### 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 $\Delta SV_{100}$) 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 $\Delta SV_{100}$ 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, $\Delta SV_{100}$ 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 ($SV_{100}b$), $\Delta SV_{100}$ 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' $\Delta SV_{100}$ is 30% of 'true' $\Delta SV_{500}$). 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 ($\Delta SV_{500}$), classification accuracy is high. This physiological circumstance (unmodelled in our simulation)

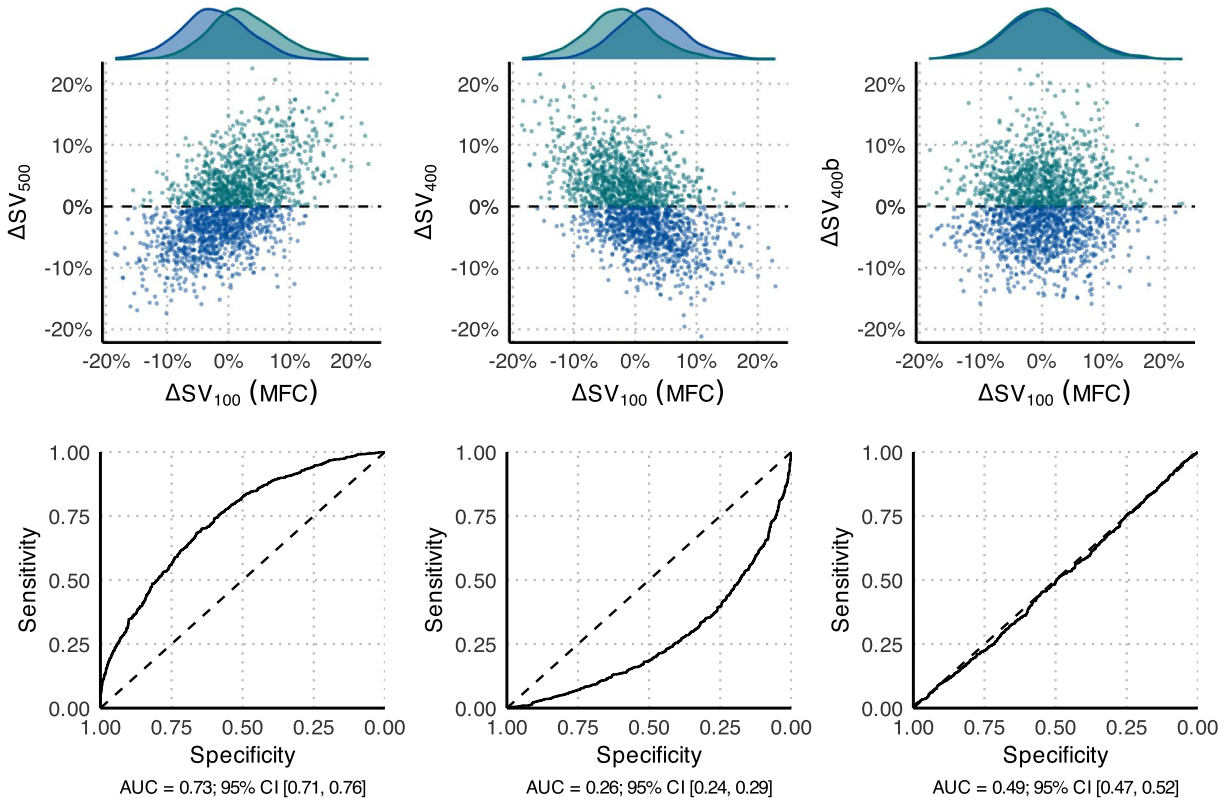


**FIGURE 4** Results of simulation 1. Upper panels are scatter plots of the simulated data (n=2000) along with distributions of the responder and non-responder subpopulations. Lower panels are the corresponding ROC classification curves of $\Delta SV_{100}$ predicting fluid responsiveness ($\Delta SV_{500}$, $\Delta SV_{400}$ and $\Delta SV_{400}b > 0\%$). The changes in stroke volume (ΔSV) are only random variation, so any correlation is a statistical artefact