

ANESTHESIOLOGY

Artifact Processing Methods Influence on Intraoperative Hypotension Quantification and Outcome Effect Estimates

Wietze Pasma, D.V.M., Linda M. Peelen, Ph.D.,
Stef van Buuren, Ph.D., Wilton A. van Klei, M.D., Ph.D.,
Jurgen C. de Graaff, M.D., Ph.D.

ANESTHESIOLOGY 2020; 132:723–37

EDITOR'S PERSPECTIVE

What We Already Know about This Topic

- Electronically collected data are increasingly used for clinical research, but include artifacts and other errors
- How best to filter electronically recorded intraoperative blood pressure remains unknown

What This Article Tells Us That Is New

- The authors identified 38 papers describing blood pressure artifact-handling methods, and applied eight methods to nearly 3,000 anesthetics
- The amount of observed hypotension at various thresholds varied considerably depending on the filtration method
- Investigators need to carefully consider artifact handling, and fully describe their methodology

Physiologic data that is automatically collected are widely used for medical record keeping and clinical research; yet, not every stored vital sign in these data represents the actual physiologic state of the patient at the time of measurement. For example, when a person leans against a patient's blood pressure cuff, this causes artifacts in blood pressure data, *i.e.*, the value that is registered does not equal the true blood pressure of the patient. Similarly, electrocautery disturbs electrocardiogram readings and causes artifacts

ABSTRACT

Background: Physiologic data that is automatically collected during anesthesia is widely used for medical record keeping and clinical research. These data contain artifacts, which are not relevant in clinical care, but may influence research results. The aim of this study was to explore the effect of different methods of filtering and processing artifacts in anesthesiology data on study findings in order to demonstrate the importance of proper artifact filtering.

Methods: The authors performed a systematic literature search to identify artifact filtering methods. Subsequently, these methods were applied to the data of anesthesia procedures with invasive blood pressure monitoring. Different hypotension measures were calculated (*i.e.*, presence, duration, maximum deviation below threshold, and area under threshold) across different definitions (*i.e.*, thresholds for mean arterial pressure of 50, 60, 65, 70 mmHg). These were then used to estimate the association with postoperative myocardial injury.

Results: After screening 3,585 papers, the authors included 38 papers that reported artifact filtering methods. The authors applied eight of these methods to the data of 2,988 anesthesia procedures. The occurrence of hypotension (defined with a threshold of 50 mmHg) varied from 24% with a median filter of seven measurements to 55% without an artifact filtering method, and between 76 and 90% with a threshold of 65 mmHg. Standardized odds ratios for presence of hypotension ranged from 1.16 (95% CI, 1.07 to 1.26) to 1.24 (1.14 to 1.34) when hypotension was defined with a threshold of 50 mmHg. Similar variations in standardized odds ratios were found when applying methods to other hypotension measures and definitions.

Conclusions: The method of artifact filtering can have substantial effects on estimates of hypotension prevalence. The effect on the association between intraoperative hypotension and postoperative myocardial injury was relatively small. Nevertheless, the authors recommend that researchers carefully consider artifacts handling and report the methodology used.

(ANESTHESIOLOGY 2020; 132:723–37)

in heart rate data. These artifacts are not relevant in clinical care, as they are easily recognized and subsequently ignored, but may influence research results.^{1,2} Consequently, before using the physiologic data for research, one should consider how to process artifacts, and this should be reported when presenting the results of the study.^{1,3–5}

To make artifactual data more suitable for retrospective analyses, the researcher has several options, varying from manual data cleaning *via* simple filtering methods to more advanced methods. Methods will vary in their ability to identify or correct artifacts in the data.^{3,6} More importantly, different filter methods can lead to different study results. Previous studies have shown that artifact occurrence depends on several factors, which could lead to incorrect

Submitted for publication May 31, 2019. Accepted for publication December 17, 2019. Published online first on February 3, 2020. From the Medical Center Utrecht (W.P., L.M.P., W.A.v.K.), and the Department of Epidemiology, Julius Center for Health Sciences and Primary Care (L.M.P.), Utrecht University, Utrecht, The Netherlands; The Netherlands Organization for Applied Scientific Research and Department of Methodology and Statistics, Faculty of Social and Behavioural Sciences, University of Utrecht (S.v.B.), Utrecht, The Netherlands; and the Department of Anesthesiology, Erasmus MC – Sophia Children's Hospital, Rotterdam, The Netherlands (J.C.d.G.).

Copyright © 2020, the American Society of Anesthesiologists, Inc. All Rights Reserved. Anesthesiology 2020; 132:723–37. DOI: 10.1097/ALN.0000000000003131

estimation of associations in clinical research, *i.e.*, a form of misclassification bias.^{4,5,7,8} We hypothesized that these filter methods used on the data would have a significant influence on the findings of the study.

The aim of this study was to explore the effect of different methods of filtering and processing artifacts in anesthesiology data on study findings. In doing so, we aimed to demonstrate the importance of proper artifact filtering. To this aim, we performed a systematic literature search to identify artifact filtering methods actually being used in practice, and subsequently used these filters on an existing perioperative dataset. For this dataset, we selected surgical patients in whom both invasive blood pressure measurements and an example outcome (myocardial injury) were measured. We examined the effects of artifact filtering methods on the quantification of intraoperative hypotension measures and subsequently its effect on the association between intraoperative hypotension and postoperative myocardial injury. The determinant invasive blood pressure measurement was chosen because it has a known high incidence of artifacts and is often used in intraoperative hypotension research.^{2,5}

Materials and Methods

Systematic Search for Artifact Filtering Methods

To find relevant artifact filtering methods, we performed a systematic literature search. No formal review protocol was developed before beginning the review. We started the search with the following search query on PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/>) on June 8, 2018: (anesthesia [tiab] OR aims [tiab] OR intraoperative [tiab] OR “Monitoring, Intraoperative”[Mesh] OR “Anesthesia”[Mesh] OR “Monitoring, Physiologic”[Mesh] OR “Heart Rate”[Mesh] OR “Oximetry”[Mesh] OR “Blood Pressure”[Mesh] OR “Arterial Pressure”[Mesh] OR “blood pressure” [tiab] OR “arterial pressure” [tiab] OR “oxygen saturation” [tiab] OR “heart rate” [tiab]) AND (artifact [tiab] OR artifact [tiab] OR “Artifacts”[Mesh] OR “measurement error” [tiab]).

First, two researchers (W.P. and L.P.) screened every title for possible inclusion. Exclusion reasons were noted and categorized. Conflicts between the reasons of exclusion were reviewed and the first reviewer made a definitive choice based on the title or abstract. Papers were included when: (1) it was published after 2000; (2) it used vital signs data for anesthesia (oxygen saturation, blood pressure, or heart-rate); and (3) a process of artifact filtering was described. Papers about the development of a filtering method were included, as well as research papers in which a specific filtering method was applied and review papers that discussed different methods. We excluded case reports, letters to editors, editorials, and studies that considered nonhuman subjects. Apart from these, no other limits were placed on the type of study design. Papers in languages other than English were also not considered. We only considered papers that

came up with our search query, we did not review reference sections of papers to identify additional candidates. No effort was made to consider unpublished studies, conference abstracts, or proceedings.

After title screening, all papers selected by one or both of the reviewers underwent abstract screening and, if necessary, full-text screening using the aforementioned approach. Abstracts were extracted from PubMed and displayed in a review form (programmed in R shiny), accompanied by the title and a link to the manuscript. The final decision and exclusion reason were filled in using dropdown fields. The form also included the choices that both reviewers made regarding title screening. If it was unclear from the abstract text, if any artifact filtering methods were described in the paper, the manuscript was screened to make this decision. From the remaining papers, any method for artifact filtering was collected and categorized. The first author screened the abstracts and identified and classified the artifact filtering methods within the selected papers. The three artifact filtering method categories described in the results section were identified and refined during the review process. Some papers included multiple types or a combination of artifact filtering methods. Therefore, one paper could fall under several artifact filtering methods. We did not contact authors of original reports to clarify the methods that they used. The information in the included papers was sufficient to categorize the methods.

Cohort

To evaluate the influence of the obtained methods of artifact filtering on the intraoperative hypotension measures, we selected a group of patients from a prospectively-defined and previously-described cohort.⁹ In short, in this ongoing cohort patients are included if they are 60 yr of age or older and undergo intermediate- to high-risk noncardiac surgical procedures at the University Medical Center Utrecht. We selected only new procedures, *i.e.*, reoperations were not included. A first procedure for a patient was defined as a procedure with an available troponin measurement during the first 3 postoperative days and not preceded in the previous 365 days by another eligible procedure. We further refined the cohort by selecting patients in whom sufficient postinduction invasive blood pressure measurements were available, *i.e.*, we excluded procedures without a known time of induction and without at least 15 invasive blood pressure measurements after the start of induction throughout the entire procedure. For the current study, procedures from January 1, 2011 to December 31, 2014 were selected. The patient cohort selection and further analysis described in the current study were designed for illustration purposes only. For collection of these data, the local ethics committees approved the protocol and waived the need for informed consent (University Medical Center Utrecht Medical Research Ethics Committee, protocol no. 18-261).

We collected invasive mean blood pressure measurements from our anesthesia information management

system (Anstat, version 2.0.4, 2015; Carepoint, Ede, The Netherlands). Invasive blood pressure is measured during anesthesia with an IntelliVue monitoring system (type MP70, X2 multimeasurement module; Philips, Germany) with a built-in filter for artifacts (*i.e.*, a 12-Hz filter is applied) that reduces resonant effects of the tubing system. Our anesthesia information management system stores one value each minute in its database, which is the median of 12 consecutive measurements with a 5-s interval supplied by the anesthesia monitor. We determined the time of induction with an algorithm that was previously published.¹⁰ Blood pressure measurements before time of induction were excluded from analysis.

Artifact Filtering Methods

From the artifact filtering methods resulting from the systematic search, we included methods into our comparison based on their frequency of use (frequently used methods were more likely to be included), the need for annotated data for the purpose of algorithm training, and whether they are applicable to anesthesia information management system data (*i.e.*, minute to minute vital sign data). In addition, as many of the artifact filtering methods contain cut-off values, we varied such cut-off values to assess differences within the same type of filter. For example, if filters are based on distribution, we used “more than [two] times the interquartile range” and “more than [three] times the interquartile range.” A detailed description of the filtering methods compared can be found in the Results section, after the results of the systematic search.

Outcome

The primary outcome was the incidence and severity of intraoperative hypotension when applying different artifact filtering methods found in the review step. We considered four measures to quantify hypotension; the presence of hypotension, the total duration of hypotension, the total area under threshold, and the maximum deviation below the threshold if hypotension was present. Hypotension was quantified by an algorithm, which took every measurement and made a linear interpolation between every data point. The interpolation was performed between subsequent measurements, regardless of the period between both measurements. The area between the threshold and this blood pressure curve was then quantified.^{11,12} Each time the blood pressure curve dropped under a specified threshold, this was identified as the start of an hypotension episode, and as soon as the blood pressure went over the threshold again, this point in time was marked as the end of that episode. The total area under the threshold is the summation of all episodes. We used four different thresholds (mean blood pressure of 50, 60, 65, and 70 mmHg) to explore whether artifact filters can have a different effect under different hypotension definitions.¹³ Altogether we compared

the artifact filtering methods for four different hypotension measures and four different thresholds.

The secondary outcome was the association between hypotension and postoperative myocardial injury, when applying different artifact methods. Postoperative myocardial injury was defined as a postoperative elevation of troponin I. According to local protocol, troponin was measured routinely for the first 3 postoperative days. Troponin elevation was present when at least one of the postoperative troponin values was more than a predefined clinical cut-off value of greater than 60 ng/l.⁹

Statistical Analysis

For each of the artifact filtering methods, we first describe the incidence or extent of hypotension that would result from applying the method to the dataset. We describe intraoperative hypotension as a proportion, and use median values with interquartile ranges for the continuous hypotension measures (total area under threshold, total duration, and maximum deviation from threshold). This is done for each of the thresholds for intraoperative hypotension.

To allow for a comparison of the effect estimate in the association between hypotension and postoperative myocardial injury, we standardized hypotension measures per measure type and threshold before estimating the effect of hypotension on postoperative myocardial injury. In order to do so, we first used a logarithmic transformation for total duration and total area under threshold, since these data were skewed to the right. After this transformation, the hypotension measures were standardized by calculating a z-score. With these standardized hypotension measures as the only explaining variable, we fitted a model for postoperative myocardial injury using logistic regression analysis. This was done again for each combination of artifact filtering method, hypotension measure, and hypotension threshold. We expressed effect estimates of hypotension measures as standardized odds ratios with CIs, using a level of significance of $\alpha=0.05$.

We collected and deidentified the data from the enterprise data warehouse with SAS software (Version 9.4, SAS Institute Inc). After deidentification, the data was further processed and analyzed in R (R Foundation for Statistical Computing, Vienna, Austria; <https://www.R-project.org>, R version 3.5.1 [2018-07-02]).

Results

Systematic Search for Artifact Filtering Methods

The systematic search for artifact filtering methods resulted in 3,585 papers. After abstract and full-text screening, 3,300 papers were excluded and 285 papers mentioning artifact filtering methods remained—of which 247 papers described methods used on high resolution data such as electrocardiogram, photoplethysmogram, and arterial blood pressure

waveform data (fig. 1). These methods often rely on the repeatable patterns in these signals, and therefore are not applicable to anesthesia information management system data which is used for clinical research. In addition, a number

of filters on these high-density data rely on additional sensor data such as accelerometer data, which are not commonly collected in an anesthesia information management system database. After excluding this group, 38 papers remained that

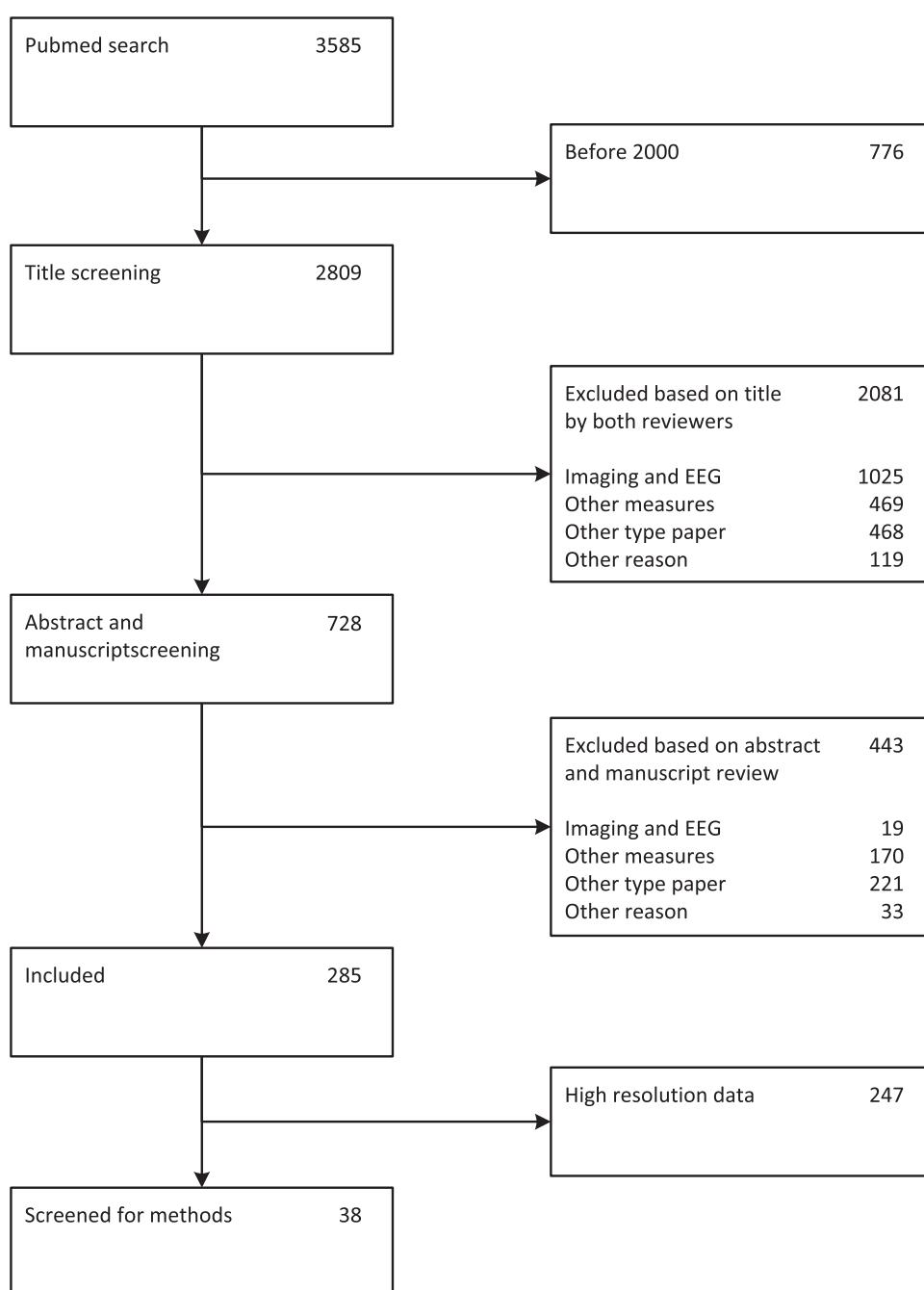


Fig. 1. Literature review flowchart. Flow diagram illustrating systematic search and review process, starting with 3,585 papers. Eventually 38 papers remained for identification of artifact filtering methods. Papers about imaging included papers of magnetic resonance imaging, computed tomography echo, and other imaging techniques. Papers about other measures include nonstandard vital sign monitors such as camera-, textile-, or wristwatch-based monitors. Other types of papers were methodologic papers or original research in which physiological data processing was not used or mentioned. Other reasons for exclusion were for example papers on pathology or ophthalmology. EEG, electroencephalography.

report methods applicable to 1-min resolution anesthesia information management system data (fig. 1).

We divided the artifact filtering methods into three basic categories. The first category consisted of methods that identify measurements as an artifact, such as limit methods that use biologic plausible blood pressure boundaries (mentioned in 27 papers). The second group consisted of methods that can alter the vital sign signal by applying a filter to extract the true signal of interest from the raw artifactual data (mentioned in 16 papers), for example a method that calculates the median over neighboring values. This new acquired data point is then used to calculate hypotension. The third category contained methods that use model strategies that take artifacts or measurement errors into account (mentioned in six papers). This can be done by applying a model to the data, such as a spline function of the blood

pressure, which is subsequently fit into a model instead of using the actual data. All 38 papers included, and the methodologies found in these papers are listed in figure 2.

Based on the results of the systematic search we chose three methods to handle artifacts in invasive blood pressure data: a limit filter, a moving median filter, and a likelihood filter based on median and interquartile range. For each of the methods we used two or three different settings for the parameters. Altogether, we compared eight different filtering approaches (including applying no artifact filtering method), which are described in detail in figure 3.

Cohort and Hypotension Measures

We included 2,988 anesthetic procedures in our analysis (fig. 4), of which the baseline characteristics are listed in

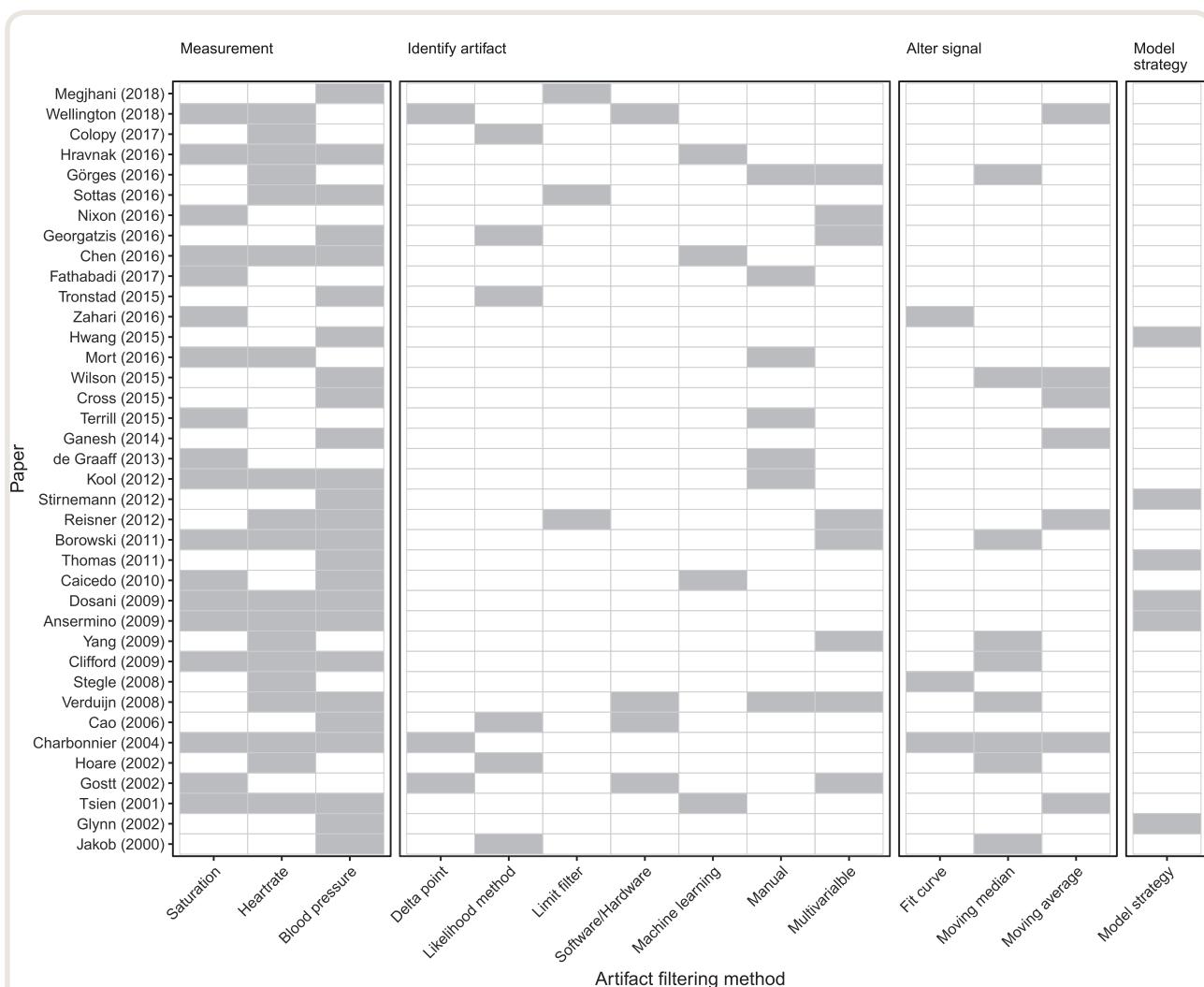


Fig. 2. Literature overview artifact methods. Methods within papers (*vertical*) found with systematic search (38 papers). The first panel describes which measurements are described in the paper. The other three panels represent the method groups that are described in this paper. Each method (*horizontal*) that was described in the paper was marked. One paper can contain more than one method.

First, **no filter** was applied, *i.e.*, all measurements after time of induction were used to calculate hypotension.

Second, we used a **limits filter** by excluding invasive blood pressure readings with a corresponding pulse pressure lower than 20 or higher than 150. In addition, the mean blood pressure had to be between 40 and 160 mmHg.¹

Third, we applied a **moving median** to the blood pressure curve that replaced every blood pressure with the median value of n readings, n being the window size, of which the current reading is the center measurement.²

$$\chi'_i = \text{med} [\chi_{i-(n-1)/2}, \dots, \chi_i, \dots, \chi_{i+(n-1)/2}]$$

Finally we applied a **likelihood filter** based on a median and variation within a time window. We divided the blood pressure signal into subsequent windows of 10 measurements. For each window the median and IQR were calculated. If a blood pressure deviated more than $n \times \text{IQR}$, but at least 10 mmHg from the median, the value was identified as an artifact and was removed from the signal.³

Fig. 3. List of artifact filtering methods applied to blood pressure data. All blood pressure measurements were used to quantify hypotension. Four artifact filtering methods were chosen, which were applied on blood pressure data. *As a variation, the lower limits for pulse pressure and mean blood pressure were varied, to 10 mmHg and 20 mmHg, respectively.^{17–19} †Two different values for n (5 and 7) were used.^{3,20–27} ‡Three different values for n (1, 2, and 3) were used.^{15,16} IQR, interquartile range.

table 1. In this cohort the occurrence of postoperative myocardial injury was 807 (27%), and 1,563 procedures were classified as high-risk surgeries.

Table 2 describes the estimated values for the different hypotension measures when using different hypotension thresholds for each of the artifact filtering methods. Different artifact filtering methods resulted in different estimates of the occurrence of hypotension. For example, when hypotension was defined as mean blood pressure below 50 mmHg, the occurrence varied from 24% with a moving median filter of seven measurements to 55% without an artifact filtering method. When a threshold of 65 mmHg was used, the presence of hypotension varied between 76 and 90%. Similarly, other hypotension measures varied among different artifact filtering methods. For example, within the definition of hypotension as blood pressure less than 65 mmHg, the total area under threshold varied between 81

mmHg \times min (interquartile range, 3 to 311) and 129 mmHg \times min (interquartile range, 25 to 383), the maximum deviation below threshold varied between 8 mmHg (interquartile range, 1 to 15) and 17 mmHg (interquartile range, 9 to 31), and the total duration of hypotension between 18 min (interquartile range, 2 to 52) and 22 min (interquartile range, 6 to 58).

Association between Hypotension and Postoperative Myocardial Injury

Figure 4 depicts the standardized odds ratios for the association between intraoperative hypotension and postoperative myocardial injury when using different artifact filtering methods for blood pressure. Again, this is shown for different hypotension measures and different thresholds. Standardized odds ratios ranged from 1.16 (95% CI, 1.07 to 1.26) for not using an artifact filter method to 1.24 (95% CI, 1.14 to 1.34) for using a

Table 1. Baseline Data Included Procedures (n = 2,988)

	n or Median	Percentage or Interquartile Range
Postoperative myocardial injury (%)	807	(27.0%)
Number of blood pressure measurements [IQR]	203	[143–297]
Age at time of surgery, yr [IQR]	70.3	[65.6–76.5]
Male (%)	1861	(62.3%)
ASA classification (%)		
I	266	(8.9%)
II	1,635	(54.7%)
III	978	(32.7%)
IV–V	109	(3.6%)
Surgical specialty (%)		
General	750	(25.1%)
Orthopedic	142	(4.8%)
Urological	106	(3.5%)
ENT and dental	287	(9.6%)
Vascular	848	(28.4%)
Neurological	716	(24.0%)
Gynecological	53	(1.8%)
Other	86	(2.9%)
Emergency surgery (%)	579	(19.4%)
High risk surgery (%)	1,563	(52.3%)
Ischemic heart disease (%)	470	(15.7%)
History of cerebrovascular disease (%)	710	(23.8%)
Chronic heart failure (%)	113	(3.8%)
Peripheral vascular disease (%)	446	(14.9%)
Hypertension (%)	1,670	(55.9%)
Use of beta blockers (%)	1,041	(34.8%)
Use of calcium antagonists (%)	587	(19.6%)
Use of ACE inhibitors or AT2 blockers (%)	1,223	(40.9%)
Use of statins (%)	1,345	(45.0%)
Use of insulin/Insulin dependency (%)	200	(6.7%)

Data are presented as number of procedures and percentage of total cohort, except age and number of blood pressure measurements, which are presented as median and IQR. ACE, angiotensin-converting enzyme; ASA, American Society of Anesthesiologists; AT2, angiotensin II receptor; ENT, ear, nose, and throat; IQR, interquartile range.

limit filter (pulse pressure greater than 20; mean blood pressure greater than 40) within the 50-mmHg threshold for presence of hypotension. For the 65-mmHg thresholds, estimates ranged from 1.14 (95% CI, 1.04 to 1.26) for using a limit filter (pulse pressure greater than 20; mean blood pressure greater than 40) to 1.21 (95% CI, 1.11 to 1.33) for using a likelihood filter method ($1 \times \text{IQR}$ above median). Within the 65-mmHg threshold, the odds ratios for maximum deviation ranged from 1.20 (95% CI, 1.10 to 1.31) to 1.30 (95% CI, 1.19 to 1.42); total area under threshold, 1.27 (95% CI, 1.18 to 1.37) to 1.36 (95% CI, 1.24 to 1.49); and total duration, 1.26 (95% CI, 1.17 to 1.37) to 1.32 (95% CI, 1.21 to 1.44).

Discussion

Effect of Artifact Filtering on Hypotension Measures

Different methods for processing artifacts result in different estimates of hypotension measures. We saw a change in the number of patients who were identified as having intraoperative hypotension when different artifact filtering methods were applied. The intraoperative hypotension

ranged from 24 to 55% when hypotension was defined as mean blood pressure less than 50 mmHg and from 76 to 90% when the defined threshold was 65 mmHg. Although we found this clear effect on hypotension measures, the resulting effect on the association between determinant and outcome (*i.e.*, intraoperative hypotension and postoperative myocardial injury) was less profound than expected.

From previous studies we learned that the occurrence of artifacts in physiologic data are related to patient and procedure characteristics,^{2,5} hence we expected changes in estimates when these artifacts were dealt with differently. In the current analyses, removing artifactual data did indeed change the associations, but overall these changes were smaller than the variation in estimates due to the choice of hypotension threshold or choice of hypotension quantity (fig. 5). We found filtering methods to have less of an effect on the association between duration and outcome than between depth of hypotension and outcome. Filtering methods for artifacts are designed to correct extreme values or outliers in the data, resulting in adjustments in the depth of hypotension domain rather than the duration of hypotension.

Of note, the odds ratios for hypotension defined according to 50- and 70-mmHg thresholds were low, in contrast to those for 60- and 65-mmHg thresholds. This is explained by the low or high intraoperative hypotension when a low or a high threshold is chosen, respectively. In both situations, the variance of the hypotension measure in the data will be low, compared to choosing an intermediate threshold (*i.e.*, 60 or 65 mmHg). With a higher variance, the distinctive power of the data used to model a given outcome, increases.

We studied the association between intraoperative hypotension and postoperative myocardial injury, because there is an extensive body of clinical research on this topic and on the methodology of analyzing hypotension and outcome due to hypotension. Significant effort has been put into determining a consensus on the hypotension threshold that should be used and on the kind of hypotension measure that should be analyzed. From our analysis we deduce that these choices are probably more important (as they yield more variation) than the choice of artifact filtering method. Nevertheless, we think that artifact filtering methods are an additional source of variance in studying the relationship between intraoperative hypotension and outcomes, which was already muddled with inconsistencies in methodology.^{10,11,13,14}

Systematic Search

We performed a systematic search to make sure that the artifact filtering methods we chose corresponded to the manner in which researchers handle artifacts. We divided the methods we found in three categories. The first category includes methods that identify artifacts. These methods include a bandwidth filter with a minimum and maximum allowed value for blood pressure or pulse pressure and a likelihood filter based on the distribution of the

Table 2. Overview of Hypotension Measures

Measure	Hypotension Definition	No Filter		Limits		Moving Median		Likelihood	
		PP > 20, mABP > 40	PP > 10, mABP > 20	Five-measure Window	Seven-measure Window	3xIQR	2xIQR	1xIQR	
Presence of hypoten-sion (yes/no)	50 60 65 70 70	1,654 (55%) 2,451 (82%) 2,699 (90%) 2,831 (95%) 2,831 (95%)	1,201 (40%) 2,288 (77%) 2,639 (89%) 2,793 (94%) 2,793 (94%)	876 (29%) 1,979 (66%) 2,401 (80%) 2,652 (89%) 2,652 (89%)	705 (24%) 1,783 (60%) 2,278 (76%) 2,581 (86%) 2,581 (86%)	1,394 (47%) 2,337 (78%) 2,638 (88%) 2,794 (94%) 2,794 (94%)	1,284 (43%) 2,289 (77%) 2,614 (87%) 2,778 (93%) 2,778 (93%)	1,058 (35%) 2,155 (72%) 2,528 (85%) 2,758 (92%) 2,758 (92%)	
Maximum decrease (mmHg)	50 60 65 70 70	0 (0–4) 12 (4–26) 17 (9–31) 22 (14–36) 1 (0–24)	0 (0–7) 8 (1–14) 13 (6–19) 18 (11–24) 0 (0–4)	0 (0–2) 5 (0–12) 10 (3–17) 15 (8–22) 0 (0–5)	0 (0–0) 3 (0–10) 8 (1–15) 13 (6–20) 0 (0–0)	0 (0–8) 0 (0–0) 0 (0–0) 20 (12–28) 0 (0–14)	0 (0–6) 0 (0–6) 15 (7–23) 19 (11–26) 0 (0–11)	0 (0–4) 0 (0–4) 14 (6–21) 17 (10–24) 0 (0–5)	7 (0–14) 7 (0–16) 12 (5–19) 17 (10–24) 0 (0–5)
Total area under threshold (mmHg × min)	50 60 65 70 70	43 (4–153) 60 (16–250) 128 (25–383) 326 (86–805) 0 (0–4)	25 (0–104) 30 (1–119) 96 (16–291) 104 (18–316) 276 (68–705)	21 (0–121) 11 (0–103) 95 (10–344) 81 (3–311) 279 (58–761)	11 (0–146) 36 (1–146) 81 (3–311) 307 (76–794) 259 (44–734)	33 (0–41) 33 (0–41) 114 (18–364) 114 (18–364) 298 (72–785)	33 (0–41) 33 (0–41) 97 (11–342) 97 (11–342) 284 (62–765)		
Total duration (min)	50 60 65 70	0 (0–4) 8 (1–25) 22 (6–58) 48 (14–110)	0 (0–2) 7 (0–23) 20 (5–55) 47 (13–109)	0 (0–3) 6 (0–22) 19 (4–57) 45 (12–107)	0 (0–0) 5 (0–20) 18 (2–52) 44 (10–106)	0 (0–4) 0 (0–3) 0 (0–3) 0 (0–3)	0 (0–3) 8 (1–25) 21 (5–57) 48 (14–110)	0 (0–3) 8 (1–25) 21 (4–57) 47 (13–110)	

Intraoperative hypotension measures after application of different artifact filtering methods. Continuous data (maximum decrease below threshold, total AUT, and total duration) are presented as median with IQR, and binomial data (presence of hypotension) are presented as number of procedures and percentage of the total procedures. AUT, area under threshold; IQR, interquartile range; mABP, mean arterial blood pressure; PP, pulse pressure.

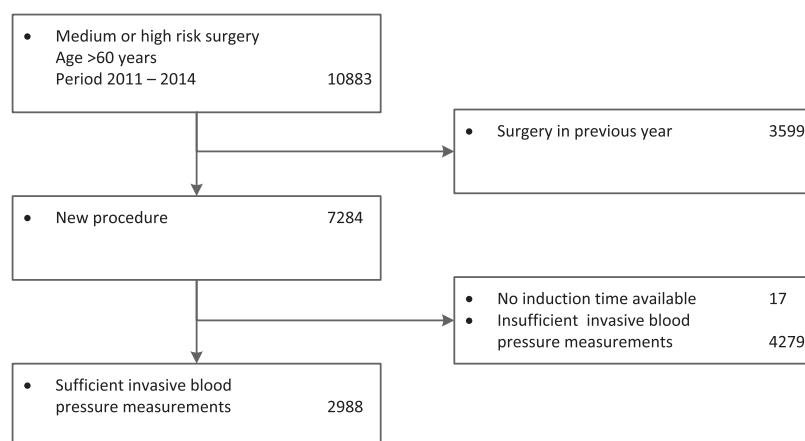


Fig. 4. Flowchart procedure inclusion. All procedures identified in the anesthetic information management system, and those included in our analyses. Included procedures were required to be medium- to high-risk procedures between 2011 and 2014 on patients who were 60 yr or older. Procedures should be new (no procedures on the same patient in the previous yr) and have sufficient (at least 15) invasive blood pressure measurements available after time of induction.^{2,3,15–27,30–53}

data, as used in this paper.^{15–19} More advanced methods exist (*e.g.*, models or machine learning algorithms that identify artifacts); although, the downfall of these more advanced methods is that they require a training set to train an algorithm, of which obtaining is labor intensive. When artifacts are identified and subsequently removed, there are several options that can be used to fill in the gaps. In the current paper, we used linear interpolation between existing data points,¹² but other methods are also available, such as last known value carried forward.

The second category is comprised of methods that replace the data signal with a new and refined signal. An example of these methods is a running median filter, as we used in this paper.^{3,20–27} Other possibilities are methods that fit more complex curves over the existing data. After applying the methods, a new dataset is created, and used in further analytical steps.

The last category consists of methods that replace the data signal with a function that is used directly in a statistical model. For example, spline functions are used instead of raw data in a model, or a joint model is constructed. This third type of method was not usable in this paper, because we required the processed blood pressure data to quantify hypotension in order to study the relation between hypotension and postoperative myocardial injury.

Despite our efforts to find artifact filtering methods in literature, the number of papers including methods is limited (38 papers) and the methods were quite heterogeneous. We looked for papers that mentioned artifacts in the title or abstract, which resulted in papers of which artifact filtering was an important element of the study (*e.g.*, the development of an artifact filter). How often these methods were used in

practice cannot be extrapolated from this systematic search. It should be noted that in contrast with the title screening, the final identification and classification of the artifact filtering methods was performed by one author only. The selection of methods applicable to anesthesia information management system data was done by the research team, but was not predefined in the review protocol.

Alternatively, we could have looked for the filters being applied in research practice, for example by searching for all papers studying intraoperative hypotension and outcomes, and then carefully studying the Methods sections. However, we expect that this would have resulted in a lower yield of artifact filtering methods. From personal experience as researchers in this field, we noticed that the methods for handling artifacts are typically not (extensively) reported.

Strengths and Limitations

Our study is one of the first to explore the impact of artifacts on clinical research using anesthesia information management system data by applying different artifact filtering methods on real clinical data. We focused on different artifact filter methods and how they influence estimated associations. These methods were identified with a systematic search. Not only did we vary artifact filtering methods, we also studied different hypotension measures and varying hypotension definitions to get a complete picture. Our detailed analysis places the issue of artifacts into perspective, and the reader can base the methodologies for filtering artifacts in future work on these findings.

One limitation of our study is that we have chosen only one type of measurement (blood pressure) and one type

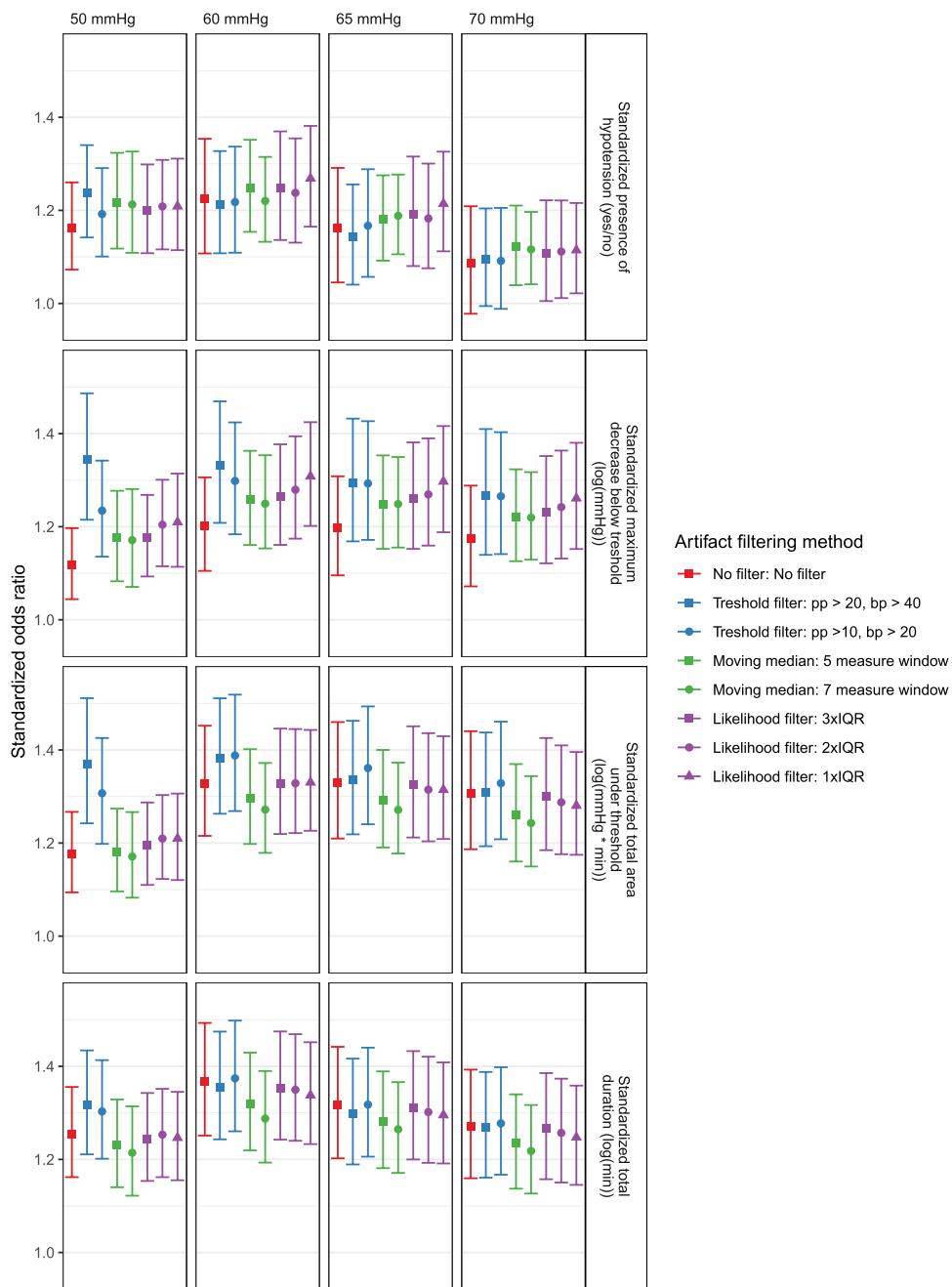


Fig. 5. Effect of artifact filtering methods on the association between intraoperative hypotension and postoperative myocardial injury. Standardized odds ratios for the association between intraoperative hypotension and postoperative myocardial injury and for different artifact filtering methods (*color and shapes*). Different hypotension measures (*vertical*) and hypotension definitions (*horizontal*) were compared.

of outcome. It is unclear to which extent we can generalize these findings to other physiologic data or research on other subjects. Second, we could not use every artifact filtering method found. More methods would have been applicable, if the anesthesia information management system data was more granular than data with 1-min intervals.

However, the vital signs in our anesthesia information management system are based on raw data measured at 5-s intervals, of which the median per min is stored. This meant the data we used was already processed for artifacts by the anesthesia information management system software. Third, we could not compare the artifact filtering methods with

a manual identification and exclusion of artifacts. For the purpose of our methods, comparison of a large dataset was required, but this hindered manual annotation of the data. Additionally, it would be questionable to do this in a retrospective manner.^{2,5} In the current study, we only included patients in whom invasive blood pressure monitoring was applied, resulting in a high proportion of high-risk surgery patients (table 1). In this subpopulation of the entire cohort, the incidence of postoperative myocardial injury was greater than previously reported.^{9,11} Nevertheless, this data were used as an illustration and were not used to estimate the true association between intraoperative and postoperative myocardial injury. Fourth, the effect of filtering artifacts could in theory be different within subgroups of procedures or patients. We did not adjust the association between intraoperative hypotension and postoperative myocardial injury for confounding, nor have we performed subgroup analyses. We considered this beyond the scope of the study. Consequently, as we have studied the effect of artifact filtering methods in a cohort with strict inclusion criteria, one cannot simply generalize these findings to a broader selection of procedures.

Fifth, we made other methodologic choices that could have influenced the association between hypotension and postoperative myocardial injury. For example, we chose linear interpolation between measurements,¹² instead of other methods such as last value carried forward. In a *post hoc* analysis we excluded cases with big gaps (greater than 15 min without data), which resulted in a tiny—although systematic—increase (about 0.02) in the odds ratio estimate. We therefore decided to not exclude any cases with gaps in blood pressure measurements, or cases in which a significant amount of measurements were removed by artifact filtering. These explorations suggested that the effect of bias (of gap removal) on the estimated odds ratio is minimal. In general, the variance of blood pressure measurements will be underestimated in linear interpolation, so if the primary interest is to estimate variability, then one should be more careful with interpolation.

Finally, this study did not cover all possible artifact filtering methods or combinations of methods. The aim of this study was to illustrate what happens when different filters are chosen, rather than to identify the best method. This best method is highly situational, depending on the type of data, the protocols used during anesthesia, the hardware used, and the way data is stored in an anesthesia information management system. Future (experimental) studies should aim to find reliable, generalizable methods for filtering artifacts in large anesthesia information management system databases.

Implications

It is hard to estimate the true association between hypotension and postoperative myocardial injury in our cohort, because we do not know which artifact filtering method is closest to the truth.¹³ Even if we would have had a fully

annotated dataset where artifacts are marked manually, one might still question whether this gives the correct estimate of the effect of intraoperative hypotension on postoperative myocardial injury.^{28,29} Despite the fact that we cannot identify the best method for artifact handling, the use of different artifact methods in one paper, like we have done, is not generally advised. A researcher should choose one method, thereby explicitly defining the outcome of interest. Preferably, the outcome should be comparable to similar research, but this does not necessarily imply that the same artifact filtering should be used. The choice of artifact filtering method will depend on the nature of the data and the type of hypotension measure, as some measures are more sensitive to artifacts than others.

Over time, improvements in artifact handling of anesthesia monitoring systems may result in better research data that contains fewer artifacts. This will decrease the need for postprocessing data for research. Currently, monitor systems cannot prevent all types of artifacts in anesthesia information management system data. For example, an anesthesia monitor has no information available on the correct placement of sensors and the resulting blood pressure will not be recognized as an artifact by the anesthesia monitor artifact algorithm.

Overall Conclusion

Although different artifact filtering methods yielded important differences in the quantification of intraoperative hypotension, we did not see a profound effect of these methods on the effect measures of the association between intraoperative hypotension and postoperative myocardial injury. It seems that the variation resulting from artifacts is smaller than the effect of the choice of hypotension measure or the chosen hypotension threshold. Nevertheless, the way one deals with artifacts may add to the reproducibility and comparability of intraoperative hypotension research. It seems wise to carefully consider how to handle artifacts in research using intraoperative physiologic data obtained from anesthesia information management systems. Authors should describe the chosen methodology for artifact filtering in detail.

Research Support

Support was provided solely from institutional and/or departmental sources.

Competing Interests

The authors declare no competing interests.

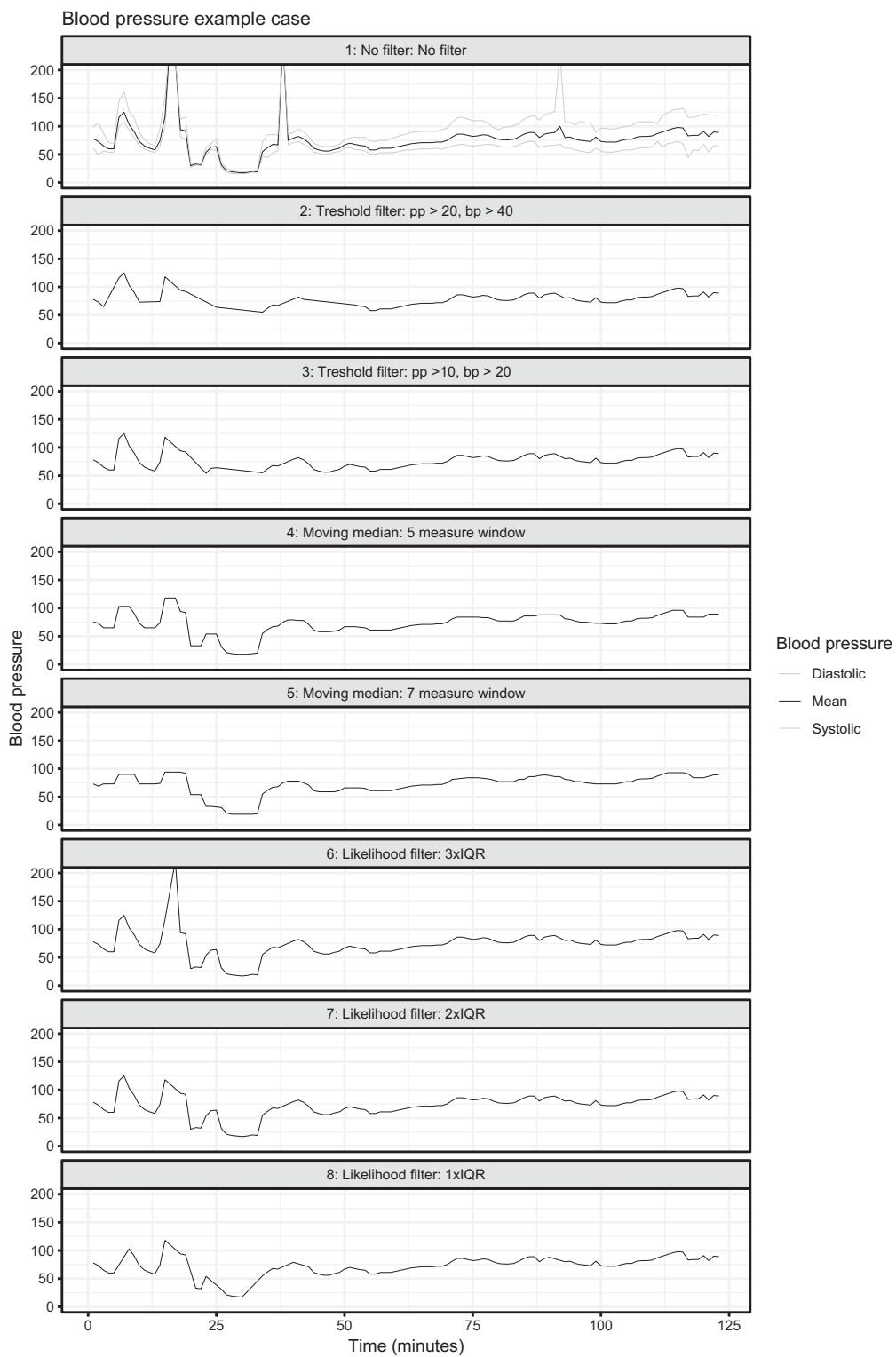
Correspondence

Address correspondence to Mr. Pasma: Department of Anesthesiology, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, 3584 CX Utrecht,

Netherlands. w.pasma@umcutrecht.nl, Information on purchasing reprints may be found at www.anesthesiology.org or on the masthead page at the beginning of this issue. ANESTHESIOLOGY's articles are made freely accessible to all readers, for personal use only, 6 months from the cover date of the issue.

Appendix: Examples of Cases with Different Filters Applied

This example illustrates the effect of different artifact filtering methods on a blood pressure signal of one of the procedures in our cohort. Eight different methods are illustrated (including no filter). These methods are also described in the paper.



References

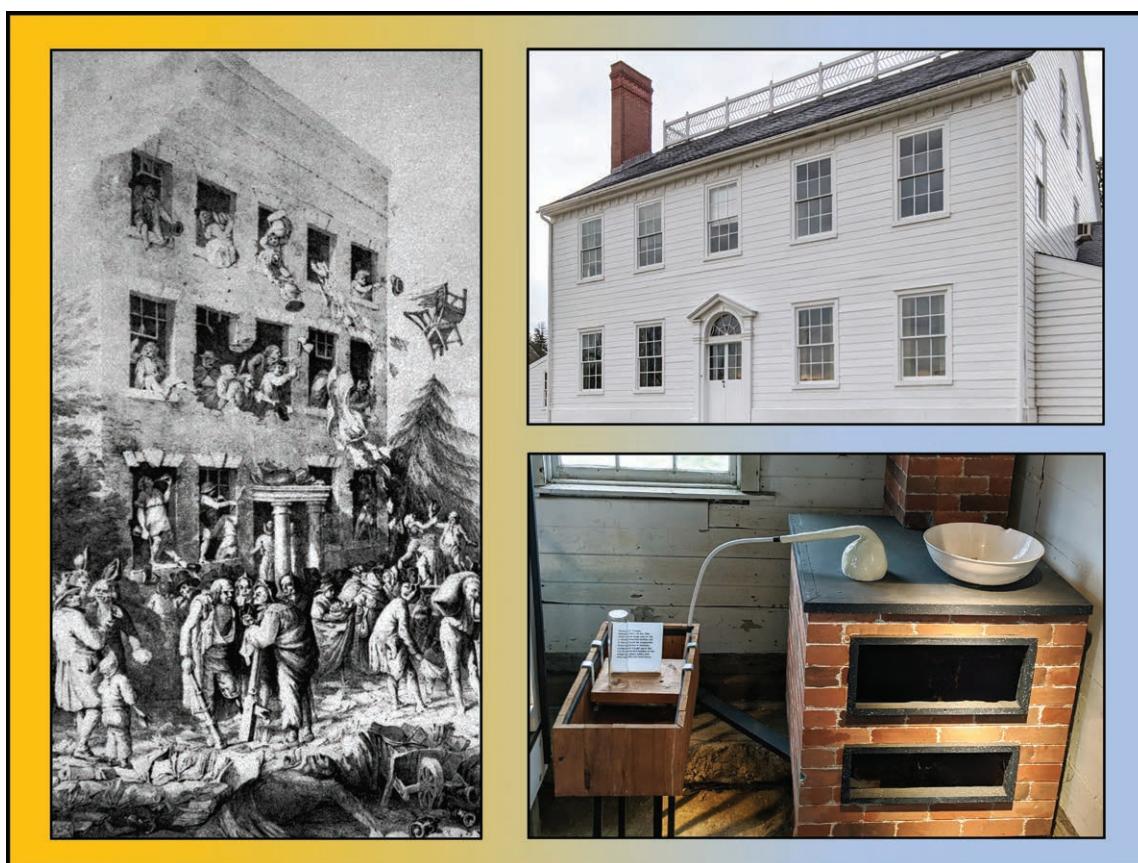
1. Martin M, Kheterpal S: Monitoring artifacts and large database research: What you don't know could hurt you. *Can J Anaesth* 2012; 59:823–6
2. Kool NP, van Waes JA, Bijker JB, Peelen LM, van Wolfswinkel L, de Graaff JC, van Klei WA: Artifacts in research data obtained from an anesthesia information and management system. *Can J Anaesth* 2012; 59:833–41
3. Verduijn M, Peek N, de Keizer NF, van Lieshout EJ, de Pont AC, Schultz MJ, de Jonge E, de Mol BA: Individual and joint expert judgments as reference standards in artifact detection. *J Am Med Inform Assoc* 2008; 15:227–34
4. Cunningham S, Symon AG, McIntosh N: The practical management of artifact in computerised physiological data. *Int J Clin Monit Comput* 1994; 11:211–6
5. Hoorweg AJ, Pasma W, van Wolfswinkel L, de Graaff JC: Incidence of artifacts and deviating values in research data obtained from an anesthesia information management system in children. *ANESTHESIOLOGY* 2018; 128:293–304
6. Cao C, Kohane IS, McIntosh N: Artifact detection in cardiovascular time series monitoring data from preterm infants. *Proc AMIA Symp* 1999:207–11
7. Batista-Foguet JM, Coenders G, Ferragud MA: Using structural equation models to evaluate the magnitude of measurement error in blood pressure. *Stat Med* 2001; 20:2351–68
8. Maslove DM, Dubin JA, Shrivats A, Lee J: Errors, omissions, and outliers in hourly vital signs measurements in intensive care. *Crit Care Med* 2016; 44:e1021–30
9. van Waes JA, Nathoe HM, de Graaff JC, Kemperman H, de Borst GJ, Peelen LM, van Klei WA: Myocardial injury after noncardiac surgery and its association with short-term mortality. *Circulation* 2013; 127:2264–71
10. Bijker JB, van Klei WA, Kappen TH, van Wolfswinkel L, Moons KG, Kalkman CJ: Incidence of intraoperative hypotension as a function of the chosen definition: literature definitions applied to a retrospective cohort using automated data collection. *ANESTHESIOLOGY* 2007; 107:213–20
11. Vernooij LM, van Klei WA, Machina M, Pasma W, Beattie WS, Peelen LM: Different methods of modelling intraoperative hypotension and their association with postoperative complications in patients undergoing non-cardiac surgery. *Br J Anaesth* 2018; 120:1080–9
12. Salmasi V, Maheshwari K, Yang D, Mascha EJ, Singh A, Sessler DI, Kurz A: Relationship between intraoperative hypotension, defined by either reduction from baseline or absolute thresholds, and acute kidney and myocardial injury after noncardiac surgery: A retrospective cohort analysis. *ANESTHESIOLOGY* 2017; 126:47–65
13. Wesselink EM, Kappen TH, Torn HM, Slooter AJC, van Klei WA: Intraoperative hypotension and the risk of postoperative adverse outcomes: A systematic review. *Br J Anaesth* 2018; 121:706–21
14. Sun LY, Chung AM, Farkouh ME, van Diepen S, Weinberger J, Bourke M, Ruel M: Defining an intraoperative hypotension threshold in association with stroke in cardiac surgery. *ANESTHESIOLOGY* 2018; 129:440–7
15. Tronstad C, Omenås IN, Rosseland LA: An improved artifact removal algorithm for continuous cardiac output and blood pressure recordings. *Conf Proc IEEE Eng Med Biol Soc* 2015; 2015:5968–72
16. Cao H, Norris P, Ozdas A, Jenkins J, Morris JA: A simple non-physiological artifact filter for invasive arterial blood pressure monitoring: A study of 1852 Trauma ICU patients. *Conf Proc IEEE Eng Med Biol Soc* 2006:1417–20
17. Megjhani M, Terilli K, Martin A, Velazquez A, Claassen J, Roh D, Agarwal S, Smielewski P, Boehme AK, Schmidt JM, Park S: Deriving the PRx and CPPopt from 0.2-Hz data: Establishing generalizability to Bedmaster users. *Acta Neurochir Suppl* 2018; 126:179–82
18. Sottas CE, Cumin D, Anderson BJ: Blood pressure and heart rates in neonates and preschool children: An analysis from 10 years of electronic recording. *Paediatr Anaesth* 2016; 26:1064–70
19. Reisner AT, Liangyou C, Reifman J: The association between vital signs and major hemorrhagic injury is significantly improved after controlling for sources of measurement variability. *J Crit Care* 2012; 27:533.e1–533.e10
20. Görge M, West NC, Karlsdóttir E, Ansermino JM, Cassidy M, Lauder GR: Developing an objective method for analyzing vital signs changes in neonates during general anesthesia. *Paediatr Anaesth* 2016; 26:1071–81
21. Wilson S, Bowyer A, Harrap SB: Characterizing the continuously acquired cardiovascular time series during hemodialysis, using median hybrid filter preprocessing noise reduction. *Med Devices (Auckl)* 2015; 8:65–70
22. Borowski M, Siebig S, Wrede C, Imhoff M: Reducing false alarms of intensive care online-monitoring systems: an evaluation of two signal extraction algorithms. *Comput Math Methods Med* 2011; 2011:143480
23. Yang P, Dumont GA, Ansermino JM: Sensor fusion using a hybrid median filter for artifact removal in intraoperative heart rate monitoring. *J Clin Monit Comput* 2009; 23:75–83
24. Clifford GD, Long WJ, Moody GB, Szolovits P: Robust parameter extraction for decision support using multimodal intensive care data. *Philos Trans A Math Phys Eng Sci* 2009; 367:411–29
25. Charbonnier S, Becq G, Biot L: On-line segmentation algorithm for continuously monitored data in intensive care units. *IEEE Trans Biomed Eng* 2004; 51:484–92

26. Hoare SW, Asbridge D, Beatty PC: On-line novelty detection for artefact identification in automatic anaesthesia record keeping. *Med Eng Phys* 2002; 24:673–81
27. Jakob S, Korhonen I, Ruokonen E, Virtanen T, Kogan A, Takala J: Detection of artifacts in monitored trends in intensive care. *Comput Methods Programs Biomed* 2000; 63:203–9
28. Reich DL, Wood RK Jr, Mattar R, Krol M, Adams DC, Hossain S, Bodian CA: Arterial blood pressure and heart rate discrepancies between handwritten and computerized anesthesia records. *Anesth Analg* 2000; 91:612–6
29. Wax DB, Beilin Y, Hossain S, Lin HM, Reich DL: Manual editing of automatically recorded data in an anesthesia information management system. *ANESTHESIOLOGY* 2008; 109:811–5
30. Glynn RJ, L'Italien GJ, Sesso HD, Jackson EA, Buring JE: Development of predictive models for long-term cardiovascular risk associated with systolic and diastolic blood pressure. *Hypertension* 2002; 39:105–10
31. Tsien CL, Kohane IS, McIntosh N: Building ICU artifact detection models with more data in less time. *Proceedings AMIA Symp* 2001:706–10
32. Gostt RK, Rathbone GD, Tucker AP: Real-time pulse oximetry artifact annotation on computerized anaesthetic records. *J Clin Monit Comput* 2002; 17:249–57
33. Stegle O, Fallert SV, MacKay DJ, Brage S: Gaussian process robust regression for noisy heart rate data. *IEEE Trans Biomed Eng* 2008; 55:2143–51
34. Ansermino JM, Daniels JP, Hewgill RT, Lim J, Yang P, Brouse CJ, Dumont GA, Bowering JB: An evaluation of a novel software tool for detecting changes in physiological monitoring. *Anesth Analg* 2009; 108: 873–80
35. Dosani M, Lim J, Yang P, Brouse C, Daniels J, Dumont G, Ansermino JM: Clinical evaluation of algorithms for context-sensitive physiological monitoring in children. *Br J Anaesth* 2009; 102:686–91
36. Caicedo A, Van Huffel S: Weighted LS-SVM for function estimation applied to artifact removal in bio-signal processing. *Conf Proc IEEE Eng Med Biol Soc* 2010; 2010:988–91
37. Thomas L, Stefanski L, Davidian M: A moment-adjusted imputation method for measurement error models. *Biometrics* 2011; 67:1461–70
38. Stirnemann JJ, Comte F, Samson A: Density estimation of a biomedical variable subject to measurement error using an auxiliary set of replicate observations. *Stat Med* 2012; 31:4154–63
39. de Graaff JC, Bijker JB, Kappen TH, van Wolfswinkel L, Zutthoff NPA, Kalkman CJ: Incidence of intraoperative hypoxemia in children in relation to age. *Anesth Analg* 2013; 117:169–75
40. Ganesh SK, Chasman DI, Larson MG, Guo X, Verwoert G, Bis JC, Gu X, Smith AV, Yang ML, Zhang Y, Ehret G, Rose LM, Hwang SJ, Papanicolaou GJ, Sijbrands EJ, Rice K, Eiriksdottir G, Pihur V, Ridker PM, Vasan RS, Newton-Cheh C, Raffel LJ, Amin N, Rotter JI, Liu K, Launer LJ, Xu M, Caulfield M, Morrison AC, Johnson AD, Vaidya D, Dehghan A, Li G, Bouchard C, Harris TB, Zhang H, Boerwinkle E, Siscovick DS, Gao W, Uitterlinden AG, Rivadeneira F, Hofman A, Willer CJ, Franco OH, Huo Y, Witteman JC, Munroe PB, Gudnason V, Palmas W, van Duijn C, Fornage M, Levy D, Psaty BM, Chakravarti A; Global Blood Pressure Genetics Consortium: Effects of long-term averaging of quantitative blood pressure traits on the detection of genetic associations. *Am J Hum Genet* 2014; 95:49–65
41. Terrill PI, Dakin C, Hughes I, Yuill M, Parsley C: Nocturnal oxygen saturation profiles of healthy term infants. *Arch Dis Child* 2015; 100:18–23
42. Cross TJ, Keller-Ross M, Issa A, Wentz R, Taylor B, Johnson B: The impact of averaging window length on the “desaturation indexes during overnight pulse oximetry at high-altitude.” *Sleep* 2015; 38:1331–4
43. Mort AJ, Fitzpatrick D, Wilson PM, Mellish C, Schneider A: Lightweight physiologic sensor performance during pre-hospital care delivered by ambulance clinicians. *J Clin Monit Comput* 2016; 30:23–32
44. Hwang YT, Wang CC, Wang CH, Tseng YK, Chang YJ: Joint model of multiple longitudinal measures and a binary outcome: An application to predict orthostatic hypertension for subacute stroke patients. *Biom J* 2015; 57:661–75
45. Zahari M, Lee DS, Darlow BA: Algorithms that eliminate the effects of calibration artefact and trial-imposed offsets of Masimo oximeter in BOOST-NZ trial. *J Clin Monit Comput* 2016; 30:669–78
46. Fathabadi OS, Gale T, Wheeler K, Plottier G, Owen LS, Olivier JC, Dargaville PA: Hypoxic events and concomitant factors in preterm infants on non-invasive ventilation. *J Clin Monit Comput* 2017; 31: 427–33
47. Chen L, Dubrawski A, Wang D, Fiterau M, Guillame-Bert M, Bose E, Kaynar AM, Wallace DJ, Guttendorf J, Clermont G, Pinsky MR, Hravnak M: Using supervised machine learning to classify real alerts and artifact in online multisignal vital sign monitoring data. *Crit Care Med* 2016; 44:e456–63
48. Georgatzis K, Lal P, Hawthorne C, Shaw M, Piper M, Tarbert C, Donald R, Williams CK: Artefact in physiological data collected from patients with brain injury: Quantifying the problem and providing a solution using a factorial switching linear dynamical systems approach. *Acta Neurochir Suppl* 2016; 122:301–5
49. Nixon GM, Davey MJ, Weichard AJ, Horne RS: Oximetry for suspected obstructive sleep apnea—Does removal of awake data affect the result? *Pediatr Pulmonol* 2016; 51:1409–13
50. Hravnak M, Chen L, Dubrawski A, Bose E, Clermont G, Pinsky MR: Real alerts and artifact classification in

- archived multi-signal vital sign monitoring data: implications for mining big data. *J Clin Monit Comput* 2016; 30:875–88
51. Kim H, Lee SB, Son Y, Czosnyka M, Kim DJ: Hemodynamic instability and cardiovascular events after traumatic drain injury predict outcome after artifact removal with deep belief network analysis. *J Neurosurg Anesthesiol* 2018; 30:347–53
52. Colopy GW, Tingting Zhu, Clifton L, Roberts SJ, Clifton DA: Likelihood-based artefact detection in continuously-acquired patient vital signs. *Conf Proc IEEE Eng Med Biol Soc* 2017; 2017:2146–9
53. Wellington G, Elder D, Campbell A: 24-hour oxygen saturation recordings in preterm infants: Editing artefact. *Acta Paediatr Int J Paediatr* 2018; 107: 1362–9

ANESTHESIOLOGY REFLECTIONS FROM THE WOOD LIBRARY-MUSEUM

Riots of Discovery: Joseph Priestley's Home Laboratories in England and America



An English theologian and natural philosopher, Joseph Priestley (1733 to 1804) discovered many gases, including “dephlogisticated air” or oxygen in 1774. However, his sympathies toward revolutionaries abroad and religious “Dissenters” at home soon made Priestley unpopular with the Crown and the Church of England. In 1791, he abandoned his English home in Birmingham just before an angry mob looted (*left*) and then burned it to the ground. Political and professional outcasts, Joseph and his wife Mary joined their children in America, settling in rural Northumberland, Pennsylvania, in 1794. Wasting no time, Joseph’s laboratory was part of the first-constructed wing of the house, which was finished in 1797 (*lower right*). By equipping the Priestleys’ new home laboratory with a fume hood to exhaust both smoke and toxic, asphyxiating, or flammable gases, Joseph spared the house that Mary had designed (*upper right*) from the fiery fate that befell their Birmingham home. Venting fumes would also help shield the Priestleys from the asphyxiating carbon monoxide that Joseph would discover in 1799. (Copyright © the American Society of Anesthesiologists’ Wood Library-Museum of Anesthesiology.)

Melissa L. Coleman, M.D., Penn State College of Medicine, Hershey, Pennsylvania, and George S. Bause, M.D., M.P.H., Case Western Reserve University, Cleveland, Ohio.