JAMIA Open, 4(3), 2021, 1–6
doi: 10.1093/jamiaopen/ooab057
Brief Communications

Brief Communications

An exploratory data quality analysis of time series physiologic signals using a large-scale intensive care unit database

Ali S. Afshar[1], Yijun Li[2], Zixu Chen[3], Yuxuan Chen[3], Jae Hun Lee[3], Darius Irani[3], Aidan Crank[3], Digvijay Singh[3], Michael Kanter[4], Nauder Faraday[1] and Hadi Kharrazi[5,6]

[1]Department of Anesthesiology and Critical Care Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland USA, [2]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA, [3]Department of Computer Science, Johns Hopkins Whiting School of Engineering, Baltimore, Maryland, USA, [4]Department of Clinical Science, Kaiser Permanente Bernard J. Tyson School of Medicine, Pasadena, California, USA, [5]Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA, and [6]Division of Health Sciences Informatics, Johns Hopkins School of Medicine, Baltimore, Maryland, USA

Corresponding Author: Ali Sobhi Afshar, PhD, Department of Anesthesiology and Critical Care Medicine, Johns Hopkins School of Medicine, 600 N Wolfe St, Baltimore, MD 21205, USA (afshar@jhu.edu)

Received 27 March 2021; Revised 4 June 2021; Editorial Decision 30 June 2021; Accepted 2 July 2021

ABSTRACT

Physiological data, such as heart rate and blood pressure, are critical to clinical decision-making in the intensive care unit (ICU). Vital signs data, which are available from electronic health records, can be used to diagnose and predict important clinical outcomes; While there have been some reports on the data quality of nurse-verified vital sign data, little has been reported on the data quality of higher frequency time-series vital signs acquired in ICUs, that would enable such predictive modeling. In this study, we assessed the data quality issues, defined as the completeness, accuracy, and timeliness, of minute-by-minute time series vital signs data within the MIMIC-III data set, captured from 16009 patient-ICU stays and corresponding to 9410 unique adult patients. We measured data quality of four time-series vital signs data streams in the MIMIC-III data set: heart rate (HR), respiratory rate (RR), blood oxygen saturation (SpO2), and arterial blood pressure (ABP). Approximately, 30% of patient-ICU stays did not have at least 1 min of data during the time-frame of the ICU stay for HR, RR, and SpO2. The percentage of patient-ICU stays that did not have at least 1 min of ABP data was ~56%. We observed ~80% coverage of the total duration of the ICU stay for HR, RR, and SpO2. Finally, only 12.5%%, 9.9%, 7.5%, and 4.4% of ICU lengths of stay had $\geq$ 99% data available for HR, RR, SpO2, and ABP, respectively, that would meet the three data quality requirements we looked into in this study. Our findings on data completeness, accuracy, and timeliness have important implications for data scientists and informatics researchers who use time series vital signs data to develop predictive models of ICU outcomes.

Key words: physiologic monitoring, data quality, intensive care unit

INTRODUCTION

Physiological data represent a major source of information in intensive care units (ICUs). In clinical practice, it is common for vital signs data to be continuously displayed by monitoring devices; however, these vital signs data are only intermittently recorded—typically hourly—by nursing personnel in the clinically deployed

LAY SUMMARY

In this study, using a large-scale intensive care database, we looked into the quality of time series vital signs data acquired in intensive care settings from the perspective of three data quality metrics: timeliness, completeness, and accuracy. Approximately, 30% of patient stays did not have at least 1 min of data during the time-frame of the intensive care unit stay for heart rate, respiratory rate, or blood oxygen saturation. The percentage of patient stays that did not have at least 1 min of arterial blood pressure data was ~56%. In addition, ~1–2% of the stays were completely removed due to a lack of accurate heart rate, respiratory rate, blood oxygen saturation signals. This percentage for arterial blood pressure signals was 4.1%. Finally, only 12.5%, 9.9%, 7.5%, and 4.4% of the lengths of intensive care stays had ≥99% data available for heart rate, respiratory rate, blood oxygen saturation, and arterial blood pressure higher frequency time series, respectively, that would meet the three data quality requirements we looked into in this study. These findings would be of particular interest to informatics researchers interested in analyzing time series vital signs data acquired in inpatient settings, and to develop predictive models for inpatient outcomes.

electronic health record (EHR). Thus, clinical decision-making is largely based on these intermittent recordings. Increasingly, continuous recordings (eg, minute-to-minute) from vital signs monitors are being captured and stored within EHRs of ICU patients. These higher frequency vital signs recordings are a rich data source that can be harnessed to predict high-impact outcomes in the context of clinical decision support tools.[1–3]

Data quality poses an important challenge to the use of continuously streaming physiological data. Previous studies have addressed data quality issues for vital signs routinely entered in EHRs by healthcare practitioners.[4–6] However, those studies have not assessed data quality for large-scale capture of higher frequency physiologic signals that is now possible within ICU settings.

Completeness, accuracy, and timeliness of clinical data have been suggested as fundamental data quality measures for operational purposes.[7,8] Past studies have investigated data quality issues such as completeness and timeliness for physiologic signals, although not within ICU settings.[9,10] Completeness refers to the presence of values in data fields[7,8,11] and the availability of information for all relevant individuals.[12–16] Accuracy refers to the degree that data "correctly" represents the parameter being represented or described.[7,8] Biologically impossible values represent one instance for "incorrect" values. Timeliness is defined as the degree to which data represent reality from the required point in time.[7,8]

In this study, we explore the data quality, that is, completeness, accuracy, and timeliness, of higher frequency (minute-by-minute) time-series vital signs data captured in ICU settings.

METHODS

We used the Medical Information Mart for Intensive Care (MIMIC-III) database. MIMIC-III is a large, single-center database comprising information on patients admitted to critical care units at Beth Israel Deaconess Medical Center (BIDMC).[17] MIMIC-III is a fully deidentified data archive that contains demographic and detailed clinical data associated with patients admitted to critical care units at BIDMC between 2001 and 2012.[17] MIMIC-III data were downloaded from the curator[17] after completion of the applicable data use agreement. The study was reviewed by the Johns Hopkins Institutional Review Board and qualified for exemption of informed consent.

MIMIC-III Waveform Database contains thousands of recordings of multiple physiologic signals ("waveforms") and time series of vital signs ("numerics") collected from bedside patient monitors in ICUs. The MIMIC-III Waveform Database Matched Subset[18] contains 22 247 time-series vital signs records which have been matched to the records in the Clinical Database of MIMIC-III. We developed a process to harmonize the physiological record frequencies and to match the records with individual ICU stays (see Supplementary Sections A, B, and C for details).

We measured the completeness, accuracy, and timeliness of four time-series (minute-by-minute) vital signs data streams: heart rate (HR), respiratory rate (RR), blood oxygen saturation (SpO2), and arterial blood pressure (ABP). We did not include data streams from blood pressure cuffs since they were not recorded minute-by-minute. We included a vital sign data stream in the analysis of a patient if at least 1 minute of the data (ie, represented by one data recording in a minute-by-minute data stream) was captured in the record for any of the four vital signs.

To assess the data quality of time series vital signs data, we applied three filtering steps based on three data quality metrics. Completeness of time series vital signs data was measured: (1) by the presence of a non-NULL record for a given physiologic variable (HR, RR, SpO2, ABP) corresponding to an ICU stay for a patient and (2) by the proportion of nonmissing data for each physiologic variable (HR, RR, SpO2, and ABP) in a given ICU stay for a patient. If an entire record corresponding to a physiologic variable was nonexistent or NULL, then the corresponding physiologic variable for that ICU stay for the patient was flagged as incomplete. Accuracy of the vital signs was measured by a series of rules examining the physiological limits of the vital signs (see Supplementary Section D). Among the methods discussed in references 7,8 to assess accuracy, applying validity checks for biologically impossible values would allow all time series data elements with higher-frequency measurements such as in this study, to be assessed. This is because the clinician-verified measurements considered as the gold standard would not be available at all these higher-frequency time points obtained by bedside monitors. Timeliness of the vital signs was assessed as the overlap of the minute-by-minute records with the ICU stay (see Supplementary Section E). Admission to and discharge from ICU was timestamped in the database allowing us to measure the temporal overlap of vital signs data with the ICU stay.

The average age of the population was calculated using the average age of each patient corresponding to each patient-ICU stay between 2001 and 2012. Our inclusion criteria were the patients had to be of age 16 or older and have an ICU stay with higher frequency time series vital signs data for at least one of these physiologic sig-

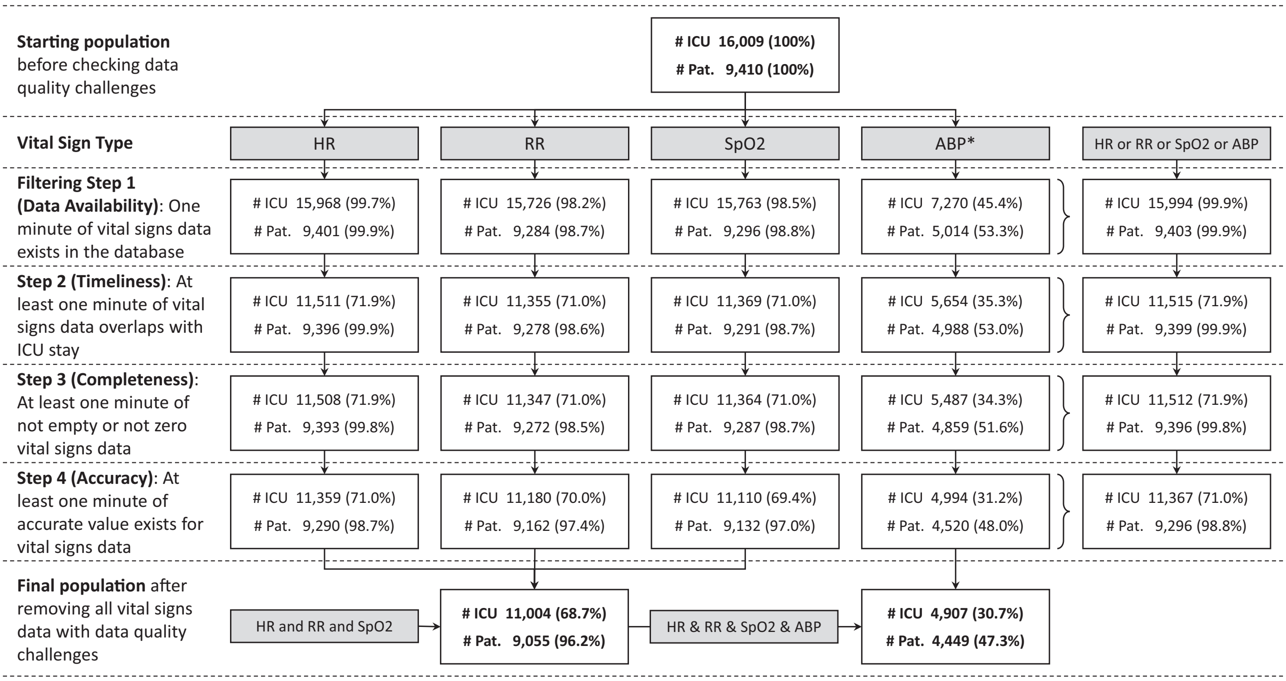
Starting population before checking data quality challenges
# ICU 16,009 (100%)
# Pat. 9,410 (100%)
Vital Sign Type
HR
RR
SpO2
ABP*
HR or RR or SpO2 or ABP
Filtering Step 1 (Data Availability): One minute of vital signs data exists in the database
# ICU 15,968 (99.7%) # Pat. 9,401 (99.9%)
# ICU 15,726 (98.2%) # Pat. 9,284 (98.7%)
# ICU 15,763 (98.5%) # Pat. 9,296 (98.8%)
# ICU 7,270 (45.4%) # Pat. 5,014 (53.3%)
# ICU 15,994 (99.9%) # Pat. 9,403 (99.9%)
Step 2 (Timeliness): At least one minute of vital signs data overlaps with ICU stay
# ICU 11,511 (71.9%) # Pat. 9,396 (99.9%)
# ICU 11,355 (71.0%) # Pat. 9,278 (98.6%)
# ICU 11,369 (71.0%) # Pat. 9,291 (98.7%)
# ICU 5,654 (35.3%) # Pat. 4,988 (53.0%)
# ICU 11,515 (71.9%) # Pat. 9,399 (99.9%)
Step 3 (Completeness): At least one minute of not empty or not zero vital signs data
# ICU 11,508 (71.9%) # Pat. 9,393 (99.8%)
# ICU 11,347 (71.0%) # Pat. 9,272 (98.5%)
# ICU 11,364 (71.0%) # Pat. 9,287 (98.7%)
# ICU 5,487 (34.3%) # Pat. 4,859 (51.6%)
# ICU 11,512 (71.9%) # Pat. 9,396 (99.8%)
Step 4 (Accuracy): At least one minute of accurate value exists for vital signs data
# ICU 11,359 (71.0%) # Pat. 9,290 (98.7%)
# ICU 11,180 (70.0%) # Pat. 9,162 (97.4%)
# ICU 11,110 (69.4%) # Pat. 9,132 (97.0%)
# ICU 4,994 (31.2%) # Pat. 4,520 (48.0%)
# ICU 11,367 (71.0%) # Pat. 9,296 (98.8%)
Final population after removing all vital signs data with data quality challenges
HR and RR and SpO2
# ICU 11,004 (68.7%) # Pat. 9,055 (96.2%)
HR & RR & SpO2 & ABP
# ICU 4,907 (30.7%) # Pat. 4,449 (47.3%)

Figure 1. Measuring data quality of vital signs data, as defined in the Methods section, and associated with ICU stays for which at least 1 min of vital sign data exists. *In contrast to HR, RR, and SpO2, capturing ABP is not clinically required for all ICU stays.

nals (HR, RR, SpO2, and ABP) (see Supplementary Sections A, B, and C). We flagged patient-ICU stays for any completeness, accuracy, and timeliness issues in their time series vital signs records for HR, RR, SpO2, and ABP. We then populated patient denominator specifications and data quality summary tables. We summarized the starting population, removed population (ie, not meeting the minimum data quality requirements as earlier defined by the application of filtering steps corresponding to timeliness, completeness, and accuracy metrics), and the remaining population for various levels of vital signs data completeness. This manuscript reports the results of our analysis when at least 1 min of complete and accurate vital sign data overlaps with an ICU stay (ie, the smallest unit of analysis for a minute-by-minute data stream). Population analysis for other rates of acceptable data quality, namely minimum thresholds of 25%, 50%, 75%, or 100% coverage of an ICU stay, were also calculated separately (see Supplementary Sections F and G). Finally, to summarize various thresholds of acceptable data quality on the number of ICU stays, we plotted the proportion of ICU stays against percentage of ICU length of stays covered by acceptable vital signs data. All analysis was performed using R version 3.5.[19]

RESULTS

Of the 22 247 time-series vital sign records in the MIMIC-III Waveform Database Matched Subset,[18] 21 326 records were: (1) mappable to distinct ICU stays (see Supplementary, Sections B and C for details); and (2) associated with patients age 16 or older. The 21 326 vital signs records corresponded to 16 009 distinct ICU stays and 9410 unique patients.

We applied certain filtering steps to determine the ICU stays with at least 1 min vital signs data associated with the stay, and to assess time-series vital signs data quality from the perspective of timeliness, completeness and accuracy metrics (Figure 1). When this data availability step was performed, the minimum 1 min HR, RR, and SpO2 records were found for most patients (>98%), but only 7270 (45.4%) of patient-ICU stays had ABP records available. This finding was clinically anticipated, but we wanted to quantify the record missingness using this large-scale data set, which would have important implications for clinical data scientists and informatics researchers. We also observed a ∼28% drop in the number of HR, RR, and SpO2 records per ICU stays after applying the filtering step for timeliness (step 2 in Figure 1, ie, having at least 1 min overlap with the vital sign's associated ICU stay). This percentage for ABP time series records was ∼13%. Within the 1 min timely vital signs, a small percentage of HR, RR, and SpO2 records (0.9–1.6%) contained incomplete or inaccurate data. For ABP records, this percentage was 4.1%. Overall, 4907 patient-ICU stays (30.7% of the starting population) had at least 1 min of accurate data for all four physiologic vital signs data after applying completeness, accuracy, and timeliness filtering steps.

We summarized the population demographics (eg, age and gender), utilization (eg, ICU length of stay (LOS)), and the statistical summaries associated with vital signs after applying all four filtering steps listed in Figure 1 (see Table 1, and the Supplementary Section F for percentage of ICU LOS thresholds). On average, we have ∼80% coverage for HR, RR, and SpO2 for the duration of ICU stay.

After identifying the patient-ICU stays with at least 1 min overlapping vital signs data within the timeframe of the ICU stay (filtering steps 1 and 2), we quantified the effect of applying filtering steps 3 (completeness) and 4 (accuracy) for this specific group of patient-ICU stays (Table 2), by categorizing the patient-ICU stay population into the "starting," "removed" (not meeting the completeness and accuracy data quality measures) and "remaining" groups for each time series vital sign variable (ie, HR, RR, SpO2, and ABP), and the patient-ICU stays that have all four vital signs (AND). In other words, for the patient-ICU stays that have at least 1 min overlapping

Table 1. Demographics and utilization rates of patients with at least 1-min of acceptable vital signs data during an ICU-stay (passing all four filtering steps in Figure 1)

	HR	RR	SpO2	ABP[a]	OR	AND
ICU stays	11 359	11 180	11 110	4994	11 367	4907
Age	64.34	64.38	64.39	64.81	64.34	64.89
Mean (SD)	(16.38)	(16.40)	(16.37)	(14.83)	(16.37)	(14.82)
Gender	4953	4882	4858	2056	4959	2019
Female *n* (%)	(43.6)	(43.7)	(43.7)	(41.2)	(43.6)	(41.1)
ICU LoS[b]	3337	3339	3349	4591	3336	4601
Median (IQR25–IQR75)	(1842–7161)	(1845–7149)	(1850–7177)	(2313–10 929)	(1842–7149)	(2347–10 882)
Mean	6776	6761	6816	9056	6773	9055
HR length[b]	2643	2645	2654	3228	2640	3252
Median (IQR25–IQR75)	(1442–5072)	(1444–5086)	(1453–5118)	(1612–6873)	(1441–5069)	(1622–6880)
Mean	4322	4322	4349	5479	6773	5493
RR length[b]	2542	2581	2574	3071	2541	3120
Median (IQR25–IQR75)	(1380–4763)	(1412–4837)	(1405–4829)	(1541–6399)	(1379–4760)	(1571–6521)
Mean	4118	4184	4171	5186	4115	5257
SpO2 length[b]	2297	2322	2358	2911	2293	2941
Median (IQR25–IQR75)	(1258–4461)	(1275–4490)	(1302–4550)	(1473–6028)	(1258–4459)	(1488–6054)
Mean	3881	3908	3968	5001	3878	5034
ABP length[b]	0[c]	0[c]	0[c]	2353	0[c]	2363
Median (IQR25–IQR75)	(0–1840)	(0–1867)	(0–1912)	(1200–5155)	(0–1838)	(1201–5152)
Mean	1895	1902	1928	4311	1894	4313

ABP: arterial blood pressure; AND: HR and RR and SpO2 and ABP; ICU: intensive care unit; IQR: interquartile range; HR: heart rate; LoS: length of stay; OR: HR or RR or SpO2 or ABP; RR: respiratory rate; SD: standard deviation; SpO2: blood oxygen saturation.

[a] ABP is not clinically required to be captured for all ICU stays;

[b] Reported in rounded minutes; and

[c] ABP time series had a median length of 0 due to the fact that more than ~56% of patient ICU stays did not have ABP records.

Table 2. Specification of the starting population (ie, output of timeliness filtering step listed in Figure 1), population removed due to not having at least 1-min of acceptable vital signs data (ie, passing completeness and accuracy measures or in other words, filtering steps 3 and 4) during an ICU-stay, and the remaining population

Patient population	Vital sign	No. of patients	No. of ICU stays	Avg length (Min)	% LoS[a]
Starting	HR	9396	11 511	4398	0.801
	RR	9278	11 355	4410	0.800
	SpO2	9291	11 369	4416	0.801
	ABP	4988	5654	5596	0.779
	AND	4955	5620	5607	0.779
Removed	HR	149	152	1370	0.724
	RR	167	175	2701	0.827
	SpO2	239	259	2352	0.818
	ABP	607	660	5699	0.872
	AND	43	43	1502	0.862
Remaining	HR	9290	11 359	4322	0.772
	RR	9162	11 180	4184	0.753
	SpO2	9132	11 110	3968	0.702
	ABP	4520	4994	4311	0.576
	AND	4449	4907	5024	0.683

We should note that the Avg Length and LoS values corresponding to AND in each group are not necessarily expected to have the lowest value in each group as these values are calculated for the group of ICU stays that have all the four vital signs (ie, AND category).

ABP: arterial blood pressure; AND: HR and RR and SpO2 and ABP; ICU: intensive care unit; HR: heart rate; LoS: length of stay; RR: respiratory rate; SpO2: blood oxygen saturation.

[a] % LoS = Duration of Vital Sign/ICU LOS.

vital signs data within the timeframe of the ICU stay (output of filtering steps 1 and 2), Table 2 sheds more light on the effect of applying completeness and accuracy filtering steps on the vital signs data in terms of the change in the average length of the vital sign record corresponding to each vital sign. The percentage length of stay (LoS) was calculated as the ratio of each vital signs record's length to the length of the ICU stay. We also calculated the specifications of the starting, removed, and remaining populations for various minimum percentages of ICU length of stay (see Supplementary Section G).

As shown in Table 2, vital signs data had different shares of the "removed" population (eg, ABP affecting 660 patient-ICU stays while HR affecting only 152 stays). We explored the overlap of the vital signs data quality challenges on the total number of "removed" patient-ICU stays ($n = 939$; not shown in Table 2). Figure 2 depicts

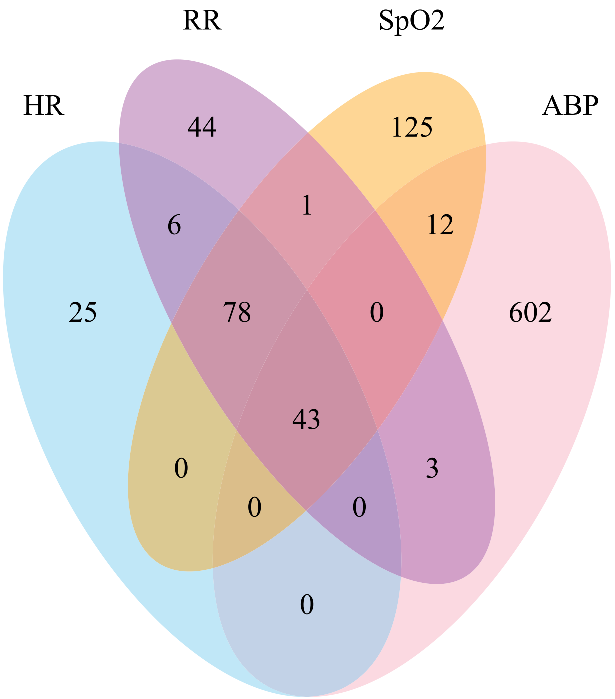

Figure 2. Venn diagram of patient-ICU stays removed when at least 1-min of acceptable vital signs data are required (ie, passing completeness and accuracy filtering steps listed in Figure 1).

the overlapping contribution of completeness and accuracy data quality issues (i.e. filtering steps 3 and 4 in Figure 1) for each physiological variable in the removed patient-ICU stays. In other words, for the "removed" population in Table 2, Figure 2 shows how many patient-ICU stays were affected by the quality of each of the vital signs corresponding to those stays, and mutually, which vital signs bore a higher burden in terms of completeness and accuracy data quality issues.

We finally explored the proportion of ICU stays (0–1) against the percentage of ICU length of stays (0% to 100%) that were covered by acceptable vital signs data (ie, complete, accurate, and timely, or in other words, passing all four filtering steps listed in Figure 1) in the study population (Figure 3). The proportion of patient-ICU stays that had $\geq$ 99% of their length covered by an acceptable vital sign data were 12.5% for HR, 9.9% for RR, 7.5% for SpO2, and 4.4% for ABP (represented by the lowest point, before zero, of each vital sign line in Figure 3). Inflection points of each physiologic variable were also extracted which corresponds to the specific percentage of ICU length of stay where the quality of vital signs data (a combination of timeliness, accuracy, and completeness) starts to degrade considerably (Figure 3).

DISCUSSION

Data quality can significantly affect the analysis of healthcare data that informs clinical decisions. Previous studies have investigated data quality issues for vital signs recorded intermittently and manually into EHRs. Other studies have addressed data quality issues for physiologic signals acquired in out-of-the-hospital settings. In this study, we measured the completeness, accuracy, and timeliness of time series physiological data (ie, HR, RR, SpO2, and ABP) using the MIMIC-III database.

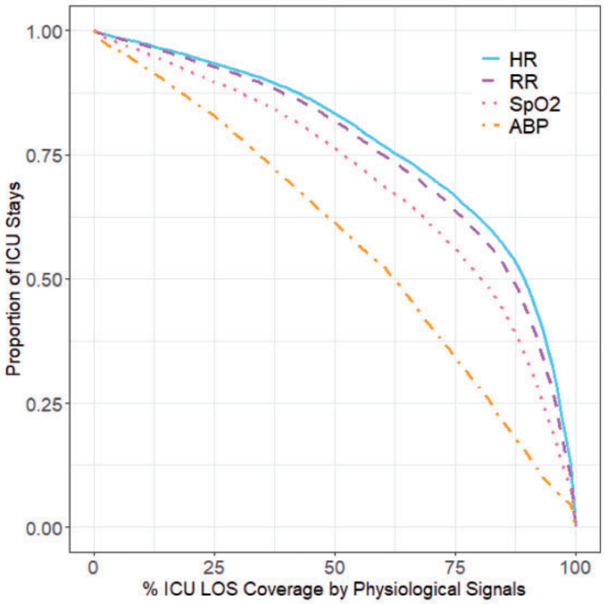

Figure 3. Percentage of ICU length of stays covered by acceptable vital sign data (ie passing all four filtering steps listed in Figure 1) versus proportion of ICU stays*. Inflection points of each physiologic variable where the quality of vital signs data (a combination of timeliness, accuracy and completeness) starts to degrade considerably are also represented: HR at 75% of the average ICU length of stays (representing 66.7% of ICU stays); RR at 73% of length of stays (65.8% of ICU stays); SpO2 at 69% of length of stays (62% of ICU stays); and, ABP at 60% of length of stays (52.7% of ICU stays). *X-axis shows the percentage of ICU length of stays covered by acceptable vital sign data. Y-axis shows the proportion of ICU stays. The top left corner represents all ICU stays with 0% requirement for vital signs data quality during an ICU stay. The bottom right corner represents proportion of ICU stays with 100% acceptable vital signs throughout the ICU length of stay.

Our results show that most ICU stays had at least 1-min of complete, accurate, and timely HR, RR, and SpO2 data; however, the availability of these physiological data is dramatically reduced when the minimum coverage of ICU length of stay is set at 75% or higher (Figure 3). Thus, analytic models requiring a high coverage of acceptable HR, RR, or SpO2 vital signs during an ICU stay, especially a 75%+ coverage, may produce poor results if applied to all ICU stays.

ABP vital sign is not required to be collected for all patients during an ICU stay, hence we anticipated, and the results confirmed, a lower coverage of ICU length of stays by acceptable ABP data. Indeed, while more than 96% of the patients had an ICU stay with at least 1-min of acceptable HR, RR, and SpO2 data, only ~48% of patients had an ICU stay with at least 1-min of acceptable ABP data. This finding would have important implications for clinical data scientists and informatics researchers as the low rates of acceptable ABP data during an ICU stay may limit the development and deployment of analytic models using ABP data for all patients.

Our study has several limitations. First, our analysis may not generalize to other ICU settings because MIMIC-III data is a preprocessed database, and certain data quality issues may have been ratified prior to our study (see Supplementary Sections A–C for details). Second, we made several assumptions of temporal patterns and thresholds with regards to accuracy and timeliness (see Supplementary Sections A–E) which may not be applicable to all settings and/or vital signs. Future research should examine the optimal temporal patterns and thresholds of time series vital signs data recorded in an

ICU setting. Third, our results have limited applicability for patients 89+ years old, as MIMIC-III is a fully deidentified database and lumps together all ages above 89, as 89 years old. Fourth, we examined the quality of vital signs data using a minute-by-minute scale and measured the coverage of ICU length of stay by either a 1-min overlap or a specific percentage overlap of acceptable data (see Supplementary Sections F and G). Future studies should explore other minimum temporal thresholds of acceptable vital signs during an ICU stay (eg, having at least 60 min) as well as different temporal frequencies of vital signs (eg, second-by-second data).

Despite these limitations, our findings would be of particular importance to clinical data scientists and informatics researchers who would be interested in analyzing time series vital signs data from inpatient settings, extracting features, and developing statistical models for outcome prediction.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

CONTRIBUTORS

A.A., N.F., and H.K. conceptualized the work. Z.C., Y.C., J.L., D.I., A.C., D.S. conducted data processing with direct supervision from A.A. Y.L. generated the tables and figures with supervision from A.A. M.K. reviewed the work as an independent reviewer and provided feedback. A.A., N.F., and H.K. wrote the manuscript by taking all the input from coauthors into account.

FUNDING

This work was funded in part by TEDCO Maryland Innovation Initiative grant 2019-MII-5066 and also by the Johns Hopkins Center for Population Health IT (CPHIT).

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY

The data underlying this article are available in the MIMIC-III Waveform Database Matched Subset, at https://dx.doi.org/doi:10.13026/C2294B.

REFERENCES

1. Sun Y, Guo F, Kaffashi F, Jacono FJ, DeGeorgia M, Loparo KA. INSMA: an integrated system for multimodal data acquisition and analysis in the intensive care unit. *J Biomed Inform* 2020; 106: 103434.
2. Gutierrez G. Artificial intelligence in the intensive care unit. *Crit Care* 2020; 24 (1): 101.
3. Hyland SL, Faltys M, Hüser M, *et al.* Early prediction of circulatory failure in the intensive care unit using machine learning. *Nat Med* 2020; 26 (3): 364–73
4. Skyttberg N, Chen R, Blomqvist H, Koch S. Exploring vital sign data quality in electronic health records with focus on emergency care warning scores. *Appl Clin Inform* 2017; 8 (3): 880–92.
5. Reimer AP, Milinovich A, Madigan EA. Data quality assessment framework to assess electronic medical record data for use in research. *Int J Med Inform* 2016; 90: 40–7.
6. Genes N, Chandra D, Ellis S, Baumlin K. Validating emergency department vital signs using a data quality engine for data warehouse. *Open Med Inform J* 2013; 7: 34–9.
7. Nasir A, Gurupur V, Liu X. A new paradigm to analyze data completeness of patient data. *Appl Clin Inform* 2016; 7 (3): 745–64..
8. Weiskopf NG, Hripcsak G, Swaminathan S, Weng C. Defining and measuring completeness of electronic health records for secondary use. *J Biomed Inform* 2013; 46 (5): 830–6.
9. Zhang J, Tüshaus L, Nuño Martínez N, *et al.* Data integrity-based methodology and checklist for identifying implementation risks of physiological sensing in mobile health projects: quantitative and qualitative analysis. *JMIR Mhealth Uhealth* 2018; 6 (12): e11896.
10. Sweeney KT, Ward TE, McLoone SF. Artifact removal in physiological signals–practices and possibilities. *IEEE Trans Inf Technol Biomed* 2012; 16 (3): 488–500.
11. Welch G, von Recklinghausen F, Taenzer A, Savitz L, Weiss L. Data cleaning in the evaluation of a multi-site intervention project. *J Electron Health Data Methods* 2017; 5 (3): 1–7.
12. Warsi A, White S, McCulloch P. Completeness of data entry in three cancer surgery databases. *Eur J Surg Oncol* 2002; 28 (8): 850–6.
13. Barlow L, Westergren K, Holmberg L, Tälback M. The completeness of the Swedish Cancer Register – a sample survey for year 1998. *Acta Oncologica* 2009; 48 (1): 27–33.
14. Lamberg AL, Cronin-Fenton D, Olesen AB. Registration in the Danish Regional Nonmelanoma Skin Cancer Dermatology Database: completeness of registration and accuracy of key variables. *Clin Epidemiol* 2010; 2: 123–36.
15. Sigurdardottir LG, Jonasson JG, Stefansdottir S, *et al.* Data quality at the Icelandic Cancer Registry: comparability, validity, timeliness and completeness. *Acta Oncol* 2012; 51 (7): 880–9.
16. Kharrazi H, Wang C, Scharfstein D. Prospective EHR-based clinical trials: the challenge of missing data. *J Gen Intern Med* 2014; 29 (7): 976–8.
17. Johnson AEW, Pollard TJ, Shen L, *et al.* MIMIC-III, a freely accessible critical care database. *Sci Data* 2016; 3 (1): 1–9. doi:10.1038/sdata.2016.35.
18. PhysioNet. Overview of databases. https://physionet.org/about/database/. Accessed May 21, 2021.
19. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria. https://www.r-project.org/. Accessed May 21, 2021.