Natural Language Processing to Assess Documentation of Features of Critical Illness in Discharge Documents of Acute Respiratory Distress Syndrome Survivors

Gary E. Weissman[1,2], Michael O. Harhay[2,3], Ricardo M. Lugo[4], Barry D. Fuchs[1], Scott D. Halpern[1,2,3], and Mark E. Mikkelsen[1,3]

[1]Pulmonary, Allergy, and Critical Care Division, Hospital of the University of Pennsylvania, [2]Leonard Davis Institute of Health Economics, and [3]Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, Pennsylvania; and [4]Division of Cardiovascular Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee

ORCID ID: 0000-0001-9588-3819 (G.E.W.).

Abstract

Rationale: Transitions to outpatient care are crucial after critical illness, but the documentation practices in discharge documents after critical illness are unknown.

Objectives: To characterize the rates of documentation of various features of critical illness in discharge documents of patients diagnosed with acute respiratory distress syndrome (ARDS) during their hospital stay.

Methods: We used natural language processing tools to build a keyword-based classifier that categorizes discharge documents by presence of terms from four groups of keywords related to critical illness. We used a multivariable modified Poisson regression model to infer patient- and hospital-level characteristics associated with documentation of relevant keywords. A manual chart review was used to validate the accuracy of the keyword-based classifier, and to assess for ARDS documentation during the hospital stay.

Measurements and Main Results: Of 815 discharge documents, ARDS was identified in only 111 (13%). Mechanical ventilation was identified in 770 (92%) and intensive care unit (ICU) admission in 693 (83%) of discharge documents. Symptoms or recommendations related to post–intensive care syndrome were included in 306 (38%) of discharge documents. Patient age (older; relative risk [RR] = 0.97/yr, 95% confidence interval [CI] = 0.96–0.98) and higher Pa_{O_2}:FI_{O_2} (decreasing illness severity; RR = 0.96/10-unit increment, 95% CI = 0.93–0.98) were associated with decreased documentation of ARDS. Being discharged from a surgical (RR = 0.33, 95% CI = 0.22–0.50) compared with a medicine service was also associated with decreased rates of ARDS documentation. The manual chart review revealed 98% concordance between ARDS documentation in the discharge summary and during the hospital stay. Accuracy of the document classifier was 100% for ARDS and mechanical ventilation, 98% for ICU admission, and 95% for symptoms of post–intensive care syndrome.

Conclusions: In the discharge documents of survivors of ARDS, ARDS itself is rarely mentioned, but mechanical ventilation and ICU stay frequently are. The low rates of documentation of ARDS appear to be concordant with low rates of documentation during the hospital stay, consistent with known underrecognition in the ICU. Natural language processing tools can be used to effectively analyze large numbers of discharge documents of patients with critical illness.

Keywords: acute respiratory distress syndrome; clinical informatics; patient discharge; critical care

(Received in original form February 19, 2016; accepted in final form May 22, 2016)

Supported by NHLBI/NIH training grants T32 HL098054 (G.E.W.) and F31 HL127947 (M.O.H.).

Author Contributions: conception and design of the work—G.E.W., S.D.H., and M.E.M.; acquisition of data for the work—G.E.W., R.M.L., and B.D.F.; analysis of data for the work—G.E.W. and M.O.H.; revision of the work for critically important intellectual content, final approval of the version submitted for publication, and accountability for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved—G.E.W., M.O.H., R.M.L., B.D.F., S.D.H., and M.E.M.

Correspondence and requests for reprints should be addressed to Gary E. Weissman, M.D., Hospital of the University of Pennsylvania, 3400 Spruce Street, Gibson 5002, Philadelphia, PA 19104. E-mail: gary.weissman@uphs.upenn.edu

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Ann Am Thorac Soc Vol 13, No 9, pp 1538–1545, Sep 2016

DOI: 10.1513/AnnalsATS.201602-131OC
Internet address: www.atsjournals.org

The acute respiratory distress syndrome (ARDS) is a devastating condition characterized by lung injury associated with bilateral pulmonary infiltrates and hypoxemia (1). With in-hospital mortality rates declining in recent years for the estimated 190,600 cases of ARDS nationally (2, 3), there are now more patients living with the long-term sequelae of ARDS. Because outpatient providers are increasingly caring for survivors of ARDS, "optimization of handoffs between levels of care and practitioners" (4) represents a target for improving postdischarge care transitions for survivors of critical illness (5).

The hospital discharge summary plays an important role in promoting safe transitions of care after hospital discharge (6). The Joint Commission has recommended that all hospital discharge summaries include the "reason for hospitalization" and "significant findings" (7). This recommendation aligns with qualitative studies of survivors of critical illness that revealed that survivors demand, yet are often unaware of, the details of their critical illness (8, 9). Despite the Joint Commission's recommendations, a review found that a primary diagnosis and a hospital course were present in only 17.5 and 14.5% of discharge summaries, respectively (10).

In addition to ARDS, there are other features of critical illness relevant to communicate to outpatient providers. Mechanical ventilation, for example, is itself associated with long-term neuropsychological and physical impairments, disability, and increased mortality (11–15). Similarly, survivors frequently experience decreased quality of life after critical illness (16–18). The post–intensive care syndrome (PICS), defined as new or worsening impairments in cognitive, psychiatric, or physical function after critical illness (4, 9), is common (5). Survivors of mechanical ventilation and ARDS, specifically, are at particularly high risk of developing PICS (19–23).

Care for patients after critical illness frequently falls on outpatient providers. To understand the patient and family experience during the hospital stay, enumerate specific diagnoses and their severity, and anticipate postdischarge sequelae, the hospital discharge summary must correctly document salient features of critical illness. We hypothesized that many features of critical illness in survivors of ARDS would be infrequently documented at time of discharge, as ARDS itself is commonly underrecognized and underdocumented during hospitalization (24–26). Natural language processing (NLP) is a novel set of tools for examining large numbers of electronic health records, yet its use in patients with critical illness has been extremely limited (27). Using NLP, we sought to determine the frequency of documentation of multiple features of critical illness at the time of discharge among patients diagnosed with ARDS during their hospital course. Some of the results of this investigation have been previously reported in the form of an abstract (28).

Methods

An expanded METHODS section with detailed code examples and supplementary analysis is available in the online supplement.

Population

We identified patients with ARDS using a prospectively captured cohort based on a previously validated early-detection algorithm (29). A stand-alone computer uses this algorithm to automatically review all radiology reports, vital signs, patient location data, and results of arterial blood gases in real time for all inpatients in the University of Pennsylvania Health System (UPHS). Prior work demonstrated that the screening system had a sensitivity of 97.6% and specificity of 97.6% (29). A determination of the presence or absence of ARDS is made from these clinical data based on the Berlin criteria (1).

Between January 1, 2013 and November 12, 2015, this detection algorithm identified 1,797 patients with ARDS among UPHS inpatients. After excluding readmissions, discharges to an inpatient hospice facility, admissions with length of stay less than 1 day or greater than 1 year, patients at high risk of cardiac failure, neurosurgical patients, those with $F_{I_{O_2}}$ less than 0.5 at the time of diagnosis, and patients who died in the hospital, we identified electronic discharge documents for 815 unique survivors of ARDS. *See* Figure 1 for details of study enrollment.

Natural Language Processing

The free-text components of each discharge document were combined, then preprocessed using removal of numbers, punctuation, and superfluous whitespace. Then stopwords, commonly appearing English words (e.g., "a," "if," "of," "is," "the") that offer little in document meaning, but can significantly increase computational complexity, were removed. We used a standard stemming algorithm to capture similar word roots (e.g., "intubated" and "intubation") (30).

We used an iterative process based on the authors' clinical judgment, local experience with health system–specific jargon (e.g., MICU, NICU, CTSICU), and relevant literature, to build four groups of keywords related to different aspects of critical illness (Table 1). These groups contain keywords that directly identify ARDS (group A), suggest the use of mechanical ventilation (group B), suggest

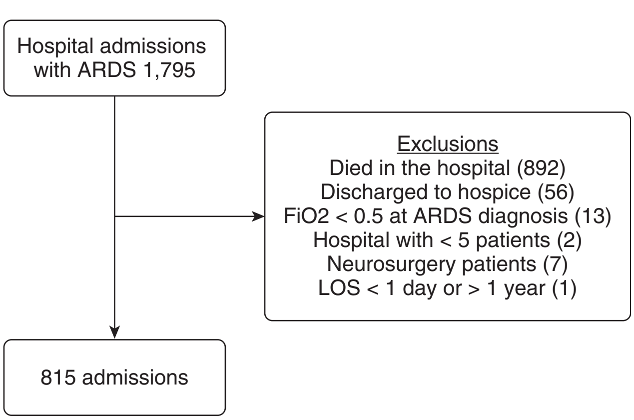

Figure 1. Selection and exclusions of the sample population. ARDS = acute respiratory distress syndrome; LOS = length of stay.

Table 1. Keyword groups that describe acute respiratory distress syndrome (group A), mechanical ventilation (group B), intensive care unit admission (group C), and symptoms related to the post–intensive care syndrome (group D)

Group A	Group B	Group C	Group D
Acute lung injury	Extubate	CCU	Anxiety
Acute respiratory distress syndrome	Intubate	Critical care	Brain dysfunction
ALI	Mechanical ventilation	Critical illness	Cognitive impairment
ALI/ARDS	VDRF	Critically ill	Confusion
ARDS	Vent	CTICU	Delirium
	Vent dependent respiratory failure	CTSICU	Depression
	Ventilator	ICU	Executive dysfunction
		Intensive care	ICU Delirium
		Intensive care unit	Immobility
		MICU	Memory dysfunction
		NICU	Memory impairment
		SICU	Mobility impairment
		TSICU	Physical impairment
			PICS
			Post–intensive care syndrome
			Post-traumatic stress disorder
			PTSD
			Weak
			Weakness

Definition of abbreviations: ALI = acute lung injury; ARDS = acute respiratory distress syndrome; CCU = coronary care unit; CTICU = cardiothoracic intensive care unit; CTSICU = cardiothoracic surgical intensive care unit; ICU = intensive care unit; MICU = medical intensive care unit; NICU = neuro–intensive care unit; PICS = post–intensive care syndrome; PTSD = post-traumatic stress disorder; SICU = surgical intensive care unit; TSICU = trauma surgical intensive care unit; VDRF = vent-dependent respiratory failure.
Notes: a keyword-based document classifier searched for all keywords in each tier to determine if any were present in each discharge summary of patients with ARDS. All searches were case insensitive. A sensitivity analysis that employed "fuzzy matching" to account for spelling errors did not find any difference in rates of documentation of these keywords.

an intensive care unit (ICU) admission (group C), and identify symptoms associated with PICS (group D). Although interrelated, the groups were selected to understand how the critical illness was communicated at the time of hospital discharge and, through the process, elucidate what "significant findings" were most frequently documented among survivors of ARDS. In addition, we piloted group D to understand how frequently PICS terms were included in discharge summaries, recognizing that documentation could reflect symptoms experienced or, as anticipatory guidance for patients and providers, symptoms encountered after discharge. A keyword-based document classifier, using the R statistical computing language (31) with the *tm* (32) and *RWeka* (33) packages, with exact matching determined if each document contained any keywords in each group.

Because the content of these documents is generated by clinicians and entered by electronic keyboard, we employed a "fuzzy matching" search using the *stringdist* package (34) in a sensitivity analysis to account for spelling errors in the text. A detailed code example for the primary and sensitivity analyses is found in the online supplement.

Statistical Analysis

The unit of analysis was the discharge summary. We report summary statistics for observed characteristics in the study population (Table 2). A multivariable, modified Poisson regression with robust error variance estimation (35, 36) was used to infer patient- and hospital-level factors associated with documentation of keywords from each group. Covariates were chosen from data available from the UPHS electronic database based on clinical judgment of relevance, with removal of collinear terms determined by measurement of the variance inflation factor for each. We report the relative risk (RR) for each model parameter. Statistical analyses were performed using Stata/IC 14.1 (College Station, TX) (37). The study was approved by the Institutional Review Board of the University of Pennsylvania (Philadelphia, PA).

Validation

We conducted a manual chart review of the text of 40 (5%) randomly selected discharge documents to assess for accuracy of our keyword-based classifier. Keywords were considered appropriately identified if their presence in the discharge document indicated that the condition or diagnosis occurred during the hospital stay. For PICS-related terms (group D), we also allowed for the use of the term in the context of anticipatory guidance for the patient. In addition, we reviewed the daily progress notes from the primary team and any critical care consult service documentation for 7 days before and after the date of electronic diagnosis of ARDS, or for the complete ICU stay, whichever was shorter, to assess for physician documentation of ARDS. All manual reviews were conducted by the same investigator (G.E.W.).

Results

Of 815 discharge documents, group A terms were present in only 111 (13%), group B and C terms were found in 770 (92%) and 693 (83%) of discharge documents, respectively, and group D terms were present in 306 (38%) of discharge documents.

For ARDS terms, patient-level factors associated with decreased documentation included patient age (older; RR = 0.97/yr, 95% confidence interval [CI] = 0.96–0.98), male sex (RR = 0.66, 95% CI = 0.47–0.92),

Table 2. Characteristics of the study cohort

Variable	Count
Discharge summaries, n	815
Documentation rate, n (%)	
Group A	109 (13)
Group B	755 (93)
Group C	677 (83)
Group D	306 (38)
Sex, n (%)	
Female	335 (41)
Male	480 (59)
Race, n (%)	
White	493 (60)
Black	214 (26)
Other	108 (14)
Median age at discharge, yr (IQR)	61 (49–71)
Median length of stay, d (IQR)	25 (14–40)
Median time from diagnosis to discharge, wk (IQR)	2.7 (1.5–3.6)
Median word count (100s) (IQR)	6.1 (4.4–8.6)
Median P:F at diagnosis (IQR)	134 (88–198)
ARDS severity* at diagnosis	
Mild	196 (24)
Moderate	365 (45)
Severe	254 (31)
Hospital, n (%)	
HUP	690 (85)
PPMC	125 (15)
Service type at discharge, n (%)	
Medicine	420 (52)
Surgery	395 (48)
Year, n (%)	
2013	330 (40)
2014	333 (41)
2015	152 (19)

Definition of abbreviations: ARDS = acute respiratory distress syndrome; HUP = Hospital of the University of Pennsylvania; IQR = interquartile range; PPMC = Penn Presbyterian Medical Center; P:F = Pa_{O_2}:Fi_{O_2} ratio.
*ARDS severity is defined by the Berlin criteria such that the Pa_{O_2}:Fi_{O_2} ratio for each category is mild (200–300), moderate (100–200), and severe (≤100).

and higher Pa_{O_2}:Fi_{O_2} (decreasing illness severity; RR = 0.96/10-unit increment, 95% CI = 0.93–0.98). Increased delay between diagnosis of ARDS and hospital discharge was associated with higher rates of documentation of ARDS (RR = 1.39/log wk, 95% CI = 1.05–1.84) and PICS (RR = 1.24/log wk, 95% CI = 1.11–1.39) terms. The only hospital-level factor independently, and inversely, associated with documentation of ARDS terms was being discharged from a surgical (RR = 0.33, 95% CI = 0.22–0.50) compared with a medicine service. The most frequently occurring keywords are presented in Table 3. The most frequently occurring terms across the entire corpus are found in Figure 2, and by document prevalence in Figure 3.

Increasing length of discharge summaries was associated with higher rates of documentation of mechanical ventilation (RR = 1.01/100-word increment, 95% CI = 1.01–1.02), ICU admission (RR = 1.02, 95% CI = 1.01–1.03), and PICS terms (RR = 1.05, 95% CI = 1.03–1.07). There were no differences in rates of documentation between blacks and whites. The full results of the primary analysis are found in Table 4.

The results did not differ significantly in the sensitivity analysis when accounting for misspelled keywords (*see* Part 1 of the online supplement). The manual chart review (Table 5) revealed documentation of ARDS during the hospital stay in only 7 of 40 (18%) charts. Furthermore, the manual chart review demonstrated that ARDS documentation, or lack thereof, aligned between the hospital stay and discharge summary in 39 of 40 (98%) charts reviewed. Specifically, in only one instance wherein ARDS was not documented in the discharge summary, an attending physician had documented ARDS during the hospitalization. In the manually reviewed sample, we found perfect accuracy of the keyword-based classifier for terms describing ARDS (group A) and mechanical ventilation (group B). We detected one false negative for terms describing ICU admission (group C), where the classifier missed the presence of the term "MICU" (it was written "MICU. MICU"), because the term was at both the end and beginning of two sentences without a space between them. We discovered two false-positive results for PICS-related terms (group D), where the

Table 3. Occurrences of stemmed keywords appearing in the corpus

Keyword	Count (*n*)	Percentage (%)
Intub	533	65.4
Extub	521	63.9
Ventil	342	42.0
Micu	304	37.3
Icu	221	27.1
Sicu	214	26.3
intens care	160	19.6
intens care unit	152	18.7
Vent	145	17.8
Delirium	130	16.0
mechan ventil	112	13.7
Ards	104	12.8
weak*	96	11.8
Depress	81	9.9
Vdrf	56	6.9
Anxiety	47	5.8
Confus	43	5.3
critic ill[†]	36	4.4
Ccu	23	2.8
icu delirium	16	2.0
vent depend respiratori failur	11	1.3
Ctsicu	11	1.3
Immobl	8	1.0
critic care	6	1.0
Nicu	5	1.0
acut respiratori distress syndrom	4	<0.5
acut lung injuri	3	<0.5
Ptsd	3	<0.5
Cticu	1	<0.5
cognit impair	1	<0.5

Each keyword of interest is "stemmed" and matched against an equivalently "stemmed" corpus in order to capture different inflections of the word. For example, "intub" will match against words that were originally "intubate" and "intubated." Keywords that did not appear at all are not shown.
*Both "weak" and "weakness" share the stem, "weak."
[†]Both "critical illness" and "critically ill" share the stem, "critic ill."

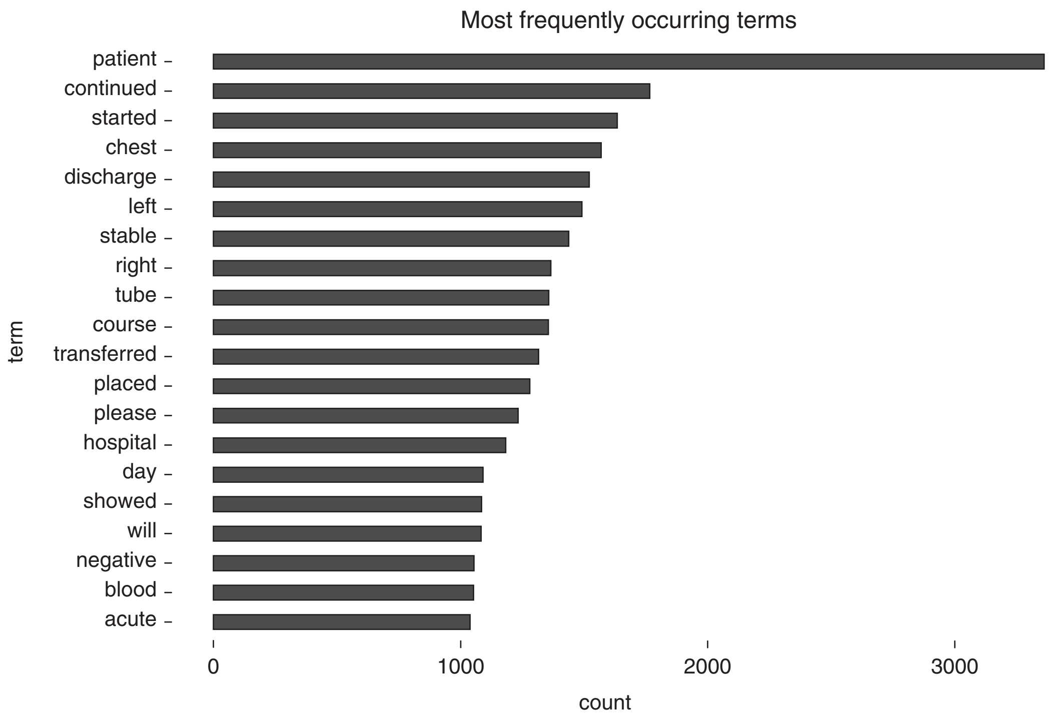

Figure 2. Raw count of the number of times the most common terms appeared in the corpus.

term was a modifier of something unrelated to PICS symptoms ("*depression* of left ventricular systolic function" and "*weak* gag").

Discussion

At the time of hospital discharge, we found that older, male survivors of ARDS, and those discharged from a surgical service, were less likely to have ARDS documented in the discharge summary. For unclear reasons, potentially explained by residual

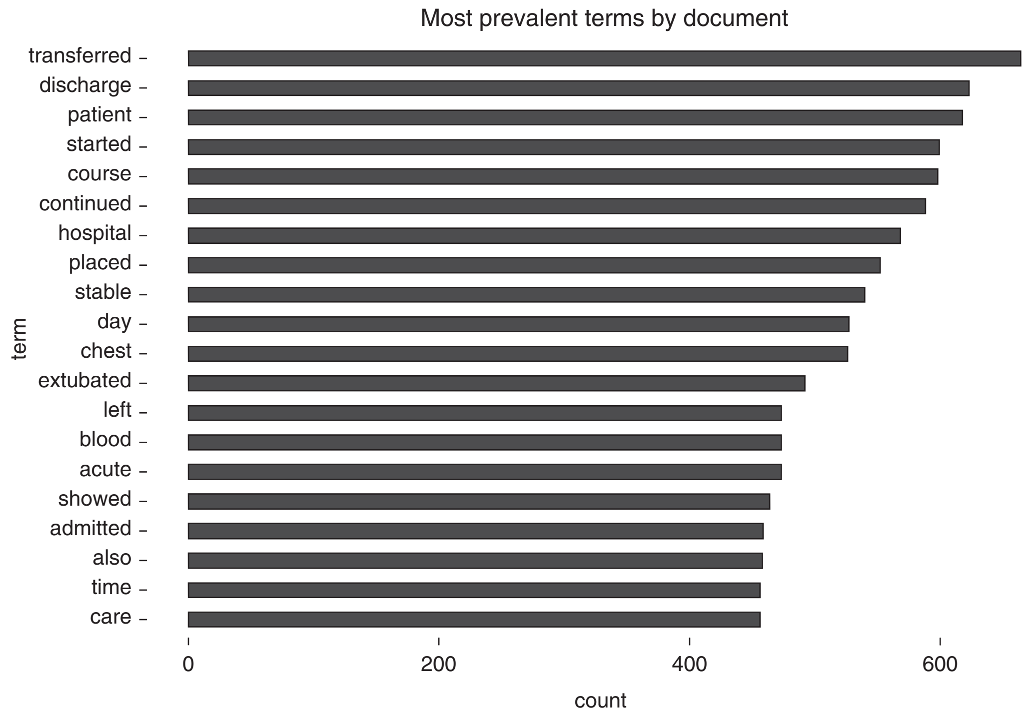

Figure 3. Count of the most common terms appearing anywhere in a document.

Table 4. Primary analysis—multivariable modified poisson regression model results

Variable	Group A Model			Group B Model			Group C Model			Group D Model		
	RR	95% CI	*P* Value	RR	95% CI	*P* Value	RR	95% CI	*P* Value	RR	95% CI	*P* Value
Age	0.97	0.96–0.98	<0.001	1.00	0.99–1.00	0.694	1.00	1.00–1.01	0.593	1.01	1.00–1.01	0.013
Time lag*	1.39	1.05–1.84	0.021	0.98	0.95–1.01	0.217	0.96	0.92–1.00	0.052	1.24	1.11–1.39	<0.001
Pa_{O_2}:FI_{O_2}†	0.96	0.93–0.98	0.002	1.00	0.99–1.00	0.049	1.00	0.99–1.00	0.929	0.99	0.98–1.01	0.258
Word count‡	1.00	0.96–1.05	0.832	1.01	1.01–1.02	<0.001	1.02	1.01–1.03	<0.001	1.05	1.03–1.07	<0.001
Facility												
HUP	1											
PPMC	0.78	0.45–1.34	0.359	0.92	0.86–0.99	0.033	0.99	0.90–1.08	0.749	1.06	0.84–1.35	0.617
Year§	0.99	0.79–1.24	0.908	1.01	0.98–1.03	0.583	1.01	0.97–1.05	0.649	0.95	0.84–1.08	0.446
Sex												
Female	1											
Male	0.66	0.47–0.92	0.014	0.96	0.92–0.99	0.021	0.97	0.91–1.03	0.337	1.02	0.86–1.21	0.808
Race												
White	1											
Black	0.88	0.59–1.32	0.543	1.03	0.98–1.08	0.280	1.00	0.93–1.08	0.989	1.00	0.80–1.24	0.976
Other	1.01	0.63–1.61	0.963	0.99	0.94–1.05	0.815	0.92	0.83–1.02	0.108	1.17	0.93–1.48	0.181
Discharging service												
Medicine	1											
Surgery	0.33	0.22–0.50	<0.001	1.06	1.01–1.10	0.009	0.98	0.92–1.05	0.565	0.70	0.59–0.84	<0.001

Definition of abbreviations: ARDS = acute respiratory distress syndrome; CI = confidence interval; HUP = Hospital of the University of Pennsylvania; PPMC = Penn Presbyterian Medical Center; RR = relative risk.
*The log of the time delay in weeks between electronic diagnosis of ARDS and hospital discharge.
†The Pa_{O_2}:FI_{O_2} ratio in 10-unit increments.
‡The word count of the free text of the discharge summary in 100-word increments.
§The year specified as a continuous variable.

confounding in terms of illness severity, we also found that a longer delay between ARDS diagnosis and discharge was associated with an increased rate of ARDS documentation. Consistent with studies examining documentation of ARDS at the time of onset (26), we found that increased illness severity was associated with increased documentation at hospital discharge. Whether these associations reflect systematic differences in timing of ARDS or differences in pattern recognition among these patients requires further investigation.

Although ARDS was infrequently documented, mechanical ventilation and admission to an ICU were accounted for in 92 and 83% of discharge documents, respectively. Through chart review, we found that the low rates of ARDS documentation at the time of discharge were concordant with low rates of documentation during the hospital stay, consistent with multiple prior studies on low recognition of ARDS (24–26). These findings suggest that the low rates of ARDS documentation at discharge are not due to information loss or recall bias as the patient transitions through the hospital. Rather, the data suggest that the plague of ARDS underrecognition (24–26) persists throughout the hospitalization, an observation that has potential implications for quality of care (e.g., use of lung-protective ventilation). In addition, as survivors of ARDS crave information about their acute illness, long-term consequences, and recovery process (38), our findings signify a potential quality gap that requires attention to better educate and prepare survivors.

Table 5. Results of manual chart review of 40 (5%) discharge documents

Feature	Count n (%)
ARDS in daily progress note	7 (18)
ARDS in discharge document concordant with progress notes	39 (98)
Accuracy of keyword-based document classifier by keyword group	
Group A	40 (100)
Group B	40 (100)
Group C	39 (98)
Group D	38 (95)

Definition of abbreviation: ARDS = acute respiratory distress syndrome.
The review examined presence of ARDS documentation in the daily progress notes and the accuracy of the findings of the keyword-based classifier.

Our chart review also provides useful results for future investigations that employ NLP to analyze discharge documents of patients with critical illness. Given the near-perfect concordance for terms related to ARDS and mechanical ventilation, a keyword-based classifier is likely sufficient for their identification during the hospital course in a discharge document. Symptoms of PICS, or the recommendation to assess for such symptoms, however, should be analyzed with some combination of tools for named entity recognition and information extraction (e.g., to distinguish between "*depressed* ventricular function" and "*depressed* mood"). The use of individual keywords in these cases is more variable, and therefore more prone to false-positive results in the context of discharge documents. The one false-negative result for

ICU admission terms suggests that human-generated, keyboard-entered text can produce errors, not just in spelling, but also in punctuation, that may be better recognized using sentence boundary detection tools before removal of punctuation.

We found that symptoms associated with PICS were mentioned in approximately one-third of discharge documents. As issues of survivorship are rarely addressed during critical illness (39), it was not surprising that our analyses and chart review revealed that guidance for patients and providers, in terms of preparing for the possibility of experiencing PICS, was uncommon at the time of discharge. However, symptoms present during the acute illness that could foreshadow PICS were documented. Specifically, delirium and confusion were documented in 16 and 5% of discharge summaries, respectively, followed by weakness (12%), depression (10%), and anxiety (6%). Although likely grossly underestimating the prevalence of delirium in this population, which has been estimated to be from 75% (16) to 84% (15), the documentation of delirium in the discharge summary is noteworthy, given its salience to patients and their loved ones (40) and its association with long-term cognitive and physical impairment (15, 16).

Strengths and Limitations

Strengths of this study include the use of a highly sensitive and specific electronic detection algorithm to prospectively diagnose ARDS, the manual chart review to examine nuances of documentation practices, and the transparency of NLP methods to aid in reproducibility.

The results should be interpreted in the context of some limitations. We did not include any individual provider identifiers in our analysis, thus limiting assessment of the contribution of provider-level variation in documentation practices. We did not account for punctuation errors in the text, or include relation extraction tools to improve classification of documents using context for keywords. Our chart review was performed by only one author, thus limiting a quality assessment of the chart review through observation of interrater reliability. Through this chart review, we found that terms related to PICS were more difficult to detect accurately using only a keyword-based classifier, and that information extraction tools should be used to distinguish between present on admission, hospital course, and anticipatory guidance contexts for the use of these terms. Further studies will need to be designed to account for these findings.

In addition, future work should include analysis of co-occurrence patterns or clustering of terms to identify latent patterns in themes of discharge documents, and potential dependence between mention of different groups of keywords, which was not addressed in the current study. In addition, more precise measures of keyword relevance, such as the term frequency–inverse document frequency, may be used to build more quantitative or automated approaches to keyword selection.

Conclusions

We found that ARDS was rarely documented at discharge, posing a threat to adherence to established recommendations for quality post–acute care transitions. Mechanical ventilation and ICU admission, however, were commonly mentioned in the discharge documents of survivors of ARDS. Although symptoms that presage PICS were occasionally documented in the discharge summary, an opportunity exists to more effectively educate and prepare survivors of ARDS and their providers for potential long-term consequences. Patient- and hospital-level factors associated with documentation may suggest targets for future interventions to improve quality of care for survivors of ARDS and critical illness. Finally, this study demonstrates the feasibility of using NLP to analyze discharge documents in survivors of ARDS. ■

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

Acknowledgment: The authors are grateful to the Acute Care Health Services Research Group at the University of Pennsylvania (Philadelphia, PA) for thoughtful comments and critiques. They are also grateful for the support of the analysts at the Data Analytics Center of the University of Pennsylvania Health System for their efforts in retrieving the electronic health information used in this study.

References

1 Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, Fan E, Camporota L, Slutsky AS; ARDS Definition Task Force. Acute respiratory distress syndrome: the Berlin definition. *JAMA* 2012;307: 2526–2533.
2 Rubenfeld GD, Caldwell E, Peabody E, Weaver J, Martin DP, Neff M, Stern EJ, Hudson LD. Incidence and outcomes of acute lung injury. *N Engl J Med* 2005;353:1685–1693.
3 Abel SJC, Finney SJ, Brett SJ, Keogh BF, Morgan CJ, Evans TW. Reduced mortality in association with the acute respiratory distress syndrome (ARDS). *Thorax* 1998;53:292–294.
4 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.
5 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.
6 Snow V, Beck D, Budnitz T, Miller DC, Potter J, Wears RL, Weiss KB, Williams MV. Transitions of Care Consensus policy statement: American College of Physicians, Society of General Internal Medicine, Society of Hospital Medicine, American Geriatrics Society, American College of Emergency Physicians, and Society for Academic Emergency Medicine. *J Hosp Med* 2009;4:364–370.
7 Kind AJ, Smith MA. Documentation of mandated discharge summary components in transitions from acute to subacute care. In: Henriksen K, Battles JB, Keyes MA, Grady ML, editors. Rockville, MD: Agency for Healthcare Research and Quality; 2008. pp. 1–10.
8 Gallop KH, Kerr CEP, Nixon A, Verdian L, Barney JB, Beale RJ. A qualitative investigation of patients' and caregivers' experiences of severe sepsis. *Crit Care Med* 2015;43:296–307.
9 Maley JH, Brewster I, Mayoral I, Siruckova R, Adams S, McGraw KA, Piech AA, Detsky M, Mikkelsen ME. Resilience in survivors of critical illness in the context of the survivors' experience and self-reported neuropsychological and physical function. *Ann Am Thorac Soc* (In press)
10 Kripalani S, LeFevre F, Phillips CO, Williams MV, Basaviah P, Baker DW. Deficits in communication and information transfer between hospital-based and primary care physicians: implications for patient safety and continuity of care. *JAMA* 2007;297:831–841.

11 Wunsch H, Guerra C, Barnato AE, Angus DC, Li G, Linde-Zwirble WT. Three-year outcomes for Medicare beneficiaries who survive intensive care. *JAMA* 2010;303:849–856.
12 Combes A, Costa M-A, Trouillet J-L, Baudot J, Mokhtari M, Gibert C, Chastre J. Morbidity, mortality, and quality-of-life outcomes of patients requiring >or=14 days of mechanical ventilation. *Crit Care Med* 2003;31:1373–1381.
13 Barnato AE, Albert SM, Angus DC, Lave JR, Degenholtz HB. Disability among elderly survivors of mechanical ventilation. *Am J Respir Crit Care Med* 2011;183:1037–1042.
14 Wunsch H, Christiansen CF, Johansen MB, Olsen M, Ali N, Angus DC, Sørensen HT. Psychiatric diagnoses and psychoactive medication use among nonsurgical critically ill patients receiving mechanical ventilation. *JAMA* 2014;311:1133–1142.
15 Brummel NE, Jackson JC, Pandharipande PP, Thompson JL, Shintani AK, Dittus RS, Gill TM, Bernard GR, Ely EW, Girard TD. Delirium in the ICU and subsequent long-term disability among survivors of mechanical ventilation. *Crit Care Med* 2014;42:369–377.
16 Pandharipande PP, Girard TD, Jackson JC, Morandi A, Thompson JL, Pun BT, Brummel NE, Hughes CG, Vasilevskis EE, Shintani AK, *et al*.; BRAIN-ICU Study Investigators. Long-term cognitive impairment after critical illness. *N Engl J Med* 2013;369:1306–1316.
17 Jackson JC, Pandharipande PP, Girard TD, Brummel NE, Thompson JL, Hughes CG, Pun BT, Vasilevskis EE, Morandi A, Shintani AK, *et al*.; BRAIN-ICU study investigators. Depression, post-traumatic stress disorder, and functional disability in the BRAIN-ICU study: a longitudinal cohort study. *Lancet Respir Med* 2014;5:369–379.
18 Oeyen SG, Vandijck DM, Benoit DD, Annemans L, Decruyenaere JM. Quality of life after intensive care: a systematic review of the literature. *Crit Care Med* 2010;38:2386–2400.
19 Hopkins RO, Weaver LK, Pope D, Orme JF, Bigler ED, Larson-LOHR V. Neuropsychological sequelae and impaired health status in survivors of severe acute respiratory distress syndrome. *Am J Respir Crit Care Med* 1999;160:50–56.
20 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.
21 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.
22 Mikkelsen ME, Christie JD, Lanken PN, Biester RC, Thompson BT, Bellamy SL, Localio AR, Demissie E, Hopkins RO, Angus DC. The adult respiratory distress syndrome cognitive outcomes study: long-term neuropsychological function in survivors of acute lung injury. *Am J Respir Crit Care Med* 2012;185:1307–1315.
23 Needham DM, Dinglas VD, Bienvenu OJ, Colantuoni E, Wozniak AW, Rice TW, Hopkins RO; NIH NHLBI ARDS Network. One year outcomes in patients with acute lung injury randomised to initial trophic or full enteral feeding: prospective follow-up of EDEN randomised trial. *BMJ* 2013;346:f1532.
24 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.
25 Weiss CH, Baker DW, Weiner S, Bechel M, Ragland M, Rademaker A, Weitner BB, Agrawal A, Wunderink RG, Persell SD. Low tidal volume ventilation use in acute respiratory distress syndrome. *Crit Care Med* (In press).
26 Kalhan R, Mikkelsen M, Dedhiya P, Christie J, Gaughan C, Lanken PN, Finkel B, Gallop R, Fuchs BD. Underuse of lung protective ventilation: analysis of potential factors to explain physician behavior. *Crit Care Med* 2006;34:300–306.
27 Abhyankar S, Demner-Fushman D, Callaghan FM, McDonald CJ. Combining structured and unstructured data to identify a cohort of ICU patients who received dialysis. *J Am Med Inform Assoc* 2014;21: 801–807.
28 Weissman GE, Harhay MO, Lugo RM, Fuchs BD, Halpern SD, Mikkelsen ME. Using natural language processing to assess documentation of the acute respiratory distress syndrome in hospital discharge documents. Presented at the AcademyHealth Annual Research Meeting. May 10, 2016; Boston, MA.
29 Koenig HC, Finkel BB, Khalsa SS, Lanken PN, Prasad M, Urbani R, Fuchs BD. Performance of an automated electronic acute lung injury screening system in intensive care unit patients. *Crit Care Med* 2011;39:98–104.
30 Cohen KB, Demner-Fushman D. Biomedical natural language processing. Philadelphia: John Benjamins Publishing Company; 2014.
31 R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2013. Available from: http://www.r-project.org/
32 Feinerer I, Hornik K, Meyer D. Text mining infrastructure in R. *J Stat Softw* 2008;25:1–54.
33 Hornik K, Buchta C, Zeileis A. Open-source machine learning: R meets Weka. *Comput Stat* 2009;24:225–232.
34 van der Loo M. stringdist: an R package for approximate string matching. *R J* 2014;6:111–122.
35 Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol* 2004;159:702–706.
36 Cummings P. Methods for estimating adjusted risk ratios. *Stata J* 2009;9:175.
37 StataCorp. Stata statistical software: release 14. College Station, TX: StataCorp LP; 2015.
38 Lee CM, Herridge MS, Matte A, Cameron JI. Education and support needs during recovery in acute respiratory distress syndrome survivors. *Crit Care* 2009;13:R153.
39 Govindan S, Iwashyna TJ, Watson SR, Hyzy RC, Miller MA. Issues of survivorship are rarely addressed during intensive care unit stays. Baseline results from a statewide quality improvement collaborative. *Ann Am Thorac Soc* 2014;11:587–591.
40 Freiman DB, Freiman AO, Meyer N, Fuchs B. A most irritating awakening. *Ann Am Thorac Soc* 2013;10:175–177.