Is comparison of changes in cardiac output, assessed by different methods, better than only comparing cardiac output to the reference method?

Editor—In the last few years, there have been several papers describing continuous cardiac output (CO) estimation by analysis of the arterial pressure waveform. These pulse contour cardiac output (PCO) methods require calibration for each patient using a method such as indicator dilution.

![Figure 1](image)

**Fig 1** Bland–Altman analysis comparing calibrated PCO with CO (simulated data, n=250). Least squares regression gave y=0.97x+0.22; \(r^2=0.89\) (95% confidence interval, 0.87–0.92); limits of agreement, 1.59 to –1.60 litres min\(^{-1}\); bias, –0.01 litres min\(^{-1}\). Data from two patients have been plotted using large open symbols. From this analysis, it is not possible to assess the ability of PCO to follow changes in CO within individual patients.
After calibration, the purpose of PCO methods is to track ensuing changes in CO. As part of a recent report in this journal, a PCO method was calibrated with transpulmonary thermodilution. Changes in CO were analysed but there was no corresponding graph and no discussion of these results. Unfortunately, many reports evaluating PCO do not provide any statistical comparison of changes in CO, assessed by the different methods. Instead they provide a 'Bland–Altman analysis' comparing calibrated PCO with the reference method: this can be misleading.

To demonstrate the problem, consider a hypothetical study with 50 patients comparing a PCO method with a reference cardiac output (refCO) method. After the initial calibration, CO is measured five times in each patient. The results, as typically presented, are shown in Figure 1.

The results presented in Figure 1 are misleading because the data have been pooled inappropriately. The impression given is that PCO estimates CO with reasonable accuracy over a large range. However, with these simulated data, the large variability of CO is mainly due to differences between patients rather than changing CO within individual patients. The apparent agreement and correlation ($r^2=0.89$) between PCO and refCO is entirely due to the initial calibrations; the subsequent changes in PCO were simulated by random number generation and are independent of changes in refCO! The poor correlation in two individual patients can be seen because the points have been plotted using different symbols.

PCO methods are designed to track changes in CO. Therefore, only changes in CO should be pooled from different patients. This is important from a clinical perspective as well as from a statistical one; a moderate CO that is falling may require more urgent attention than a low CO that is stable. Preferably the data should be analysed after logarithmic transformation because proportionate errors (rather than absolute errors) are clinically important. Bland and Altman provide further details for using logarithmic data.

In Figure 2, the percentage change in CO following calibration was calculated for each data point. This reveals the lack of agreement between PCO and refCO. Currently, thermodilution cardiac output (TDCO) is frequently used as the reference method. Typically, the error of TDCO measurements is $10\pm2$% and so the error of calculated changes in TDCO is $\sqrt{2}\times10=14\%$. In order to demonstrate agreement between PCO and TDCO, the actual CO must change by more than these measurement errors.

It is our view that all studies of PCO methods should include analyses similar to those used for Figure 2. It is important that consensus is achieved so that different studies can be compared. The response to vasoactive drugs and other interventions also forms an important part of PCO assessment.

N. W. F. Linton
R. A. F. Linton
London, UK

Editor—We appreciate the interest of Drs Linton and Linton in our article, and the opportunity to reply. We agree with them over the necessity for consensus about data analysis to compare the results of different studies. In our article we used the COpa technique, which involves bolus thermodilution in the pulmonary artery, as the reference method for all comparisons, because it is the current clinical standard. For statistical analysis, we used the method described by Bland and Altman. Additionally, correlation analysis was performed between methods. As stated in our article, changes (Δ) with each technique in COpa, aortic transpulmonary thermodilution (COart), continuous thermodilution (CCO) and PCCO were analysed using Bland–Altman plots and linear regression, and the results were reported in Table 7. Changes in the variables were calculated by subtracting the first measurement from the second one ($Δ=T2−T1$), and the second from the third ($Δ=T3−T2$).

We now report these changes diagrammatically in Figures 3 and 4 (linear regression) and Figures 5 and 6 (bias±2 sd). No significant difference in the agreement and precision of PCCO, CCO, or COart against COpa was found during the study period. As previously reported, during the clamping and reperfusion phases we observed the development of a non-significant inaccuracy in these measurements. We agree that changes in CO

Declaration of interest. The authors have previously worked for LiDCO Ltd (manufacturer of monitoring equipment), and RL has an equity interest in this company.
and the accuracy of these measurements when they are changing rapidly are clinically more important than the accuracy of stable values. For this reason we compared the differences as the bias ± 2 SD and gave the correlation coefficients (Table 6). In our study the analysis of changes (Δ) confirmed the efficacy of these techniques even in stressful and extreme haemodynamic conditions, albeit with a lesser degree of agreement and precision (Figs 5 and 6). This may be because of the corresponding lack of accuracy during the clamping and reperfusion phases.

All papers comparing COart, PCCO or CCO with the current clinical standard COpa are based on the method described by Bland and Altman. Many authors additionally performed linear regression analyses ($r^2$). We agree with Drs Linton and Linton that a consensus is needed about the methodology used to compare different studies. The use of logarithmic data should also be considered in such a consensus.

G. Della Rocca
M. G. Costa
Rome, Italy


