OPEN

Optimized Arterial Line Artifact Identification Algorithm Cleans High-Frequency Arterial Line Data With High Accuracy in Critically Ill Patients

Jasmine M. Khan, BSc[1]

David M. Maslove, MD, MSc, FRCPC[2,3]

J. Gordon Boyd, MD, PhD, FRCPC[2,3]

OBJECTIVES: High-frequency data streams of vital signs may be used to generate individualized hemodynamic targets for critically ill patients. Central to this precision medicine approach to resuscitation is our ability to screen these data streams for errors and artifacts. However, there is no consensus on the best method for data cleaning. Our goal was to determine whether an error-checking algorithm developed for intraoperative use could be applied to high volumes of arterial line data in an ICU population.

DESIGN: Multicenter observational study.

SETTING: ICUs across Ontario, Canada.

PATIENTS: Nested cohort of ICU patients with shock and/or respiratory failure requiring invasive mechanical ventilation.

INTERVENTIONS: High-frequency blood pressure data was analyzed. Systolic, diastolic, and mean arterial pressure minute averages were calculated. For manual analysis, a trained researcher retrospectively reviewed mean arterial pressure data, removing values that were deemed nonphysiological. The algorithm was implemented and identified artifactual data.

MEASUREMENTS AND MAIN RESULTS: Arterial line data was extracted from 15 patients. A trained researcher manually reviewed 40,798 minute-by-minute data points, then subsequently analyzed them with the algorithm. Manual review resulted in the identification of 119 artifacts (0.29%). The optimized algorithm identified 116 (97%) of these artifacts. Five hundred thirty-seven data points were erroneously removed or modified. Compared with manual review, the modified algorithm incorporating absolute thresholds of greater than 30 and less than 200 mm Hg had 97.5% sensitivity, 98.7% specificity, and a Matthew correlation coefficient of 0.41.

CONCLUSIONS: The error-checking algorithm had high sensitivity and specificity in detecting arterial line blood pressure artifacts compared with manual data cleaning. Given the growing use of large datasets and machine learning in critical care research, methods to validate the quality of high-frequency data is important to optimize algorithm performance and prevent spurious associations based on artifactual data.

KEY WORDS: artifacts; critical care; data cleaning; invasive blood pressure monitoring; mean arterial pressure

DOI: 10.1097/CCE.0000000000000814

High-frequency data streams of vital signs can be used to generate individualized hemodynamic targets for critically ill patients (1). However, accuracy of this approach is dependent on adequate removal of artifacts from data streams prior to analysis (2). Arterial line data, which is used to

KEY POINTS

Question: Can an algorithm developed for intraoperative use accurately clean invasive arterial line data collected in the ICU?

Findings: The optimized algorithm identified errors with high accuracy, producing 98% sensitivity and 99% specificity when compared with manual error-checking.

Meanings: The algorithm was easy to implement with high volumes of arterial line data. It has the potential to be integrated into invasive blood pressure data analysis in the context of research and clinical care.

derive these targets, are prone to artifacts due to blood sampling, malpositioning, line flushing, and calibration procedures (3, 4). High volumes of data generated in the ICU present a logistical challenge for manual artifact identification, necessitating the automation of arterial line data cleaning. Despite this, there is no consensus on the best method for cleaning arterial line data derived from the ICU (5).

Various methods have been used to assess artifacts in arterial line traces in the ICU and can be broadly classified as machine learning, statistical, or logic-based (6). Each model presents its own challenges, including implementing arbitrary thresholds, averaging data over large time intervals, or using complex machine learning techniques (6, 7). In the ICU, most existing artifact identification algorithms require the input of waveform data (2, 4, 6, 8–11), which is much larger and more computationally demanding than basic numeric data and which may not be as readily available. Few existing algorithms use numeric data input; this approach has potential advantages but requires further validation (12, 13). More stringent methods are needed to capture changes that appear physiologic, preserve data granularity, and can be implemented with ease on invasively derived numeric blood pressure data.

The recent emergence of an arterial line error-checking algorithm developed for intraoperative use by Du et al (7) provides a promising avenue for numeric blood pressure data cleaning. The algorithm identifies artifacts by integrating minute-by-minute blood pressure changes over time as well as the difference between systolic blood pressure (SBP), diastolic blood pressure (DBP), and mean arterial pressure (MAP) values. When sufficient data is present across traces, points identified as discrepant are altered based on data from preceding and succeeding time points as well as from the other two traces. An assessment of the algorithm's performance in the ICU is warranted before use in critical care, as intraoperative conditions differ significantly. First, arterial line placement, zeroing, and stabilization usually occur once for the entire operation, whereas in the critical care setting, there may be more disruption due to patient movement, procedures, blood draws, and transport. Second, invasive arterial pressure monitoring in the OR is only done for the duration of surgery, typically a few hours, whereas ICU monitoring may extend much longer, over multiple days. Longer term recordings have the potential for a different set of morphologic variabilities and artifacts, for example, due to nonphysiologic changes attributed to longer duration of catheters use (e.g., kinking and blockages). Third, clinical and pharmacological factors predispose patients to arterial line artifacts in the ICU that may impact algorithm accuracy (14).

Our objective was to determine whether an error-checking algorithm that assesses the association between SBP, DBP, and MAP data developed for intraoperative use could be applied to high volumes of arterial line data in an ICU population to reliably detect and remove artifacts.

MATERIALS AND METHODS

Study Design

High-frequency blood pressure data from a nested cohort of ICU patients enrolled in the multicenter prospective Cerebral Oxygenation and Neurological Outcomes FOllowing CriticAL Illness-2 (CONFOCAL-2) study (15) (Clinicaltrials.gov ID: NCT03141619) from three ICUs across Ontario, Canada, were analyzed. The study was conducted in accordance with the ethical standards of Clinical Trials Ontario and with the Helsinki Declaration of 1975, as most recently amended. Ethics approval was obtained by Clinical Trials Ontario (Study Title: Correlation of Cerebral Oxygenation During Critical Illness and Neurological Outcomes, Approval no.: 0815, Approval Date: July 17, 2017), and informed consent was obtained for the CONFOCAL-2 study for each participant. Patients were enrolled if in shock and/or respiratory

failure requiring invasive mechanical ventilation for greater than 24 hours. Patients were excluded if they had less than 24 hour life expectancy, neurologic admitting diagnosis, or inability to participate in follow-up assessments (15). Invasive blood pressure monitoring was collected using GE Solar monitors and extracted with commercial software (Bedmaster, Anandic Medical Systems, Feuerthalen, Switzerland) for up to 72 hours beginning at study enrolment. SBP, DBP, and MAP values were derived from waveforms captured at 240 Hz and sampled at 0.5 Hz. Minute averages (i.e., 1/60 Hz) were calculated using the numerical data captured at 0.5 Hz without reference to the waveforms themselves.

Manual Analysis

For manual analysis, the researcher retrospectively reviewed a tabular depiction of high-frequency MAP data point-by-point and removed any values deemed nonphysiological. Artifacts were defined as values that increased or decreased rapidly (e.g., > 20 mm Hg transient fluctuations from preceding MAP values) and returned to values similar to those preceding the large change in a short period of time or were nonphysiological (i.e., values < 30 and > 200 mm Hg). An example of fluctuations deemed nonphysiological is shown in **Figure 1**. The reviewer was blinded to systolic and diastolic traces. A second researcher reviewed points deemed anomalous, and any disagreements were discussed until a consensus was reached.

Algorithm Modification

The algorithm derived by Du et al (7) was implemented using the preset parameters to identify artifactual data. **Figure S1** (http://links.lww.com/CCX/B105) describes the algorithm in detail. The algorithm contains modifiable delta values (step 2) that dictate the magnitude of change between consecutive readings for each pressure trace deemed within normal limits. For

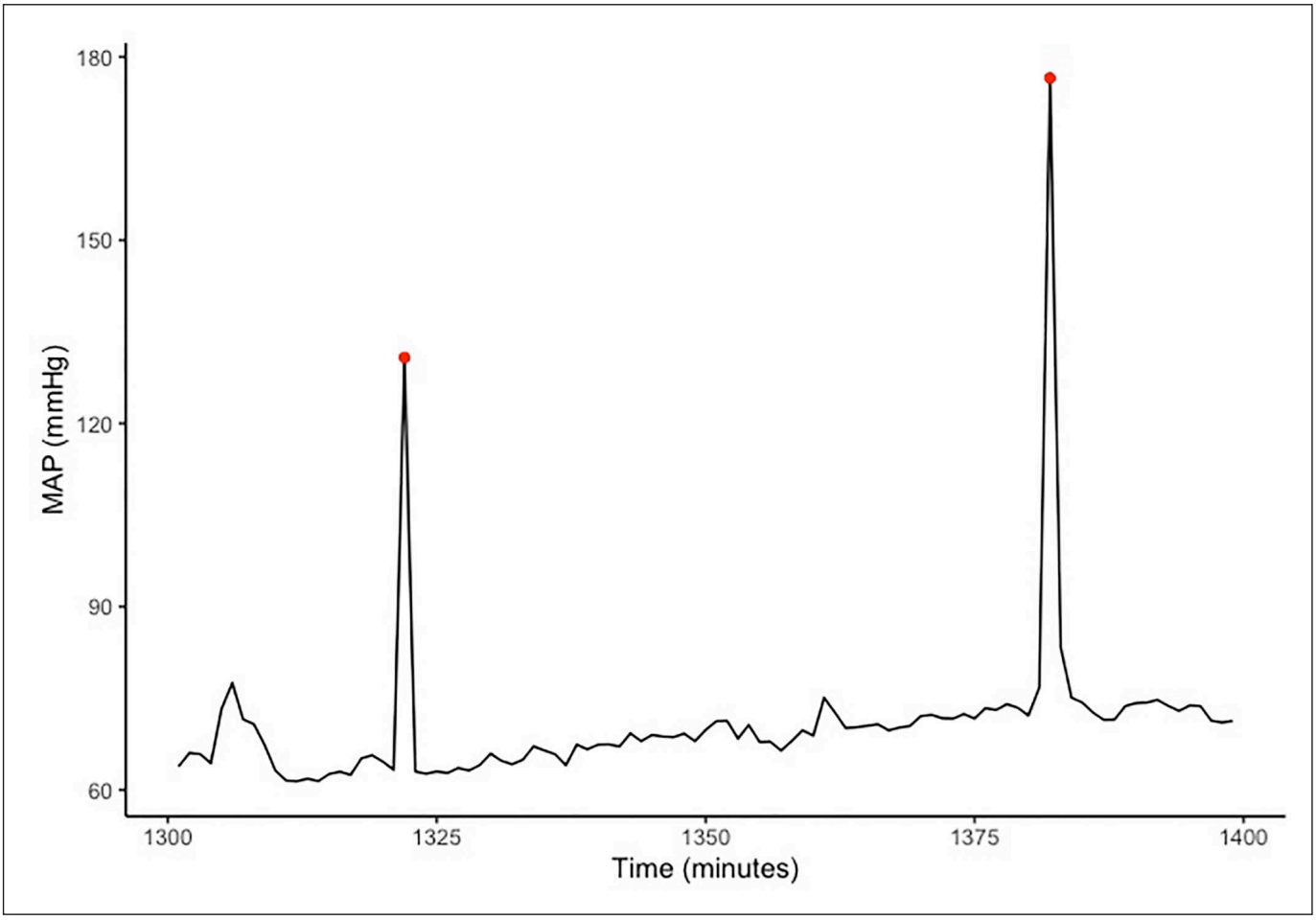

Figure 1. Representative mean arterial pressure (MAP) trace containing fluctuations deemed nonphysiological, which were subsequently removed manually (*red circles*).

values that vary significantly from acceptable limits, the data point is interpolated based on values from the other two blood pressure traces and/or from the first available data preceding and succeeding this artifact in time. If there are insufficient data to interpolate a modified value, the data point is removed and replaced with "NA."

We defined any modification greater than or equal to 5 mm Hg to be an artifactual data point based on clinical reasoning. Alternative definitions ranging from 0–10 mm Hg were also assessed (**Table S1**, http://links.lww.com/CCX/B105). Original delta values from Du et al (7) for MAP, SBP, and DBP were 6, 8, and 5 mm Hg, respectively. During algorithm optimization, we modified delta values to determine the values associated with the best performance in our critically ill population. Delta values were multiplied by one factor across all traces. For example, modification by a factor of 2 resulted in MAP, SBP, and DBP delta values of 12, 16, and 10 mm Hg, respectively. Additional modifications to the algorithm included adding upper and lower absolute limits. We removed MAP values less than 30 and greater than 200 mm Hg as these were unlikely to represent true physiological data (4, 9).

Algorithm Analysis

2 × 2 tables were generated to compare algorithm performance at different delta thresholds with and without implementation of absolute blood pressure limits. Sensitivity, specificity, Matthew correlation coefficient (MCC) (16, 17), F1 score, and Cohen kappa (18) were calculated for each algorithm iteration to evaluate performance. MCC was chosen based on its robustness to data imbalance in a binary classification model. The event rate was extremely low (0.3%), therefore warranting analysis with MCC rather than Cohen kappa (19). MCC ranges from –1 to 1, where 1 indicates complete agreement between manual and algorithm cleaning. All analyses and plots were generated in R (Version 4.1.1) (20). The modified algorithm is available for use at https://github.com/jasmine-jk/ICU-MAP-Cleaning.git.

RESULTS

Arterial line blood pressure data was extracted from 15 patients enrolled over 9 months (from April to December 2018). Forty thousand seven hundred ninety-eight minute-by-minute data points were collected with mean (± SD) individual recording length of 48 hours (± 1,243 min). Manual review resulted in the identification of 119 artifacts (0.29%), of which 47 were less than 30 mm Hg and 8 were greater than 200 mm Hg. The unmodified algorithm identified 111 (93%) of manually identified artifacts. The algorithm also identified additional 575 data points, which were classified as false positives. The unmodified algorithm had 92.3% sensitivity and 98.8% specificity compared with manual cleaning. A representative example of manual and algorithmic MAP trace cleaning is shown in **Figure 2**.

To optimize accuracy and MCC for use in critical care, we tested various multipliers of delta values. Multiplication factors ranged from 0.1 to 6. Each factor was tested with and without the addition of upper and lower MAP value limits. Further, we incorporated post hoc removal of data breaks as we found the algorithm attempted to fill in spaces inappropriately (**Fig. S2**, http://links.lww.com/CCX/B105). Metrics to analyze algorithm performance including 2 × 2 tables, MCC, F1 score, and Kappa were generated (**Table 1**; and **Table S2**, http://links.lww.com/CCX/B105).

Sensitivity and specificity without the addition of limits were optimized when delta values were multiplied by a factor of 0.4 (deltaMAP = 2.4 mm Hg, deltaSBP = 3.2 mm Hg, deltaDBP = 2 mm Hg), producing a 95.0% sensitivity and 98.0% specificity. Seven values were less than 30 mm Hg, and no values were greater than 200 mm Hg. MCC was optimized with a delta multiplication factor of 1.3. The addition of greater than 30 mm Hg and less than 200 mm Hg limits resulted in optimized accuracy when deltas were multiplied by a factor of 1.3 (deltaMAP = 7.8 mm Hg, deltaSBP = 10.4 mm Hg, deltaDBP = 6.5 mm Hg), resulting in 97.5% sensitivity and 98.7% specificity compared with manual cleaning (**Table 2**). One hundred sixteen (97%) of artifacts identified manually were removed or modified with the algorithm, with an additional 537 false positives generated. MCC was highest at a delta multiplication factor of 2. The optimized algorithm included greater than 30 mm Hg and less than 200 mm Hg limits, with a 1.3 delta factor.

Algorithm performance was also assessed on a per-patient basis. **Table 3** describes the association between manual and algorithm artifacts. The rate of true artifact identification was high, with the algorithm

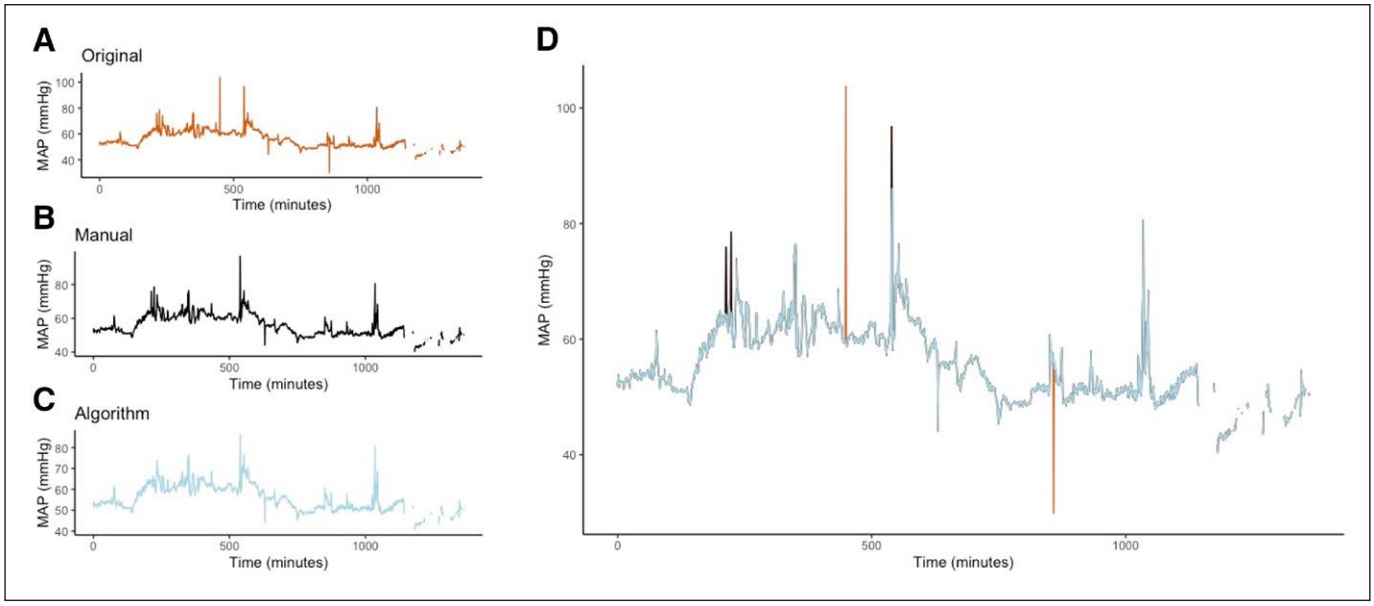

Figure 2. Visual representation of MAP data cleaning. Representative example comparing mean arterial pressure (MAP) traces before modification (**A**), after manual cleaning (**B**), and with algorithm cleaning (**C**). **D**, Overlapping raw and clean signals, where *orange lines* represent raw data identified as artifact by both manual and algorithmic cleaning, and *black lines* represent raw data identified as artifact by algorithm only.

TABLE 1.
Performance Metrics of the Algorithm With Various Modifications, Including Adjustment of the Delta Factor and the Addition of Upper and Lower Limits

	No Upper and Lower Thresholds			Lower (> 30 mm Hg) and Upper (< 200 mm Hg) Thresholds		
Delta	Sensitivity (%)	Specificity (%)	MCC	Sensitivity (%)	Specificity (%)	MCC
1× deltas	92.31	98.75	0.39	97.48	98.59	0.40
0.4× deltas	94.96[a]	98.02[a]	0.34	97.48	98.02	0.35
1.3× deltas	93.28[b]	98.68[b]	0.40	97.48[a]	98.68[a]	0.41
2× deltas	91.60	98.72	0.40	96.64[b]	98.72[b]	0.42
4× deltas	86.55	98.72	0.37	95.80	98.72	0.41

MCC = Matthew correlation coefficient.
[a]Sensitivity and specificity optimized.
[b]MCC optimized.

TABLE 2.
Characterizing Artifactual Data Identified by Human and Algorithm

Classification	Artifact (Human)	Nonartifact (Human)	Total
Artifact (algorithm)	116	537	653
Nonartifact (algorithm)	3	40,142	40,145
Total	119	40,679	40,798

2 × 2 table using the optimized algorithm with the addition of mean arterial pressure limits (> 30 and < 200 mm Hg), where delta factor = 1.3×.

identifying between 75% and 100% of artifacts manually cleaned. The majority of false positives occurred in patients 3 and 13, as shown in **Figure S3** (http://links.lww.com/CCX/B105).

Artifacts identified by the optimized algorithm were further analyzed based on how the algorithm dealt with the anomalous data. Of the 653 artifacts identified in the optimized algorithm, 91 (14%) were removed, whereas 562 (86%) were altered based on SBP and DBP data. Of the 562 modified data points, 497 (88%) were false positives, and 65 (12%) were true positives, where false positives represent data only flagged as an artifact via algorithm, whereas true positives represent

TABLE 3.
Artifacts Grouped by Patient Using the Optimized Algorithm Parameters in Table 2

Patient	True Artifacts (*n*)	Algorithm "Artifacts" (*n*)	Coverage (*n* [%])
1	1	6	1 (100)
2	11	27	11 (100)
3	16	343	16 (100)
4	3	11	3 (100)
5	1	19	1 (100)
6	20	28	18 (90)
7	2	5	2 (100)
8	4	10	3 (75)
9	2	5	2 (100)
10	2	4	2 (100)
11	4	15	4 (100)
12	9	16	9 (100)
13	41	154	41 (100)
14	3	6	3 (100)
15	0	4	-

Coverage represents the number of true artifacts also identified by the algorithm. Manual cleaning resulted in the identification and removal of 119 artifacts across 14 patients, whereas algorithm cleaning resulted in the identification and removal or modification of 653 artifacts across all patients.

data flagged as an artifact by both the algorithm and manual cleaning. Data points modified by the algorithm were parsed based on whether they were true positives (orange) or false positives (black) (**Fig. 3*A***). Agreement between human and algorithm occurred at the upper and lower extremes of MAP values, whereas the majority of false positives fell within the physiologically plausible MAP range. The distribution of MAP modifications is shown in **Figure 3*B***. Modifications ranged from 5 to 69 mm Hg, with approximately 50% of false positive modifications being less than 12 mm Hg.

DISCUSSION

As research and clinical practice in the ICU moves toward precision medicine, the analysis of high-frequency vital sign data is becoming increasingly important. Adequate tools to identify artifacts are imperative to ensure reliable data analysis. There is no consensus on the best tool to clean arterial line data in the ICU; therefore, we assessed the use of an intraoperative arterial line error-checking algorithm to detect artifacts in critically ill patients during the first 72 hours of ICU admission. Our goals were to determine whether the algorithm would provide a viable method for cleaning numeric data obtained from ICU arterial lines and whether it could be optimized for a critically ill population. Ultimately, we found that the Du et al (7) algorithm detected artifacts with high accuracy.

We demonstrated high sensitivity and specificity, and moderate MCC of the algorithm compared with manual cleaning. The unmodified algorithm produced an MCC of 0.39, with a sensitivity of 92.3% and specificity of 98.8%. Modifying the delta values by 0.4-fold led to an improved accuracy. However, the algorithm required stringent criteria of 2.4, 3.2, and 2 mm Hg for MAP, SBP, and DBP, respectively; these fluctuations are small and likely to result in the removal of true physiologic fluctuations. Further, several MAP values were below 30 mm Hg. The addition of upper and lower limits to MAP values led to higher accuracy and MCC, where a delta factor of 1.3 resulted in optimal sensitivity of 97.5% and specificity of 98.7% and MCC of 0.41. Therefore, a delta value of 1.3, along with the addition of greater than 30 mm Hg and less than 200 mm Hg thresholds, represents the ideal application of the algorithm for critically ill patients.

Despite high accuracy with each algorithm iteration, MCC was low to moderate based on the parameters used. This is in part attributed to the large number of false positives; five times more data points were deemed artifactual by algorithm than manual review. It is difficult to assess whether these modifications are warranted without live annotation. However, the majority of false positives were not removed from the data set, but rather modified via linear interpolation. In plotting the true and false positives, we showed that the majority of false positives existed within the physiologic MAP range and that modifications were commonly of a small magnitude (>50% of changes were <12 mm Hg). These extensive modifications were mostly confined to two patients, in which MAP data was noisy. Therefore, these false positives may represent artifactual data that are missed with manual review because the fluctuations appear physiologic. Overall, the majority of true artifacts were

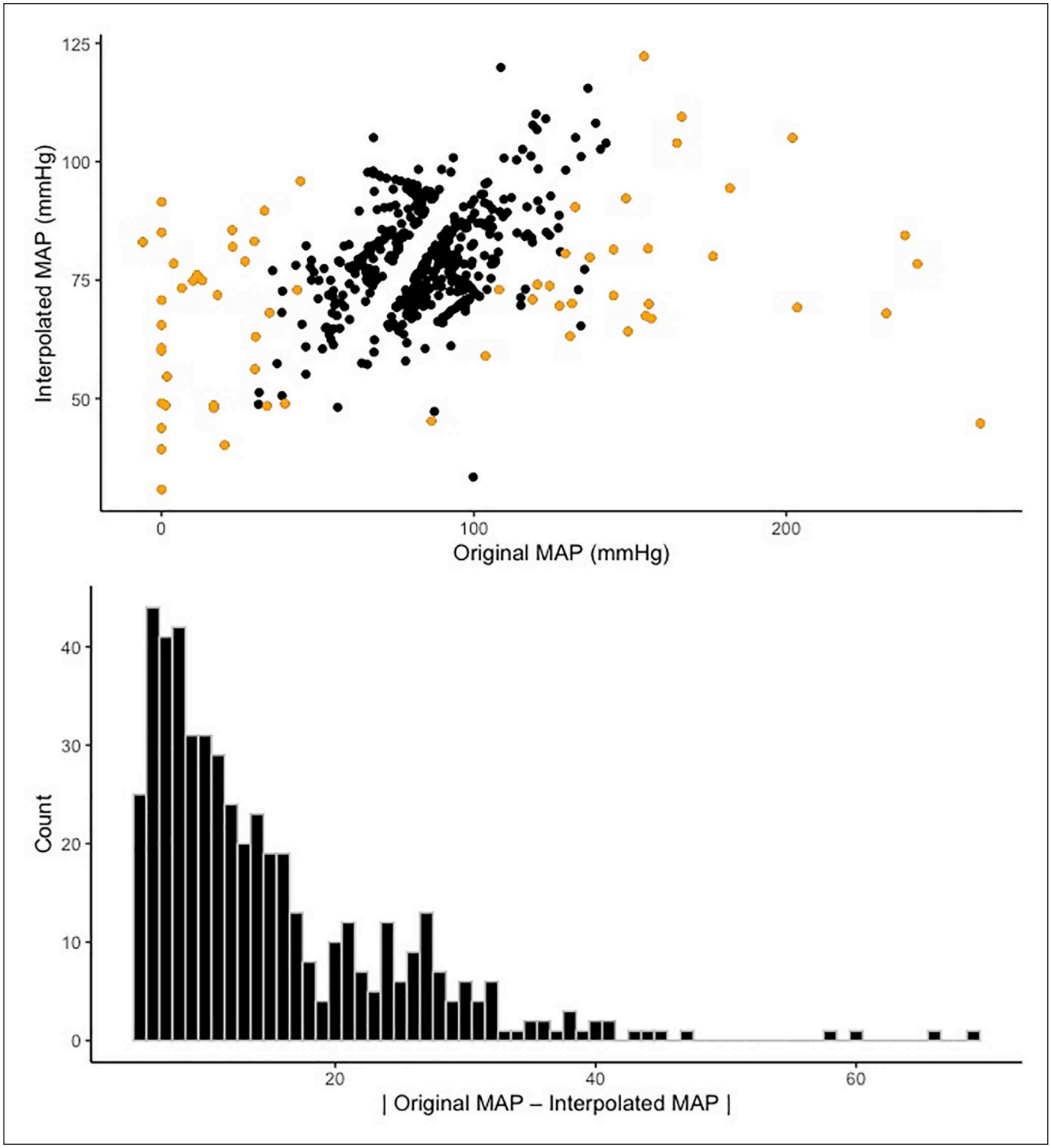

Figure 3. Characterizing data modified by the algorithm across patients. **A**, Visualization of mean arterial pressure (MAP) data that were modified by the optimized algorithm and the magnitude of change from the original values, where *orange* = true positives, *black* = false positives. **B**, Histogram of the magnitude of MAP.

either removed or modified with the algorithm, providing evidence that the algorithm is able to identify and remove erroneous data that could interfere with data analysis.

There are several limitations to this study. First, we encountered difficulties when there were long stretches of missing data for an individual patient. In these circumstances, the algorithm attempted to interpolate

and fill in data erroneously. Therefore, we incorporated an automated removal of gaps in data after the algorithm was run as a part of our modified algorithm. Further, MAP was manually cleaned without integrating associated SBP and DBP data. Manual cleaning incorporating SBP and DBP is possible, but the process is cumbersome, especially with high volumes of data (18). Therefore, the process undertaken is more likely representative of the data cleaning process in real-world conditions. Finally, we modified all three deltas by the same factor for each iteration. The algorithm could be further optimized by testing different combinations of modifications to individual MAP, SBP, and DBP deltas. However, we were able to achieve adequate results without this step.

Although algorithms to clean invasive arterial blood pressure data exist (**Table S3**, http://links.lww.com/CCX/B105), they have several limitations leading to barriers in implementation (6, 21). First, most rely on the input of waveform data. Logic-based algorithms to clean arterial line waveforms have performed well with sensitivities and specificities ranging from 94% to 100% and 84% to 98%, respectively (4, 9–11). Machine learning techniques have also been employed on waveform data with similar success (2, 6, 8). Son et al (6) used a deep belief network on arterial waveforms from traumatic brain injury patients with greater than 20 continuous hours of recording. The optimized algorithm had a net prediction rate of 95.9% and contained five hidden layers with two to 128 hidden units within each layer. Waveform data is high dimensional, making it an attractive source for signal cleaning even if numeric data is ultimately used for analyses (6). The high-frequency data provide a wealth of information for training to produce cleaning algorithms with high performance. Despite high accuracy of these algorithms, MAP waveforms can be difficult to obtain and analyze as data is large, cumbersome, and costly to store and maintain (22). Further, deep belief networks contain hidden layers with weighted edges and nonlinear transformations between input and output variables that lead to difficulty in interpretating what constitutes an erroneous data point (23). Therefore, algorithms that can be applied to numeric data may provide an avenue for cleaning that can be implemented with greater ease and increase interpretability.

Two groups have developed artifact identification methods based on numeric data. One algorithm used a statistical time-series approach with low-order autoregressive models and phase space models on arterial time-averaged 1/60-Hz data (13). They concluded that both required an experienced user with in-depth knowledge of the statistical techniques employed and were too sensitive, and therefore not feasible for bedside use. Another used logic-based rules adapted from fetal heart rate data cleaning, but only examined systolic and diastolic traces (12). Further, algorithm performance was based on the ability to predict hospital death after artifact removal rather than direct comparison with an alternate cleaning strategy. Ultimately, currently available numeric cleaning algorithms have not been sufficiently characterized and validated in critically ill adults.

Another major barrier to implementation is that many algorithms are not made publicly available. Of the nine ICU algorithms discussed above, only one made code accessible (4). Despite the existence of invasive blood pressure signal cleaning methods, researchers have resorted to more simplistic approaches to MAP cleaning, removing artifacts by employing upper and lower physiologic limits (24, 25). Although in our study the addition of upper and lower limits improved algorithm performance, 54% of artifacts fell within the defined physiologic range. Therefore, a majority of erroneous data points would escape detection using this approach.

The present algorithm requires minute-to-minute arterial blood pressure data, with transparent and modifiable equations that can be easily reproduced and/or modified to fit the population analyzed. The use of this algorithm can also reduce data loss. The majority (86%) of the artifacts identified in this data set were modified rather than deleted, representing data that would have been lost if cleaned manually. Further, this study validated the algorithm across three different centers. This is one of the first algorithms that relies solely on minute-averaged invasive arterial line data that have been validated in the ICU. Overall, the algorithm is easy to use, quick, and simple, making it a good candidate for artifact cleaning in ICU research and for use at the bedside.

CONCLUSIONS

The error-checking algorithm created by Du et al (7) provides a reasonable approach for cleaning arterial

line data from the ICU, as it was easy to employ and performed very well across all modifications. When optimized, the error-checking algorithm had high accuracy and moderate MCC in detecting arterial line blood pressure artifacts compared with manual data cleaning. Given the growing use of artificial intelligence in critical care research, methods to validate the quality of high-frequency data is important to optimize algorithm performance and prevent spurious associations based on artifactual data. Comparing algorithm performance to live annotation may be useful to further validate the algorithm. Future work incorporating the error-checking algorithm in real time is warranted, as well as quantifying the clinical impact of MAP cleaning on future data analysis.

ACKNOWLEDGMENTS

We thank the sites who took part in the data collection for this article, including site investigators Dr. Ian Ball, Dr. Shane English, Dr. Marat Slessarev, as well as research coordinators Tracy Boyd, Eileen Campbell, Miranda Hunt, Athena Ovsenek, Rebecca Porteous, and Irene Watpool.

1 *Centre for Neuroscience Studies, Queen's University, Kingston, ON, Canada.*

2 *Department of Critical Care Medicine, Queen's University, Kingston, ON, Canada.*

3 *Department of Medicine, Queen's University, Kingston, ON, Canada.*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (http://journals.lww.com/ccejournal).

Ms. Khan participated in study conception, data analysis, and article preparation. Dr. Maslove and Boyd participated in study conception, data review, and article preparation.

The remaining authors have disclosed that they do not have any potential conflicts of interest.

For information regarding this article, E-mail: gordon.boyd@kingstonhsc.ca

REFERENCES

1. Maslove DM, Elbers PWG, Clermont G: Artificial intelligence in telemetry: What clinicians should know. *Intensive Care Med* 2021; 47:150–153
2. Kim H, Lee SB, Son Y, et al: Hemodynamic instability and cardiovascular events after traumatic brain injury predict outcome after artifact removal with deep belief network analysis. *J Neurosurg Anesthesiol* 2018; 30:347–353
3. Takla G, Petre JH, Doyle DJ, et al: The problem of artifacts in patient monitor data during surgery: A clinical and methodological review. *Anesth Analg* 2006; 103:1196–1204
4. Li Q, Mark RG, Clifford GD: Artificial arterial blood pressure artifact models and an evaluation of a robust blood pressure and heart rate estimator. *Biomed Eng Online* 2009; 8:1–15
5. Simpao AF, Nelson O, Ahumada LM: Automated anesthesia artifact analysis: Can machines be trained to take out the garbage? *J Clin Monit Comput* 2021; 35:225–227
6. Son Y, Lee SB, Kim H, et al: Automated artifact elimination of physiological signals using a deep belief network: An application for continuously measured arterial blood pressure waveforms. *Inf Sci (Ny)* 2018; 456:145–158
7. Du CH, Glick D, Tung A: Error-checking intraoperative arterial line blood pressures. *J Clin Monit Comput* 2019; 33:407–412
8. Lee SB, Kim H, Kim YT, et al: Artifact removal from neurophysiological signals: Impact on intracranial and arterial pressure monitoring in traumatic brain injury. *J Neurosurg* 2019; 132:1952–1960
9. Sun JX, Reisner AT, Mark RG: A signal abnormality index for arterial blood pressure waveforms. *2006 Comput Cardiol* 2006; 33:13–16
10. Zhang P, Liu J, Wu X, et al: A novel feature extraction method for signal quality assessment of arterial blood pressure for monitoring cerebral autoregulation. *2010 4th Int Conf Bioinforma Biomed Eng* 2010; 82:21–24
11. Zong W, Moody GB, Mark RG: Reduction of false arterial blood pressure alarms using signal quality assessment and relationships between the electrocardiogram and arterial blood pressure. *Med Biol Eng Comput* 2004; 42:698–706
12. Cao H, Norris P, Ozdas A, et al: A simple non-physiological artifact filter for invasive arterial blood pressure monitoring: A study of 1852 trauma ICU patients. *Annu Int Conf IEEE Eng Med Biol Soc* 2006;2006:1417–1420
13. Imhoff M, Bauer M, Gather U, et al: Statistical pattern detection in univariate time series of intensive care on-line monitoring data. *Intensive Care Med* 1998; 24:1305–1314
14. Romagnoli S, Ricci Z, Quattrone D, et al: Accuracy of invasive arterial pressure monitoring in cardiovascular patients: An observational study. *Crit Care* 2014; 18:4–8
15. Wood MD, Khan J, Lee KFH, et al: Assessing the relationship between near-infrared spectroscopy-derived regional cerebral oxygenation and neurological dysfunction in critically ill adults: A prospective observational multicentre protocol on behalf of the Canadian Critical Care Trials Group. *BMJ Open* 2019; 9:e029189
16. Chicco D, Tötsch N, Jurman G: The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Min* 2021; 14:1–22
17. Chicco D, Warrens MJ, Jurman G: The Matthews Correlation Coefficient (MCC) is more informative than Cohen's kappa and Brier score in binary classification assessment. *IEEE Access* 2021; 9:78368–78381

18. Pasma W, Wesselink EM, van Buuren S, et al: Artifacts annotations in anesthesia blood pressure data by man and machine. *J Clin Monit Comput* 2021; 35:259–267
19. Delgado R, Tibau X-A: Why Cohen's kappa should be avoided as performance measure in classification. *PLoS One* 2019; 14:e0222916
20. R Core Team: R: A Language and Environment for Statistical Computing. 2018. https://www.R-project.org, accessed March 3, 2021
21. Nizami S, Green JR, McGregor C: Implementation of artifact detection in critical care: A methodological review. *IEEE Rev Biomed Eng* 2013; 6:127–142
22. Goodwin AJ, Eytan D, Greer RW, et al: A practical approach to storage and retrieval of high-frequency physiological signals. *Physiol Meas* 2020; 41:035008
23. Shahid N, Rappon T, Berta W: Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PLoS One* 2019; 14:e02123561–e02123522
24. Hyland SL, Faltys M, Hüser M, et al: Early prediction of circulatory failure in the intensive care unit using machine learning. *Nat Med* 2020; 26:364–373
25. Griesdale DEG, Sekhon MS, Wood MD, et al; Cerebral Oximetry to Assess Cerebral Autoregulation in Hypoxic-Ischemic Brain Injury (CONCEPT) Research Group, on behalf of the Canadian Critical Care Trials Group: Near-infrared spectroscopy to assess cerebral autoregulation and optimal mean arterial pressure in patients with hypoxic-ischemic brain injury: A prospective multicenter feasibility study. *Crit Care Explor* 2020; 2:e0217