Assessing the Diagnostic Accuracy of Pulse Pressure Variations for the Prediction of Fluid Responsiveness

A “Gray Zone” Approach

Maxime Cannesson, M.D., Ph.D.,* Yannick Le Manach, M.D., Ph.D.,† Christoph K. Hofer, M.D.,‡ Jean Pierre Goarin, M.D.,§ Jean-Jacques Lehot, M.D., Ph.D.,|| Benoît Vallet, M.D., Ph.D.,# Benoît Tavernier, M.D., Ph.D.#

ABSTRACT

Background: Respiratory arterial pulse pressure variations (PPV) are the best predictors of fluid responsiveness in mechanically ventilated patients during general anesthesia.

Despite a strong predictive value, the gray zone approach applied to pulse pressure variation may be inconclusive in approximately 25% of patients for prediction of fluid responsiveness in mechanically ventilated patients during general anesthesia.
Gray Zone Approach Applied to Pulse Pressure Variation

DYNA MIC variables relying on cardiopulmonary interactions have been shown to be the best predictors of fluid responsiveness in patients during general anesthesia and mechanical ventilation. Today, these variables can be continuously monitored using dedicated devices and they can be derived noninvasively from the pulse oximeter waveform. Several studies have suggested that these variables could be used for intraoperative goal-directed fluid management and that this approach may have a beneficial impact on patients’ outcome in terms of a reduction in postoperative complications resulting in a decrease in length of hospital stay. Thus far, studies evaluating the ability of respiratory arterial pulse pressure variations (PPV) to predict fluid responsiveness have been conducted in small samples of selected patient groups from single centers, using a receiver operating characteristic (ROC) curve approach. However, most quantitative tests do not perfectly discriminate between subjects with and without a given status (i.e., fluid responder or nonresponder in the current case), and their results do not allow certainty in the determination of this status for screening purposes. The “gray zone” approach has been proposed to avoid the binary constraint of a “black-or-white” decision of the ROC curve approach that often does not fit the reality of clinical or screening practice. The gray zone technique proposes two cutoffs that constitute the borders of the gray zone. The first cutoff allows exclusion of the diagnosis (fluid responsiveness in the current case) with near certainty (i.e., privilege sensitivity and negative predictive value), whereas the second cutoff is chosen to include the diagnosis with near certainty (i.e., privilege specificity and positive predictive value). Intermediate values included in the gray zone correspond to a prediction not precise enough for diagnostic decision. A second potential limitation of the ROC curve analysis for evaluation of dynamic variables is that most previous studies indicate it necessitates a fixed definition of fluid responsiveness (an increase of more than 10% or 15% in stroke volume [SV] or cardiac output [CO]) after a 250- or 500-ml fluid challenge, whereas according to the Frank-Starling curve, the response to a given volume load is actually a continuum of values ranging from “no increase” (or even a decrease) to a “large increase” in SV and/or CO. In addition, the benefit-risk balance of fluid administration may vary between patients, and this has never been taken into account in previous analyses of prediction of fluid responsiveness. The application of these approaches to a large population of patients would allow knowing: (1) the proportion of cases within the gray zone, (2) the range of changes in SV or CO that can be expected in these patients when a fluid challenge is presented, and (3) refinement of the definition of thresholds that should be used for PPV-guided management protocols.

Thus, the aim of our study was to assess the diagnostic accuracy of PPV for prediction of fluid responsiveness in a large, multicenter study of mixed groups of patients in the perioperative period and to express its predictive value using a gray zone approach. We also studied the impact of fluid responsiveness definition and a model for benefit-risk assessment of fluid administration on the predictive value of PPV.

Materials and Methods

Institutional review board (Comité de Protection des Personnes Hospices Civils de Lyon, Lyon, France; Comité de Protection des Personnes Paris-Ile de France, France; Comité de Protection des Personnes Nord Ouest, Lille, France; and Institutional Review Board, Triemli City Hospital, Zurich, Switzerland) approvals were obtained. Patients were included either as part of clinical trials (for whom written and informed consents were obtained) or as part of routine clinical care (for whom no randomization and only routine care was performed, so waived informed consent was authorized).

Data Collection

Four European institutions (Hôpital Louis Pradel, Lyon, France; Hôpital La Pitié-Salpêtrière, Paris, France; Triemli City Hospital, Zurich, Switzerland; and Centre Hospitalier Universitaire de Lille, Lille, France) participated in this study. We defined preload responsiveness evaluation as an intravenous volume load of 500 ml colloid solution given over 10–20 min, immediately preceded and followed (2–5 min later) by hemodynamic measurements, including PPV and CO (or SV), performed with the aim of measuring the change in CO (or SV) induced by volume expansion. Each investigator collected every sequence of preload responsiveness evaluation that was prospectively recorded and available in his own database, provided that (1) the patient was an adult, with no history of arrhythmia, right ventricular failure, valvular heart disease, or intracardiac shunt, (2) he was undergoing general anesthesia, muscle paralysis, and mechanical ventilation in the controlled volume mode, (3) measurements and volume loading were performed in the operative room or in the early postoperative period in the closed-chest condition, and (4) hemodynamic and ventilatory data chosen for analysis in the current study had been obtained in predefined conditions, as described in Data Acquisition and Experimental Protocol. When more than one sequence of preload responsiveness evaluation was available for a given patient, only the first fluid challenge per patient was used for analysis.

Data Acquisition and Experimental Protocol

All patients had a 3- or 5-French catheter inserted in radial or femoral artery for arterial blood pressure monitoring. Pres-
sure transducers were leveled at the midaxillary line and fixed to keep the transducer at the atrial level throughout the study. All transducers were zeroed to atmospheric pressure.

CO was measured in all patients (1) by thermodilution via a pulmonary artery catheter (Swan-Ganz catheter, 7.5-French; Edwards Lifesciences, Irvine, CA) with correct position of the pulmonary artery catheter in West’s zone 3 assessed using the method of Teboul et al., using the average of five successive measurements obtained by injection of 10 ml dextrose at room temperature randomly during respiratory cycle; or (2) by the pulse contour method using a 4-French thermistor-tipped arterial catheter (Pulsiocath™ thermistor catheter, Pulsion Medical Systems, München, Germany) inserted into the left femoral artery and connected to a stand-alone PiCCOplus™ or PiCCO2 monitor (Pulsion Medical Systems). In that case, continuous CO measurement was initiated after initial calibration of the system by three injections of 20 ml ice-cold normal saline into the central venous catheter (transpulmonary thermodilution); or (3) via transesophageal echocardiography: all echo-Doppler data were recorded (Vivid Q, GE Healthcare®, Milwaukee, WI) by trained observers and analyzed off-line by another investigator (JPG), who was not aware of the characteristics of the patients. The aortic valve area was calculated from the left ventricular outflow tract diameter measured at the insertion of the aortic cusp as

left ventricular outflow tract area (cm²) = \pi \times (LVOTd^2/4). \quad (1)

The left ventricular outflow tract diameter was only measured before the fluid challenge to decrease the risk of error. The mean Doppler SV was obtained by multiplying the left ventricular outflow tract area by the aortic flow time velocity integral using the formula

SV (ml) = \text{left ventricular outflow tract area (cm²)} \times \text{aortic flow time velocity integral (cm)}. \quad (2)

The mean Doppler SV (mean SV₃) was calculated as the average of the consecutive beats of a complete respiratory cycle (5–7 beats). Aortic Doppler calculation of CO was obtained by multiplying the mean SV by the heart rate; or (4) esophageal Doppler (Hemosonic® 100; Arrow, Reading, PA) allowing measurements of aortic diameter and aortic blood flow, adjusted to obtain the best aortic blood velocity signal and not moved until the end of the fluid challenge, as reported previously.¹

The following variables were collected in all patients both before and after volume expansion from the value provided by the monitor screen: heart rate, mean arterial pressure, SV, CO, and systemic vascular resistance. End-expiratory central venous pressure (CVP) was also recorded when available. PPV was obtained online using the PiCCO monitor or was calculated off-line from the arterial pressure waveform recorded on a personal computer using data acquisition software over three consecutive respiratory cycles and averaged for analysis, as described previously.²² Off-line recordings of PPV were analyzed with the reviewer unaware of the hemodynamic data. Ventilatory settings and airway pressures were also recorded.

All patients were studied after 2–3 min hemodynamic stability with no changes in anesthetic protocol or ventilator settings and no volume expansion. Baseline hemodynamic measurements were obtained and then followed by an intravenous volume expansion consisting of 500 ml colloid solution (hetastarch 6% or modified fluid gelatin) given over 10–20 min. Hemodynamic measurements were performed within 2–5 min after volume expansion.

**Statistical Analysis**

All individual data were pooled for analysis. Results are expressed as mean ± SD and median (95% CI) for nonnormally distributed variables (normality was assessed with the D’Agostino-Pearson omnibus test) or numbers (percentages). Comparison of two means was performed using unpaired Student t tests; comparison of two medians was performed using the Mann–Whitney test; and comparison of proportions was performed using the Fisher exact method.

Patients were divided into two groups according to the percentage increase in CO after volume expansion. In the first part of the analysis, responders were defined as patients with an increase of 15% or more in CO, and nonresponders were defined as patients with an increase of less than 15% in CO. The discriminative power of PPV and CVP was then assessed using ROC curves. These curves were obtained by averaging 1,000 populations bootstrapped (sampling with replacement) from the original study population. This method limits the impact of outliers and allows the provision of more robust representations. Box plots were used to depict CI of the average ROC curves. Comparison of two areas under the ROC curves was performed using the nonparametric technique described by DeLong et al.²³

**Threshold and Gray Zone Determination**

A gray zone was used to describe the value for which the variable of interest did not provide conclusive information. As mentioned, the gray zone corresponds to a range of values for which formal conclusions could not be obtained. To determine this interval of values, we first assessed the “optimal” threshold using Youden’s index (J = Sensitivity + Specificity – 1 = Sensitivity – False-Positive Rate). Maximizing J corresponds to maximizing the overall correct classification rates. It minimizes misclassification rates. We did not retain the choice of the closest point to (0,1) because it involves a quadratic term, the clinical meaning of which remains unknown. Youden’s index determination was then conducted for each bootstrapped population, resulting in a set of 1,000 “optimal” values. The mean value of these optimal values and its 95% CI were then estimated. Thus, the gray zone was defined as well as its 95% CI. This first approach was completed by an alternative approach, aiming at

---

Anesthesiology 2011; 115:231–41 Cannesson et al.
defining three classes of response: negative, inconclusive, positive. We defined inconclusive responses for values with a sensitivity lower than 90% or specificity lower than 90% (diagnosis tolerance of 10%). To better present these results, a two-curve (sensitivity, specificity) representation was provided. Nonparametric curves and values were estimated using an artificial neural network (a fast and robust method to fit the sensitivity and specificity curves). These two approaches were not exclusive but complementary; thus, the largest CI was defined as the gray zone.

### Misclassification Cost Term

In some cases, the clinician could privilege either sensitivity or specificity because the expected consequence of false-negative or false-positive results is not equivalent in terms of benefit-risk balance. For example, it is likely more important to avoid unnecessary volume loading (*i.e.*, which does not increase CO) in a patient at high risk of acute lung injury than in a patient with preserved left ventricular function with any lung pathology who has not yet received volume expansion. The previous analysis did not take into account these possibilities. Thus, we introduced a ratio of costs (R) defined as: $R = \text{cost (false positive)/cost (false negative)}$. In this study, R less than 1 denotes that not treating a false-negative result is worse than treating a false-positive one (which would characterize a “liberal” fluid control, where nonmaximized CO [potential risk: organ hypoperfusion] is considered worse than unnecessary fluid loading [potential risk: fluid overload]); R = 1 is equivalent to the previous analyses, and R more than 1 denotes that to treat a false-positive result is worse than missing a false-negative one (characterizing a “restrictive” or tight fluid control, as may apply to patients for whom unnecessary volume loading is considered potentially more deleterious than nonmaximized CO).

### Multiendpoint Analyses

Because the primary endpoint of this study (an increase of 15% or more in CO) was only one of the arbitrary endpoints chosen in previous studies, we repeated the determination of thresholds and gray zone for each increase in CO, percent by percent, between 5% and 25%. Each determination was conducted as described for the 15% increase in CO.

All $P$ values are two-tailed, and values $<0.05$ were considered to denote significant differences. Statistical analysis was performed with R software** and specific packages.24

### Results

#### General Data

We studied 413 patients (Hôpital La Pitié-Salpêtrière, Paris: 190 patients; Hôpital Louis Pradel, Lyon: 82 patients; Trémiâlis City Hospital, Zurich, Switzerland: 80 patients; Centre Hospitalier Universitaire de Lille, France: 61 patients) (table 1). Type of surgery was abdominal aorta surgery in 201 (49%) cases, cardiac surgery (before chest opening) in 162 (39%) cases, and abdominal (nonvascular) surgery in 50 (12%) cases (table 1). CO was measured using pulmonary artery catheter in 129 patients, the pulse contour method in 126 patients, and aortic echo-Doppler in 158 patients (table 1). The whole group consisted of 307 (74%) men and 106 (26%) women between 31 and 90 yr (mean age, 65 ± 9 yr; mean height, 171 ± 12 cm; mean weight, 77 ± 13 kg; body surface area, 2.4 ± 0.2 m²). Respiratory rate was 13 ± 2 breaths/min with an average heart rate/respiratory rate ratio of 5.3 ± 1.5. Only 38 patients (9%) had a heart rate/respiratory rate ratio less than 3.6. Average tidal volume was 7.9 ± 1.3 ml/kg body weight, with 212 patients (51%) being ventilated with a tidal volume of 8 ml/kg or more. Average peak airway pressure was 20 ± 5 cm H₂O (minimum 8 cm H₂O, maximum 45 cm H₂O), and positive end-expiratory pressure was 3 ± 2 cm H₂O (minimum 0 cm H₂O, maximum 10 cm H₂O).

#### Table 1. Demographic Data According to the Center

<table>
<thead>
<tr>
<th></th>
<th>Lille</th>
<th>Lyon</th>
<th>Paris</th>
<th>Zurich</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients (n)</td>
<td>61</td>
<td>82</td>
<td>190</td>
<td>80</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>60 ± 12</td>
<td>66 ± 10</td>
<td>66 ± 7</td>
<td>66 ± 9</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>169 ± 9</td>
<td>170 ± 9</td>
<td>172 ± 8</td>
<td>172 ± 8</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>70 ± 14</td>
<td>76 ± 15</td>
<td>77 ± 11</td>
<td>82 ± 12</td>
</tr>
<tr>
<td>Gender (F/M)</td>
<td>40/150</td>
<td>25/55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of surgery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiac (n)</td>
<td>0</td>
<td>82</td>
<td>0</td>
<td>80</td>
</tr>
<tr>
<td>Digestive (n)</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Vascular (n)</td>
<td>11</td>
<td>0</td>
<td>190</td>
<td>0</td>
</tr>
<tr>
<td>Cardiac output measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermodilution (S.G.)</td>
<td>47</td>
<td>82</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pulse contour (Picco)</td>
<td>0</td>
<td>0</td>
<td>46</td>
<td>80</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>14</td>
<td>0</td>
<td>144</td>
<td>0</td>
</tr>
</tbody>
</table>

Data are presented as mean ± SD.
S.G. = Swan-Ganz.

** Available at: www.r-project.org. Accessed May 6, 2011.
The ability of CVP and PPV to predict fluid responsiveness was tested using ROC curve analysis (fig. 2A). A clear significant difference of discriminative power was retrieved between them (areas under ROC curve: 0.89 [0.86; 0.92] for PPV, and 0.57 [0.54; 0.59] for CVP; all \( P < 0.001 \)), confirming the well-known superiority of PPV over CVP to predict fluid responsiveness.

The optimal threshold to predict fluid responsiveness was then determined as the one that minimized the explicit cost ratio \( (R = 1) \) (fig. 2B), which is equivalent to maximizing the Youden’s index. By determining it for all the 1,000 resampled populations, we obtained a distribution for the optimal thresholds of the 1,000 samples (fig. 2C). The median of these optimal thresholds was 12%, and the 95% CI of this distribution (the gray zone) was 8.7–12.9%. Using the alternative approach of splitting the ROC curve into two curves of sensitivity and specificity (fig. 2D), an inconclusive zone spreading from 9.3% to 12.5% was retrieved. Considering that the determination of the gray zone remains a stochastic process, which is prone to produce slightly different results between assessments, and that the precision of the measurements of PPV does not exceed one unit, the results of the two approaches were merged as a gray zone spreading from 9% to 13%. The frequencies of responders in each class of PPV values, as well as the median magnitude of the increase of CO, are presented in table 3. Gray zone limits were also calculated for each center, CO measurement method, type of surgery, and time of measurement. This analysis shows a large overlap in all gray zones depending on these parameters, as shown in figure 3.

### Changes in Gray Zone Limits According to the Cost Ratio

To depict specific clinical situations in which the risk balance of fluid loading might not be equilibrated, we defined two opposite strategies: (1) tight fluid control, including patients for whom an excessive fluid loading would be two times more deleterious than nonoptimal CO maximization (cost ratio \( = 2 \)); and (2) liberal fluid control, in which nonmaximized CO would be two times more deleterious than an excessive fluid loading (cost ratio \( = 0.5 \)). The gray zones for these two alternative strategies (fig. 4) were, respectively 11–14% and 8–11%.

### Changes in Gray Zone Limits According to the Increase in CO Considered Clinically Significant after Fluid Loading

The increase in CO considered clinically significant after a fluid loading remains undetermined. In fact, the minimal increase of CO detectable by modern devices might be very low with two assessments conducted by the same observer in the same patient. Nevertheless, no direct proof of clinical benefits exists for these

---

**Table 2.** Hemodynamic Data in Responders and Nonresponders to Volume Expansion (Fluid Responsiveness Defined as More Than 15% Increase in Cardiac Output after Volume Expansion)

<table>
<thead>
<tr>
<th></th>
<th>Responders to Volume Expansion (n = 209)</th>
<th>Nonresponders to Volume Expansion (n = 204)</th>
<th>( P ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart rate (bpm)</td>
<td>70 ± 20</td>
<td>65 ± 15</td>
<td>0.003</td>
</tr>
<tr>
<td>Mean arterial pressure (mmHg)</td>
<td>70 ± 12</td>
<td>76 ± 12</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Central venous pressure (mmHg)</td>
<td>8 ± 5</td>
<td>9 ± 4</td>
<td>0.03</td>
</tr>
<tr>
<td>PPV (%)</td>
<td>16 ± 6</td>
<td>8 ± 4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Stroke volume (ml)</td>
<td>69 ± 20</td>
<td>81 ± 23</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cardiac output (l/min)</td>
<td>4.7 ± 1.4</td>
<td>5.2 ± 1.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SVR (dyn/s/cm(^5))</td>
<td>1,502 ± 731</td>
<td>1,699 ± 663</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Data are presented as mean ± SD. 

bpm = beats per minute; PPV = respiratory variations in systemic arterial pulse pressure; SVR = systemic vascular resistance.
low increases in CO. Figure 5 was constructed using the following method and reproduced for each value of increase in CO between 5% and 25%. A crescent relationship was retrieved between the percentage change in CO used to define the response to volume expansion and the median of threshold for PPV. Furthermore, the width of the gray zones was relatively conserved among the range of value tested to define fluid responsiveness. In fact, if an increase of 7% in CO was considered clinically relevant, the gray zone for PPV would no longer be 9–13% (calculated for an increase of 15%) but 5–8%.

**Discussion**

This study is the first to test the application of the gray zone concept to PPV for the prediction of fluid responsiveness and to do so in a large sample. Our results show that (1) this approach identifies a range of PPV values, between 9% and 13%, for which fluid responsiveness cannot be reliably predicted; (2) such gray zone values may be seen in approximately 25% of patients in the operating room; (3) the gray zone limits change according to the fluid management strategy to be applied (11–14% for tight fluid control strategy and 8–11% for liberal fluid control strategy); and (4) there is a crescent relationship between the percentage change in CO used to define response to volume expansion and the median threshold of PPV.

Our results confirm that PPV is an accurate predictor of fluid responsiveness and that it theoretically may be useful for perioperative hemodynamic optimization, as suggested by several recent studies. PPV has been studied for several years as a predictor of fluid responsiveness in relatively small samples of selected patients in standardized conditions. In contrast, the current results were obtained in a large pop-
ulation of mixed groups of patients, recruited in four different centers and studied in conditions of ventilation and hemodynamic management corresponding to usual clinical practice. Most of the studies evaluating the ability of PPV to predict the effects of volume expansion on CO used a ROC curve analysis to define the optimal threshold allowing for maximum sensitivity / specificity value. Although this approach has been used for years to test a diagnostic tool,\(^{18}\) it may not accurately reflect the decision-making process applied to clinical management.\(^{19}\) This approach provides a single cutoff that dichotomizes the population (responders and nonresponders in this setting). The gray zone approach provides two cutoffs. The first cutoff is chosen to exclude the diagnosis (responder to volume expansion in the current setting) with near certainty (i.e., privileges specificity), whereas the second cutoff is chosen to include the diagnosis with near certainty (i.e., privileges sensitivity). Consequently, this approach clearly outlines the lowest and highest threshold values that allow solid clinical decision-making on fluid administration in the operating room. When PPV decreases into the gray zone between the two cutoffs, uncertainty exists, and clinicians should pursue a diagnosis using additional tools. In our study, uncertainty was clearly attested, in a narrow range of PPV values (9 –13%), with an equal proportion of responders and nonresponders (50%) and significant differences between changes in CO observed in responder (me-

| PPV below the lower limit of the gray zone | 159 (38%) | 18 (11%) | 5% (−13%, 27%) | 19% (15%, 55%) | 5 (−13%; 15%) |
| PPV into the gray zone | 98 (24%) | 49 (50%) | 15% (−7%, 47%) | 24% (16%, 49%) | 4% (−14%, 14%) |
| PPV above the upper limit of the gray zone | 156 (38%) | 142 (91%) | 28% (4%; 62%) | 30% (17%, 62%) | 8% (−8%, 15%) |
| All patients | 413 (100%) | 209 (51%) | 15% (−10%; 53%) | 27% (15%, 55%) | 5% (−13%, 15%) |

CO = cardiac output; PPV = pulse pressure variation.

Table 3. Frequencies of Responders and Magnitude of the Variation of Cardiac Outputs According to the Strata of PPV Values

Fig. 3. Pulse pressure variations (PPV) a and gray zone ranges according to the center, the method used for cardiac output (CO) measurement, the time of measurement, and the type of surgery. Red circles = PPV threshold; red horizontal lines = gray zone ranges. The black diamond at the bottom of the figure shows this information for the whole population. S.G. = Swan-Ganz catheter.
Median: 24%) versus nonresponder (median: 4%) patients after the 500-ml fluid bolus, as shown in table 3.

Recently, Marik et al. conducted a meta-analysis evaluating the ability of PPV to predict fluid responsiveness. In the study, the authors found that the mean threshold for PPV was 12.5 ± 1.6%, providing a sensitivity of 89% and a specificity of 88%, with an area under the ROC curve of 0.94. It should be appreciated that this value of 12.5 ± 1.6% does not represent a range of gray zone values (with its usefulness for the decision-making process) but only the range of uncertainty of the estimated best cutoff value. In addition, despite general similarities, the conduct of meta-analysis for diagnostic tools studies differs from meta-analyses of randomized control trials in several important ways. First, the assessment of study quality for diagnostic studies varies considerably from randomized control trials. Individual studies on the same diagnostic tool can vary considerably on the choice of threshold used, the population under study, and even the measurement of the variable or reference standard. The choice of recruitment strategy for patients can also affect the assessment, with one study finding that recruiting patients and controls separately can lead to an overestimation of the test’s diagnostic accuracy. The statistical techniques used to aggregate the results of diagnostic tool studies also

![Fig. 4. Gray zone determination for pulse pressure variation (PPV) according to a tight fluid control policy (cost ratio: R = 0.5) or a large fluid control policy (cost ratio: R = 2). The average evolutions of the explicit cost in the 1,000 resampled populations are depicted for each policy (A and B). The distributions of the optimal cutoffs are depicted by histograms (C and D). Hatch rectangles = 95% CI; Red lines = medians of the optimal cutoffs.](image1)

![Fig. 5. Evolution of the gray zone determined by the resampling approach with an equilibrated cost ratio (R = 1) according to the increase of cardiac output (CO) defined as clinically significant (between 5% and 25%). PPV = pulse pressure variation.](image2)
differ from the meta-analyses of randomized control trials. The meta-analysis of diagnostic studies requires the consideration of two index measures (e.g., sensitivity and specificity), as opposed to a single index in the meta-analysis of a randomized control trial.\(^{26}\) It is also expected that heterogeneity in the indices will be observed from several different sources,\(^{29}\) and this heterogeneity must be considered in the statistical model used to pool the estimates.\(^{27}\) For these reasons, meta-analyses of diagnostic tools such as PPV are not as powerful as for randomized control trials.

Apart from the gray zone approach, we also found that PPV cutoffs vary according to the cost ratio. In essence, we found that depending on whether clinicians aim at conducting a tight fluid control or a large fluid control, PPV cutoffs and gray zone change. Consequently, clinicians should consider different thresholds, depending on whether they want to be more sensitive for fluid responsiveness detection or more specific (according to the clinical scenario).

Recently, several studies have suggested that PPV could be used for goal-directed therapy in patients undergoing surgery.\(^{13–17}\) These studies aimed at maintaining PPV under a given threshold (between 10 and 15%) by administering iterative boluses of fluid. Our results show that this approach may be improved because, according to our results, 24% of our patients presented with a PPV value in the gray zone when volume expansion was performed (table 3). Consequently, using the gray zone instead of a single threshold value for conducting goal-directed therapy may improve fluid management of these patients and make protocols more accurate. One may consider fluid titration in patients whose PPV value is in the gray zone, using a smaller amount of fluid and observing the change in PPV after titration. More specifically, it could be suggested from our results that fluid may be titrated to reach a PPV value of less than 9%, at least when the cost ratio is ≤ 1. When the cost ratio is estimated to be more than 1, no fluid may be given and spontaneous evolution of PPV observed. In other cases, when CO is monitored, we also suggest that fluid boluses may be used when the PPV value is in the gray zone. In this situation, the effects of volume expansion on CO probably would be the best way to assess the preload dependence status of the patient. This approach would reconcile the concepts of CO maximization and dynamic indicator minimization. However, one has to remember that to make this approach meaningful, the CO monitoring device used should be able to accurately track changes in CO induced by volume expansion. For this purpose and because it has been used in several outcome studies, the esophageal Doppler can be recommended. However, other devices, such as those used in the current study (thermodilution, transesophageal echocardiography, and pulse contour immediately preceded by calibration via transpulmonary thermodilution), may also be used for this purpose.\(^{28–30}\)

Finally, our results demonstrate that the definition of fluid responsiveness has an impact on the PPV threshold value and on the gray zone. As shown in figure 5, the gray zone is minimized when fluid responsiveness is defined as an increase in CO between 15% and 20%. When lower thresholds are used to define fluid responsiveness, the gray zone for PPV is increased. This is probably related to the lack of precision of CO monitoring devices for detecting such low variations in CO. Consequently, one may consider using a 15–20% increase in CO to define fluid responsiveness in studies testing the ability of a variable to predict the effects of volume expansion on CO. However, one has to remember that in the current study, volume expansion was performed using a 500-ml bolus.

Our study presents some limitations. First, PPV was used in patients during general anesthesia and mechanical ventilation who presented with no cardiac arrhythmia and no right ventricular failure. Therefore, the well-known limitations of PPV have not been challenged. However, we included some patients with a tidal volume less than 8 ml/kg body weight, and this had no major impact on the predictive value of PPV. The lowest tidal volume in our patient was 6 ml/kg. Consequently, we can postulate that if the study had been limited to patients with a tidal volume of at least 8 ml/kg body weight, the ability of PPV to predict fluid responsiveness would have been improved. Second, the discriminative power of PPV was assessed using ROC curves obtained by averaging 1,000 populations bootstrapped (sampling with replacement) from the original study population. This method limits the impact of outliers and allows the provision of more robust representations but is not equivalent to a study focusing on an original population, including the same number of patients. However, it would have been practically impossible to include the number of patients requested to obtain these results. Thus, this procedure should be considered as an internal validation but not an external validation, which requires additional studies. Third, we classified responder and nonresponder patients using various methods of CO measurement, all of which have unique errors of measurements and limited clinical agreement between them, as recently demonstrated by a meta-analysis conducted by Peyton and Chong.\(^{31}\) However, all the techniques used in our study have demonstrated high validity for monitoring changes in CO during patient management, and recent studies suggest that they can be used to track changes accurately in CO.\(^{28,30}\) Accordingly, all of them have been used for determination of fluid responsiveness in previously published studies.\(^{4,9,29}\) Our results emphasize this point because we found an overlap of gray zone ranges according to the type of CO monitor used. In addition, the meta-analysis by Marik et al. showed that previously published studies found similar threshold values for PPV with these different CO monitoring devices.\(^{2}\) Fourth, we did not adjust volume expansion to patients’ weight or body surface area to be in line with methods of previously published studies. We found approximately 50% of responders to volume expansion (presenting with a wide range of change in CO), which is consistent with the findings of most previously published studies on the topic.
Moreover, previously published studies using different volume challenges found similar thresholds for dynamic parameters of fluid responsiveness. Finally, we pooled patients from different centers who originally were included in different studies or studied in routine clinical conditions. This approach is not as strong as a single study including the same number of patients, but then again practically it would be very difficult to conduct such a study in a single center. In addition, the overlap found in gray zones according to center, type of surgery, and time of measurements may constitute a strength of the study because it reinforces its potential relevance for clinical practice in various conditions. However, additional studies focusing on well-defined patient populations, specific settings, and conducted with standardized CO monitors may help to define better the clinical applicability of PPV.

In conclusion, the gray zone approach applied to PPV for prediction of fluid responsiveness in mechanically ventilated patients during general anesthesia identifies a range of PPV values, between 9% and 13%, for which fluid responsiveness cannot be reliably predicted. Such PPV values may be seen in approximately 25% of patients. Thus, using the gray zone instead of a single-threshold value for conducting goal-directed therapy may improve perioperative fluid management.

References


ANESTHESIOLOGY REFLECTIONS

An Advertising Card for Peruvian Syrup

Not made in Lima or Cusco, so-called “Peruvian Syrup” was actually manufactured by J. P. Dinsmore of New York and distributed from Boston by Seth W. Fowl & Son. By combining cocaine and “protoxide of iron,” Dinsmore’s potion promised to transform “weakly, sickly, suffering creatures” into “strong, healthy, happy men and women.” Through advertising cards (above), the aggressive promoters of Peruvian Syrup insisted that it “vitalizes and enriches the blood, tones up the system.”

(Permission © the American Society of Anesthesiologists, Inc. This image also appears in the Anesthesiology Reflections online collection available at www.anesthesiology.org.)

George S. Bause, M.D., M.P.H., Honorary Curator, ASA’s Wood Library-Museum of Anesthesiology, Park Ridge, Illinois, and Clinical Associate Professor, Case Western Reserve University, Cleveland, Ohio. UJYOC@aol.com.