2017

Interrater reliability of surveillance for ventilator-associated events and pneumonia

Meeta Prasad Kerlin
University of Pennsylvania

William E. Trick
Cook County Health and Hospitals Systems

Deverick J. Anderson
Duke University

Hilary M. Babcock
Washington University School of Medicine in St. Louis

Ebbing Lautenbach
University of Pennsylvania

See next page for additional authors

ORIGINAL ARTICLE

Interrater Reliability of Surveillance for Ventilator-Associated Events and Pneumonia

Meeta Prasad Kerlin, MD, MSCE;[1] William E. Trick, MD;[2] Deverick J. Anderson, MD, MPH;[3] Hilary M. Babcock, MD;[4] Ebbing Lautenbach, MD, MPH, MSCE;[1] Renaud Gueret, MD;[2] Michael Klompas, MD, MPH[5,6] for the CDC Prevention Epicenters

OBJECTIVE. To compare interrater reliabilities for ventilator-associated event (VAE) surveillance, traditional ventilator-associated pneumonia (VAP) surveillance, and clinical diagnosis of VAP by intensivists.

DESIGN. A retrospective study nested within a prospective multicenter quality improvement study.

SETTING. Intensive care units (ICUs) within 5 hospitals of the Centers for Disease Control and Prevention Epicenters.

PATIENTS. Patients who underwent mechanical ventilation.

METHODS. We selected 150 charts for review, including all VAEs and traditionally defined VAPs identified during the primary study and randomly selected charts of patients without VAEs or VAPs. Each chart was independently reviewed by 2 research assistants (RAs) for VAEs, 2 hospital infection preventionists (IPs) for traditionally defined VAP, and 2 intensivists for any episodes of pulmonary deterioration. We calculated interrater agreement using κ estimates.

RESULTS. The 150 selected episodes spanned 2,500 ventilator days. In total, 93–96 VAEs were identified by RAs; 31–49 VAPs were identified by IPs, and 29–35 VAPs were diagnosed by intensivists. Interrater reliability between RAs for VAEs was high (κ, 0.71; 95% CI, 0.59–0.81). Agreement between IPs using traditional VAP criteria was slight (κ, 0.12; 95% CI, −0.05–0.29). Agreement between intensivists was slight regarding episodes of pulmonary deterioration (κ 0.22; 95% CI, 0.05–0.39) and was fair regarding whether episodes of deterioration were attributable to clinically defined VAP (κ, 0.34; 95% CI, 0.17–0.51). The clinical correlation between VAE surveillance and intensivists' clinical assessments was poor.

CONCLUSIONS. Prospective surveillance using VAE criteria is more reliable than traditional VAP surveillance and clinical VAP diagnosis; the correlation between VAEs and clinically recognized pulmonary deterioration is poor.

Infect Control Hosp Epidemiol 2017;38:172–178

Ventilator-associated pneumonia (VAP) is a complication of mechanical ventilation associated with significant morbidity, including prolongation of mechanical ventilation, increased ICU and hospital lengths of stay, and higher mortality rates.[1] For many years, the National Healthcare Safety Network (NHSN) of the Centers for Disease Control and Prevention (CDC) encouraged hospitals to conduct surveillance for VAP using a definition that required patients to fulfill radiographic, systemic, and pulmonary criteria. However, surveillance using this definition was labor-intensive, nonspecific, and highly subjective.[2,3] To address these issues, the CDC created a new definition for ventilator-associated events (VAEs), including nested categories for ventilator-associated conditions (VACs), infection-related ventilator-associated complications (IVACs), and possible pneumonia (Figure 1).[4]

VAE definitions were designed to increase the objectivity and reproducibility of surveillance by using quantitative rather than qualitative clinical data and by providing clear formulas on how to combine these data to meet VAE criteria. It is unknown, however, whether VAE surveillance in practice is more reproducible and consistent than surveillance using the NHSN's traditional VAP definitions. Some concerns about VAE surveillance have been raised. Klouwenberg et al[5] demonstrated that different strategies for defining the daily

Affiliations: 1. Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania; 2. Collaborative Research Unit, Cook County Health and Hospitals Systems, Chicago, Illinois; 3. Duke Center for Healthcare Epidemiology, Duke University, Durham, North Carolina; 4. Department of Medicine, Washington University School of Medicine, St. Louis, Missouri; 5. Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Healthcare Institute, Boston, Massachusetts; 6. Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts.

Received July 26, 2016; accepted October 7, 2016; electronically published November 8, 2016

 0899-823X/2017/3802-0007. DOI: 10.1017/ice.2016.262

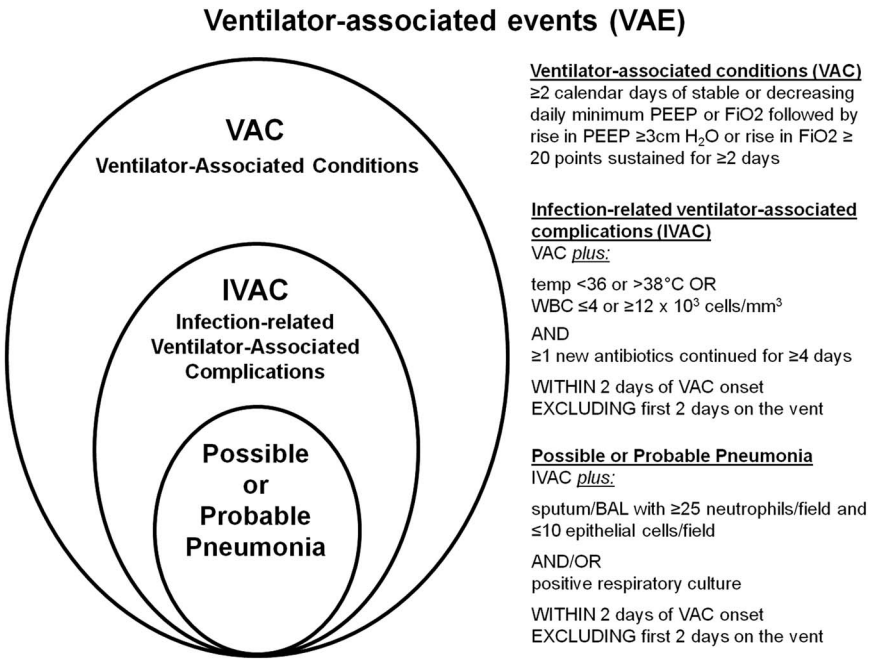

FIGURE 1. Ventilator-associated events as defined at the time of the study in 2014. The pneumonia tier has been modified slightly since then. Reprinted with permission of the American Thoracic Society. Copyright ©2016 American Thoracic Society.[7] *The American Journal of Respiratory and Critical Care Medicine* is an official journal of the American Thoracic Society.

minimum positive end expiratory pressure (PEEP) (eg, minimum per minute, minimum per hour, minimum sustained for at least 1 hour, or 10th percentile) identified different episodes as VAEs. Mann et al[6] demonstrated that manual surveyors make mistakes when applying VAE criteria and that different manual surveyors capture different cases relative to one another and to automated surveillance systems.

Given that VAE definitions have been proposed as possible quality improvement and benchmarking metrics, it is important to quantify the reliability and reproducibility of VAE surveillance. We compared the interrater reliability of VAE surveillance to the interrater reliability of NHSN's former VAP definition and to the clinical diagnosis of VAP by practicing intensivists. We hypothesized that interrater reliability for VAE would be high and better than interrater reliability for traditionally defined VAP and clinical VAP diagnosis by intensivists.

METHODS

Study Overview

This retrospective chart-review study was nested within the CDC Prevention Epicenters' Wake-Up-and-Breathe study.[7] Briefly, the Wake-Up-and-Breathe study was a multicenter prospective evaluation of a multidisciplinary opt-out protocol for daily sedative interruptions and spontaneous breathing trials conducted in 2011–2013, for which the primary study outcome was VAE rates. The study included 20 ICUs in 13 academic and community hospitals affiliated with 5 academic centers of the CDC Prevention Epicenters' Collaborative. The study tracked 5,164 consecutive episodes of mechanical ventilation over a 19-month period.

Study Population

We selected 30 patients from each of the 5 participating Epicenter sites for the current study. Given the rarity of both VAEs and traditionally defined VAPs, each site included all VAEs and all traditionally defined VAPs in their sample and then randomly selected the balance of patients (for a net total of 30 patients) from the population of patients ventilated for ≥10 days. A 10-day minimum was selected for these patients to enrich the study sample with patients more likely to have suffered respiratory complications. The sample size was selected for feasibility. Each Epicenter designated 1 or more of their ICUs enrolled in the Wake Up and Breathe Collaborative to include in this substudy based on ease of access to records.

Chart Reviews

All charts from all sites were independently reviewed by 2 research assistants (RAs) for VAEs (the primary RA for the prospective study and a second RA for this chart review), by 2 infection preventionists (IPs) for traditionally defined VAPs[8] (the primary IP for the prospective study and a second IP for

this chart review), and by 2 intensivists for clinical events compatible with VAP. All reviewers were blinded to patient VAE and VAP status as identified in the primary study and to each other's determinations. All data were collected using web-based forms that were centrally collated at the study's data coordinating center (Medical Research Analytics and Informatics Alliance [MRAIA], Chicago, IL).

The 2 RA VAE reviewers abstracted daily minimum PEEP and daily minimum FiO2s from patient charts. One RA reviewed charts as part of the primary, prospective study, and the other RA was provided the dates of mechanical ventilation and was asked to review charts retrospectively after the primary study had ended. Values were entered into the web-based data collection tool that then automatically identified VAEs using CDC criteria.[9] For both reviewers, if the website flagged a VAE, the reviewers were asked to also enter the patient's minimum and maximum daily temperatures, white blood cell count, and daily antibiotic exposure to assess for IVAC. When available, microbiology results were entered to assess for possible pneumonia. All laboratory testing was performed at the discretion of the treating clinicians subject to the individual protocols of the ICU and was not influenced by the research team.

The 2 IPs were asked to identify episodes of VAP using the NHSN then-current PNEU criteria. One IP reviewed charts as part of routine operational surveillance, and the other was provided the dates of mechanical ventilation and asked to review charts retrospectively, as with the RAs. NHSN trainers provided a webinar-based refresher course on PNEU surveillance for all IPs.

All prospective RA reviewers were trained together in data collection and VAE surveillance by the primary research staff via webinar. All retrospective RA reviewers were trained on an ad hoc basis by members of the research team at their local institutions. None of the reviewers (RAs or IPs, prospectively or retrospectively) had any contact with bedside providers during surveillance activities. All reviewers had access to the same data, including paper medical charts and electronic health records.

The 2 intensivists (board-eligible or board-certified critical care specialists) were provided with the dates for each patient's designated episode of mechanical ventilation and reviewed the charts retrospectively after the primary study had ended. They were asked to determine whether the patient had had an episode of "respiratory deterioration," and if so, to select its cause from a list of possible causes that included VAP. They were asked to use their expert opinion to define deterioration as well as the possible cause(s) of deterioration without further guidance, as our objective was to understand the practical clinical correlates of VAEs and VAPs. The intensivists were permitted to review any part of a patient's chart to the extent they felt necessary to make their assessment, including physiologic and lab data, diagnostic imaging, and clinical notes. Intensivists documented their assessment on an encrypted web form.

Statistical Analysis

We summarized the characteristics of included patients using standard descriptive statistics. We assessed interrater reliability (1) to identify VAEs according to the 2 RAs; (2) to identify VAP per the prior NHSN definition by the 2 IPs, (3) to identify any pulmonary deterioration event during the episode of mechanical ventilation according to the intensivist reviewers; (4) to determine the primary cause of pulmonary deterioration according to the intensivist reviewers; (5) to determine the primary cause of pulmonary deterioration due to VAP according to the intensivist reviewers, and (6) to identify any episode of pulmonary deterioration according to the intensivist reviewers versus VAE by the primary study RA. In this last category, we chose to compare intensivists' assessments to VAE classifications by the primary study RA because the study RAs received the most robust training and supervision on VAE surveillance; thus, we felt their assessments were closest to a gold standard for identification of VAEs.

Primary estimates of agreement were performed at the level of a respiratory event (eg, VAE according to study RAs, VAP according to IP, and respiratory deterioration according to intensivist reviewers). We considered reviewers to be in agreement if events identified by each rater were within 2 days of each other or if no events were identified for a given patient, considering that episodes of mechanical ventilation varied in duration and that multiple events could occur for a single patient. We also performed an analysis at the level of episodes of mechanical ventilation (ie, a binary measure of VAE or VAP at any time during an episode of mechanical ventilation, to address the reliability of using VAE surveillance methods for hospital reporting). We did not perform the analysis at the day level (ie, a binary measure of VAE or VAP on each day of mechanical ventilation) because the vast majority of days would be negative for VAE/VAP, which in turn would generate an overly positive estimate of agreement.

Interrater reliability was evaluated by estimating observed agreement, agreement expected by chance (calculated as the probability of random agreement based on the observed probabilities of events according to each rater), and Cohen's κ statistics for pairwise comparisons.[10] The scale used to interpret the κ estimates was as follows: <0, less than chance agreement; 0.01–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; 0.81–0.99, almost perfect agreement.[11] We estimated 95% confidence intervals using bootstrapping methods. All statistical analyses were performed using Stata 12.0 (StataCorp, College Station, TX). The study was approved by the institutional review board at each participating hospital.

RESULTS

The study included a total of 30 patients from each of 5 participating centers for a total of 150 patients with 150 episodes of mechanical ventilation spanning 2,500 ventilator days.

TABLE 1. Patient Characteristics (N = 150)

Characteristic	Data
Age, median y (IQR)	62 (49–75)
Male, %	64
Reason for intubation	
Unknown, %	32
Surgery, %	17
Cardiac arrest, %	9
Altered level of consciousness, %	8
Community-acquired pneumonia, %	5
Trauma, %	4
Sepsis, %	4
Hospital-acquired pneumonia, %	3
Asthma/COPD, %	3
Airway protection, %	3
ARDS, %	2
Other, %	10
SOFA score, median (IQR)	10 (8–13)
Outcomes	
No. of days of mechanical ventilation, median (IQR)	13 (9–20)
VAP (according to hospital IP), %	21
VAC (according to study RA), %	63
Mortality, %	51

NOTE. IQR, interquartile range; COPD, chronic obstructive pulmonary disease; ARDS, acute respiratory distress syndrome; SOFA, sequential organ failure assessment; VAP, ventilator-associated pneumonia; IP, infection preventionist, VAC, ventilator-associated condition; RA, research assistant.

Characteristics of patients and outcomes of mechanical ventilation are summarized in Table 1. The 150 episodes of mechanical ventilation included 93 VAEs and 32 traditionally defined VAPs (according to the primary study RAs and infection preventionists, respectively), as well as 38 episodes that included neither a VAE nor a VAP.

Reliability of VAE and VAP Surveillance Definitions

Interrater reliability for VAEs including the timing of the event was high (Table 2), with agreement of 86% and κ of 0.71 (95% CI, 0.59–0.81). Interrater reliability for VAE at any time during an episode of mechanical ventilation (ie, ignoring whether the identified events were within 2 days of each other) was also high, with agreement of 93% and κ of 0.86 (95% CI, 0.78–0.94). Agreement on the raw clinical data used for VAE assessment was almost perfect. Values of κ ranged from 0.90 (95% CI, 0.85–0.94) for daily minimum white blood cell count to 0.94 (95% CI, 0.92–0.95) for daily maximum temperature to 0.97 (95% CI, 0.95–0.98) for daily maximum white blood cell count. Agreement between the IPs using the NHSN traditional VAP criteria was slight (κ, 0.12; 95% CI, −0.05–0.29).

Correlation of VAEs with Clinical Assessments

Among patients who met criteria for a VAE according to the study RA, the most common causes identified by intensivists

TABLE 2. Interrater Reliability for Identification of Ventilator-Associated Events (VAEs) and Ventilator-Associated Pneumonia (VAP)[a]

A[b]	RA 2 VAE Present, No.	RA 2 VAE Absent, No.	Agreement, %	Expected Agreement, %	κ Value (95% CI)
RA 1 VAE present	83	10	86	52	0.71 (0.59–0.81)
RA 1 VAE absent	13	55	…	…	…
B[c]	IP 2 VAP Present, No.	IP 2 VAP Absent, No.	Agreement, %	Expected Agreement, %	κ Value (95% CI)
IP 1 VAP present	13	18	66	62	0.12 (−0.053–0.29)
IP 1 VAP absent	36	94	…	…	…
C[d]	Intensivist 2 VAP Present, No.	Intensivist 2 VAP Absent, No.	Agreement, %	Expected Agreement, %	κ Value (95% CI)
Intensivist 1 VAP present	14	15	83	75	0.34 (0.17– 0.51)
Intensivist 1 VAP absent	21	164	…	…	…

NOTE. RA, research assistant; IP, infection preventionist.
[a]The number of events identified by each reviewer and the agreement between reviewers are summarized.
[b]Panel A illustrates interrater reliability between 2 RAs for VAE assessment; a total of 161 events (106 VAEs identified by 1 or both RAs and 55 patients with no events) were included in κ estimates.
[c]Panel B illustrates interrater reliability between 2 IPs for traditionally defined VAP; a total of 161 events (67 VAPs identified by 1 or both IPs and 94 patients with no VAP) were included in κ estimates.
[d]Panel C illustrates interrater reliability between 2 intensivists for clinical VAP as the cause for respiratory deterioration; a total of 214 events (including 179 events of respiratory deterioration identified by 1 or both intensivists and 35 patients with no events according to either intensivist) were included in κ estimates.

TABLE 3. Primary Diagnoses According to Intensivists for 93 Ventilator-Associated Events (VAEs) Identified by Primary Study Research Assistants

Clinical Diagnosis	Intensivist 1, No. of Cases	Intensivist 2, No. of Cases	κ Value
No deterioration present	41	42	0.717
VAP	14	12	0.553
ARDS	10	11	0.410
Aspiration	3	6	0.652
Atelectasis	7	2	0.425
Cardiac arrest	1	1	1.000
CHF	2	4	−0.030
Sepsis	5	4	0.417
Cancer	0	2	0.000
Pneumothorax	1	0	0.000
Other	6	7	0.421
Unknown	3	2	−0.027

NOTE. VAP, ventilator-associated pneumonia; ARDS, acute respiratory distress syndrome; CHF, congestive heart failure.

were VAP (13%–15%), acute respiratory distress syndrome (ARDS) (11%–12%), sepsis (4%–5%), aspiration (3%–6%), and atelectasis (2%–8%) (Table 3). However, 44%–45% of VAEs identified by RAs were not recognized by intensivist reviewers as episodes of pulmonary deterioration at all. Conversely, intensivists identified a higher total number of episodes of pulmonary deterioration overall than study RAs (179 episodes vs 93 VAEs).

There was fair agreement between intensivists on whether or not patients suffered an episode of pulmonary deterioration from any cause (κ, 0.22; 95% CI, 0.05–0.39). There was also fair agreement between intensivists on the clinical diagnosis of VAP (κ, 0.34; 95% CI, 0.17–0.51) (Table 2). Agreement regarding the cause for respiratory deterioration identified by intensivists was quite variable, with κ ranging between −0.007 and 0.798 (Table 4).

TABLE 4. Primary Diagnoses for Respiratory Deterioration According to Intensivists

Clinical Diagnosis	No. of Events Identified by Intensivist 1	No. of Events Identified by Intensivist 2	No. of Events Identified by Both Intensivists	κ Value[a]
Pneumothorax	3	2	2	0.798
Cardiac arrest	2	1	1	0.665
Aspiration	10	12	6	0.521
Pulmonary hemorrhage	2	2	1	0.495
Sepsis	16	15	8	0.478
Atelectasis	16	13	6	0.371
VAP	29	35	14	0.340
ARDS	19	19	6	0.249
CHF	8	11	2	0.175
Pulmonary embolism	0	2	0	0.000
SIRS	2	1	0	−0.006
Cancer	1	3	0	−0.007
Other	24	24	9	0.296
Unknown	7	4	1	0.162

NOTE. VAP, ventilator-associated pneumonia; ARDS, acute respiratory distress syndrome; CHF, congestive heart failure; SIRS, systemic inflammatory response syndrome.

[a]A total of 214 events (including 179 events of respiratory deterioration identified by 1 or both intensivists and 35 patients with no events according to either intensivist) were included in κ estimates.

DISCUSSION

Within a multicenter prospective study of VAE surveillance, agreement between research assistants on the identification of VAEs was high overall and was also high for the individual elements used to define VAEs. Consistent with previous literature, agreement between trained IPs using traditional surveillance definitions for VAP was slight, at best.[3,5,12,13] Agreement between intensivists on the clinical diagnosis of VAP was only fair. Intensivists also disagreed with one another on whether patients suffered episodes of pulmonary deterioration at all, and there was poor overlap between intensivists' perceptions of pulmonary deterioration and VAE criteria. Many VAEs were not identified as deterioration events by intensivist reviewers, and many deterioration events identified by intensivists did not meet VAE criteria.

Given that these surveillance methods are intended for quality measurement and improvement and hospital performance evaluation, the improved objectivity, reliability, and potential ease of data capture from electronic medical records represent an improvement over the previous surveillance methods, which were more subjective and prone to disagreement.[2] That this study confirmed high agreement between RAs, even when using different surveillance strategies (eg, real-time data capture versus retrospective chart review), which provides reassurance that some of the goals of this revised algorithm have been met.

Our findings are also consistent, however, with previous literature suggesting a disconnect between objective surveillance and clinical evaluations.[2,14] The surveillance definitions both miss events that clinicians identify as pulmonary deterioration and capture events that are not identified as such. Clinicians themselves, however, frequently disagreed with each other regarding whether or not patients suffered episodes of pulmonary deterioration. These discrepancies likely have several causes, including the lack of specificity for most clinical signs of pulmonary disease. In addition, retrospective chart review rather than prospective bedside evaluation may have magnified uncertainty and differences of opinion between intensivists. Finally, intensivists' clinical assessments in general may be nuanced and inherently more uncertain than decisions made for surveillance.

Our study has several important strengths. It was nested within a large, multicenter study of VAE surveillance, allowing for a broad case mix for study. It is one of only a few studies with such a high number of charts reviewed for VAP diagnosis, and it is the only study to our knowledge that included reviews by 3 different stakeholders in VAE and VAP surveillance: RAs for VAEs, IPs for NHSN VAPs, and intensivists for clinical causes of respiratory deterioration. The current study is therefore a rich source of data on how perspectives vary between and among these key stakeholders in hospital surveillance.

Our study also has several limitations. First, as a multicenter study, each site had some liberty to perform surveillance in ways that were feasible. For example, some centers were able to collect some surveillance data elements automatically through electronic medical records, whereas others required complete hand collection by RAs. Furthermore, the conditions for reviewers sometimes differed within reviewer pairs. For example, one reviewer might have collected data in a prospective manner and possibly directly on-site in the study ICUs, whereas the other reviewer in a pair might have collected data retrospectively, with access to data from the entire admission. Despite these differences, interrater reliability between research assistants was high, indicating that the surveillance definitions do not appear to be heavily compromised by different collection strategies. A second limitation is that because events and diagnoses were generally rare, agreement expected by chance alone was high (because of the high nonevent rates); this was particularly so in the analyses of the interrater reliability among intensivists, where many events of pulmonary deterioration were identified but few were attributed to VAP. Therefore, κ estimates were frequently low even when observed agreement was relatively high; these numbers should be interpreted with caution.

In summary, the reliability of VAE assessment is significantly better than the prior CDC VAP surveillance definitions and better than interobserver agreement between intensivists. However, discrepancies between VAE surveillance and clinical impressions still exist. The clinical significance of these differences and their impact on quality improvement efforts are unclear. On the one hand, VAE was not intended for clinical diagnosis or clinical management but rather to help identify general sources of harm for patients at the population level. VAE definitions include a built-in threshold effect to detect only severe events (ie, complications severe enough to precipitate a sustained increase in ventilator settings). It is therefore understandable and predictable that VAE surveillance will miss some sources of respiratory deterioration identified by clinicians. This can be an advantage for quality improvement practitioners because it helps them to focus their efforts only on severe complications. On the other hand, clinicians find mismatches between surveillance events and clinical events to be disturbing and a cause to question the clinical relevance of surveillance. This tension between population surveillance using quality improvement metrics versus clinical diagnosis and management has not been resolved by the transition to VAE surveillance and merits further evaluation and better understanding before these measures can be deemed suitable for public reporting or benchmarking.

ACKNOWLEDGMENTS

We would like to express our gratitude to the many people who assisted us with the project. From Cook County Hospital, we thank Onofre Donceras and Stephen Littleton for performing chart reviews and Rosie Lyles and George Markovski for developing the data extraction tool and managing the study databases. From Duke University, we thank Susan Louis, Vanessa Walker, and Karen Welty-Wolf for performing chart reviews. From Harvard University, we thank Puja Kohli, Vickie Morrison, and Chanu Rhee for performing chart reviews. From Washington University, we thank Marin Kollef, Kathleen McMullen and Chad Witt for performing chart reviews. And from the University of Pennsylvania, we thank Dan Dorgan, Rachel Reis, Carol Samel, and Rupal Shah for performing chart reviews, Pam Tolomeo for project management, and Warren Bilker for statistical advice.

Financial support: This study was supported by the Centers for Disease Control and Prevention (grant no. U54CK000172-01S1).

Potential conflicts of interest: No authors have any conflicts of interest to disclose.

Address correspondence to Meeta Prasad Kerlin, Perelman School of Medicine at the University of Pennsylvania, 3600 Spruce St, Gibson 5011, Philadelphia, PA 19104 (prasadm@uphs.upenn.edu).

REFERENCES

1. Safdar N, Dezfulian C, Collard HR, Saint S. Clinical and economic consequences of ventilator-associated pneumonia: a systematic review. *Crit Care Med* 2005;33:2184–2193.
2. Klompas M. Interobserver variability in ventilator-associated pneumonia surveillance. *Am J Infect Control* 2010;38:237–239.
3. Stevens JP, Kachniarz B, Wright SB, et al. When policy gets it right: variability in US hospitals' diagnosis of ventilator-associated pneumonia. *Crit Care Med* 2014;42:497–503.
4. Surveillance for Ventilator-Associated Events. Centers for Disease Control and Prevention website. http://www.cdc.gov/nhsn/acute-care-hospital/vae/index.html. Published 2015. Accessed March 31, 2016.
5. Klein Klouwenberg PM, van Mourik MS, Ong DS, et al. Electronic implementation of a novel surveillance paradigm for ventilator-associated events: feasibility and validation. *Am J Respir Cri. Care Med* 2014;189:947–955.
6. Mann T, Ellsworth J, Huda N, et al. Building and validating a computerized algorithm for surveillance of ventilator-associated events. *Infect Control Hosp Epidemiol* 2015:1–5.
7. Klompas M, Anderson D, Trick W, et al. The preventability of ventilator-associated events. The CDC Prevention Epicenters Wake Up and Breathe Collaborative. *Am J Respir Crit Care Med* 2015;191:292–301.
8. Horan TC, Andrus M, Dudeck MA. CDC/NHSN surveillance definition of health care–associated infection and criteria for specific types of infections in the acute care setting. *Am J Infect Control* 2008;36:309–332.
9. Magill SS, Klompas M, Balk R, et al. Developing a new, national approach to surveillance for ventilator-associated events. *Crit Care Med* 2013;41:2467–2475.

10. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.
11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977:159–174.
12. Emori TG, Edwards JR, Culver DH, et al. Accuracy of reporting nosocomial infections in intensive-care–unit patients to the National Nosocomial Infections Surveillance System: a pilot study. *Infect Cont Hosp Ep* 1998;19:308–316.
13. Schurink CA, Van Nieuwenhoven CA, Jacobs JA, et al. Clinical pulmonary infection score for ventilator-associated pneumonia: accuracy and inter-observer variability. *Intens Care Med* 2004;30:217–224.
14. Tejerina E, Esteban A, Fernández-Segoviano P, et al. Accuracy of clinical definitions of ventilator-associated pneumonia: comparison with autopsy findings. *J Crit Care* 2010; 25:62–68.