Can You Read Me Now? Unlocking Narrative Data with Natural Language Processing

Michael W. Sjoding^{1,2} and Vincent X. Liu³

¹Department of Internal Medicine, and ²Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, Michigan; and ³Division of Research, Kaiser Permanente, Oakland, California

The rapid digitization of medicine brought about by widespread application of electronic health records offers the promise of access to massive new collections of data for patient care and clinical research (1). However, much of these data exist as narrative text documents; for example, as clinical notes or radiology and pathology reports. As a result, these data, often called “unstructured” because they cannot be neatly stored in the row-and-column schema of standard databases, can remain tantalizingly out of reach without the use of painstaking manual review. And yet, contained in these documents is a wealth of information about patients’ medical histories and treatments, their trajectories of illness and recovery, and their clinical outcomes—vital information for contextualizing and improving patient health.

Computer scientists and computational linguists have long tackled the challenge of extracting usable information from narrative text through steady advances in Natural Language Processing (NLP). NLP is a branch of computer science focused on developing computer algorithms that can transform written or spoken human language into a form that can be analyzed by computation. Today, NLP is used in seemingly mundane activities like auto-correcting spelling and grammatical errors or classifying certain email content as “spam.” It is also used in highly advanced applications like Apple’s Siri technology, which allows users to interact with smartphone programs through voice commands. Importantly, NLP now encompasses a broad platform of

well-established tools that enable high-throughput and scalable access to the critical information locked within text-based documents, well beyond what is feasible through manual curation. And consistent with the trends toward open-sourcing software used in even the most cutting-edge informatics applications, powerful NLP tools are now freely available and readily accessible.

Despite its promise in biomedical research, NLP has seen relatively modest use. For example, among Medline-indexed publications from 1985 to 2015, fewer than 400 articles each year used NLP (Figure 1), despite more than 800,000 indexed citations in 2015. However, use of NLP has risen steadily since 2003, likely mirroring the rapidly expanding use of advanced electronic health records.

NLP-enabled approaches have the potential to dramatically accelerate novel discovery as well as improve care delivery in healthcare, as they already have in other industries. Early examples include the use of over 9 million clinical notes (2) or web-search histories drawn from 57 million search-engine users (3) to facilitate the discovery of unrecognized adverse drug reactions or the real-time provision of patient risk adjustment based on the automated analysis of comorbid illness documented within clinical notes (4). Another high-value target for using NLP approaches is to facilitate improved care coordination between providers, for example, by ensuring that key inpatient events are conveyed to outpatient care providers.

In this issue of *AnnalsATS*, Weissman and colleagues (pp. 1538–1545) use NLP to study care “hand-offs” when patients with acute respiratory distress syndrome (ARDS) transition from inpatient to outpatient care at hospital discharge (5). A substantial body of evidence shows that survivors of ARDS suffer significant long-term consequences, including functional disability (6), cognitive deficits (7), and psychiatric symptoms (8), collectively named the post-intensive care syndrome (9). They also remain at increased risk of premature mortality (10). The lack of adequate planning and communication at the time of hospital discharge may be an important barrier to ensuring that patients who develop these complications are appropriately managed in the outpatient setting (11). Thus, conveying these crucial inpatient events, and perhaps even further providing recommendations that alert outpatient providers to underappreciated complications, may yield improved care for long-term post-ARDS morbidity.

In their study, the Weissman and colleagues examined the discharge summaries of 815 patients who developed ARDS during hospitalization to determine whether crucial events—including intensive care unit admission, receipt of mechanical ventilation, ARDS development, and the patient’s potential for complications related to post-intensive care syndrome (i.e., cognitive, psychiatric, or physical impairments)—were documented. However, instead of performing manual reviews of each

(Received in original form June 27, 2016; accepted in final form July 4, 2016)

Supported by National Institutes of Health grants T32 HL007749 (M.W.S.) and K23 GM112018 (V.X.L.).

Correspondence and requests for reprints should be addressed to Michael Sjoding, M.D., M.Sc., University of Michigan, 3916 Taubman Center, 1500 E. Medical Center Dr., SPC 5360, Ann Arbor, MI 48109-5360. E-mail: msjoding@med.umich.edu

Ann Am Thorac Soc Vol 13, No 9, pp 1443–1445, Sep 2016

Copyright © 2016 by the American Thoracic Society

DOI: 10.1513/AnnalsATS.201606-498ED

Internet address: www.atsjournals.org

Figure 1. The number of Medline-indexed citations with “natural language processing” in any search field.

discharge document, the authors used an NLP-enabled approach to automate searches of the narrative text. They developed term groups relevant to each concept of interest (e.g., ARDS development, mechanical ventilation, intensive care unit admission, and symptoms of post-intensive care syndrome). They then used open-source software to identify instances within the text. To ensure that their automated search strategy was accurate, they validated their approach by manually reviewing a 5% sample of discharge documents, finding near-perfect accuracy.

The authors found that the use of mechanical ventilation and admission to the intensive care unit were documented in 92% and 83% of discharge documents, respectively. However, the development of ARDS was documented in only 13% of discharge summaries among patients with ARDS. The lack of documentation of ARDS in discharge summaries was strongly correlated with lack of documentation in daily progress notes, a finding further supporting the strong evidence that ARDS frequently goes unrecognized, or at least under-reported (12). The documentation of symptoms potentially related to post-intensive care syndrome by patients was also present in

only a minority of charts (38%). The authors conclude that the glaring lack of documentation regarding key critical care events and potential long-term consequences of critical illness is a missed opportunity to ensure that patients receive high-quality care during this vulnerable care transition.

The study has several limitations. For example, in performing primarily keyword-based searches of the narrative text, the authors did not comprehensively address many of the challenges of extracting medical concepts from text with NLP (known as named-entity recognition tasks). Frequently, keyword searches are not sufficient to reliably extract information from narrative medical reports, as sentences containing symptoms or diagnosis are often negated (i.e., “no evidence of pneumonia”) (13), the findings are associated with a level of clinical uncertainty (i.e., “cannot rule out pneumonia”) (14), or may be only partially relevant to the patient (i.e., “father with history of colon cancer”). However, the study’s use of discharge summaries is likely to be less vulnerable to these limitations than other types of clinical notes. Further, the study’s manual validation of their strategy also demonstrated its sufficiency for extracting information from patient discharge summaries. One notable

exception was in the identification of symptoms of the post-intensive care syndrome, where an automated search of the word “depressed” identified the word as part of the phrase “depressed ejection fraction.” To effectively extract symptoms related to post-intensive care syndrome from clinical narratives, the authors conclude that more advanced NLP approaches will likely be necessary.

This study is part of a small but growing body of research using NLP to evaluate and improve the quality of medical care. Highlights include a study by Murff and others who found that the majority of post-operative complications found by NLP within narrative reports were not present in corresponding hospital diagnosis codes, even though these codes are used to evaluate overall hospital performance (15). Danforth and others derived automated methods for identifying patients with incident pulmonary nodules by analyzing a population-level sample of radiology reports (16), including patients whose nodules were ultimately diagnosed as malignancy—a technique that has the potential to ensure these patients get appropriate follow-up care. Similar approaches, building from the current study’s methods, offer the potential to facilitate seamless longitudinal care especially to prevent the “loss in translation” that can accompany patients’ transitions between care settings.

The utility of NLP approaches is likely to increase substantially, not only as a tool to drive novel discovery in massive clinical datasets but also to facilitate the delivery of high-quality care with actionable signals within the sea of noisy electronic health record data. Driving this innovation is a strong community of researchers committed to developing and sharing open-source NLP software tailored for medical applications. ■

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

References

- 1 Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA* 2013;309:1351–1352.
- 2 Wang G, Jung K, Winnenburg R, Shah NH. A method for systematic discovery of adverse drug events from clinical notes. *J Am Med Inform Assoc* 2015;22:1196–1204.
- 3 White RW, Wang S, Pant A, Harpaz R, Shukla P, Sun W, DuMouchel W, Horvitz E. Early identification of adverse drug reactions from search log data. *J Biomed Inform* 2016;59:42–48.
- 4 Health Fidelity. Technology. [accessed 2016 Jun 23]. Available from: <http://healthfidelity.com/technology/>
- 5 Weissman GE, Harhay MO, Lugo RM, Fuchs BD, Halpern SD, Mikkelsen ME. Natural language processing to assess

- documentation of features of critical illness in discharge documents of ARDS survivors. *Ann Am Thorac Soc* 2016;13:1538–1545.
- 6 Herridge MS, Tansey CM, Matté A, Tomlinson G, Diaz-Granados N, Cooper A, Guest CB, Mazer CD, Mehta S, Stewart TE, *et al.*; Canadian Critical Care Trials Group. Functional disability 5 years after acute respiratory distress syndrome. *N Engl J Med* 2011;364:1293–1304.
 - 7 Hopkins RO, Weaver LK, Collingridge D, Parkinson RB, Chan KJ, Orme JF Jr. Two-year cognitive, emotional, and quality-of-life outcomes in acute respiratory distress syndrome. *Am J Respir Crit Care Med* 2005;171:340–347.
 - 8 Huang M, Parker AM, Bienvenu OJ, Dinglas VD, Colantuoni E, Hopkins RO, Needham DM; National Institutes of Health, National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome Network. Psychiatric symptoms in acute respiratory distress syndrome survivors: a 1-year national multicenter study. *Crit Care Med* 2016;44:954–965.
 - 9 Elliott D, Davidson JE, Harvey MA, Bemis-Dougherty A, Hopkins RO, Iwashyna TJ, Wagner J, Weinert C, Wunsch H, Bienvenu OJ, *et al.* Exploring the scope of post-intensive care syndrome therapy and care: engagement of non-critical care providers and survivors in a second stakeholders meeting. *Crit Care Med* 2014;42:2518–2526.
 - 10 Wang CY, Calfee CS, Paul DW, Janz DR, May AK, Zhuo H, Bernard GR, Matthay MA, Ware LB, Kangelaris KN. One-year mortality and predictors of death among hospital survivors of acute respiratory distress syndrome. *Intensive Care Med* 2014;40:388–396.
 - 11 Needham DM, Davidson J, Cohen H, Hopkins RO, Weinert C, Wunsch H, Zawistowski C, Bemis-Dougherty A, Berney SC, Bienvenu OJ, *et al.* Improving long-term outcomes after discharge from intensive care unit: report from a stakeholders' conference. *Crit Care Med* 2012;40:502–509.
 - 12 Bellani G, Laffey JG, Pham T, Fan E, Brochard L, Esteban A, Gattinoni L, van Haren F, Larsson A, McAuley DF, *et al.*; LUNG SAFE Investigators; ESICM Trials Group. Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. *JAMA* 2016;315:788–800.
 - 13 Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. Evaluation of negation phrases in narrative clinical reports. *Proc AMIA Symp* 2001:105–109.
 - 14 Liu V, Clark MP, Mendoza M, Saket R, Gardner MN, Turk BJ, Escobar GJ. Automated identification of pneumonia in chest radiograph reports in critically ill patients. *BMC Med Inform Decis Mak* 2013;13:90.
 - 15 Murff HJ, FitzHenry F, Matheny ME, Gentry N, Kotter KL, Crimin K, Dittus RS, Rosen AK, Elkin PL, Brown SH, *et al.* Automated identification of postoperative complications within an electronic medical record using natural language processing. *JAMA* 2011;306:848–855.
 - 16 Danforth KN, Early MI, Ngan S, Kosco AE, Zheng C, Gould MK. Automated identification of patients with pulmonary nodules in an integrated health system using administrative health plan data, radiology reports, and natural language processing. *J Thorac Oncol* 2012;7:1257–1262.