeconomic status on 30-day and 1-year mortalities after intensive care unit admission in South Korea: a retrospective cohort study. Acute Crit Care 2018;33:230-7.