



The “grey zone” or how to avoid the binary constraint of decision-making

Maxime Cannesson, MD, PhD

Received: 5 August 2015 / Accepted: 16 August 2015 / Published online: 22 August 2015
© Canadian Anesthesiologists' Society 2015

In this issue of the *Journal*, Vos *et al.* report on the accuracy of pulse pressure variation (PPV), stroke volume variation (SVV), and stroke volume index (SVI) to predict fluid responsiveness at multiple thresholds.¹ In their study, they show that the threshold used to define fluid responsiveness impacts the predictive value of these indices and the range of their zones for decision-making (i.e., the “grey zone”)—the lower the threshold for the definition of fluid responsiveness, the lower the predictive value and the wider the grey zone. Although these results are not completely surprising (i.e., how one defines an outcome generally impacts the accuracy of the tools used to predict that outcome) or entirely new (i.e., the concept of using a PPV grey zone - though perhaps not these authors' precise definition - for understanding fluid responsiveness had already been reported),² they nevertheless remind us of the importance of using the grey zone methodology when approaching any diagnostic tool (such as PPV and SVV). Accordingly, this editorial serves as a reminder of the relevance, impact, and global meaning of the grey zone for diagnostic tools.

The predictive value of continuous diagnostic indices such as PPV are best evaluated using the receiver operating characteristic (ROC) curve approach. Very briefly, this approach determines an optimal threshold that provides the highest combination of sensitivity and specificity for a given diagnostic tool. In addition, the accuracy of a test depends on the ability of the test to separate the group being tested into those with and those without the disease, and it is quantified by the area under the ROC curve which

ranges from 0 to 1. An area of 1 represents an ideal test, while an area of 0.5 represents a worthless test (i.e., same predictive value as flipping a coin). While this approach has been used for years to assess the accuracy of diagnostic tools³ (such as the ability of PPV to separate responders from non-responders to fluid administration), its main limitation is that it transforms the biological nature of a continuous variable into an artificially dichotomous (i.e., “black or white”) statistical index that does not always accurately reflect the decision-making process applied to clinical management.⁴ Indeed, the very reason we originally proposed using the grey zone approach to evaluate PPV was to avoid this type of binary constraint.

The grey zone technique proposes two numerical cutoffs that constitute its borders. The first cutoff is used to exclude the diagnosis (e.g., a 9% PPV where fluid responsiveness is not present) with near certainty (i.e., privilege sensitivity and negative predictive value), whereas the second cutoff is chosen to include the diagnosis (e.g., a 13% PPV where fluid responsiveness is present) with similar near certainty (i.e., privilege specificity and positive predictive value).⁴ Intermediate values representing the grey zone correspond to a prediction that is too imprecise for a diagnostic decision, referring to what Feinstein called “*the inadequacy of binary models for the clinical reality of three-zone diagnostic decisions*”.⁵ Indeed, for PPV, the first study published on the topic found a threshold of 13% to predict fluid responsiveness, an associated sensitivity of 94%, a specificity of 96%, and an area (standard deviation) under the ROC curve of 0.98 (0.03).⁶ Most subsequently published studies found similar results.⁷ Nevertheless, a few years later when we studied the predictive value of PPV using a grey zone approach, we found a similar area under the curve, but the grey zone ranged from 9–13% and

M. Cannesson, MD, PhD (✉)
Department of Anesthesiology & Perioperative Care, University
of California, Irvine, CA, USA
e-mail: mcanness@uci.edu

about a quarter of the patients were inside the grey zone.² In the intensive care unit, a recent study by Biais *et al.* found an even wider grey zone for PPV with more patients lying in the grey zone.⁸

Despite these results, it is important to remember that the grey zone approach does not invalidate the predictive value of an index such as PPV and does not actually reflect the full “clinical relevance” of an index. Just because the PPV value is inside the grey zone does not mean that it cannot be used to predict fluid responsiveness. And contrary to what many think,⁹ the grey zone approach is not meant to invalidate PPV. Its main goal is to create a more rational, more clinically oriented, and less theoretical procedure. As a matter of fact, all clinicians working at the patients’ bedside on a daily basis know that decision-making never relies on a single parameter alone. Clinicians know that decision-making in medicine is based on an association of signs converging in a similar direction and leading to the final decision. When it comes to hemodynamic management, PPV is only one parameter among a myriad of others, and it is an illusion to think that this index alone would be sufficient to determine whether a patient should receive fluids.

The one line of reasoning we learn from the grey zone is as follows: if the PPV value is more than 13% (when recorded in appropriate conditions), it is very likely that the patient is a fluid responder, and if the PPV is less than 9%, it is very likely that the patient is not a fluid responder. When PPV is 9–13%, there are two options: wait for the PPV to move one direction or the other or stress the system (i.e., give fluid) and see how PPV varies.¹⁰ If PPV decreases, it is more likely that the patient increased their stroke volume with fluid administration.¹⁰ Fluid should be given for some when PPV is in the grey zone, and the impact of volume expansion on stroke volume should be evaluated.

In their manuscript, Vos *et al.* add another brick to understanding how PPV can be used clinically and how the grey zone helps better articulate this index in daily clinical decision-making. They show that the definition of fluid responsiveness impacts the predictive value of PPV and the range of the grey zone; that is, the higher the threshold to define fluid responsiveness, the tighter the grey zone. In any case, whatever its range, the grey zone does not inform whether PPV or SVV are clinically meaningful. Rather, the grey zone informs clinicians on how to use these indices in clinical practice and on when to gather more information in order to make a more rational decision. It brings these indices back to clinical reality and hopefully helps clinicians understand the limits and benefits of these parameters. In doing so, it aids in avoiding what can otherwise be constraining binary decisions.

La « zone grise » ou comment éviter la contrainte binaire de la prise de décision

Dans ce numéro du *Journal*, Vos *et coll.* publient un article sur l’exactitude des variations de la pression du pouls (PPV), des variations du volume d’éjection systolique (SVV) et du volume d’éjection systolique indexé (SVI) pour prédire la réactivité aux liquides à de multiples seuils.¹ Ils montrent dans leur étude que le seuil utilisé pour définir la réactivité aux liquides a un impact sur la valeur prédictive de ces indices et sur l’étendue de leurs zones pour la prise de décision (c’est-à-dire, la « zone grise »)—plus le seuil pour la définition de la réactivité aux liquides est bas, plus la valeur prédictive est faible et plus la zone grise est large. Bien que ces résultats ne soient pas totalement surprenants (la façon dont on définit un aboutissement a habituellement un impact sur la précision des outils utilisés pour prédire cet aboutissement), ni entièrement nouveaux (le concept consistant à utiliser la zone grise de la PPV, même si cela n’est pas la définition exacte de ces auteurs, pour la compréhension de la réactivité aux liquides a déjà été décrit),² ils nous rappellent néanmoins l’importance de l’utilisation de la méthodologie de la zone grise lorsque l’on aborde un outil diagnostique, quel qu’il soit (comme la PPV et la SVV). En conséquence, cet éditorial sert de rappel sur la pertinence, l’impact et la signification globale de la zone grise des outils diagnostiques.

La valeur prédictive d’indices diagnostiques continus tels que la PPV est le mieux évaluée quand on utilise l’approche de la courbe ROC (fonction d’efficacité du récepteur). Très brièvement, cette approche établit un seuil optimal qui procure la plus forte combinaison de sensibilité et de spécificité pour un outil diagnostique donné. De plus, l’exactitude d’un test dépend de la capacité du test à distinguer dans le groupe testé ceux qui ont ou n’ont pas la maladie; il est quantifié par la surface entre la courbe ROC qui est comprise entre 0 et 1. Une surface de 1 représente un test idéal, alors qu’une surface de 0,5 représente un test sans valeur, c’est-à-dire qu’il a une valeur prédictive équivalente au lancer d’une pièce de monnaie. Alors que cette approche est utilisée depuis des années pour évaluer l’exactitude des outils diagnostiques³ (tel que la capacité de la PPV à différencier les patients répondant à l’administration de liquides de ceux qui n’y répondent pas), sa principale limite réside dans le fait qu’elle transforme la nature biologique d’une variable continue en un indice statistique artificiellement dichotomique

(« noir ou blanc ») qui ne reflète pas toujours avec exactitude le processus de prise de décision appliqué à la gestion clinique.⁴ En fait, la vraie raison pour laquelle nous avons initialement proposé cette approche de la zone grise pour évaluer la PPV était d'éviter de type de contrainte binaire.

La technique de la zone grise propose deux seuils numériques qui constituent ses limites. Le premier seuil sert à exclure le diagnostic (une PPV de 9 % où il n'y a pas de réactivité aux liquides) avec une quasi-certitude (c'est-à-dire qu'on privilégie la sensibilité et la valeur prédictive négative) alors que le second seuil est choisi pour inclure le diagnostic (une PPV à 13 % où la réactivité aux liquides est présente) avec une quasi-certitude comparable (c'est-à-dire, qu'on privilégie la spécificité et la valeur de prédiction positive).⁴ Les valeurs intermédiaires représentant la zone grise correspondent à une prédiction qui est trop imprécise pour prendre une décision diagnostique, renvoyant à ce que Feinstein appelait « l'inadéquation des modèles binaires pour la réalité clinique des décisions diagnostiques à trois zones ». ⁵ En fait, pour la PPV, la première étude publiée sur le sujet a trouvé un seuil de 13 % pour prédire la réactivité aux liquides, une sensibilité associée de 94 %, une spécificité de 96 %, et une surface (écart-type) sous la courbe ROC de 0,98 (0,03).⁶ La plupart des études publiées ultérieurement ont trouvé des résultats semblables.⁷ Mais, quand nous avons étudié la valeur prédictive de la PPV quelques années plus tard en utilisant l'approche de la zone grise, nous avons trouvé une surface similaire sous la courbe, mais la zone grise était comprise entre 9 % et 13 %, et environ un quart des patients se trouvaient dans cette zone grise.² En unité de soins intensifs, une étude récente de Biais *et coll.* a trouvé une zone grise pour la PPV encore plus vaste et plus de patients reposant dans cette zone grise.⁸

Malgré ces résultats, il est important de se rappeler que l'approche de la zone grise n'invalide pas la valeur prédictive d'un indice tel que la PPV et ne reflète pas complètement la « pertinence clinique » d'un indice. Ce n'est pas parce que la valeur de la PPV se situe à l'intérieur de la zone grise qu'elle ne peut pas être utilisée pour prédire la réactivité aux liquides. Et contrairement à ce que beaucoup pensent,⁹ l'approche de la zone grise n'est pas destinée à invalider la PPV. Son principal objectif est de créer une procédure plus rationnelle, plus cliniquement orientée et moins théorique. En fait, tous les cliniciens travaillant quotidiennement au chevet des patients savent qu'une prise de décision ne repose jamais sur un seul et unique paramètre. Les cliniciens savent qu'en médecine, la prise de décision repose sur une association de signes convergeant dans la même direction et menant à la décision finale. Quand on en vient à la gestion hémodynamique, la

PPV n'est qu'un paramètre parmi une myriade d'autres et il serait illusoire de croire que ce seul indice pourrait être suffisant pour déterminer si un patient doit recevoir des liquides.

La ligne unique de raisonnement que nous pouvons tirer de la zone grise est la suivante : si la valeur de la PPV est supérieure à 13 % (quand enregistrée dans des conditions convenables), il est probable que le patient est répondeur aux liquides et si la PPV est inférieure à 9 %, il est très probable que le patient ne sera pas un répondeur aux liquides. Quand la PPV est comprise entre 9 % et 13 %, il y a deux possibilités : attendre que la PPV se déplace dans une direction ou l'autre, ou stimuler le système (c'est-à-dire, administrer des liquides) et voir comment la PPV varie.¹⁰ Si la PPV diminue, il est plus probable que le patient ait augmenté le volume d'éjection systolique avec l'administration de liquides.¹⁰ Du liquide devrait être administré à certains quand la PPV est dans la zone grise et l'impact de l'expansion du volume sur le volume d'éjection systolique devrait être évalué.

Dans leur manuscrit, Vos *et coll.* ajoutent une autre pierre à la compréhension de la raison pour laquelle la PPV peut être utilisée en clinique et comment la zone grise aide à mieux situer cet indice dans la prise de décision clinique quotidienne. Ils montrent que la définition de la réactivité aux liquides a un impact sur la valeur prédictive de la PPV et sur l'étendue de sa zone grise; c'est-à-dire que plus le seuil de définition de la réactivité aux liquides est élevé, plus la zone grise est serrée. Dans tous les cas, quelle que soit son étendue, la zone grise n'informe pas sur le fait de savoir si la PPV ou la SVV sont cliniquement pertinentes. La zone grise informe plutôt les cliniciens sur la façon d'utiliser ces indices en pratique clinique et sur quand rassembler davantage d'informations pour prendre une décision plus rationnelle. Cela ramène ces indices à la réalité clinique et, espérons-le, aide les cliniciens à comprendre les limites et avantages de ces paramètres. De cette façon, cela aide à éviter ce qui pourrait être autrement des décisions binaires contraignantes.

Conflicts of interest None declared.

Conflicts d'intérêts Aucun déclaré.

References

1. Vos JJ, Poterman M, Papineau Salm P, et al. Noninvasive pulse pressure variation and stroke volume variation to predict fluid responsiveness at multiple thresholds: a prospective observational study. *Can J Anesth* 2015; 62: this issue. DOI:10.1007/s12630-015-0464-2.
2. Cannesson M, Le Manach Y, Hofer CK, et al. Assessing the diagnostic accuracy of pulse pressure variations for the prediction

- of fluid responsiveness: a “gray zone” approach. *Anesthesiology* 2011; 115: 231-41.
3. *Feldman JM*. Is it a bird? Is it a plane? The role of patient monitors in medical decision making. *Anesth Analg* 2009; 108: 707-10.
 4. *Ray P, Le Manach Y, Riou B, Houle TT*. Statistical evaluation of a biomarker. *Anesthesiology* 2010; 112: 1023-40.
 5. *Feinstein AR*. The inadequacy of binary models for the clinical reality of three-zone diagnostic decisions. *J Clin Epidemiol* 1990; 43: 109-13.
 6. *Michard F, Boussat S, Chemla D, et al*. Relation between respiratory changes in arterial pulse pressure and fluid responsiveness in septic patients with acute circulatory failure. *Am J Respir Crit Care Med* 2000; 162: 134-8.
 7. *Marik PE, Cavallazzi R, Vasu T, Hirani A*. Dynamic changes in arterial waveform derived variables and fluid responsiveness in mechanically ventilated patients: a systematic review of the literature. *Crit Care Med* 2009; 37: 2642-7.
 8. *Biais M, Ehrmann S, Mari A, et al*. Clinical relevance of pulse pressure variations for predicting fluid responsiveness in mechanically ventilated intensive care unit patients: the grey zone approach. *Crit Care* 2014; 18: 587.
 9. *Michard F, Chemla D, Teboul JL*. Applicability of pulse pressure variation: how many shades of grey? *Crit Care* 2015; 19: 144.
 10. *Le Manach Y, Hofer CK, Lehot JJ, et al*. Can changes in arterial pressure be used to detect changes in cardiac output during volume expansion in the perioperative period? *Anesthesiology* 2012; 117: 1165-74.