On algorithms for calculating arterial pulse pressure variation during major surgery

To cite this article: Shaoxiong Sun et al 2017 Physiol. Meas. 38 2101

View the article online for updates and enhancements.
On algorithms for calculating arterial pulse pressure variation during major surgery

Shaoxiong Sun\textsuperscript{1,2}, Wouter H Peeters\textsuperscript{2}, Rick Bezemer\textsuperscript{2}, Xi Long\textsuperscript{1,2}, Igor Paulussen\textsuperscript{2,3}, Ronald M Aarts\textsuperscript{1,2} and Gerrit J Noordergraaf\textsuperscript{2,3}

\textsuperscript{1} Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, Netherlands
\textsuperscript{2} Philips Research, Eindhoven, Netherlands
\textsuperscript{3} Elisabeth-Tweesteden Hospital, Tilburg, Netherlands

E-mail: shaoxiong.sun@philips.com and sunshxbear@gmail.com

Received 18 May 2017, revised 27 September 2017
Accepted for publication 24 October 2017
Published 28 November 2017

Abstract

Objective: Arterial pulse pressure variation (PPV) is widely used for predicting fluid responsiveness and supporting fluid management in the operating room and intensive care unit. Available PPV algorithms have been typically validated for fluid responsiveness during episodes of hemodynamic stability. Yet, little is known about the performance of PPV algorithms during surgery, where fast changes of the blood pressure may affect the robustness of the presented PPV value. This work provides a comprehensive understanding of how various existing algorithmic designs affect the robustness of the presented PPV value during surgery, and proposes additional processing for the pulse pressure signal before calculating PPV. Approach: We recorded arterial blood pressure waveforms from 23 patients undergoing major abdominal surgery. To evaluate the performance, we designed three clinically relevant metrics. Main results and Significance: The results show that all algorithms performed well during episodes of hemodynamic stability. Moreover, it is demonstrated that the proposed processing helps improve the robustness of PPV during the entire course of surgery.

Keywords: fluid responsiveness, pulse pressure variation, baseline extraction, adaptive peak filter, major surgery

(Some figures may appear in colour only in the online journal)
1. Introduction

Optimal circulatory volume status of patients in the operating room (OR) and intensive care unit (ICU) is of key importance since hypovolemia can lead to inadequate organ perfusion, and hypervolemia can cause cardiac, renal, and pulmonary injury (Cannesson 2010). Hemodynamic optimization, with the goal of maintaining optimal circulatory conditions, has been shown to improve postoperative outcome and reduce the cost of surgery (Gan et al 2002, Wakeling et al 2005, Donati et al 2007, Cannesson et al 2011a).

Before subjecting a patient to fluid therapy, one should assess whether the heart would actually increase its output upon volume loading. This is called fluid responsiveness (Feissel et al 2004, 2005, Michard 2005). To assess fluid responsiveness, static and dynamic indicators have been proposed (Gödde et al 1998, Michard et al 2000, Marik et al 2008, Avolio et al 2010, Alian et al 2014, Pinsky 2014). Static indicators, such as central venous pressure (CVP), pulmonary artery occlusion pressure, and left ventricular end-diastolic area, have been demonstrated to have poor performance (Michard 2005, Vincent and Weil 2006, Osman et al 2007, Marik et al 2008). In contrast, dynamic indicators, relying on cardiopulmonary interactions, have been shown to be better predictors in patients undergoing mechanical ventilation (Michard 2005, Vincent and Weil 2006, Marik et al 2008, Monnet and Teboul 2013).

During mechanical ventilation, the cyclic changes in the intrathoracic pressure induce cyclic changes in the venous return and therefore the preload (end-diastolic blood volume) of the heart. These cyclic changes in the preload induce cyclic changes in the stroke volume, which will appear as cyclic changes in pulse pressure (PP) (Michard 2005). During positive pressure ventilation, pulse pressure increases during inspiration and decreases during expiration. The parameter pulse pressure variation (PPV) is designed to quantify such cardiopulmonary interactions (Michard et al 2000). It is defined as the ventilation-induced variation in pulse pressure normalized by the mean pulse pressure (Michard et al 2000). PPV has been shown to be useful for making therapeutic choices in mechanically ventilated patients with acute circulatory failure related to sepsis (Michard et al 2000) as well as in patients who have undergone coronary artery bypass grafting (Kramer et al 2004). In general, the use of PPV in clinical decision making has been proven to decrease the length of hospitalization (Lopes et al 2007).

Originally, Michard et al (2000) specified the PPV as the mean of three consecutive raw PPV values, each of which is derived from a single breath. The formula is given by

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

(1)

where \(PP_{\text{max}}\) and \(PP_{\text{min}}\) are the maximum and minimum pulse pressures sampled from a sample window with a duration of a single respiratory cycle. Before calculating PPV, the interpolated pulse pressure (IPP) signal is obtained by using different methods. The original specification is equivalent to linear interpolation of the pulse pressure, whereas Aboy et al (2009, 2004) applied kernel smoothing to derive the IPP. When calculating PPV, various sample durations were adopted: one ventilation cycle (Michard et al 2000), three or five ventilation cycles (Kim and Pinsky 2008), two ventilation cycles (Aboy et al 2009), 8 s (Derichard et al 2009), 10 s (Addison et al 2015). Extended sample durations reduce detrimental effects originating from the asynchrony of the cardiac and respiratory cycle (Kim and Pinsky 2008), and eliminate the need for accurate estimation of the ventilation frequency. Different ways to smooth raw PPV values in order to filter out fast PPV fluctuations were also investigated: three-point mean filter (Michard et al 2000), three-point median filter (Aboy et al 2004), and Kalman filter (Aboy et al 2009). Although not clearly published, commercial patient monitors seem to employ 8–12 s sample durations and to subsequently use a three- or four-point mean filter, thereby
evaluating in total about 30 s of the ABP waveform for each presented PPV value (Kim and Pinsky 2008, Derichard et al 2009).

The validity of PPV is well established, although the reliability of PPV has only been verified under relatively stable hemodynamic conditions (Michard et al 2000, Michard 2005, Solus-Biguenet et al 2006, Cannesson et al 2007). Little is known about the difference in the performance of algorithmic designs in the face of dynamically changing blood pressures, which may occur during surgery.

In this work, we evaluated five key algorithms existing in literature, for patients during ongoing surgery and present a comprehensive understanding of how different algorithmic designs affect the performance robustness. We designed three PPV performance metrics: the hourly occurrence of unstable PPV calculation episodes, the hourly occurrence of short-term elevations, and discrepancy compared to the original specification (Michard et al 2000) during periods of hemodynamic stability. Inspired by existing algorithms and our observations, we propose additional IPP post-processing before computing PPV to improve the robustness of PPV during surgery.

The paper is organized as follows: section 2 describes the acquired data, the general framework for the analysis of the algorithms, and comparison criteria for performance of algorithms during surgery; section 3 describes five existing algorithms in the framework described in section 2; section 4 presents our algorithm in the same framework; section 5 compares the performance of existing algorithms and our proposed algorithm; section 6 discusses our findings; section 7 concludes the paper.

2. Materials and methods

2.1. Protocol and data acquisition

The study protocol was reviewed and approved by the regional medical ethics committee (METC Brabant, The Netherlands, NL48421.028.14-P1409). Written informed consents were obtained from 29 patients scheduled for major abdominal surgery consisting mainly of open, radical, prostatectomy cases. Anesthesia was induced by propofol (2 mg kg$^{-1}$), sufentanil (0.5 mcg kg$^{-1}$), and rocuronium (0.6 mg kg$^{-1}$), and later maintained by means of continuous infusion of sufentanil and propofol. The depth of anesthesia was assessed using bispectral index (BIS) with a target of 40–55. The patients were ventilated in a volume-controlled mode with tidal volume of 6–10 ml kg$^{-1}$ at a frequency of 10–14 times min$^{-1}$, and adjusted to maintain normocapnea. The positive end-expiratory pressure was set at 6 cm H$_2$O and adjusted as needed. Fluid management was at the discretion of the physician. During surgery, continuous electrocardiogram (ECG) and invasive arterial blood pressure (ABP) signals (Philips Heartstart MRx monitor) were collected.

The analysis of ABP and PPV was confined to the period of mechanical ventilation. Signal segments with low signal quality or with severe cardiac arrhythmia were excluded by manual selection and a dedicated program. To determine the quality of each pulse, this program calculated distances between neighboring detected peaks, distances between neighboring detected valleys, and the amplitude of each detected pulse. For each of these parameters derived from every single pulse, if the difference between the present value and the extrema (maximum or minimum) in the 30 s history window was larger than a threshold that was defined as the discrepancy between the maximum and minimum values, this pulse would be rejected. Data from six patients were removed entirely due to signal quality or continuous cardiac arrhythmia. In the remaining 23 patients, 91.2 h data were found eligible for the further analysis of PPV (9.8% data were excluded).
2.2. General framework for the analysis of existing and proposed algorithms

In this paper, we provide insights into how the algorithmic designs in the existing algorithms impacted the performance of a full algorithm during surgery. To this end, we analyzed five key algorithms existing in the literature and investigated the impact of the processing steps of those algorithms on their overall performance.

The algorithms selected for comparison are (see also table 2):

(1) An algorithm applying the original specification of the PPV by Michard et al. (2000), using linear interpolation of the pulse pressure to derive the IPP. It uses a sample duration of a single ventilation cycle and subsequently averages three subsequent raw PPV values.

(2) An algorithm similar to (1) using a sample duration of three ventilation cycles, and no subsequent smoothing (Kim and Pinsky 2008).

(3) An algorithm using kernel smoothing to derive the IPP, using a sample duration of a single ventilation cycle for the raw PPV values, and subsequently using a three-point median filter. The ventilation frequency is determined from the IPP (Aboy et al. 2004).

(4) An algorithm that is adapted from (2), but using a longer sample duration for calculation of the raw PPV values, and using an advanced Kalman filter for PPV post-processing to deal with dynamic blood pressure changes (Aboy et al. 2009).

(5) A fixed-time-window algorithm with kernel smoothing to derive the IPP. This fixed-time-window type of algorithm is relevant, because such algorithms omit the need of estimating the ventilation frequency (Cannesson et al. 2008, Derichard et al. 2009, Addison et al. 2015). We implemented such an algorithm, as mentioned by Derichard et al. (2009).

The procedure of PPV calculation is described using a block diagram (see figure 1). Different options for each block can be found in sections 3 and 4. IPP is first derived from the ABP signal. After post-processing the IPP, raw PPV is calculated. Finally, the PPV values are post-processed to generate the presented PPV.

2.3. Performance indicators of PPV algorithms during surgery

In this paper, we employed three performance indicators to quantify the performance of the PPV algorithms. These performance indicators address both the performance during episodes of hemodynamic stability and the performance during the whole time span of surgery including more dynamic hemodynamic episodes. An elaborate discussion on the selection of these three indicators is given in appendix A, and this section provides a summary.

The first indicator is the closeness of the generated PPV values to those calculated using its original specification during periods of hemodynamic stability, as can be seen figure 2(a). We manually selected a segment of 2–3 min without significant changes in the heart rate or ABP from each patient as stable hemodynamic episodes. In total, 64 min of recorded data were selected as stable periods with PPV values ranging from 0.4% to 13.3%.

The other two performance indicators were applied to the entire course of the surgery, and aimed to quantify the performance during more dynamic hemodynamic episodes. These performance indicators are the hourly occurrences of two false PPV patterns that are confusing and not representative of the underlying hemodynamic status. The lower the hourly occurrence of these patterns, the better the performance.

One false PPV pattern, which is the second performance indicator, is an episode of unstable PPV calculation as illustrated in figure 2(b). Such an episode was defined as a time window of 40 s (i.e. approximately 10 ventilation cycles), in which the interquartile range of PPV values is larger than 2% (PPV values typically range in the order of 5–20% and change rather slowly). PPV algorithms are prone to producing this false pattern if the subsequent pulse
pressure fluctuates due to motion artifacts, irregular beats, dyssynchrony of the heart rate to the ventilation rate.

The other false PPV pattern, which is the third performance indicator, is a short-term elevation, as illustrated in figure 2(c). It was defined as a 10–40 s increase in median values larger than 2% compared to the values in preceding 40 s and following 40 s. PPV algorithms are prone to producing this false pattern if the pulse pressure gradually changes over a time span of several ventilations.

A summary of all these parameters can be found in table 1.

2.4. Statistical analysis

To investigate our proposed IPP post-processing, we applied a Wilcoxon signed-rank test on the hourly occurrence of unstable PPV calculation episodes and short-term elevations for

Addison et al (2012) used the baseline of the PPG signal to estimate respiration frequency and we applied a similar approach to the ABP signal.
each sample durations and each IPP derivation method after a Kolmogorov–Smirnov test for distribution normality (Hollander et al 2014). The p values were corrected according to the Bonferroni method using the number of comparisons (in this case 6). To investigate the impact of sample durations, we applied Friedman’s one-way test on the hourly occurrence of unstable PPV calculation episodes and short-term elevations after a Kolmogorov–Smirnov test for distribution normality (Hollander et al 2014). Following Friedman’s test, pairwise comparisons were carried out according to Tukey’s honestly significant difference criterion. To study the influence of IPP derivation, we applied a Wilcoxon signed-rank test on the hourly occurrence of unstable PPV calculation episodes and short-term elevations for each sample durations after a Kolmogorov–Smirnov test for distribution normality. The p values were corrected according to the Bonferroni method using the number of comparisons (in this case 6). For the final comparison between existing algorithms and our proposed algorithms, we applied Friedman’s one-way test after a Kolmogorov–Smirnov test for distribution normality. Again, following Friedman’s test, pairwise comparisons were carried out according to Tukey’s honestly significant difference criterion. A p < 0.05 was regarded statistically significant.

3. Description of existing algorithms

We analyzed and implemented five existing algorithms including Michard et al (2000) (the original specification of PPV), Aboy et al (2004, 2009), Derichard et al (2009) and Kim and Pinsky (2008). A summary of the processing steps of these existing algorithms and our
4. Description of our algorithm

4.1. Step 1: IPP derivation

For our new algorithm, we first obtained the baseline of the ABP signal by applying an LPF at a cutoff frequency of 0.4 Hz. The reason is that, in addition to the PP of the ABP signal, the baseline of the ABP signal is also modulated by ventilation and can be used for ventilation frequency estimation. Thus, we stored this baseline at this step for estimating ventilation frequency, such that it can be used in step 2 (section 4.2).

Next, we removed high-frequency components in the ABP signal by applying an LPF at a cutoff frequency of 15 Hz. To detect peaks and valleys in the signal, we used a dedicated method based on the first derivative of the ABP signal. Adaptive thresholds based on weighted averages of upper and lower envelope of the first derivative of the ABP signal were used to find the steepest slopes and thereby define the pressure pulse. The peaks and valleys were detected at the first zero-crossing in the first derivative after the steepest slope and the last zero-crossing before the steepest slope, respectively.

After deriving PP from the peaks and valleys, we separately used linear interpolation and kernel smoothing to derive IPP as comparison. For the kernel smoothing, as indicated by Aboy et al. (2004), it was followed by low-pass filtering at a cutoff frequency of 0.45 Hz. This filtering achieved a comparable filtering effect as the method described in appendix B, while reducing computational complexity. Note that all the filtering was done bi-directionally to avoid phase delay.

4.2. Step 2: IPP post-processing

Unlike the other algorithms which incorporated post-processing of PPV values, we post-processed the intermediate IPP signal. This method dealt more effectively with noise and also avoided excessive delay in displaying PPV values. To post-process IPP, we began by extracting the baseline of the IPP. This was implemented by bi-directionally applying a six-order Butterworth LPF at a cutoff frequency of 0.1 Hz. The extracted baseline was then stored to allow, in the next step, the computation of the mean IPP during one ventilation cycle. The detrended IPP, obtained by the subtraction of the baseline from the IPP, was used in the next step to determine the difference between the maximum and minimum IPP in a sample window. This method aimed to make the algorithm robust against the problem illustrated in figure 2(c).

In addition to extracting the baseline of the IPP, we applied an adaptive peak filter to the detrended IPP, such that we focused on the variations in the IPP for calculating PPV that were in the narrow frequency band centered at the ventilation frequency. This was targeted at alleviating the problem illustrated in figure 2(b). A classic second order adaptive peak filter (Orfanidis 2010) was used in this work. The transfer function of this filter in the z-domain is given by

$$H(z) = (1 - b) \frac{1 - z^{-2}}{1 - 2b \cos w_0 z^{-1} + (2b - 1)z^{-2}}$$

$$b = \frac{1}{1 + \tan(\Delta w/2)}$$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1: IPP derivation</td>
<td>Linear interpolation</td>
<td>Linear interpolation</td>
<td>Kernel smoothing + LPF$^a$</td>
<td>Kernel smoothing + LPF</td>
<td>Kernel smoothing + LPF</td>
<td>Linear interpolation/Kernel smoothing + LPF</td>
</tr>
<tr>
<td>Step 2: IPP post-processing</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>Baseline extraction + adaptive peak filter</td>
</tr>
<tr>
<td>Step 3: Calculation of Raw PPV (sample duration)</td>
<td>One ventilation cycle derived from the airway pressure</td>
<td>Three ventilation cycles derived from the ventilator</td>
<td>One ventilation cycle derived from the IPP</td>
<td>Two ventilation cycles derived from the IPP</td>
<td>8 s window</td>
<td>One ventilation cycle derived from the ABP baseline</td>
</tr>
<tr>
<td>Step 4: Calculation of presented PPV (smoothing of raw PPV)</td>
<td>Mean of three consecutive raw PPV values</td>
<td>---</td>
<td>Median of three consecutive raw PPV values</td>
<td>Kalman filter</td>
<td>Mean of four consecutive raw PPV values</td>
<td>Median of three consecutive raw PPV values</td>
</tr>
</tbody>
</table>

$^a$ IPP: interpolated pulse pressure.

$^b$ LPF: low-pass filter.
where $w_0$ is the center frequency and $\Delta w$ is the bandwidth.

The center frequency of this filter (i.e. the ventilation frequency) was estimated by applying a fast Fourier transform (FFT) on the baseline of the ABP signal of the previous 60s. The estimated ventilation frequency was updated every 20s. The bandwidth of the filter was set at 0.04 Hz, which was empirically found to be an effective bandwidth as it stabilized the detrended IPP in a 30s window.

4.3. Step 3: calculation of raw PPV

For our new algorithm, we used the ventilation frequency determined from the ABP baseline modulation derived in step 2 (section 4.2) to set our sample duration, and we used one ventilation cycle for calculating PPV. Furthermore, we used the extracted baseline of the IPP to compute the mean IPP. This approach was less dependent on sample durations compared to the methods described by Aboy et al (2009) and (2004). In these methods, if the length of ventilation cycle (the sample duration) was not determined correctly, the mean IPP would alter significantly because of the cyclic components in the IPP. In contrast, the mean IPP in our method, the extracted baseline of the IPP, was free of cyclic components.

4.4. Step 4: calculation of presented PPV

To suppress the influence of irregular beats and noise, we adopted the same three-point median filter as Aboy et al (2004).

5. Results

5.1. Step 1: the influence of different methods for IPP derivation

An example of the results of various methods for deriving IPP can be found in figure 3. In figure 3(a), it can be seen that there was no significant change in volume status. However, figure 3(b) shows that linear interpolation generated different extrema values for neighboring ventilation cycles. In contrast, the methods based on kernel smoothing constructed extrema of more similar values. This similarity in extrema values offered a promise of reconstructing the actual maximum and minimum of IPP. Nevertheless, figure 4 shows that the algorithms based on kernel smoothing produced more unstable PPV calculation episodes and short-term elevations than those based on linear interpolation. After IPP post-processing, the performance of the two algorithms was comparable.

5.2. Step 2: effect of our IPP post-processing

In order to isolate the influence of IPP post-processing from other confounding factors, we did not post-process the raw PPV values for all the compared algorithms. Figure 4(a) shows that the proposed IPP post-processing significantly reduced the hourly occurrence of unstable PPV calculation episodes independently of the method for IPP derivation and sample durations ($p < 0.001$). Figure 4(b) shows that the proposed IPP post-processing significantly reduced the hourly occurrence of short-term elevations independently of the method for IPP derivation and sample durations ($p < 0.001$).

By extracting the baseline in the IPP, we reduced the hourly occurrence of short-term elevations. Figure 5 shows the effect of extracting the baseline in the IPP. It can be seen that there was an increasing trend in the IPP. The method without baseline extraction suffered from a
false temporary elevation in the PPV values. On the contrary, the method with baseline extraction showed no such short-term elevations in the PPV values and gave stable (physiologically-sensible) output.

By applying an adaptive peak filter, we achieved fewer unstable PPV calculation episodes. Figure 6 shows the effect of applying an adaptive peak filter to the de-trended IPP. It can be seen that the filtering led to more stable de-trended IPP. For example, the PPV values around 3860 s and 3920 s, which seemed to be outliers, were successfully suppressed using the adaptive filter.

The performance of the adaptive peak filtering depended on the accuracy with which the ventilation frequency was estimated. We compared our estimated frequency with the frequency provided by the ventilator. The mean difference was 0.003 Hz (i.e. 0.18 min⁻¹) with a standard deviation (SD) of 0.01 Hz.

5.3. Step 3: the influence of the sample duration

Figure 4 shows the influence of the sample duration for the different algorithms. We compared three sample durations: one ventilation cycle, two ventilation cycles (50% overlapping), three ventilation cycles. We did not include 8 s for comparison, as this sample duration is often between two and three ventilation cycles, so the performance is accordingly between that of two ventilation cycles and that of three ventilation cycles. In order to isolate the influence of sample duration from other confounding factors, we did not post-process the PPV values for all the compared algorithms. Again, we compared the performance using the hourly occurrence of unstable PPV calculation episodes and short-term elevations. It can be seen that the hourly occurrence of short-term elevations increased with the sample duration. Algorithms using two ventilation cycles and three ventilation cycles generated significantly more short-term elevations than those using one ventilation cycle independently of IPP derivation (step 1) and IPP post-processing (step 2). When it comes to hourly occurrence of unstable PPV calculation episodes, the influence of sample duration was not very apparent. Note that our proposed IPP post-processing method helped reduce number of unstable PPV calculation episodes and
short-term elevations regardless of the sample duration and that the sample duration had a stronger effect on the occurrence of short-term elevations than on the occurrence of unstable PPV calculation episodes.

Figure 4. Boxplots of two performance indicators for different configurations of the first three algorithmic steps of the PPV algorithm without PPV post-processing. (a) Hourly occurrence of unstable PPV calculation episodes. (b) Hourly occurrence of short term elevations. IPP post-processing refers to our proposed baseline extraction and adaptive peak filter. The sample durations in step 3 are given as the number of ventilation cycles. The lower the occurrence rates, the better the performance of the algorithms. Significant comparisons are indicated by the horizontal lines where * means $p < 0.05$, ** means $p < 0.01$, and *** means $p < 0.001$. 

short-term elevations regardless of the sample duration and that the sample duration had a stronger effect on the occurrence of short-term elevations than on the occurrence of unstable PPV calculation episodes.
Figure 5. The effect of extracting the IPP’s baseline from the IPP. (a) A segment of the ABP signal. (b) The IPP and its baseline. (c) The de-trended IPP. (d) PPV values. The two compared PPV algorithms in panel (d) both used kernel smoothing to derive IPP, applied sample durations of 8 s, and did not post-process PPV: the only difference was the application of baseline extraction.

Figure 6. The effect of applying an adaptive peak filter to the de-trended IPP. (a) A segment of the ABP signal. (b) The IPP and its baseline. (c) The de-trended IPP and the de-trended IPP processed using an adaptive peak filter. (d) PPV values. The two compared PPV algorithms in the panel (d) both used kernel smoothing to derive IPP, post-processed IPP with baseline extraction, applied sample durations of 8 s, and did not post-process PPV: the only difference was the application of an adaptive peak filter.
Table 3. Performance of existing and proposed algorithms during episodes of hemodynamic stability.

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Difference from the original PPV specification$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kim and Pinsky (2008)</td>
<td>1.15% (0.93%)$^b$</td>
</tr>
<tr>
<td>Aboy et al (2004)</td>
<td>−0.19% (0.75%)</td>
</tr>
<tr>
<td>Aboy et al (2009)</td>
<td>0.42% (0.76%)</td>
</tr>
<tr>
<td>Derichard et al (2009)</td>
<td>0.47% (0.76%)</td>
</tr>
<tr>
<td>Proposed algorithm based on linear interpolation</td>
<td>−0.64% (0.49%)</td>
</tr>
<tr>
<td>Proposed algorithm based on kernel smoothing</td>
<td>−0.11% (0.56%)</td>
</tr>
</tbody>
</table>

$^a$ Applied to episodes of hemodynamic stability.
$^b$ Mean (standard deviation).

5.4. Overall comparisons

The results of overall comparisons can be found in table 3 and figure 7. Table 3 shows that during selected episodes of hemodynamic stability, all existing and the proposed algorithm were in good agreement with the original specification by Michard et al (2000). The mean of difference between the original specification and the other algorithms was between $−0.64\%$ and $1.15\%$, where the proposed algorithm based on kernel smoothing achieved $−0.11\%$, followed by $−0.19\%$ from the work by Aboy et al (2004). The SD of difference between the original specification and the other algorithms was between $0.49\%$ and $0.93\%$, where the proposed algorithm based on linear interpolation achieved $0.49\%$, followed by $0.56\%$ from the proposed algorithm based on kernel smoothing.

Figure 7 shows the performance of existing algorithms and the proposed algorithms during the entire course of the surgery. During the entire course of surgery, the proposed algorithms based on linear interpolation performed the best in avoiding unstable PPV calculation episodes, followed by the proposed algorithm based on kernel smoothing. The proposed algorithm based on linear interpolation showed significantly better performance compared to existing algorithms, except for the algorithm by Aboy et al (2004). When it comes to the hourly occurrence of short-term elevations, the best performance was achieved by the proposed algorithm based on linear interpolation, followed by the proposed algorithm based on kernel smoothing. Both proposed algorithms performed significantly better than existing algorithms, except for the algorithm by Aboy et al (2004).

6. Discussion

In this study, we evaluated five key PPV calculation algorithms existing in the literature on the data obtained in ventilated patients undergoing surgery. To understand the role of each processing step in the performance of the algorithms, we first decomposed the existing algorithms into four steps and analyzed different methods and effects for each of them. To enable this, we developed new clinically relevant measures of PPV calculation performance. Insights into the behavior of each algorithm and the role of each algorithmic step has enabled us to propose an additional IPP post-processing approach, improving the robustness during surgery. The results show that all evaluated algorithms showed good agreement with the original PPV specification proposed by Michard et al (2000) during hemodynamically stable phases in surgery. In more dynamic phases, the proposed IPP post-processing algorithm helped improve the performance.
In the past, little attention has been focused on post-processing IPP. However, various disturbances may exist in the IPP. These include the pulse pressure trend, dyssynchrony of the heart rate to the ventilation rate, and irregular beats. In this light, merely post-processing raw PPV values may perform inadequately, whereas post-processing IPP can help to improve the performance. The results show that our proposed baseline extraction and adaptive peak filtering significantly improved the robustness of the PPV algorithm during surgery.

We found it is preferred to use a sample duration of a single ventilation cycle instead of extended sample durations as using longer sample durations resulted in a larger hourly occurrence of short-term elevations. This is a consequence of the fact that the blood pressure baseline changes had a larger impact on the maximum and minimum of IPP when using longer sample durations. In addition, the PPV values computed using longer sample durations were consistently higher than those computed using shorter windows. This finding is also in line with the work by Kim and Pinsky (2008), in which the authors suggested that separate validations are needed to define the decision threshold. Interestingly, Cannesson et al (2011b) also identified varied thresholds when calculating PPV using commercial patient monitors or when calculating PPV according to the original specification. They further defined a gray zone where PPV is inconclusive.

We also compared two different methods of deriving IPP: linear interpolation and kernel smoothing. As is shown, during periods of hemodynamic stability, the kernel smoothing helped to reconstruct the actual extrema of IPP. However, the algorithm based on linear interpolation outperformed that based on kernel smoothing over the entire course of surgery. It could be that kernel smoothing was less stable in dynamic periods. Nevertheless, the algorithms based on both methods achieved improved robustness when applying the proposed IPP post-processing and the performance differed marginally.

![Boxplots of two performance indicators for comparisons between existing algorithms and the proposed algorithms during the entire course of the surgery. (a) Hourly occurrence of unstable PPV calculation episodes. (b) Hourly occurrence of short-term elevations. The compared algorithms were Michard et al (2000), Aboy et al (2004), Kim and Pinsky (2008), Aboy et al (2009) and Derichard et al (2009), proposed algorithm based on linear interpolation (Prop1), and proposed algorithm based on kernel smoothing (Prop2). Significant comparisons are indicated by the horizontal lines where * means $p < 0.05$, ** means $p < 0.01$, and *** means $p < 0.001$.](image)
There are several limitations in this work. Firstly, here we could only investigate the role of blood pressure changes on PPV calculating performance during surgery. While it would have been useful to understand how cardiac output may affect the PPV calculation performances, cardiac output was however not monitored in these patients. The lack of measurement of cardiac output also hindered us from investigating the predictability of PPV for fluid responsiveness. Future study may include the measurement of cardiac output to test the predictability of all the PPV algorithms for fluid responsiveness during surgery. Secondly, our data size was limited. A sample size of this magnitude made it challenging to demonstrate a small significantly higher performance for a particular technique. Despite this, we still found some significant improvements in the performance indicators when using IPP post-processing. We also found significant difference in generating short-term elevations when using different sample durations. Thirdly, when it comes to the sample window according to the original specification, we defined a sample window rather than using the precise ventilation timing derived from the airway pressure or capnography signals. We believe both methods of obtaining the sample window would not differ too much, as the signal in such short time scale is often assumed to be stable. Finally, when post-processing the IPP, we applied a classic filter in a bi-directional manner to avoid phase delay and distortion; in practice this would require a latency of approximately 4 s.

7. Conclusion

All evaluated algorithms showed good agreement with the original PPV specification proposed by Michard et al (2000) during hemodynamically stable phases in surgery, but the overall performance during surgery depended on specific algorithmic choices. Appropriate interpolation for deriving IPP was helpful for reconstructing actual maximum and minimum of the PP within each ventilation cycle. Extracting the baseline in the IPP can greatly reduce the hourly occurrence of short-term elevations. Applying an adaptive peak filter to the detrended IPP can reduce the hourly occurrence of unstable PPV calculation episodes. The sample duration should not be longer than one ventilation cycle as longer windows generate an increased hourly occurrence of short-term elevations as well as consistently higher PPV values. Finally, our proposed IPP post-processing improved the performance independent of the interpolation method and sample duration during surgery.

Acknowledgments

This work was financially supported by China Scholarship Council (CSC). This work was conducted as part of a collaborative PhD research program, called the Healthcare Flagship, in which the Eindhoven University of Technology, Royal Philips, the Maxima Medical Center Veldhoven, the Catharina Hospital Eindhoven, and the Kempenhaeghe Sleep Center Heeze join forces to create impactful innovations in healthcare. The authors wish to express their gratitude to Arthur Bouwman for providing clinical insights and reviewing the manuscript. The authors wish to extend their gratitude to Jan-Willem Bikker for his advice on statistical analysis.

Appendix A

The validity of PPV, when calculated using its original specification, under relatively stable hemodynamic conditions is well-established (Michard et al 2000, Cannesson et al 2005, Michard 2005, Solus-Biguenet et al 2006). This means that a PPV calculation algorithm should generate values close to the original specification during relatively stable hemodynamic
episodes. This comparative analysis, performed under stable hemodynamic episodes, is illustrated in figure 2(a). We manually selected a segment of 2–3 min without significant changes in the heart rate or ABP from each patient as stable hemodynamic episodes. In total, 64 min data were included as stable periods with PPV values ranging from 0.4% to 13.3%.

The above performance indicator (i.e. the difference from the original specification, in episodes without sudden hemodynamic changes), is not supposed to be extended to the episodes in which acute blood pressure changes occur. To date, there are no existing PPV performance indicators addressing PPV performance under these conditions.

To address this issue, new PPV performance indicators need to be designed to assess the performance in more dynamic hemodynamic episodes. We identified two common patterns in the PPV time-trace that are, as we will argue, not representative of the true, underlying, hemodynamic status. The occurrence of each of these two false patterns were used as two additional PPV performance indicators during surgery.

The first false pattern is unstable PPV values over a short period of time. To quantify such fluctuations, we heuristically defined a time window of 40 s (i.e. approximately 10 ventilation cycles), in which the interquartile range of PPV values should be no larger than 2% (PPV values typically range in the order of 5–20% and change rather slowly). If the fluctuations exceeded this, the episode was labeled as an ‘unstable PPV calculation episode’. The time windows where PPV values consistently increased or decreased were excluded from this category. An illustration of a typical unstable PPV calculation episode can be found in figure 2(b).

Figure A1. Simulated IPP with an increasing trend that leads to artificially elevated PPV values during an episode of gradual changes of the pulse pressure (see also figure 5 for a real clinical example). (a) Simulated ABP, with an episode where the pulse pressure gradually changes over a time span of several ventilation cycles. (b) Derived IPP exhibiting a gradual increase from 40 mmHg to 70 mmHg. The detected extrema in IPP are marked with a circle. (c) Calculated PPV values according to the original PPV specification (using a sample duration of one single ventilation cycle, and a subsequent three-point mean filter (Michard et al 2000)), and an algorithm with an extended sample duration (using a sample duration of three ventilation cycles and no subsequent averaging (Kim and Pinsky 2008)).
We computed the hourly occurrence of these unstable PPV calculation episodes as an indicator inversely related to performance.

The second false pattern is short-term elevations of the PPV and a subsequent decrease (figure 2(c)). To demonstrate that this artifact is inherent to the PPV calculation algorithm, figure A1 shows a simulated example where the pulse pressure gradually rises from 40 to 70 mmHg while the PPV is kept constant to 7.5% before and after the transition. This simulated example reflects a gradual change in pulse pressure while having an unaltered volume status of the patient. However, during the gradual transition from lower to higher pulse pressure, significantly elevated PPV values are computed. Such pattern is unlikely to reflect true volume status: the underlying physiological consideration is that a patient’s volume status cannot change back and forth rapidly within a short time frame. This false pattern can be confusing and misleading to clinicians in that the PPV can jump temporarily to a level higher than a decision boundary (e.g. for determining fluid requirements) while later decreasing back to a level comparable to baseline. If the physician responds by volume loading, then the patient will receive excessive volume, which may lead to complications. In fact, this pattern originates from the mathematical equation for calculating PPV. When pulse pressure is gradually increasing or decreasing, the difference between maximum and minimum pulse pressure is amplified (see figure A2). The effect becomes more pronounced for extended sample durations. We heuristically defined a short-term elevation as a 10–40 s increase in median values larger than 2% compared to the values in preceding 40 s and following 40 s (figure 2(c)). We computed the hourly occurrence of such elevations as an indicator inversely related to performance. The performance indicators, their valid periods, and their clinical relevance are all summarized in table 1.

Appendix B

Step 1: IPP derivation

When calculating PPV, it is important to derive the continuous IPP from the ABP signal. The derivation of the IPP facilitates the analysis and processing of PP.
In order to derive the IPP, peaks and valleys in the ABP signal should be detected. The detected peak and valley locations were denoted by \( p \) and \( v \), respectively:

\[
p = (p_1, p_2, \cdots, p_N)^T \quad (B.1)
\]

\[
v = (v_1, v_2, \cdots, v_N)^T. \quad (B.2)
\]

With these detected extrema, one can readily obtain the beat-wise PP by subtracting the valley values from the peak values. Note that the method of deriving the IPP in the original specification (Michard et al. 2000) is equivalent to the linear interpolation of the PP to obtain the IPP.

The IPP can also be derived using the upper and lower envelopes of the ABP signal based on these detected extrema. Aboy et al. (2004, 2009) proposed a method based on kernel smoothing to obtain the upper and lower envelope and computed the IPP from the difference between these envelopes. In this method, the upper envelope was constructed at the sampling frequency. For each data point of the upper envelope, a weighted average of nearby peak values was computed. The lower envelope was computed in the same way using the valley locations and values. The upper envelope denoted by \( u_e(n) \) and the lower envelope denoted by \( l_e(n) \) are given by

\[
u_e(n) = \frac{\sum_{k=1}^{N} x(p) * g \left( \frac{|nT_s - t(k)|}{\sigma_g} \right)}{\sum_{k=1}^{N} g \left( \frac{|nT_s - t(k)|}{\sigma_g} \right)} \quad (B.3)
\]

\[
l_e(n) = \frac{\sum_{k=1}^{N} x(v) * g \left( \frac{|nT_s - t(k)|}{\sigma_g} \right)}{\sum_{k=1}^{N} g \left( \frac{|nT_s - t(k)|}{\sigma_g} \right)} \quad (B.4)
\]

where \( T_s = 1/f_s \) is the sampling period with \( f_s \) corresponding to the sampling frequency of the ABP signal, \( t(k) \) is the \( k \)th peak/valley locations, \( \sigma_g \) is the Gaussian kernel width, and \( g(u) \) is a clipped Gaussian kernel function, given by

\[
g(u) = \begin{cases} 
\exp \left( -\frac{u^2}{2} \right), & \text{if } -5 \leq u \leq 5 \\
0, & \text{otherwise} \end{cases} \quad (B.5)
\]

The kernel width determines the smoothing effect and it depends on the heart rate, which can be derived from the ABP signal. Experimentally, a kernel width of 0.2 s was found to work well for heart rates up to 4 Hz (240 bpm). After constructing the upper and lower envelopes, the IPP was obtained by subtracting the upper envelope \( u_e(n) \) by the lower envelope \( u_l(n) \). To reject noise in the process of generating the IPP, the 2004 algorithm by Aboy et al. (2004) applied a non-casual elliptic low-pass filter (LPF). The cutoff frequency of the filter was chosen empirically as 1.75 \times \text{respiratory frequency}.

**Step 2: IPP post-processing**

The existing algorithms discussed in this paper did not apply any post-processing to the IPP (see table 2). To the best of our knowledge, our algorithm is the first algorithm including post-processing of the IPP (see table 2 and section 4).
**Step 3: calculation of raw PPV**

According to the original specification, each PPV value was computed over the sample window from the onset of each breath till the end of this breath. However, this method requires a continuous recording of the airway pressure or capnography signal and signal synchronization, which is often not available in practice. As a surrogate, Aboy et al. (2004, 2009) estimated ventilation cycles from the IPP. These cycles were determined by two neighboring local valleys in the IPP. After deriving ventilation cycles, one such cycle was selected as the sample window in their original algorithm (Aboy et al. 2004), whereas two such cycles with 50% overlap were chosen as the sample window in their updated algorithm (Aboy et al. 2009). Other approaches have also been developed in the absence of an airway pressure signal. Kim and Pinsky (2008) used the knowledge of the ventilator to set the period of three ventilation cycles, over which PPV was calculated. Sample windows of fixed length such as 8 or 10 s were used by Derichard et al. (2009) and Addison et al. (2015), respectively.

In the PPV formula, the difference between the PP$_{\text{max}}$ and PP$_{\text{min}}$ is normalized by their average. This normalization is designed to compute the relative variation of PP. The average of PP$_{\text{max}}$ and PP$_{\text{min}}$ is equivalent to the mean IPP during the ventilation cycle. During episodes of dynamic blood pressure changes, the mean IPP is more robust than the average of PP$_{\text{max}}$ and PP$_{\text{min}}$. Thus, both of algorithms by Aboy et al. (2004, 2009) used the mean IPP as the normalization factor.

**Step 4: calculation of presented PPV**

In order to obtain reliable and robust PPV values, several methods have been proposed to post-process PPV values. The original specification by Michard et al. (2000) applied a three-point mean filter, while the 2004 algorithm by Aboy et al. (2004) applied a three-point median filter. The 2009 algorithm by Aboy et al. (2009) applied a Kalman filter working as follows. The underlying PPV value was regarded as the state variable, and the most recently measured PPV value was taken as the measurement. The algorithm updated the PPV value using the domain knowledge that PPV evolves slowly over time. The simplified formula is given by

$$\hat{p}_{n+1|n+1} = \hat{p}_{n+1|n} + K_{n+1} (y_{n+1} - \hat{p}_{n+1|n})$$  \hspace{1cm} (B.6)

$$p_{n+1} = p_n + u_n$$  \hspace{1cm} (B.7)

$$K_{n+1} = \begin{cases} k_1, & \text{if } |e_{n+1}| = |y_{n+1} - \hat{p}_{n+1|n}| \leq \epsilon_1 \\ k_2, & \text{if } \epsilon_1 \leq |e_{n+1}| = |y_{n+1} - \hat{p}_{n+1|n}| \leq \epsilon_2 \\ k_3, & \text{if } |e_{n+1}| = |y_{n+1} - \hat{p}_{n+1|n}| \geq \epsilon_2 \end{cases}$$  \hspace{1cm} (B.8)

where $\hat{p}_{n+1|n}$ is the predicted PPV value at time $n + 1$, $\hat{p}_{n+1|n+1}$ is the updated PPV value at time $n + 1$ using the measurement $y_{n+1}$, $K = (k_1, k_2, k_3)$ and $T = (\epsilon_1, \epsilon_2)$ are determined empirically to be $K = (1, 0.5, 0)$ and $T = (1\%, 25\%)$.

If the difference between the predicted PPV value and the calculated PPV value was beyond physiological variability of 25%, the calculated value was discarded and the predicted value was used as the state at that moment. If the difference was between 1% and 25%, an average of the predicted and calculated values was used. If the difference was smaller than 1%, the calculated PPV value was taken as the true PPV value.

Another post-processing method using a four-point mean filter on the PPV values derived using 8 s sample duration was mentioned in the algorithm of Derichard et al. (2009).
The algorithm by Addison et al. (2015) included intensive post-processing (e.g. selecting values in the interquartile range in a 120 s smoothing window), improving stability but resulting in a latency of up to 2 min.

References

Avolio A P, Butlin M and Walsh A 2010 Arterial blood pressure measurement and pulse wave analysis— their role in enhancing cardiovascular assessment Physiol. Meas. 31 R1–47
Kim H K and Pinsky M G 2008 Effect of tidal volume, intrathoracic pressure, and cardiac contractility on variations in pulse pressure, stroke volume, and intrathoracic blood volume Crit. Care Med. 36 255–62
Michard F 2005 Changes in arterial pressure during mechanical ventilation Anesthesiology 103 419–28
Monnet X and Teboul J L 2013 Assessment of volume responsiveness during mechanical ventilation: recent advances Crit. Care 17 217
Osman D, Ridel C, Ray P, Monnet X, Anguel N, Richard C and Teboul J L 2007 Cardiac filling pressures are not appropriate to predict hemodynamic response to volume challenge Crit. Care Med. 35 64–8
Vincent J L and Weil M H 2006 Fluid challenge revisited Crit. Care Med. 34 1333–7