Existing fluid responsiveness studies using the mini-fluid challenge may be misleading

Enevoldsen, Johannes; Scheeren, Thomas W L; Berg, Jonas M; Vistisen, Simon T

acta Anaesthesiologica Scandinavica

Existing fluid responsiveness studies using the mini-fluid challenge may be misleading: Methodological considerations and simulations

Johannes Enevoldsen[1,2] | Thomas W. L. Scheeren[3] | Jonas M. Berg[1,2] | Simon T. Vistisen[1,2]

[1]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

[2]Department of Anaesthesiology and Intensive Care, Aarhus University Hospital, Aarhus, Denmark

[3]Department of Anesthesiology, University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands

Correspondence

Johannes Enevoldsen, Department of Clinical Medicine, Aarhus University, Palle Juul-Jensens Boulevard 82, 8200 Aarhus N, Denmark.

Email: enevoldsen@clin.au.dk

Abstract

Background: The mini-fluid challenge (MFC) is a clinical concept of predicting fluid responsiveness by rapidly infusing a small amount of intravenous fluids, typically 100 ml, and systematically assessing its haemodynamic effect. The MFC method is meant to predict if a patient will respond to a subsequent, larger fluid challenge, typically another 400 ml, with a significant increase in stroke volume.

Methods: We critically evaluated the general methodology of MFC studies, with statistical considerations, secondary analysis of an existing study and simulations.

Results: Secondary analysis of an existing study showed that the MFC could predict the total fluid response (MFC + 400 ml) with an area under the receiver operator characteristic curve (AUROC) of 0.92, but that the prediction was worse than random for the response to the remaining 400 ml (AUROC = 0.33). In a null simulation with no response to both the MFC and the subsequent fluid challenge, the commonly used analysis could predict fluid responsiveness with an AUROC of 0.73.

Conclusion: Many existing MFC studies are likely overestimating the classification accuracy of the MFC. This should be considered before adopting the MFC into clinical practice. A better study design includes a second, independent measurement of stroke volume after the MFC. This measurement serves as reference for the response to the subsequent fluid challenge.

1 | INTRODUCTION

The term *mini-fluid challenge* (MFC) was coined by Muller et al about a decade ago[1] as a new way to predict fluid responsiveness. At the time, common fluid infusion practice consisted of 'let's give some fluid and see what happens' as highlighted by the accompanying editorial.[2] That 'some fluid' was a fluid challenge of around 500 ml as identified by the FENICE study.[3] Motivated by the finding that fluid is not harmless and may induce fluid overload, Muller et al suggested the MFC: the haemodynamic effect of a rapid infusion of a small

amount of fluid could guide whether or not a larger amout of fluid should be given. The authors tested whether the change in aortic velocity time integral (VTI; an echocardiographic measure correlated with stroke volume [SV]) induced by the MFC (100 ml within 1 min) could predict the effect of a 'normal' fluid challenge of 500 ml, specifically, the combined effect of the MFC and another 400 ml. The method was highly predictive (area under the receiver operating characteristic [ROC] curve, AUROC, of 0.92).[1] Others have since investigated and validated the MFC, and a recent systematic review including seven MFC studies (368 fluid challenges in 324 patients)[1,4-9]

identified a pooled AUROC of 0.91 for the MFC method.[10] Since the systematic review, more MFC studies have been published, all pointing to the same compelling conclusion: that the method is accurate in predicting fluid responsiveness.[11-15]

In 2018, we published a correspondence debating the way MFC studies were designed.[16] The correspondence raised clinical and statistical issues with the most adopted methodology. Yet, the notion that optimal MFC methodology may not be completely settled has hardly influenced methodology in subsequent publications. In this paper, we will:

- explain in simple terms the problems with the most frequently used MFC method
- demonstrate, by secondary analysis of an existing study and by simulations, the potential magnitude of the problem
- discuss strengths and limitations of less frequently used designs
- give recommendations on the way forward for researching this otherwise compelling method.

1.1 | A representative MFC study design

To simplify the key message, we will consider and discuss a representative MFC study design as depicted in Figure 1: 100 ml fluid is infused within 1 min (the MFC), the haemodynamic response (relative SV change) of that MFC is evaluated, and subsequently another 400 ml fluid (totalling 500 ml) is infused over 15 min. The final response (outcome) is evaluated as a relative SV change from baseline (i.e. before any fluid administration) to after the full amount of 500 ml. While we use SV in the examples, the arguments can be generalised to any method for estimating SV or cardiac output.

2 | METHOD

Figure 1 identifies that calculations of the haemodynamic response to the MFC (ΔSV_{100}) and the response to the *full* fluid challenge (ΔSV_{500}) both include the haemodynamic variable measured at *baseline*, that is before the MFC. Specifically, ΔSV_{100} and ΔSV_{500} are calculated as

$$\Delta SV_{100} = \frac{SV_{100} - SV_{baseline}}{SV_{baseline}}$$

and

$$\Delta SV_{500} = \frac{SV_{500} - SV_{baseline}}{SV_{baseline}}.$$

This shared baseline causes the problem.[1] It introduces two effects that, in addition to a true classification accuracy, can explain the high classification accuracy found in several MFC studies:

1. The predictor and the outcome share measurement error, creating a spurious correlation.

2. The predictor (ΔSV_{100}) is also a part of the outcome we try to predict (ΔSV_{500}).

2.1 | Shared error

Any measurement is associated with uncertainty (error). This can be subdivided into a systematic error (often referred to as bias) and a random error (often referred to as variance and defining *precision*).[17,18] It is useful to think of a 'true' SV and a random error around this value. The 'true' SV is what the clinician wants to measure, and what they hope to increase with a fluid infusion. The random error comprises both the imprecision of the monitoring equipment *and* minor temporal (minute-wise) physiologic changes in haemodynamics that are effectively noise in the context of evaluating a fluid response. It is the random error on the baseline measurement that causes the problem. In the following equations, each measured SV is divided into a 'true' SV and a random measurement error.

$$\Delta SV_{100} = \frac{(SV_{100} + \varepsilon_{SV,100}) - (SV_{baseline} + \varepsilon_{SV,baseline})}{SV_{baseline} + \varepsilon_{SV,baseline}}$$
$$= \frac{SV_{100} + \varepsilon_{SV,100}}{SV_{baseline} + \varepsilon_{SV,baseline}} - 1,$$

and

$$\Delta SV_{500} = \frac{SV_{500} + \varepsilon_{SV,500}}{SV_{baseline} + \varepsilon_{SV,baseline}} - 1.$$

These two equations essentially depict the problem: the random error ($\varepsilon_{SV,baseline}$) is part of the denominator in the calculation of both the predictor (ΔSV_{100}) and the outcome (ΔSV_{500}), making them spuriously correlated, and therefore more likely to agree.[19]

2.2 | The predictor (ΔSV_{100}) is also a part of the outcome we try to predict (ΔSV_{500})

The MFC should be used as a *predictive* method, that is to decide whether to administer the remaining 400 ml fluid or not. Thus,

FIGURE 1 Representation of the design used in most mini-fluid challenge (MFC) studies. In this design, stroke volume (SV) is measured three times: (1) At baseline, (2) after the MFC and (3) after the full fluid challenge

$$\Delta SV_{100} = \frac{SV_{100} - SV_{baseline}}{SV_{baseline}} \qquad \Delta SV_{400} = \frac{SV_{500} - SV_{100}}{SV_{100}}$$

$$\Delta SV_{500} = \frac{SV_{500} - SV_{baseline}}{SV_{baseline}}$$
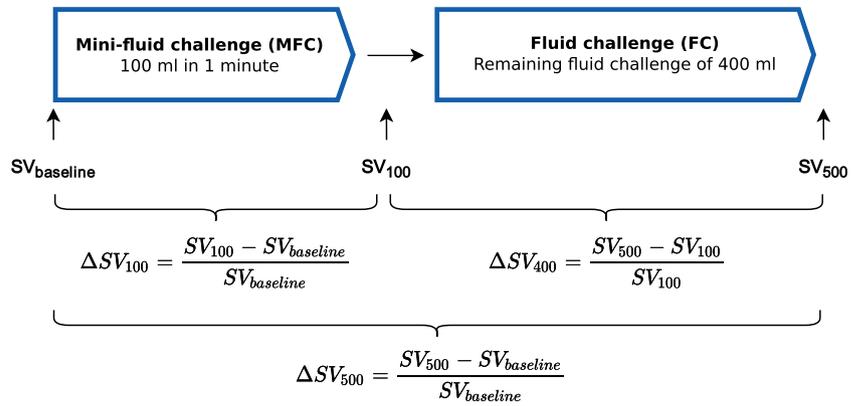
when evaluating the accuracy of the MFC as a predictor of fluid responsiveness, only the effect of the last 400 ml should define the outcome. The naive solution would be to use the MFC (ΔSV_{100}) to predict ΔSV_{400} (see Figure 1). Unfortunately, this does not solve the shared error problem. The random variation of the SV_{100} measurement will introduce a similar problem, since that variable is now a constituent of both the predictor (ΔSV_{100}) and outcome (ΔSV_{400}) variables (see Figure 1). In this case, the random variation in SV_{100} will make the predictor and outcome variables *less likely* to agree, by creating a spurious, negative, correlation, leading to an underestimation of the true classification accuracy.

Both the problems described above arise from *mathematical coupling* of the predictor and outcome.[20,21]

2.3 | Secondary analysis of an existing study

To illustrate what happens with classification, if we try to predict ΔSV_{400} instead of ΔSV_{500}, we extracted ΔSV_{100} and ΔSV_{500} from plot 3A in the pioneering study by Muller et al and calculated the corresponding ΔSV_{400}.[1] Data were captured using DataThief III (version 1.7, datathief.org). Although the study reported relative VTI changes (ΔVTI), we will continue to use the SV term for consistency.

ΔSV_{400} is defined as

$$\Delta SV_{400} = \frac{SV_{500} - SV_{100}}{SV_{100}}.$$

If ΔSV_{100} and ΔSV_{500} are known, we can calculate ΔSV_{400}:

$$\Delta SV_{500} + 1 = (\Delta SV_{100} + 1) \cdot (\Delta SV_{400} + 1).$$

Therefore,

$$\Delta SV_{400} = \frac{\Delta SV_{500} + 1}{\Delta SV_{100} + 1} - 1.$$

We then analysed ΔSV_{100}'s (MFC) ability to predict $\Delta SV_{400} > 15\%$.

2.4 | Simulations

Simulations can reveal how shared error can introduce a significant bias to the result of MFC studies. The magnitude of the problem in existing studies is impossible to calculate exactly, since some relevant variables have to be estimated, but a simulation can provide a ballpark estimate.

Using R (4.0.4) and R packages, *pROC* and *Tidyverse*,[22-24] we simulated SV measurements at all three measurement points in Figure 1 (baseline, after 100 ml and after 500 ml fluid) for 2000 subjects. Annotated code generating the simulations is available from the digital Supplementary Material S1, and an interactive tool that allows changing simulation parameters is available from https://johannesne.shinyapps.io/mini-fluid-challenge-simulation/.

2.4.1 | Simulation 1

First, we simulated how the MFC methodology performs in virtual patients whose SV are entirely unresponsive to fluid, but with random variation in SV measurements. Since there is nothing to predict, any apparent predictive ability is a statistical artefact. Each patient was assigned a constant 'true' SV for all three windows (mean = 75 ml, SD = 10 ml), with an additional random variation (mean = 0, SD = 3 ml) that was independent between time windows (see Figure 2). A random error with a SD of 3 ml gives an 8% precision at 75 ml SV. This was chosen to match the between examination variability in VTI measurements performed by the same observer (although the magnitude of this variation will only effect the results of simulation 2).[25] From these three simulated measurements of a 'constant' SV (but with random measurement error added), we calculated ΔSV_{100}, ΔSV_{400} and ΔSV_{500}. We also simulated a second independent SV_{100} measurement ($SV_{100}b$) to serve as the reference for an independent outcome measure ($\Delta SV_{400}b = (SV_{500} - SV_{100}b)/SV_{100}b$). In this initial simulation, we regarded any increase in SV as a positive fluid response. Using ROC analysis, we showed how well ΔSV_{100} predicted an increase in SV with either ΔSV_{500}, ΔSV_{400} or $\Delta SV_{400}b > 0\%$ as the outcome measure. Since SV varies randomly, half of patients should be responders by this definition, and because the variation is independent between the time windows, it should
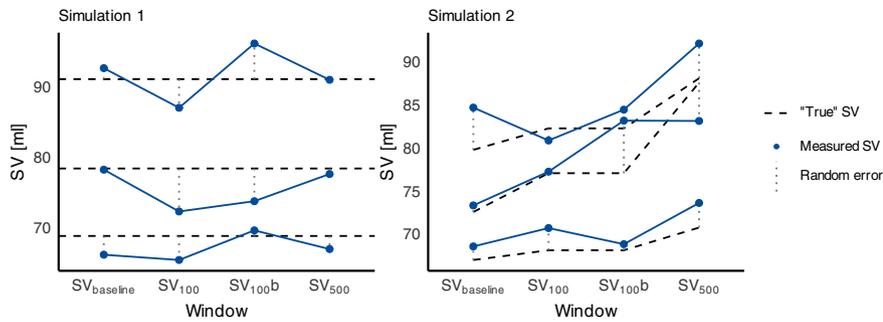
FIGURE 2 Illustration of how stroke volume (SV) measurements were simulated. Each panel shows three of 2000 simulated subjects. The dotted lines indicate the added random error at each time point
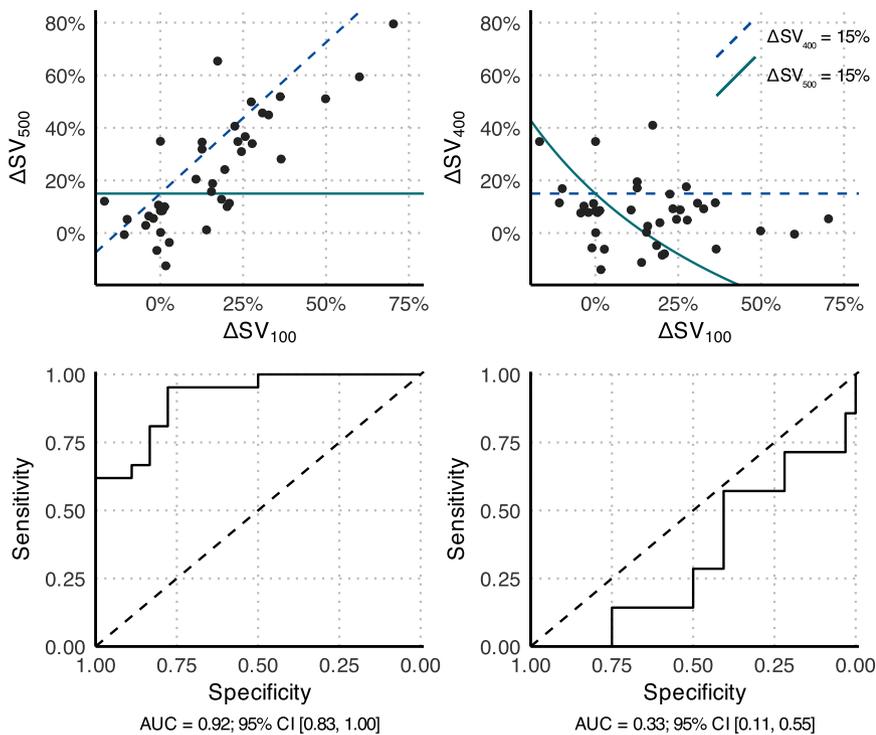

FIGURE 3 Reconstruction of data from figure 3A from Muller et al. (2011).[1] Upper panels: Scatter plots of the relation between ΔSV_{500} and ΔSV_{100} (left) and the relation between ΔSV_{400} (derived) and ΔSV_{100} (right). The full line represents the level at which $\Delta SV500$ is 15% and the dashed line represents the level at which ΔSV_{400} is 15%. Lower panels: Corresponding ROC classification curves of ΔSV_{100} predicting $\Delta SV_{500} > 15\%$ and $\Delta SV_{400} > 15\%$ respectively

be impossible to predict which patients will have an increase in SV after 500 ml.

2.4.2 | Simulation 2

In a second, more realistic, simulation we simulated a 'true' response, still with additional random variation. Each subject was assigned an individual fluid response, which is the 'true' relative change from $SV_{baseline}$ to SV_{500} (the 'true' ΔSV_{500}). The simulated fluid response was drawn from a normal distribution (mean change = 15%, SD = 10%). To keep the simulation simple, the 'true' ΔSV_{100} was defined as 30% of this 'true' ΔSV_{500}:

'True' $SV_{baseline}$ was drawn from a normal distribution (mean = 75 ml, SD = 10 ml).

'True' SV_{500} = 'true' $SV_{baseline}$ · (1 + individual fluid response).

'True' SV_{100} = 'true' $SV_{baseline}$ · (1 + 0.3 individual fluid response).

Independent random variation was subsequently added to each of these three 'true' measurements (mean = 0, SD = 3 ml) (see Figure 2). Again, we also simulated a second independent SV_{100} measurement ($SV_{100}b$) to serve as the reference measurement for an independent outcome measure ($\Delta SV_{400}b$). An increase in SV of >15% was considered a significant positive fluid response in this *clinical* simulation.

3 | RESULTS

3.1 | Secondary analysis of an existing study

In Figure 3, plots are shown for ΔSV_{100}'s ability to predict $\Delta SV_{500} > 15\%$ (left panels) and ΔSV_{100}'s ability to predict $\Delta SV_{400} > 15\%$ (right panels). It is evident from Figure 3 that the classification goes from excellent (AUROC: 0.92) to worse than random (AUROC: 0.33) if SV_{100} is used as the reference value for the subsequent fluid response (ΔSV_{400}).

3.2 | Simulations

3.2.1 | Simulation 1

In a simulated population with no 'true' response to fluid, the commonly used MFC methodology (prediction of $\Delta SV_{500} > 0\%$ using ΔSV_{100}) predicted a fluid response with an AUROC of 0.73 (see Figure 4). Conversely, the prediction of $\Delta SV_{400} > 0\%$ (AUROC = 0.26) showed an equally large underestimation of the expected AUROC of 0.5. The independent outcome $\Delta SV_{400}b > 0\%$ was predicted by ΔSV_{100} with an AUROC of ~0.5, appropriately matching that variation in SV was random in this simulation.

3.2.2 | Simulation 2

In this simulation of a 'true' fluid response, ΔSV_{100} predicted $\Delta SV_{500} > 15\%$ with an AUROC of 0.78, and $\Delta SV_{400} > 15\%$ with an AUROC of 0.47 (see Figure 5). With a new, independent measurement after 100 ml ($SV_{100}b$), ΔSV_{100} predicted $\Delta SV_{400}b > 15\%$ with an AUROC of 0.65.

4 | DISCUSSION

This study demonstrates that the MFC study design most widely used in the literature (Figure 1) is problematic. Results from studies with such problematic designs may overestimate the true classification accuracy of an MFC. This should be considered before adopting the MFC into clinical practice. Still, there are aspects of the above simulations that are worth discussing, and other study designs that should be considered in the search for the optimal MFC methodology.

4.1 | Simulations vs secondary analysis of an existing study

The simulations above were designed to illustrate only the shared error problem that arises, when the same random error is included in both predictor and outcome variables. Simulation 2 assumes a proportional relationship between the 'true' MFC response and the 'true' full response ('true' ΔSV_{100} is 30% of 'true' ΔSV_{500}). Translated into physiology, the model implies a straight Frank–Starling curve, that never plateaus. A real patient, on the other hand, can have a 'true' response to the MFC, but no 'true' response to the subsequent fluid administration, because the plateau of the Frank–Starling curve was already reached with the MFC. Indeed, in the study by Muller et al., most of the fluid response took place with the MFC, indicating that many patients were no longer fluid responsive after the MFC. But since the MFC response is also a part of the outcome (ΔSV_{500}), classification accuracy is high. This physiological circumstance (unmodelled in our simulation)

FIGURE 4 Results of simulation 1. Upper panels are scatter plots of the simulated data (n=2000) along with distributions of the responder and non-responder subpopulations. Lower panels are the corresponding ROC classification curves of ΔSV_{100} predicting fluid responsiveness (ΔSV_{500}, ΔSV_{400} and $\Delta SV_{400}b > 0\%$). The changes in stroke volume (ΔSV) are only random variation, so any correlation is a statistical artefact
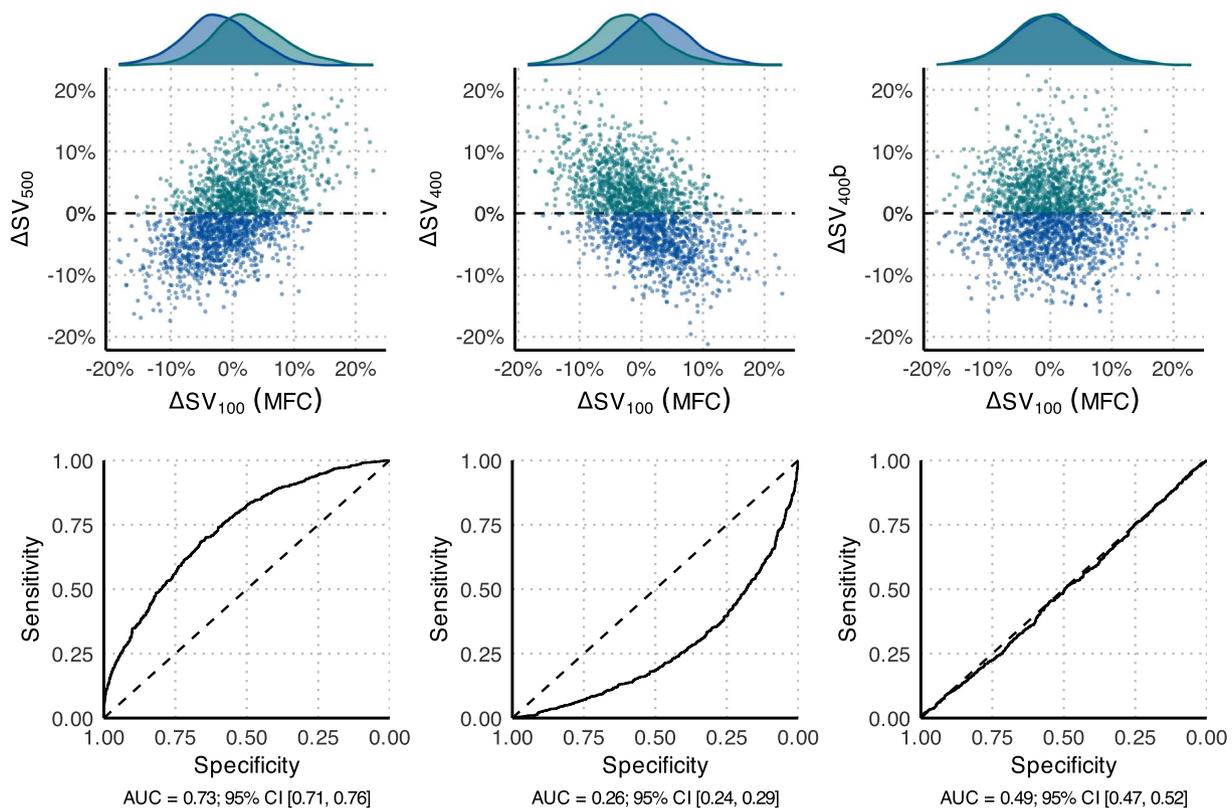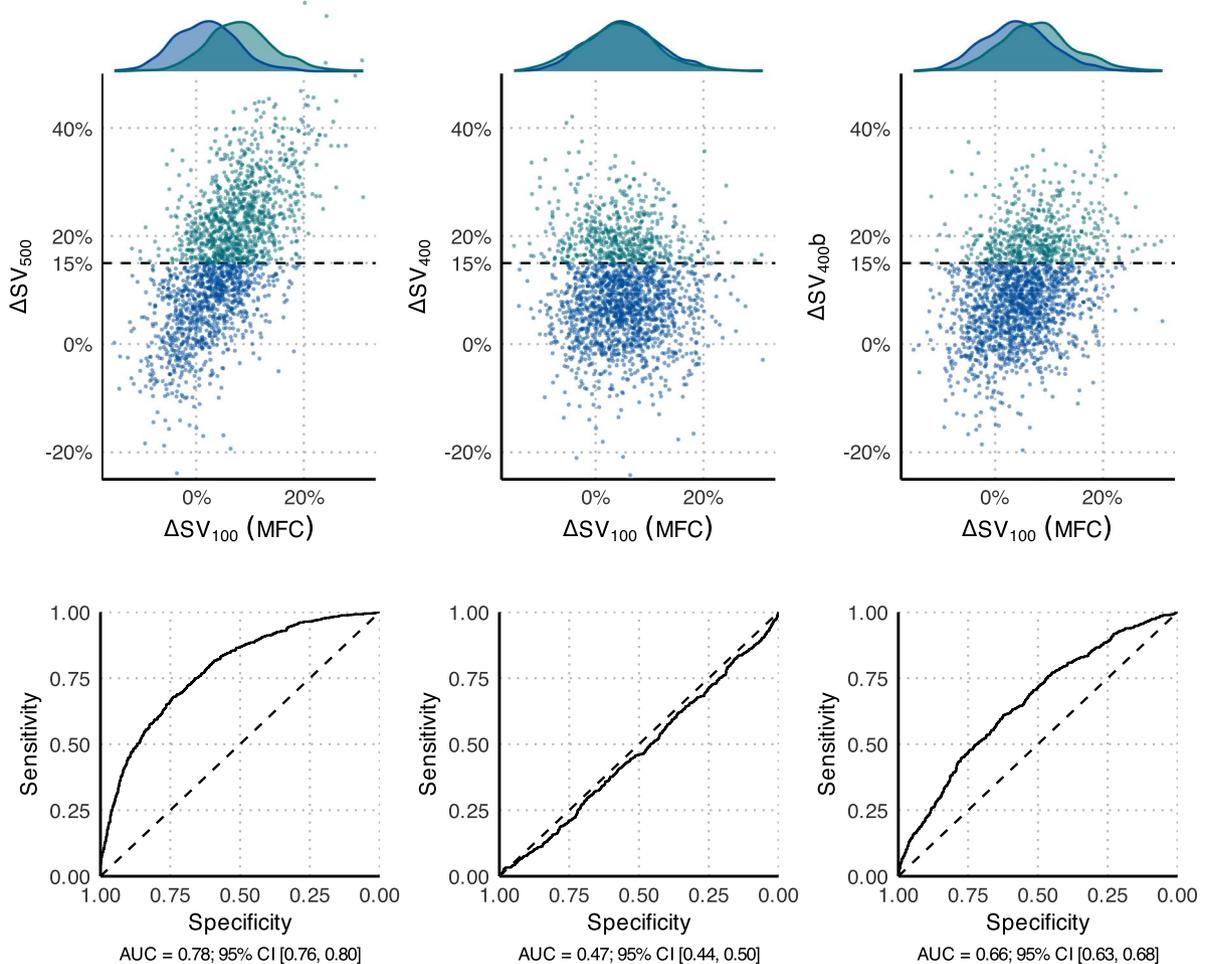
FIGURE 5 Results of simulation 2. Upper panels are scatter plots of the simulated data (n=2000) along with distributions of the responder and non-responder subpopulations. Lower panels are the corresponding ROC classification curves of ΔSV_{100} predicting fluid responsiveness (ΔSV_{500}, ΔSV_{400} and $\Delta SV_{400}b > 15\%$). The simulation identifies the same problem highlighted in Figure 3, although at a lower magnitude, indicating that the assumptions for the statistical modelling may be too conservative in comparison with the behaviour of real-world data

theoretically gives rise to a further overestimation of the classification accuracy in studies using the problematic MFC design, compared to the demonstrated overestimation in our simulation. The difference between predicting $\Delta SV_{500} > 15\%$ and $\Delta SV_{400} > 15\%$ is larger in the study by Muller et al than that in our simulations (see Figures 3–5). This can be explained by the combination of the shared error problem and the relatively large MFC response in the study by Muller et al. It is important to note that while ΔSV_{400} is a more *clinically* meaningful outcome to predict, we discourage using ΔSV_{400} as the outcome given the *mathematical coupling* still present due to the *shared constituent value* (SV_{100}). Neither of the two ROC curves in Figure 3 reveal the 'truth'.

4.2 | Designs with different monitoring modalities for predictor and outcome variables

In one study, authors used different monitoring modalities for predictor and outcome variables: changes in pulse pressure variation (ΔPPV) predicting fluid responsiveness (defined as change in cardiac output).[8] This approach has the advantage that baseline measurements of PPV and thermodilution-derived cardiac output (CO_{TD}) have separate measurement errors:

$$\text{Predictor: } \Delta PPV = PPV_{100} - PPV_{baseline},$$

$$\text{Outcome: } \Delta CO_{TD} = CO_{TD,500} - CO_{TD, baseline}.$$

This reduces the concern about spurious correlation/mathematical coupling. However, while measurement errors are no longer shared, fluctuating physiology over time may still couple different haemodynamic modalities measured simultaneously. Also, this design still includes the response to the MFC in the outcome. Unlike other fluid responsiveness approaches such as the passive leg raising (PLR) manoeuvre, the MFC induces an irreversible physiologic change (because 100 ml fluid is not subsequently removed from the bloodstream).
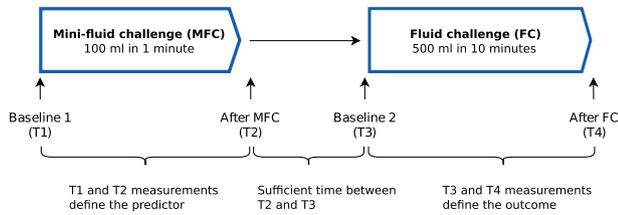
FIGURE 6 An illustration of the MFC study design used by Guinot et al.[5] In this design, the predictor and outcome are NOT *mathematically coupled*

4.3 | A new reference measurement after the MFC

To date, the study with the most appropriate design is that by Guinot et al.[5] Importantly, these authors incorporated an additional SV measurement 5 min after the MFC, to serve as reference for defining the outcome (see conceptual design in Figure 6). In that study, all four SV measurements were obtained by thoracic impedance cardiography (NICCOMO, Imedex, France). A spurious (negative) correlation could, in theory, remain, provided that the error (measurement and physiological) at T2 is correlated with the error at T3. However, it seems plausible that a 5-min window is sufficient to consider errors independent between T2 and T3. This is supported by the data, since any spurious correlation should theoretically reduce classification accuracy, which was probably not encountered in the study by Guinot et al,[5] reporting an AUROC of 0.93. Other monitoring modalities than NICCOMO may have data-stabilising moving-average algorithms implemented making a 5-min window insufficient. An extreme case of this is the *continuous* cardiac output (CCO) measurement from thermodilution pulmonary artery catheters that is only (truly) updated every 4–12 min due to a moving-average algorithm.[26,27]

The time window between T2 and T3 is not without concern though. On average, the effect of the MFC is likely to subside during this period, making the patients more fluid responsive at T3 than at T2. Essentially, this design is using an MFC to predict the response to fluid given 5 min later. In clinical practice, the remaining fluid will likely be given immediately if the MFC response is above a certain threshold. While it may be reasonable to give the fluid right away, if the patient will respond in 5 min, this discrepancy between the study design and clinical practice should be kept in mind. This may be a necessary trade-off to avoid the statistical problems described in this paper.

4.4 | Additional considerations

Infusion rates and timing of the SV measurements can impact the results. Most MFC studies infuse the MFC in 1–2 min and the remaining fluid in 10–30 min, making the infusion rate considerably higher during the MFC.[10] Prather et al show, from fluid expansions of dogs, that cardiac output returns to baseline faster than circulatory volume, and note that rapid infusion results in markedly higher peak cardiac output compared to slower infusion.[28] In a human study, 250 ml crystalloid was infused over 5 min and cardiac output had largely returned to baseline 10 min after end infusion.[29] The effect

on circulating volume is longer: it takes about 30 min before infused crystalloid is distributed between plasma and interstitial fluid, and the elimination half-life is around 20–40 min in conscious humans and several times longer during general anaesthesia.[30,31] Because of the different infusion rates and durations, the MFC is not simply a 'mini' version of the full fluid challenge. It is possible that most healthy hearts will respond to a rapid fluid infusion, while some degree of hypovolaemia may be necessary for a lasting response to a slow infusion. Thus, infusion rates and timing of the SV measurements should be carefully considered in the design of an MFC study.

In most fluid responsiveness studies (incl. MFC studies), the outcome (e.g. ΔSV_{500}) is dichotomised into 'responder' (e.g. $\Delta SV_{500} \geq 15\%$) or 'non-responder'. While this approach simplifies analysis and interpretation, the threshold is more-or-less arbitrary. Dichotomisation of continuous variables is generally not recommended.[32,33] For normally distributed data, it results in a loss of power equivalent to at least a 36% reduction in sample size, and considerably more if the split is not balanced.[34] MFC studies, and fluid responsiveness studies in general, would benefit from keeping variables on a continuous scale.

Lastly, it may be possible to do a statistically valid analysis on data from a study with only three SV measurements (as in Figure 1). Unfortunately, we have not yet seen an example of this, nor found a satisfactory solution ourselves.

5 | CONCLUSION AND RECOMMENDATIONS

The vast majority of published MFC studies used designs that are problematic. These probably overestimate the accuracy of using MFC to guide fluid therapy.

We strongly recommend that a study design separating the predictor from the outcome is applied in the future studies. This is exemplified by the study by Guinot et al as depicted in Figure 6. Here, two separate measurements were obtained after the MFC—one to evaluate the MFC response and one to serve as a new reference for the remaining fluid infusion.

We recommend that specific attention is paid to ensure that outcome and predictor variables are indeed separated by a sufficient time window between the T2 and T3 measurements (see Figure 6). An appropriate time window will depend on the used monitoring modality and its underlying algorithms and time resolution.

Researchers should strongly consider keeping both the predictor and outcome on a continuous scale, and be cautious of spurious correlations when analysing changes.

AUTHOR CONTRIBUTIONS

JE: Conception, manuscript preparation, simulation, artwork and revision. TWLS: Conception, manuscript preparation and revision. JMB: Manuscript preparation and revision. STV: Conception, manuscript preparation, simulation, artwork and revision.

ORCID

Johannes Enevoldsen https://orcid.org/0000-0002-9190-6566
Thomas W. L. Scheeren https://orcid.org/0000-0002-9184-4190
Jonas M. Berg https://orcid.org/0000-0001-9056-7470
Simon T. Vistisen https://orcid.org/0000-0002-1297-1459

REFERENCES

1. Muller L, Toumi M, Bousquet P-J, et al. An increase in aortic blood flow after an infusion of 100 ml colloid over 1 minute can predict fluid responsiveness: the mini-fluid challenge study. *Anesthesiology.* 2011;115:541-547.
2. Vincent JL. "Let's Give Some Fluid and See What Happens" versus the "Mini-fluid Challenge". *Anesthesiology.* 2011;115 455-456.
3. Cecconi M, Hofer C, Teboul JL, et al. Fluid challenges in intensive care: the FENICE study: a global inception cohort study. *Intensive Care Med.* 2015;41:1529-1537.
4. Smorenberg A, Cherpanath TGV, Geerts BF, et al. A mini-fluid challenge of 150mL predicts fluid responsiveness using Modelflow (R) pulse contour cardiac output directly after cardiac surgery. *J Clin Anesth.* 2018;46:17-22.
5. Guinot P-G, Bernard E, Defrancq F, et al. Mini-fluid challenge predicts fluid responsiveness during spontaneous breathing under spinal anaesthesia: an observational study. *Eur J Anaesthesiol.* 2015;32:645-649.
6. Biais M, de Courson H, Lanchon R, et al. Mini-fluid challenge of 100 ml of crystalloid predicts fluid responsiveness in the operating room. *Anesthesiology.* 2017;127:450-456.
7. Xiao-ting W, Hua Z, Da-wei L, et al. Changes in end-tidal CO_2 could predict fluid responsiveness in the passive leg raising test but not in the mini-fluid challenge test: a prospective and observational study. *J Crit Care.* 2015;30:1061-1066.
8. Mallat J, Meddour M, Durville E, et al. Decrease in pulse pressure and stroke volume variations after mini-fluid challenge accurately predicts fluid responsiveness. *Br J Anaesth.* 2015;115:449-456.
9. Wu Y, Zhou S, Zhou Z, Liu B. A 10-second fluid challenge guided by transthoracic echocardiography can predict fluid responsiveness. *Crit Care.* 2014;18:R108.
10. Messina A, Dell'Anna A, Baggiani M, et al. Functional hemodynamic tests: a systematic review and a metanalysis on the reliability of the end-expiratory occlusion test and of the mini-fluid challenge in predicting fluid responsiveness. *Crit Care.* 2019;23:264.
11. Ali A, Dorman Y, Abdullah T, et al. Ability of mini-fluid challenge to predict fluid responsiveness in obese patients undergoing surgery in the prone position. *Minerva Anestesiol.* 2019;85:981-988.
12. Mukhtar A, Awad M, Elayashy M, et al. Validity of mini-fluid challenge for predicting fluid responsiveness following liver transplantation. *BMC Anesthesiol.* 2019;19:56.
13. Lee C-T, Lee T-S, Chiu C-T, Teng H-C, Cheng H-L, Wu C-Y. Mini-fluid challenge test predicts stroke volume and arterial pressure fluid responsiveness during spine surgery in prone position: A STARD-compliant diagnostic accuracy study. *Medicine (Baltimore).* 2020;99:e19031.
14. Fot EV, Izotova NN, Smetkin AA, Kuzkov VV, Kirov MY. Dynamic tests to predict fluid responsiveness after off-pump coronary artery bypass grafting. *J Cardiothorac Vasc Anesth.* 2020;34:926-931.
15. Messina A, Lionetti G, Foti L, et al. Mini fluid chAllenge aNd End-expiratory occlusion test to assess flUid responsiVEness in the operating room (MANEUVER study): a multicentre cohort study. *Eur J Anaesthesiol EJA.* 2021;38:422-431.
16. Vistisen ST, Scheeren TWL. Challenge of the mini-fluid challenge: filling twice without creating a self-fulfilling prophecy design. *Anesthesiology.* 2018;128:1043-1044.
17. Squara P, Scheeren TWL, Aya HD, et al. Metrology part 1: definition of quality criteria. *J Clin Monit Comput.* 2021;35:17-25.
18. Squara P, Scheeren TWL, Aya HD, et al. Metrology part 2: procedures for the validation of major measurement quality criteria and measuring instrument properties. *J Clin Monit Comput.* 2021;35:27-37.
19. Fuguitt GV, Lieberson S. Correlation of ratios or difference scores having common terms. *Sociol Methodol.* 1973;5:128.
20. Archie J. Mathematic coupling of data: a common source of error. *Ann Surg.* 1981;193:296-303.
21. Stratton HH, Feustel PJ, Newell JC. Regression of calculated variables in the presence of shared measurement error. *J Appl Physiol.* 1987;62:2083-2093.
22. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. 2021 Available from: https://www.r-project.org/
23. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* 2011;12:77.
24. Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. *J Open Source Softw.* 2019;4:1686.
25. Jozwiak M, Mercado P, Teboul J-L, et al. What is the lowest change in cardiac output that transthoracic echocardiography can detect? *Crit Care.* 2019;23:116.
26. Haller M, Zollner C, Briegel J, Forst H. Evaluation of a new continuous thermodilution cardiac output monitor in critically ill patients: a prospective criterion standard study. *Crit Care Med.* 1995;23:860-866.
27. Bootsma IT, Boerma EC, Scheeren TWL, de Lange F. The contemporary pulmonary artery catheter. Part 2: measurements, limitations, and clinical applications. *J Clin Monit Comput.* 2021;1-15.
28. Prather J, Taylor A, Guyton A. Effect of blood volume, mean circulatory pressure, and stress relaxation on cardiac output. *Am J Physiol-Leg Content.* 1969;216:467-472.
29. Aya HD, Ster IC, Fletcher N, Grounds RM, Rhodes A, Cecconi M. Pharmacodynamic analysis of a fluid challenge. *Crit Care Med.* 2016;44:880-891.
30. Hahn RG, Lyons G. The half-life of infusion fluids: an educational review. *Eur J Anaesthesiol EJA.* 2016;33:475-482.
31. Hahn RG. Understanding volume kinetics. *Acta Anaesthesiol Scand.* 2020;64:570-578.
32. Dawson NV, Weiss R. Dichotomizing continuous variables in statistical analysis: a practice to avoid. *Med Decis Making.* 2012;32:225-226.
33. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ.* 2006;332:1080.
34. Fedorov V, Mannino F, Zhang R. Consequences of dichotomization. *Pharm Stat.* 2009;8:50-61.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.