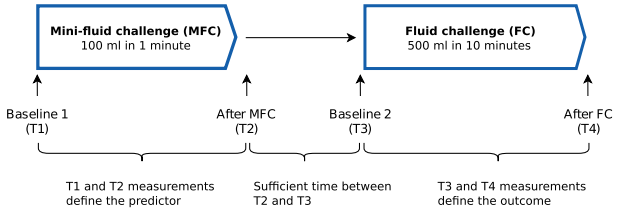

FIGURE 6 An illustration of the MFC study design used by Guinot et al.[5] In this design, the predictor and outcome are NOT *mathematically coupled*

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.

CONFLICT OF INTEREST

TWLS received research grants and honoraria from Edwards Lifesciences (Irvine, CA, USA) and Masimo Inc. (Irvine, CA, USA) for consulting and lecturing, and from Pulsion Medical Systems SE (Feldkirchen, Germany) for lecturing. JE, JMB and STV have no conflict of interests to declare.