Assessing the measurement error of different methods used to calculate Pulse Pressure Variation

S. NJUGUNA¹, P. WYFFELS¹, P. WOUTERS¹

¹Department of Anesthesiology, Ghent University Hospital, Corneel Heymanslaan 10, 9000 Ghent, Belgium.

Corresponding author: Njuguna S. K., Department of Anesthesiology, Ghent University Hospital, Corneel Heymanslaan 10, 9000 Ghent, Belgium. E-mail: smn.njuguna@yahoo.co.uk

Abstract

Objective: To assess the measurement error of different methods used to calculate Pulse Pressure Variation.

Background: Many studies have demonstrated the use of pulse pressure variation (PPV) as a predictor of fluid responsiveness as long as the limitations to its use are understood and respected. These limitations have proven a constraint in the use of PPV and various researchers have published methods of overcoming these constraints in daily practice. Different methods also exist to calculate PPV. This study aims to systematically calculate the measurement error of the different methods used to calculate PPV and compare them.

Methods: After approval of the institutional trial board and ethics committee of Ghent University Hospital, Ghent, Belgium, and registration with the local code number B670201629642 (intern:EC/2016/1113), postoperative measurements of invasive arterial pressure and ECG were simultaneously recorded over 1 minute between 29/11/2016 to 16/11/2018. Data was then analyzed using different methods, namely (i) individual PPV averaged over fixed number of respiratory cycles (iPPV family), (ii) pooled PPV over fixed number of respiratory cycles (pPPV family) and (iii) methods over fixed window in terms of time (Aboy and Lansdorp). The Taffe extension of the Bland Altman method was used to compare and determine the measurement error of these four different methods. IPPV1 was chosen as the common reference. Differential en proportional bias and precision are reported as the intercept and the slope respectively of the models studied.

Results: Data from 27 subjects were collected. The iPPV showed minimal bias and improved precision. pPPV showed increasing bias (0.879 - 1.999) with the increase in the respiratory cycles as well as precision (0.633-1.08). The Aboy algorithm model showed reduction in bias (-0.473 - -0.139) and precision (0.235-0.146) by the larger fixed windows. Bias increases from the smaller windows to the larger windows in the Lansdorp method. Precision improves over the same range.

Conclusions: Every method has its own measurement error. There is a proportionality in the measurement error in the methods we compared for calculating PPV. The bias is variable by each method we studied.

Mesh terms: Hemodynamics, Pulse Pressure Variation, Measurement Error, Blood Pressure Physiology, Predictive Values of Tests.

Introduction

Since its first description, many studies have demonstrated the value of pulse pressure variation (PPV) as a predictor of fluid responsiveness. Studies have shown that intraoperative optimization of cardiac output (CO) by repeated volume loading reduces postoperative morbidity after major surgery. However, unnecessary intravenous fluids may be deleterious, and intraoperative fluid restriction has also been shown to improve clinical outcome, thus the importance of accurate predictors of fluid responsiveness. A meta-analysis published in 2014 concluded that PPV predicts fluid responsiveness accurately (sensitivity 88%, specificity 89%), as long as limitations to its use are understood and respected. The most common physiological constraints can be summarized as: general anesthesia, sinus rhythm and full mechanical ventilation with (a) a tidal volume (TV) ≥8 mL/kg of body weight and...
(b) a positive end-expiratory pressure (PEEP) <5 cm H₂O, closed chest conditions and a high HR/RR (Heart rate/respiratory rate) ratio³,⁶,⁷.

Even though PPV use gained popularity in the ICU and the operating room, its clinical use in predicting fluid responsiveness has been questioned. Fewer patients than expected meet the criteria necessary for its use⁴ and mechanical ventilation with lower tidal volumes (<8ml/kg body weight) is recommended to prevent acute lung injury⁵. In addition, cardiac arrhythmias can occur on an irregular basis coupled with an expected increase in the incidence of atrial fibrillation with an ageing population⁶. These factors can negatively influence the predictive value of dynamic indices⁷.

Research into overcoming these constraints is promising, with the tidal volume challenge⁸,⁹ providing an alternative for use of PPV in low tidal volume situations. Wyffels et al recently reported an alternative algorithm that makes it feasible to determine isolated ventilation induced Pulse Pressure Variation (VPPV) in patients with atrial fibrillation¹⁰. The mini fluid challenge showed reliable predictability in fluid responsiveness with high-sensitivity and specificity in patients under operative conditions¹¹. Passive Leg Raise (PLR) test has also been extensively studied and proved to be an excellent predictor of fluid responsiveness, also in patients with arrhythmias or spontaneously breathing¹², but this is not applicable in the operating room.

Different methods have been used to calculate Pulse Pressure Variation. Derichard et al¹³ studied automated PPV using pooled data over 8 seconds and compared this with an offline calculated PPV as reference during abdominal surgery. Their findings showed a sufficient degree of diagnostic accuracy between the automated method and the offline method. This correlation possibly increased the use of automated PPV use perioperatively¹⁴,¹⁵.

On the other hand, Lansdorp et al¹⁶ compared different methods of calculating PPV. They argued that increasing the number of breaths over which the dynamic indices are calculated can increase the calculated values, because larger and smaller pulse pressures are more likely to occur during a longer observation period, resulting in a more pronounced variation in the dynamic indices¹⁷.

Furthermore, clinical devices such as the Philips IntelliVue Patient Monitor (Philips Medical Systems, Eindhoven, The Netherlands), the PiCCO² (Pulsion Medical Systems, Munich, Germany), and the Dräger infinity (Dräger Medical, Lübeck, Germany) use software that samples a defined time interval without identifying the number of breaths. During low respiratory rates, this leads to significant variation in PPV, which has no physiological background and should therefore be discarded. In other cases, the time window of 30 s leads to an unnecessary high number of breaths, which increases the chance to include cardiac arrhythmias such as premature ventricular contractions. This negatively influences the predictive value of dynamic indices¹⁸,¹⁹.

These different methods of calculation may negatively influence the reliability of PPV to predict fluid responsiveness¹⁰ Furthermore, studies over the measurement error of these different methods are scarce.

The aim of this study is to systematically calculate the measurement error of the different methods used to calculate PPV and compare them. We assessed this measurement error in terms of systematic error (bias) as well as the random error (precision) related to the measurement.

Methods

Compliance with Ethical standards

After approval of the institutional trial board and ethics committee of Ghent University Hospital, Ghent, Belgium, this study was registered with the local code number B670201629642 (intern:EC/2016/1113). Informed consent was obtained from all participants according to the Declaration of Helsinki and International Conference on Harmonization/ Good Clinical Practice. The study took place between 29/11/2016 to 16/11/2018. Patients were eligible for the study if; (i) age >18 years, (ii) had sinus rhythm, and (iii) scheduled to undergo Robotic assisted surgery (RAS) in extreme trendelenburg position. Exclusion criteria were; (i) atrial fibrillation, (ii) COPD, (iii) aortic valve stenosis or insufficiency, (iv) Right ventricular failure and (v) involvement in a clinical trial within the past 30 days.

Study Procedure

On the day of intervention, all patients were equipped with the standard monitoring for the planned elective surgery; 5 lead ECG, pulse oximeter and an arterial catheter in the right or left radial artery, which was placed after induction of anesthesia.

Induction protocol was standardized as per local guidelines. After intubation, the patients were ventilated using standard ventilator settings (FiO₂-40%, tidal volume of 8ml/kg, frequency-12/minute and PEEP-5). Adjustments were allowed based on measured saturation and end tidal CO₂ level. Maintenance of anesthesia was standardized with
sevoflurane and boluses of sufentanil for analgesia when required. After surgery, the patients were transported to the post anesthesia care unit (PACU) under target controlled infusion anesthesia (TCI) with remifentanil and propofol and further mechanically ventilated in anti-trendelenburg position for at least one hour, which was standard procedure at the institution at the time of the study as a preventive measure for facial/laryngeal edema and agitation after prolonged RAS in extreme trendelenburg position.

Perioperatively, a 3F catheter (Arterial Leadercath; Vygon, Écouen, France) was placed in the radial artery. The transducer was levelled at the mid-axillary line and zeroed to atmospheric pressure. At start of the study in the PACU (Post Anesthesia Care Unit), TV was set at 8ml/kg and ventilation rate was adjusted for an HR/MVR around 5 (ventilatory frequency of 12 bpm with a 1:E ratio of 1:2 and PEEP set at 5 cm H2O). After stabilization and a check to assure there was no spontaneous breathing activity, invasive arterial pressure and ECG were simultaneously recorded during 1 minute, with a sampling rate of 200 and 100 Hz respectively. ECG (Lead II and V2) and arterial pressure signals were stored using LabSystem™ Pro version 2.4a (BARD Electrophysiology, Lowell, MA, USA).

Data Analysis

All data strips were stored and analyzed offline. For each beat Systolic (SBP), Diastolic blood pressure (DBP) and corresponding Pulse Pressure (PP max and PP min) were determined, using a personal MATLAB® script based on the methods described by Li and colleagues.

The general formula for calculating PPV is;

\[ PPV (\%) = 100 \times \frac{(PP_{\text{max}} - PP_{\text{min}})}{(PP_{\text{max}} + PP_{\text{min}})/2} \]

With \( PP_{\text{max}} \) and \( PP_{\text{min}} \) the maximal and minimal values of pulse pressure, respectively. This general formula forms the basis of a wide array of methods that is used in research and clinical practice. We divide these methods in 3 groups:

1. individual PPV over fixed number of respiratory cycles (iPPV family)
2. pooled PPV over fixed number of respiratory cycles (pPPV family)
3. method over fixed window in terms of time (Aboy and Lansdorp)

Individual PPV is the PPV over a number of respiratory cycles, averaged over the number of respiratory cycles. For example, iPPV2 is the mean of PPV calculated over 2 respiratory cycles and iPPV3 over 3 respiratory cycles. Pooled PPV is the PPV generated using pooled data over a number of respiratory cycles. E.g., pPPV3 is the PPV generated from pooled data over 3 respiratory cycles. This is further represented in Fig 1.

IPPV was generated for up to 5 consecutive cycles (iPPV1 to iPPV5 respectively). Data for pPPV was pooled over 1 respiratory cycle up to 5 respiratory cycles (pPPV1 up to pPPV5).

For the fixed window method of calculating PPV, the Aboy and Lansdorp methods were used. For these methods we used the following fixed windows; 12 seconds, 15 seconds, 20 seconds, 30 seconds and 60 seconds, to generate PPV values.

For every method a maximum of independent measurements for each data strip were done according to the pattern in Figure 2.

Statistics

To determine the measurement error of the different methods, the Taffe modification of the Bland Altman method was used. This method is especially useful when comparing two methods.
Bias

The bias is represented as the deviation from the horizontal line in the models shown in Figure 4. Interception of the x axis by the red line represents the differential bias while the slope represents the proportional bias.

In all the iPPV models, both the differential bias and precision bias show minimal change when comparing the smaller windows and the larger windows (-0.047 - -0.099) and (1.011-1.018) respectively.

Pooled PPV models demonstrate that the differential bias increases from the smaller windows to the larger windows (0.879 - 1.999) while the proportional bias does not increase as substantially from the smaller windows to the large ones (1.013 - 1.001).

In the Aboy series, the differential bias is greater by the smaller windows, generally reducing by the larger windows (-0.473 - -0.139). Proportional bias remains almost equal across the series (0.904 - 0.989).

By the Lansdorp model, there is a significant increase in the differential bias from the smaller windows to the larger windows. (0.233 – 2.33). Proportional bias does show an increase, but not as substantial as by the differential bias (0.758-0.996).

Table II shows the details of the differential bias and the proportional bias for each of the methods.

Precision

The precision in the models is once again demonstrated as the deviation from the x axis, which represents the point of no standard deviation. For the methods studied, this is shown in Fig 5.

When comparing the smaller windows to the larger ones in the iPPV model, the general trend of the differential precision from the smaller windows to the large windows is a trend towards zero (0.529-0.065). The proportional precision is also reducing as the number of respiratory cycles increase. (0.056-0.031).

In the pPPV model, the differential precision increased as the number of respiratory cycles increased (0.633-1.08), whereas the proportional precision reduces slightly (0.042 - 0.012).

The Aboy model showed a decrease in differential precision (0.235-0.146) as the windows increased. This was once again echoed by the proportional precision, reducing from 0.107 to 0.013.

Differential precision in the Lansdorp model generally shows a reduction in trend, except for the 60s window that was the highest measurement in the trend. The proportional precision also
follows this trend, decreasing as the fixed windows increase. These trends are represented numerically in Table III.

### Discussion

The main findings of our study are that every method we compared for calculating PPV has its own measurement error and that there is a proportionality in the measurement error in these methods. The bias is variable by each method we studied.

The iPPV model showed minimal bias over the measured respiratory cycles, with the precision improving as the number of respiratory cycles used increased. This falsely suggest that these methods provide the most accurate PPV calculation. It should be noted that the calculated bias is the relative difference in comparison to the ‘arbitrarily’ chosen common reference. In the literature, iPPV appears to be the most used method in studies where PPV is calculated manually. When looking at the pooled PPV models, the bias increases as the respiratory number increases, meaning it is more accurate by lower respiratory cycles. The precision does not seem improve with an increase in respiratory cycles.

The Aboy model shows a general trend of underestimation of PPV by the smaller windows, with bias reduction by the larger fixed windows. Its precision also improved by the larger time frames, meaning this method shows its least measurement error by the larger fixed windows when compared to the common reference. Lansdorp model suggests that this method underestimates PPV by the lower fixed windows and overestimates it as the fixed windows increase when compared to iPPV.

These findings could have some influence when interpreting some studies around PPV. In the gray zone approach, PPV accuracy was considered inconclusive within certain limits in about 25% of patients undergoing general anesthesia. This is after the authors used two cut off thresholds (an upper and a lower) within which the accuracy of PPV as a diagnostic tool was inconclusive. The measurement error demonstrated by our results may partially explain the limits to this gray zone.

In the tidal volume challenge conducted by Myatra et al, the hemodynamic data was collected over a time period of 30 seconds and then averaged out for calculation of dynamic indices, including PPV. Messina et al also conducted a similar study. The monitoring systems used in these studies differed, with Myatra’s group using a Philips monitoring system.
However, these measurements are at the expense of missed events that can occur between these large windows.

In clinical practice, devices such as the Philips IntelliVue Patient Monitor (Philips Medical Systems, Eindhoven, The Netherlands), the PiCCO2 (Pulsion Medical Systems, Muenchen, Germany), and the Dräger infinity (Dräger Medical, Lübeck, Germany) use software that sample at defined time while Messina’s used Mindray (BeneView T8; Soma Technology, Inc., Bloomfield, Connecticut, USA). Our findings may be one of the reasons for the difference in their findings.

While trying to ascertain which method one should consider for measuring PPV, certain factors should be considered. In terms of measurement error (in terms of bias and precision), the larger windows provide the most accurate measurement of PPV. However, these measurements are at the expense of missed events that can occur between these large windows.
Derichard et al. noted that automated devices can give falsely raised readings of PPV and advised cautious interpretation, especially in cases of unexpected or unnoticed cardiac events during these defined time intervals.

The need for predicting fluid responsiveness is important so as to achieve optimal fluid therapy, which is one of the components for improving perioperative outcomes. We as clinicians should be aware of the algorithms/methods used to calculate PPV in the monitoring systems that we use in our daily practice. This may help acknowledge the possibility of a measurement error and in some instances, correct for it where possible.

**Fig. 4** — Bias plots of the studied methods

The X axis represents the BLUP while the Y axis represents the bias of measured PPV%. No bias is represented at the zero line on the x axis. The red line represents the actually measured bias per method. Deviation from zero is a representation of bias. The gray zone represents the uncertainty of the model (95 CI). A represents the methods using respiratory cycles while B represents methods using fixed windows.
Some shortcomings to this study should be noted. Only data from 27 subjects was used and this can under power our findings. Furthermore, the range of PPV we studied was at the lower limits. Higher PPV limits, have not really been assessed here. Also, data from 2 patients with early COPD that was well controlled, was included. We only studied one patient population and this could have an influence on the generalizability of our findings.

Bland and Altman’s limits of agreement (LoA) have traditionally been used in clinical research to assess the agreement between different methods of measurement for quantitative variables. However, when the variances of the measurement errors of the two methods are different, Bland and Altman’s plot may be misleading. Taffé’s modification offers more insight in bias and precision when studying heteroscedastic dataset, like was the case in this study. The Taffe modification was however tested with sample sizes of 100 individuals, while we had just 27.

We can therefore conclude that there is a proportionality to the measurement error in the methods used to calculate PPV and each method used to calculate PPV has its own characteristics.

Acknowledgements: No sources of funding to declare and no external contributors other than the aforementioned authors. There are no conflict of interests to report.

References


doi.org/10.56126/73.S1.24