Prediction of fluid responsiveness in spontaneously breathing patients

Xavier Monnet1,2, Jean-Louis Teboul1,2

1Hôpitaux Universitaires Paris-Saclay, Assistance Publique - Hôpitaux de Paris, Hôpital de Bicêtre, Service de Médecine Intensive-Réanimation, Le Kremlin-Bicêtre, France; 2Inserm UMR S_999, Univ Paris-Saclay, Le Kremlin-Bicêtre, France

Contributions: (I) Conception and design: None; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: None; (V) Data analysis and interpretation: None; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Prof. Xavier Monnet. Service de médecine intensive-réanimation, Hôpital de Bicêtre, 78 rue du Général Leclerc, Le Kremlin-Bicêtre, F-94270 France. Email: xavier.monnet@aphp.fr.

Abstract: In patients with acute circulatory failure, the primary goal of volume expansion is to increase cardiac output. However, this expected effect is inconstant, so that in many instances, fluid administration does not result in any haemodynamic benefit. In such cases, fluid could only exert some deleterious effects. It is now well demonstrated that excessive fluid administration is harmful, especially during acute respiratory distress syndrome and in sepsis or septic shock. This is the reason why some tests and indices have been developed in order to assess “fluid responsiveness” before deciding to perform volume expansion. While preload markers have been used for many years for this purpose, they have been repeatedly shown to be unreliable, which is mainly related to physiological issues. As alternatives, “dynamic” indices have been introduced. These indices are based upon the changes in cardiac output or stroke volume resulting from various changes in preload conditions, induced by heart-lung interactions, postural manoeuvres or by the infusion of small amounts of fluids. The haemodynamic effects and the reliability of these “dynamic” indices of fluid responsiveness are now well described. From their respective advantages and limitations, it is also possible to describe their clinical interest and the clinical setting in which they are applicable.

Keywords: Volume expansion; passive leg raising (PLR); heart-lung interactions


doi: 10.21037/atm-2020-hdm-18

View this article at: http://dx.doi.org/10.21037/atm-2020-hdm-18

Introduction

Fluid administration is the first therapeutic measure in the majority of cases of acute circulatory failure. It is intended to increase cardiac output and oxygen delivery (1). However, apart from cases in which the absolute or relative hypovolemia is very deep, such as during haemorrhage, significant extracellular dehydration or septic shock in the initial phase, the administration of fluid only leads to a significant increase in cardiac output in half of the cases (2). In cases where preload dependence is absent, volume expansion does not have a beneficial effect, but exerts deleterious effects which are today well demonstrated (3). Therefore, before administering it, predicting whether or not a bolus of fluid will increase cardiac output is a strategy that reasonably reduces fluid administration and avoids the harmful effects of fluid when it has no beneficial effect.

To detect preload dependence and predict fluid responsiveness, several tests and indices have been developed over the past twenty years (4). They all consist of inducing or observing variations in cardiac preload, and measuring the resulting changes in cardiac output or stroke volume. The magnitude of these changes allows one to predict which changes will be induced by the fluid infusion.

What are the tests and indices that can be used in patients who are not intubated, or intubated but who have
some spontaneous breathing activity? How should these tests be performed in practice? What are their limitations and the conditions under which they can be used? What is the overall strategy in which they should be implemented? These are the questions we will answer in this review, based on the most recent data in the literature.

**What is the concept of fluid responsiveness prediction?**

Fluid administration in patients with acute circulatory failure is intended to increase the mean systemic pressure, which is the upstream pressure of systemic venous return, to increase cardiac preload, and therefore to improve cardiac output (5). However, in the early 1980s, it appeared that the increase in stroke volume and cardiac output only occurs in half of the patients who receive a fluid bolus (6). The explanation is probably physiological. The Frank-Starling relationship between cardiac preload and stroke volume is curvilinear, and the response to volume expansion can only occur in a state of preload dependence, i.e., if the slope of the relationship is sufficiently steep (4). However, this slope, which is determined by the contractile function of the two ventricles, varies from patient to patient, and it is impossible to determine it from basic clinical data. This is the first part of the problem.

The second part, highlighted a little later (7), is that the administration of fluid in critical patients is deleterious. Fluids have a multitude of deleterious effects, particularly in intensive care patients with sepsis, lung damage, kidney failure or abdominal problems (3). The fluid balance is even a factor which is linked to the mortality of these patients independently of the other gravity factors (7). Therefore, it appears that administering fluid to a patient who is not “responder” to fluid is not only ineffective, but harmful. Like any drug with inconsistent efficacy and significant side effects, it appears necessary to administer fluid only if one is almost certain that it will be effective. The concept of prediction of response to volume expansion is based on the idea that fluid should only be administered in critically ill patients in a state of dependence preload. This should reduce the fluid balance, and ultimately improve the prognosis for these patients.

**The static approach must be abandoned!**

For years, it was on central venous pressure (CVP) and other “static” markers of cardiac preload that the decision was made whether or not to give boluses. However, it has been clearly demonstrated in a considerable number of studies that this strategy does not work (8). Perhaps outside of extreme values, a given level of cardiac preload does not predict the response to volume expansion. It is clear from a figure on which are superimposed curves of cardiac function with different slopes, that a given value on the abscissa axis does not determine the slope of the curve and therefore the degree of preload dependence. In this spontaneously breathing population, to which we are interested in this review, another reason is that barometric markers are difficult to measure in the event of irregular and rapid breathing. On a CVP or pulmonary artery occlusion pressure curve, it can be difficult to distinguish expiration, during which the intravascular pressure should be measured to overcome the influence of intrathoracic pressure.

Despite physiological reality and the large number of studies that have demonstrated it, it is very surprising to find that CVP still guides many intensivists in their decision to administer fluid or not (9). In this regard, it must be said that this strategy has been recommended for septic patients for many years by the Surviving Sepsis Campaign (10,11).

In contrast to this “static” method, the “dynamic” approach for detecting dependence preload is based on the observation of changes in stroke volume or cardiac output which result from changes in cardiac preload, observed spontaneously or induced by specific tests. Some of these tests and indices cannot be used in spontaneously breathing patients. The respiratory variation of arterial pulse pressure is very reliable, but can be used only in case of regular mechanical ventilation with no spontaneous cycle (12). This is also the case of the variation of the diameter of the inferior or superior venae cavae, which is in anyway a less reliable index of preload responsiveness (13). Nevertheless, in spontaneously breathing patients, a study has suggested that the changes in inferior vena cava diameter induced by a standardised deep inspiration predict fluid responsiveness reliably (14).

**Fluid challenge**

**What is it?**

The easiest way to test dependence preload *a priori* is to administer fluid and measure the increase in cardiac output it induces. This “fluid challenge” method has been used *de...*
facto for many years.

**How to do it in practice?**

The “classic” fluid challenge consists of the administration of 300 to 500 mL of fluid over 30 minutes. The crystalloid and colloid solutions may be suitable (15), but the recommendation not to use colloid should be reminded here in patients with septic shock (16). The effectiveness of the fluid challenge must be evaluated above all on the reversion of the criterion which initiated it: drop in diuresis, skin mottling, increase in lactate, etc. (15). However, the effects on cardiac output should be estimated on a direct measurement of cardiac output as they are poorly reflected by simultaneous changes in blood pressure or in heart rate (17). This is the case even when the arterial pulse pressure, physiological reflection of the stroke volume, is considered (18,19).

Finally, it is recommended by the proponents of the method to set safety limits, in particular to avoid fluid overload. A CVP limit of 15 mmHg can be used (15).

**What are the limitations?**

The most obvious limitation of the fluid challenge is that it is not a “test”, but the treatment itself. In the event of fluid unresponsiveness, it is impossible to withdraw from the patient the fluid which has been administered but which is ineffective. Inherently, managing the fluid therapy with the fluid challenge induces fluid overload. This is particularly the case if the test must be repeated several times, which may occur in the first hours of acute circulatory failure in many patients.

**Mini-fluid challenge**

**What is it?**

The principle is based on the very limitation of the “standard” fluid challenge. The idea is to administer not 300–500 mL of fluid, but only a few tens of millilitres of colloid or crystalloid. The response of cardiac output to this low volume is used to predict the effects of more important volumes of fluid. The test is based on the hypotheses that a small volume of fluid can significantly increase cardiac preload and that this increase in preload is sufficient to test the preload dependence of the two ventricles (20). Of course, the test assumes that changes in cardiac output can be measured despite their small amplitude.

**How to do it in practice?**

The first study which tested the validity of a “mini fluid challenge” reported the injection of 100 mL of hydroxyethyl starch over a few minutes, the effects of which were measured by changes in velocity-time integral (VTI) in transthoracic echocardiography (21).

In the studies that followed, volumes of 50 to 150 mL were tested, with colloids and crystalloids (22,23). However, some studies using lower volumes have shown less reliability (24).

**What are the limitations?**

The first limitation of the mini fluid challenge is that, even if its volume is less than that of the classic fluid challenge, it persists that repeating it several times in a few hours in a patient can only lead to an increase in the total fluid balance.

The second limitation is related to the technique used to measure changes in cardiac output. Indeed, small volumes of fluid can only induce small changes in cardiac preload, which can only induce small increases in cardiac output even in the case of preload responsiveness. The threshold reported to define the positivity of the test is also low (21). This implies that the technique that measures cardiac output must be very precise. From this point of view, it should be remembered that the smallest significant change in VTI that ultrasound can measure is only 10% (25). In comparison, the pulse contour analysis can detect changes in cardiac output as low as 1.3% (26), and may be more suitable for the mini-fluid challenge (24). It has been demonstrated that the decrease in pulse pressure variation induced by a mini fluid challenge could detect preload responsiveness, but the study was performed in deeply sedated patients (27). Whether the method is possible in spontaneously breathing patients should be investigated.

**End-expiratory occlusion test**

**What is it?**

This test is based on the haemodynamic effects of mechanical ventilation and can only be used in intubated patients. However, unlike the variations induced in pulse
pressure, stroke volume or diameter of the venae cavae, it can be used in patients who have slight spontaneous respiratory activity.

Under positive pressure ventilation, each insufflation of the ventilator increases the intrathoracic pressure, and this increase is transmitted to the right atrium through its thin and compliant free wall. This cyclically causes a drop in the pressure gradient of systemic venous return (mean systemic pressure - right atrial pressure) and a drop in cardiac preload. Thus, stopping ventilation at end-expiration stops the cyclic impediment in systemic venous return (28). During the expiratory pause, the right cardiac preload increases. The increase in preload is transmitted from the right to the left side. If, in response, stroke volume and cardiac output increase, the two ventricles are preload dependent (28).

Since the first study published in 2009 (29), a substantial number of other publications have come to support the validity of the end-expiratory (EEXPO) test. Several of them have been included in recent meta-analyses, concluding that the test is valid (23). The threshold for increase in cardiac output that defines positivity is 5%.

How to do it in practice?

In a patient under mechanical ventilation, a first value of the cardiac output is measured. The patient’s condition must be stable enough for this value to be considered a reliable reference. Ventilation is stopped at the end of expiration, with the same procedure as that usually used to measure intrinsic positive end-expiratory pressure. Importantly, the hold should be at least 15 seconds (Figure 1). The reason is that devices that measure cardiac output continuously do so on a moving average of several seconds, tending to delay the on-screen appearance of the maximum value reached by cardiac output. This maximum value of cardiac output measured at the end of the 15 seconds is noted, and the percentage of change compared to the pre-pause value is calculated (28).

The technique for measuring cardiac output must meet two criteria (28). The first is that it must be able to detect rapid changes in cardiac output occurring in real time. The second criterion is that the technique must be precise enough to measure changes in cardiac output of only a few percent.

Pulse contour analysis, which is a very precise technique, is perfectly suited for the test, and has been used in several studies. On the contrary, cardiac echography and other ultrasonic techniques may be less reliable for this test because they lack precision. From this point of view, our group has shown that, to be used with echocardiography, the 15-second EEXPO test must be associated with a 15-second end-inspiratory occlusion (EIXPO) test, separated from the first by a few seconds, the while the patient’s condition stabilises again (30). In the case of preload-dependence, the subaortic VTI increases during the EEXPO and decreases during the EIXPO pause in a greater manner than in the case of preload-independence (Figure 1). If the effect (in absolute value) of these two manoeuvres is added, the test detects preload dependence with a positivity threshold of 15%. If, on the other hand, only the changes in VTI are considered during the EEXPO pause, the sensitivity and specificity are correct, but the diagnostic threshold is only 4%, which is too low compared to the accuracy of the echocardiography (30). Similar results have been demonstrated with oesophageal Doppler, which suffers from the same lack of precision as echocardiography (31).

What are the limitations?

Some studies have suggested that the EEXPO test was less reliable in patients with a tidal volume of 6 mL/kg rather than 8 mL/kg (32,33). However, this result was not found in all the studies that demonstrated the reliability of the test when they included a large number of ventilated patients with low tidal volume. It seems that the level of positive end-expiratory pressure does not influence the reliability of the EEXPO test (33,34).

Of course, the test cannot be used in patients who are unable to support a breathing pause as long as 15 seconds, that is, if the patients have too much spontaneous breathing activity. Also, as stated above, the test requires direct measurement of cardiac output. Indeed, if the changes in arterial pulse pressure can detect the changes in cardiac output in this circumstance, these changes cannot be easily assessed on the bedside monitors.

Passive leg raising (PLR)

What is it?

When transferring a patient from the semi-recumbent position at 30–45°, to a position in which the trunk is horizontal and the lower limbs raised at 30–45°, a portion of the venous blood stagnating
in the lower limbs and in the vast splanchnic territory is transferred to the heart chambers (35). The resulting increase in cardiac preload mimics the effects of a fluid challenge. PLR has in fact been shown to cause a significant increase in mean systemic filling pressure (5). Unlike a fluid challenge, however, the PLR test has the major advantage of being reversible when the patient is returned to the semi-recumbent position (36). Compared to tests using heart-lung interactions, the PLR test has the advantage that it can be used also in patients without mechanical ventilation or ventilated but with spontaneous breathing cycles.

A now large number of studies have shown that the test is reliable. The threshold for increasing cardiac output used for positivity is 10% (35,37). The last version of the Surviving Sepsis Campaign recommends to use PLR for guiding fluid therapy in patients with septic shock (16).

**How to do it in practice?**

A first measurement of cardiac output, stroke volume, or their surrogate is collected at the base, making sure that the value considered is fairly stable. Simple rules must then be followed for the test to be reliable (Figure 2). First, it must be started from the semi-recumbent position (38). Started from the horizontal position, the test does not mobilize the vast reservoir of splanchnic venous blood, making the test less sensitive than when it begins with the trunk tilted at 30–45° (39). Rather than performing the test by lifting manually lower limbs holding the heels of the patient, it is better to use the automatic movements of the electric bed (38). This prevents possible pain from creating false positives. The maximum value of cardiac output or stroke volume, which generally appears in less than a minute, is noted and the percentage increase compared to the value measured before the test is calculated. After the test, it must be checked that the cardiac output or stroke volume return to its baseline value when the patient has been returned to the semi-recumbent position. This makes sure that the baseline value considered was indeed stable (38).

Above all, the test must be performed by measuring cardiac output or stroke volume directly (38). When its effects are measured on blood pressure, the test is less reliable, with a significant proportion of false negatives. This has been reported by several studies (35).

The technique used to measure cardiac output must be able to measure changes that occur over a short period of time. Indeed, the effects may decrease after reaching
the maximum value, in certain particularly vasodilated patients and in whom capillary leakage is significant. Thus, pulmonary or transpulmonary thermodilution is not appropriate because of the time required to repeat cold boluses injection. The pulse wave contour analysis is particularly simple to use. With echocardiography, the changes in VTI are proportional to those in cardiac output, without the need to measure the area of the left ventricle outflow tract.

Bioreactance can be used when the last version of the system is used, because it is more reactive to changes in cardiac output than the previous one (40). There is some doubt that the effects of the test can be detected via changes in carotid or femoral flow measured by arterial Doppler, because positive and negative results are found in the literature (41-43).

Capnography, which can detect changes in cardiac output using those of carbon dioxide at the end of expiration, is an interesting technique because it is not invasive. However, in this case the ventilation must be perfectly stable (44), so that changes in carbon dioxide in exhaled gas are only due to changes in cardiac output. The method is not suitable for patients with spontaneous ventilation, to whom this review is devoted.

The assessment of the haemodynamic status through clinical examination is gaining more and more interest (45,46). One study reported that the test could be performed by measuring cardiac output from capillary refill time instead (47). However, to ensure the reproducibility of the measurement, the capillary refill time was measured according to a standardised method which cannot be used in current practice (47). Automating the method may make the test easier.

Recently, our group reported that changes in the plethysmography perfusion index (ratio between the pulsatile portion and the non-pulsatile portion of the signal) are able to track changes in cardiac output during the PLR test, so that the changes in this index induced by the PLR test are capable of detecting the preload-dependence (48). These results undoubtedly need to be confirmed. Also, the perfusion index signal was unstable in some patients, but it is an interesting opportunity (48).

**Limitations**

First of all, as we have seen, performing the PLR test...
requires a direct measurement of cardiac output, which is the main limitation in practice (38). Then, the test is contraindicated in case of intracranial hypertension. Intra-abdominal hypertension compresses the splanchnic territory, probably hinders blood transfer from the lower limbs to the thorax and certainly reduces the splanchnic blood volume that can be mobilised by the manoeuvre. Indeed, intra-abdominal hypertension seems to be a condition in which false negatives appear on the PLR test (49).

The test is not feasible in prone positioning patients, and a reverse Trendelenburg manoeuvre has been described as a reliable alternative to the PR test in these cases (50).

**When and how to use these tests and indices?**

It should always be remembered that preload responsiveness is a physiological state and that patients should not receive fluid for the sole reason that the indices or tests of preload dependence are positive. Two questions must be asked beforehand. The first is whether there are arguments to believe that the cardiac output is too low compared to the oxygen requirements of the organism. Hyperlactatemia, a lowering of central venous oxygen saturation, a decrease in urine flow or an increase in the veno-arterial carbon dioxide gap are, for example, arguments for this.

The second question is whether the risks of fluid administration are not greater than the benefits that can be expected. The increase in the fluid balance is a deleterious condition which must be avoided (3). Fluids are treatments that are both inconsistent and dangerous. As with all treatments of this type used in frail patients, we must carefully predict their effectiveness and estimate the risk associated with their use. The level of extravascular lung water and pulmonary permeability, the level of CVP, the ratio between the arterial partial pressure of oxygen and the inspired fraction of oxygen, the level of intra-abdominal pressure are undoubtedly indices which help to assess the risk of volume expansion (3).

It should be borne in mind that there are conditions in which preload responsiveness is obviously constant, and in which the fluid must be administered without considering any of the indices and tests that we have described. In the event of obvious fluid or blood loss, or in the initial phase of sepsis, when no fluid has yet been administered, the implementation of tests and indices of fluid responsiveness would only lead to a dangerous delay in treatment (4).

Finally, if tests are used in the resuscitation phase to decide whether to administer fluid or not, they can be very useful in the de-escalation phase. In this phase, removing fluid is often a therapeutic objective, and the question that arises in this context is to know what volume should be removed without causing hemodynamic impairment. Testing preload responsiveness allows one to undertake depletion only in the case where it is certain that the decrease in preload will not reduce the cardiac output. At the time of weaning from mechanical ventilation, the PLR test has been shown to reliably predict that a spontaneous breathing trial will lead to weaning-induced pulmonary oedema (51).

**Positive test, negative test?**

No index or any of the tests that detect preload responsiveness is perfect. First, they all have their own terms and conditions of use, as we have seen. Then, even under the optimal conditions of their use, their sensitivity and specificity is not perfect.

In addition, none of the diagnostic thresholds proposed should be considered in an absolute manner. There is necessarily a grey zone, in which the sensitivity and specificity are not absolute (52). This may be linked to the unreliability of the test, but also to the lack of precision of the measurement method used to estimate its effects. Finally, it should be borne in mind that the relationship between cardiac preload and cardiac output is curvilinear, and is not a biphasic relationship. There is a continuum between the state of preload responsiveness and that of preload responsiveness, and patients in whom the preload responsiveness or responsiveness is weak. Therefore, the decision whether or not to administer fluid based on the test result should be made with more confidence if the changes observed are far from the threshold value reported in the literature (4). In the future, it is very likely that this rather rough way of predicting treatments effects will be replaced, at least in part, by more sophisticated predictive analytics based on machine learning (53).

**Conclusions**

It is now clearly demonstrated that volume expansion is a dangerous treatment whose efficacy is inconstant. It is therefore reasonable to predict its effects, in order to avoid administering fluid to a patient who is not dependent on cardiac preload. Several tests are currently available to do this. The advantage of having several tests is to be able to bypass the limits of each, and to base the diagnosis when their result is close to the recommended threshold value.
In addition to a proper evaluation of the fluid efficacy once it has been administered (1,17), the attitude of detecting preload responsiveness may contribute to guide the fluid strategy in a safer and more precise way.

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors (Dr. Glenn Hernández and Dr. Guo-wei Tu) for the series “Hemodynamic monitoring in critically ill patients” published in Annals of Translational Medicine. The article was sent for external peer review organized by the Guest Editors and the editorial office.

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-2020-hdm-18). The series “Hemodynamic monitoring in critically ill patients” was commissioned by the editorial office without any funding or sponsorship. XM and JLT report personal fees from Pulsion Medical Systems, during the conduct of the study. XM and JLT report personal fees from Getinge/Pulsion Medical System, outside the submitted work. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Monnet X, Téboul JL. My patient has received fluid. How to assess its efficacy and side effects? Ann Intensive Care 2018;8:54.


breathing patients with severe sepsis or acute pancreatitis. Crit Care Med 2010;38:819-25.

Cite this article as: Monnet X, Teboul JL. Prediction of fluid responsiveness in spontaneously breathing patients. Ann Transl Med 2020. doi: 10.21037/atm-2020-hdm-18