BMJ Open  Pulse oximetry and supplemental oxygen use in nationwide Veterans Health Administration hospitals, 2013–2017: a Veterans Affairs Patient Database validation study

Xiao Qing Wang,1 Theodore Iwashyna,1,2 Hallie Prescott,1,2 Valeria Valbuena,3 Sarah Seelye © 2

ABSTRACT
Objective  Extraction and standardisation of pulse oximetry and supplemental oxygen data from electronic health records has the potential to improve risk-adjustment, quality assessment and prognostication. We develop an approach to standardisation and report on its use for benchmarking purposes.

Materials and methods Using electronic health record data from the nationwide Veteran’s Affairs healthcare system (2013–2017), we extracted, standardised and validated pulse oximetry and supplemental oxygen data for 2765 446 hospitalisations in the Veteran’s Affairs Patient Database (VAPD) cohort study. We assessed face, concurrent and predictive validities using the following approaches, respectively: (1) evaluating the stability of patients’ pulse oximetry values during a 24-hour period, (2) testing for greater amounts of supplemental oxygen use in patients likely to need oxygen therapy and (3) examining the association between supplemental oxygen and subsequent mortality.

Results  We found that 2 700 922 (98%) hospitalisations had at least one pulse oximetry reading, and 864 605 (31%) hospitalisations received oxygen therapy. Patients monitored by pulse oximetry had a reading on average every 6 hours (median 4; IQR 3–7). Patients on supplemental oxygen were older, white and male compared with patients not receiving oxygen therapy (p<0.001) and were more likely to have diagnoses of heart failure and chronic pulmonary diseases (p<0.001). The amount of supplemental oxygen for patients with at least three consecutive values recorded during a 24-hour period fluctuated by median 2 L/min (IQR: 2–3), and 81% of such triplicates showed the same level of oxygen receipt.

Conclusion  Our approach to standardising pulse oximetry and supplemental oxygen data shows face, concurrent and predictive validities as the following: supplemental oxygen clusters in the range consistent with hospital wall-dispersed oxygen supplies (face validity); there are greater amounts of supplemental oxygen for certain clinical conditions (concurrent validity) and there is an association of supplemental oxygen with in-hospital and postdischarge mortality (predictive validity).

Strengths and limitations of this study
► This study focuses on a large dataset of 2 765 446 hospitalisations at >130 different Veterans Affairs hospitals.
► Detailed extraction and standardisation process for complex pulse oximetry and supplemental oxygen data in >17.5 million observations.
► Standardisation process assessed for face validity, concurrent validity and predictive validity.
► This standardisation was developed for use in pulse oximetry and supplemental oxygen outside the intensive care unit, and therefore did not include recordings of invasive mechanical ventilation, non-invasive positive pressure ventilation or heated high-flow nasal cannula.

INTRODUCTION
Supplemental oxygen is among the most administered medical therapies in hospitalised patients.1–3 In many hospitals, supplemental oxygen can be administered and titrated without a physician order, in contrast to most medications.4 5  New hypoxemia is a cardinal bedside sign of acute or impending decompensation as well as a marker of disease severity for several chronic comorbidities.6

Despite the importance of hypoxemia, analyses of hospital data often exclude data on the respiratory system. Respiratory rate is routinely collected as a vital sign, but may be recorded as ‘18’ or ‘20’ regardless of actual respiratory rate. Furthermore, respiratory rate is often influenced by fear, anxiety, pain and sleep, reducing its reliability as a measure of pulmonary function.7 In contrast, pulse oximetry, which measures the fraction of haemoglobin in the blood-carrying oxygen, has become ubiquitous, but it can be interpreted only in the context of knowing how
much supplemental oxygen a patient is on at the same time. Adding further complexity, there are diverse modes of oxygen delivery and variable charting conventions within and across hospitals. Although bedside clinicians can routinely interpret such diversely formatted data, the absence of standardisation limits the ability of health services researchers to use large datasets to study processes of care and outcomes for conditions where clinicians routinely define severity in part in terms of degree of hypoxemia—conditions such as pneumonia, acute exacerbation of chronic obstructive pulmonary disease (COPD) and congestive heart failure.

Extraction and standardisation of pulse oximetry and supplemental oxygen data from electronic health record data has the potential to improve risk-adjustment, quality assessment and prognostication. We therefore sought to identify the ways in which pulse oximetry and supplemental oxygen use were recorded in the nationwide Veterans Health Administration (VA) system, develop an approach to standardise these data and report on its use for benchmarking purposes. This study makes possible an infrastructure for future studies to explore the advantages and potential overuse of pulse oximetry. We focus on the data systems used in VA hospitals outside of the intensive care unit (ICU) and operating room (OR) context—the so-called hospital ‘floors’—because a single system is used nationwide for recording vital signs on the floors, while each VA hospital may use one of several different ICU and OR data systems, including paper charting during the period under study.

METHODS
Study setting and data source
The nationwide VA healthcare system is an integrated system that provides comprehensive medical care to Veterans and operates >130 acute care hospitals. All VA hospitals use the same electronic health record platform, the Computerized Patient Record System, which archives data to a single central repository—the Corporate Data Warehouse (CDW). In this study, we extracted vital sign data from the CDW for all 2765 446 hospitalisations in the Veteran’s Affairs Patient Database (VAPD) for Veterans followed from 1 January 2013 through 31 December 2017; construction of the VAPD has been described previously. Comorbidities were assessed using Elixhauser comorbidities and diuretic use was identified from Bar Code Medication Administration data within the VAPD, as previously described. Mortality data were extracted from the CDW’s Vital Status File, which contains dates of death from sources including the VA’s Beneficiary Identification Records Locator Subsystem File, Medicare Vital Status File and the Social Security Administration Death Master File.

Extracting pulse oximetry and supplemental oxygen
All numeric pulse oximetry measurements are archived in the CDW Vital Signs Domain, where they can be readily used for analysis. However, because supplemental oxygen values are hand-entered by healthcare staff, with raw values such as ‘3.0 L/min’, ‘2.0 L/min’, or ‘35%’, additional cleaning and standardisation by clinicians was required. (See online supplemental appendices 1, 2 and tables 1,2A,B for a step-by-step description of our approach to cleaning and standardisation.)

There is not a uniformly accepted way to report the amount of supplemental oxygen administered. There are two complexities here. The first is reporting: different institutions have different conventions of reporting supplemental oxygen when the same therapy is administered. The second is technical: oxygen can be delivered at a fixed flow rate or at a fixed percentage of inspired oxygen. Standard nasal cannula delivers a fixed flow, but the percentage inspired will vary as a function of the patient’s respiratory mechanics. The greater the inspiratory flow generated by the patient, the more likely the flow of oxygen delivered will be overcome, and the fraction of inspired air that is oxygen (FiO₂) received will be reduced by entrainment of room air. Fixed percentage inspired oxygen devices (eg, Venturi masks) attempt to overcome this variation by delivering high gas flow rates that are less likely to be overcome, although there is variation in effective FiO₂ delivered depending on a patient’s respiratory pattern, positioning and anatomy. We therefore used a common clinical ‘rule of thumb’ formula to approximate the oxygen flow in litres per minute (LPM) as a function of the fixed inspired oxygen fraction (FiO₂, when that is used), as LPM = (FiO₂ – 0.21) / 0.03, to standardise all modes of supplemental oxygen delivery as if they were fixed flow devices. We standardised to LPM rather than to FiO₂ because the imprecision of fixed flow devices for FiO₂ received is generally known clinically and may thereby prevent false precision in the interpretation of the data.

Evaluation of validity
We assessed the validity of SpO₂ and supplemental oxygen data using a three-criteria approach to validation: face validity, concurrent validity and predictive validity. Validity describes the accuracy of a given method to measure a given parameter. When a method uses measurements that closely reflect the true value of a given parameter, the method is considered valid. There are several types of validity by which a measurement can be examined, which is particularly useful when a gold standard comparison is not available. Face validity is established when a method ‘appears’ to measure the parameter of interest. Concurrent validity is established when higher levels of the measurement are more common in populations where the higher levels of the true parameter are expected to be more common. Predictive validity is established when the measurement is associated with future outcomes in a manner consistent with the true parameter. In the case of this investigation and pulse oximetry, we consider the following.
Face validity
Clinical experience suggests that (1) pulse oximetry readings will be overwhelmingly 88% and higher because values lower than this indicate clinically significant hypoxemia that clinicians will promptly correct, (2) supplemental oxygen is most commonly administered in whole LPM and (3) supplemental oxygen is most often titrated to even numbers. We assessed these elements by looking at the distribution of pulse oximetry and supplemental oxygen values. We further assessed face validity by examining consecutive triplets of supplemental oxygen readings within 24 hours to assess the stability of values and the frequency at which the second reading had a similar flow rate as the first and third.

Concurrent validity
We assessed this by testing for greater amounts of supplemental oxygen in patients likely to need supplemental oxygen based on their clinical situation—for example, those with cardiopulmonary disease, those receiving diuretic therapy and those whose hospitalisation involves ICU stays.

Predictive validity
We examined whether greater amounts of supplemental oxygen use on the first day of hospitalisation were associated with higher rates of in-hospital mortality, and whether greater amounts of supplemental oxygen on the last day of hospitalisation were associated with higher rates of mortality in the subsequent year. To assess each of these, we used logistic regression analysis with patient-level random effects to examine the association between supplemental oxygen and mortality. Patient-level random effects were included to account for cases in which multiple admissions occurred during the study period.

Data management and analyses
Data management and analyses were conducted using SQL and SAS (SAS Institute), V.9.4. The analytic code is available online to allow others to extract and clean pulse oximetry and supplemental oxygen values as we have (https://github.com/CCMRcodes/PulseOx).

Data are presented as mean values, medians and IQRs at the hospitalisation-day and patient-day level. We assessed the proportion of a hospitalisation with pulse oximetric coverage as the time between the first and last pulse oximetry reading during a hospitalisation (in hours), divided by the time from admission to discharge (in hours).

Patient and public involvement
Research was completed without patient or public involvement; patients were not invited to contribute to the design of the study, interpretation of results, or the writing or editing of the article.

---

**Table 1** Seven general forms for recording the amount of supplemental oxygen and our approach to standardisation

| Type | Form | Example | Number of values in CDW vital signs domain | Percent of total values (%) | Number of hospitals at which this format was found | Approach to standardisation |
|------|------|---------|-------------------------------------------|-----------------------------|----------------------------------------------------|-----------------------------|
| 1    | nn%  | 35%     | 712568                                    | 4.1                         | 129                                                | Interpreted as fixed percentage inspired oxygen device; converted to LPM according to equation LPM = (FiO2–0.21)/0.03 |
| 2    | nnL/min | 2 L/min | 652303                                    | 37.2                        | 129                                                | Interpreted as nn LPM fixed flow |
| 3    | L/min% | L/min%  | 9461987                                   | 54.0                        | 42                                                 | Interpreted as zero LPM |
| 4    | L/min% | L/min95%| 20041                                     | 0.1                         | 37                                                 | Excluded because nn were not clinically interpretable values |
| 5    | nn L/min% | 2.0 L/min% | 461408                                   | 2.6                         | 42                                                 | Interpreted as nn LPM fixed flow |
| 6    | nn L/min% | xx% where xx%=Pulse Ox reading at same time | 2 L/min99% when pulse ox=99% | 1.3                         | 129                                                | Interpreted as nn LPM (eg, 2 LPM) |
| 7    | nn L/min% | xx% where xx% not equal to pulse Ox reading at same time | 2 L/min99% when pulse ox=93% | 0.7                         | 129                                                | Excluded |

Total 17 537 684

See also online supplemental appendix.

CDW, Corporate Data Warehouse; LPM, litres per minute.
RESULTS

The VAPD consists of a total of 15 437 270 patient-days for 2 765 446 hospitalisations at 134 VA hospitals from 2013 to 2017. Among all hospitalisations, 2 700 922 (98%) hospitalisations had at least one pulse oximetry reading, and 864 605 (31%) hospitalisations received oxygen therapy. Pulse oximetry data were present for 12 369 844 days not in an ICU, and 391 330 days on which there were transitions into an ICU.

Patterns of pulse oximetry and supplemental oxygen data

The amount of supplemental oxygen that patients received were reported in one of seven forms (table 1). Of the 17 537 684 supplemental oxygen readings in the CDW, every hospital used at least four different reporting conventions; indeed, not a single hospital relied on only one reporting method. Data demonstrating the rationale for each of these interpretations are presented in the supplementary file (online supplemental appendix 1, figure 1).

Among 2 700 922 (97.7%) hospitalisations with one or more pulse oximetry values, 1 836 317 (66.4%) hospitalisations received a pulse oximetry reading without supplemental oxygen. For these hospitalisations, the number of pulse oximetry readings was median 10 (IQR: 6–18) (table 2). There were median 3 (IQR: 2–4) pulse oximetry readings per patient-day. On average, patients received a pulse oximetry reading every 6 hours (median 5; IQR 3–8), with 63% of the hospital stay (defined in hours) under pulse oximetry monitoring (median 66; IQR 50–81). The distribution of pulse oximetry readings is

| Table 2 | Pulse oximetry measurements |
| --- | --- |
| Receiving supplemental oxygen and pulse oximetry | Receiving pulse oximetry but not supplemental oxygen |
| Hospitalisations with pulse oximetry, N (%) | 864 605 (31.3) | 1 836 317 (66.4) |
| Number of pulse oximetry measurements during hospitalisations, if at least one | | |
| Mean | 26 | 15 |
| Median | 17 | 10 |
| IQR | 10–31 | 6–18 |
| Number of pulse oximetry measurements per day, if at least one | | |
| Mean | 4.0 | 3.5 |
| Median | 4 | 3 |
| IQR | 2–5 | 2–4 |
| Number of pulse oximetry measurements per calendar day, if at least one N (%) | 5 716 352 | 7 709 751 |
| One time | 513 532 (9%) | 843 037 (11%) |
| Two times | 939 268 (16%) | 1 592 787 (21%) |
| Three times | 1 363 715 (24%) | 2 114 788 (27%) |
| >4 times | 2 899 837 (51%) | 3 159 139 (41%) |
| Hours between pulse oximetry measurements within a hospitalisation | | |
| Mean | 5.9 | 5.9 |
| Median | 4 | 5 |
| IQR | 3–7 | 3–8 |
| Proportion of hospitalisation with pulse oximetry monitoring, by hour | | |
| Mean | 0.72 | 0.63 |
| Median | 0.78 | 0.66 |
| IQR | 0.63–0.88 | 0.50–0.81 |
shown in figure 1, with 99% reading at ≥88%, supporting face validity of the data.

As part of our validation process for face validity, we examined patients with triplets of at least three consecutive supplemental oxygen values during a 24-hour period to ensure that supplemental oxygen values were clinically stable and that changes in LPM were reasonable. The amount of supplemental oxygen for patients with at least three consecutive values recorded during a 24-hour period fluctuated by median 2 LPM (IQR: 2–3). Furthermore, 81% of such triplets showed the same level of oxygen receipt throughout.

**Patterns of supplemental oxygen use**

Between 2013 and 2017, 864605 (31.3%) hospitalisations received supplemental oxygen (table 2). Patients on supplemental oxygen received more frequent pulse oximetry monitoring during the hospitalisation (mean 26; median 17; IQR 10–31 vs mean 15; median 10; IQR 6–18 for patients not on oxygen) and throughout the day (mean 4.0; median 4; IQR 2–5 vs mean 3.5; median 3; IQR 2–4 for patients not on oxygen) (table 2). Patients on supplemental oxygen had a pulse oximetry reading on average every 6 hours (median 4; IQR 3–7). As shown in figure 2, oxygen was typically given in integer—often even integer—flow rates, consistent with the face validity of the data.

Consistent with the concurrent validity of the data, patients who received oxygen therapy at VA hospitals were older compared with patients who did not receive oxygen therapy (table 3, p<0.001). They were also more likely to have diagnoses of heart failure and chronic pulmonary diseases (p<0.001). Patients with hospitalisations that included stays in the ICU as well as those on mechanical ventilation were also more likely to receive oxygen therapy recorded in the CDW Vital Signs domains (hence from the general care floors) at some point during their hospitalisation (p<0.001).

In an unadjusted model with patient-level random effects, greater amounts of supplemental oxygen use on the first day of hospitalisation were associated with higher rates of in-hospital mortality (OR 1.16 per LPM-equiv, 95% CI 1.15 to 1.16, p<0.001). Similarly, greater amounts of supplemental oxygen on the last day of hospitalisation were associated with higher rates of 1-year post-discharge mortality (OR 1.55 per LPM-equiv, 95% CI 1.54 to 1.56, p<0.001). These are consistent with predictive validity of the data.

**DISCUSSION**

In this study, we examine 2.7 million hospitalisations across >100 hospitals over 5 years. We show that pulse
oximetry and supplemental oxygen data can be usefully extracted from a large, unstandardised national database of inpatient vital sign recordings. The extracted data have plausible face, content and predictive validities, suggesting their potential usefulness in analyses. We find that pulse oximetry and supplemental oxygen readings can be used in secondary data analyses to identify non-ICU patients at risk of adverse outcomes, such as in-hospital mortality and one-year postdischarge mortality. We also find that pulse oximetry and supplemental oxygen readings can be used in retrospective studies to define clinically relevant subgroups of patients with disease for targeted study, including pneumonia patients with a new inpatient oxygen requirement and patients with a newly hypoxic exacerbation of COPD. We believe these results demonstrate that valid pulse oximetry and supplemental oxygen data can be recovered from the VA electronic health record despite the diverse and unstandardised ways in which data are recorded by bedside clinicians. These data can be considered for use in large-scale, system-wide risk-adjustment, assessments of processes of care and in defining patient subgroups of interest.

While efforts at ensuring a standard approach to recording these data—as has been done by creating separate systolic, diastolic, and mean arterial blood pressure fields in the recording of that vital sign—may be of value in the future, science using pulse oximetry data in the VA need not necessarily wait for widescale adoption. The work necessary to recover these data for large-scale analyses may also serve as a caution to those designing electronic health records, because even for seemingly simple things, bedside clinicians often carry diverse representations of information, and those diverse representations will find their way into free-text fields unless (and perhaps even if) there is a clear and easy-to-use indication of the standardised, structured approach for recording the data.

The face validity of the data is supported by the clustering of pulse oximetry values in a plausible physiological range that is compatible with life. Face validity is also supported by intervals that align with conventional shift durations and vital sign measurement timing, with integer values of 2–15 LPM—values which most hospital wall-dispensed oxygen supplies provide. The concurrent validity of supplemental oxygen rates is supported by greater amounts of supplemental oxygen in clinical conditions for which supplemental oxygen use is more common. The predictive validity of the supplemental oxygen data is shown by its association with in-hospital and postdischarge mortality.

Our study should be interpreted in the context of some limitations. First, our analysis did not include direct validation of the data against the ground truth of what patients were receiving, as this was not feasible. Individual clinician notes typically cover shifts, days, or even longer time-periods, whereas pulse oximetry and supplemental oxygen can change second to second, and therefore cannot be confirmed via chart review. Second, our approach to data extraction and standardisation was developed for use outside of the ICU, and hence did not include recordings of invasive mechanical ventilation, non-invasive positive pressure ventilation, or heated high flow nasal cannula. Future work is needed to expand our methods to include these advanced modes of respiratory support.

In conclusion, the recording of pulse oximetry and supplemental oxygen within the VA varies across sites. However, our approach to standardising these data yielded results with face, concurrent, and predictive validity, indicating that the data are valid and can be used for secondary data analysis as well as for operational purposes within the VA.

Contributors TI, HP, XQW and SS were involved in study design and conception. XQW, TI and HP were involved in data acquisition. XQW performed data analysis. TI, HP, XQW and SS were involved in data interpretation. TI, W, XQW and SS drafted the manuscript. TI, HP, VV, XQW and SS were involved in critical manuscript review. All authors participated in final manuscript revision and take responsibility for the integrity of the data and the accuracy of the data analysis.

Funding This study was supported by VA HSR&D IIR 17–045, VA IIR 17–2019 and AHRQ R01 HS026725.

Disclaimer The views in this article do not necessarily represent the view of the US Government, Department of Veterans Affairs, the University of Michigan or the Agency for Healthcare Research and Quality.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Analyses were reviewed and approved with waiver of informed consent by the VA Ann Arbor Institutional Review Board (1597242-3).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD
Sarah Seelye http://orcid.org/0000-0001-5098-7470

REFERENCES
1 Cousins JL, Wark PAB, McDonald VM. Acute oxygen therapy; a review of prescribing and delivery practices. Int J Chron Obstruct Pulmon Dis 2016;11:1067–75.
2 Blakeman TC. Evidence for oxygen use in the hospitalized patient: is more really the enemy of good? Respir Care 2013;58:1679–93.
3 Grainge C. Breath of life: the evolution of oxygen therapy. J R Soc Med 2004;97:489–93.
4 Devoe NC, Kyriazis P, Eltanbedawi A. An audit of oxygen supplementation in a large tertiary hospital - we should treat oxygen as any other drug. Hosp Pract 2020;1–4.
5 Small D, Duha A, Wieskopf B, et al. Uses and misuses of oxygen in hospitalized patients. Am J Med 1992;92:591–5.
Bedoya AD, Bhavsar NA, Adagarla B, et al. Unanticipated respiratory compromise and unplanned intubations on general medical and surgical floors. *Respir Care* 2020;65:1233–40.

Lovett PB, Buchwald JM, Stürmann K, et al. The vexatious vital: neither clinical measurements by nurses nor an electronic monitor provides accurate measurements of respiratory rate in triage. *Ann Emerg Med* 2005;45:68–76.

DeMeulenaere S. Pulse oximetry: uses and limitations. *J Nurse Pract* 2007;3:312–7.

Gruber P, Kwiatkowski T, Silverman R, et al. Time to equilibration of oxygen saturation using pulse oximetry. *Acad Emerg Med* 1995;2:810–5.

Pruitt WC, Jacobs M. Breathing lessons: basics of oxygen therapy. *Nursing* 2003;33:43–5.

Fihn SD, Francis J, Clancy C, et al. Insights from advanced analytics at the Veterans health administration. *Health Aff* 2014;33:1203–11.

US Department of Veterans Affairs. 172VA10P2: VHA corporate data Warehouse, VA. 79 FR 4377.

Wiitala WL, Vincent BM, Burns JA, et al. Variation in laboratory test naming conventions in EHRs within and between hospitals: a nationwide longitudinal study. *Med Care* 2019;57:e22–7.

Wang XQ, Vincent BM, Wiitala WL, et al. Veterans Affairs patient database (VAPD 2014–2017): building nationwide granular data for clinical discovery. *BMC Med Res Methodol* 2019;19:94.

Vincent BM, Wiitala WL, Burns JA, et al. Using Veterans Affairs corporate data Warehouse to identify 30-day hospital readmissions. *Health Serv Outcomes Res Method* 2018;18:143–54.

Elixhauser A, Steiner C, Harris DR, et al. Comorbidity measures for use with administrative data. *Med Care* 1998;36:8–27.

Thanasegaran G. Reliability and validity issues in research. *Integ Dissemin* 2009;4:35–40.

Price P, Jhangiani R, Chiang I. *Research methods of psychology*. 2nd Canadian Edn, 2015.