Reply to Ranjbar et al.

Satoshi Numata*, Keiichi Itatani, Keiichi Kanda and Hitoshi Yaku
Department of Cardiovascular Surgery, Kyoto Prefectural University of Medicine, Kyoto, Japan

Received 25 July 2016; accepted 28 July 2016

Keywords: Aneurysm • Dissection • Fluid dynamics • Simulation

We thank Ranjbar et al. for their comments on our recently published manuscript [1, 2]. This manuscript showed flow dynamics inside the aortic arch using computer simulation. We have chosen several different aortic arch anatomies for this study and evaluated wall shear stress (WSS), oscillatory shear index (OSI) and flow velocity in one cardiac cycle. Ranjbar et al. pointed out that a couple of natural parameters such as the role of helical orientation of the left ventricle on the elasticity of the aortic arch was not considered in our study [2].

In this study, we have focused on the near wall flow characteristics based on the computational fluid dynamics (CFD) with rigid wall assumption and with straightforwardly developed inflow. As Ranjbar et al. pointed out, ascending aortic flow is helical due to twisting motion of the left ventricular wall; however, it is dependent on the ventricular wall motion and difficult to predict, so it is one of the CFD model limitation. 4D-flow MRI can detect the helical inflow toward the descending aorta, however predictive parameters including WSS and OSI are not always accurate in 4D-flow MRI due to insufficient spatial and temporal resolution. Of course, we have performed many validation studies of our CFD analysis [3] with 4D-flow MRI, and our boundary conditions provide sufficiently accurate velocity fields at least aortic flow in elderly patients, who do not have high twisting LV power and who have stiff sclerotic vascular walls.

As we all know, fluid dynamics parameters are not the only factors causing aortic diseases. For example, entry of the type B acute aortic dissection and pseudoaneurysm formation by chest blunt trauma could be related with ductus arteriosus. It has been reported that movement of aortic annulus is related with acute type A dissection. These mechanical factors could be essential for aortic disease. Therefore, fluid-structure interaction (FSI) analysis which combines CFD and structural analysis could increase the accuracy. As we have described in our limitation section, future studies should be performed based on the FSI modelling considering accurate elastic property distribution from the patient specific images [4], and based on the accurate boundary layer turbulence flow simulation and on the accurate OSI calculation that truly makes impact on the endothelium functions and intimal damage [5]. The important fact is that aortic wall elasticity might not be uniform with atherosclerotic plaque or calcification distribution we often see in clinical conditions providing sufficiently accurate velocity fields at least aortic flow in elderly patients, who do not have high twisting LV power and who have stiff sclerotic vascular walls.

In their recent paper, Kieser et al. [1] compared the predictive performance of EuroSCORE against its successor EuroSCORE II in a consecutive series of isolated coronary artery bypass graft patients with total arterial grafting by a single surgeon. Although comparative validation studies such as these are extremely important, we have a number of concerns on the study design and analysis, for which we will highlight only a couple of issues, that question how anyone can meaningfully interpret their findings.

Validation studies are an important aspect of evaluating a risk score, and methodological rigor and transparent reporting are keys to ensure the results are meaningful and interpretable. An important aspect often overlooked in validation studies is study design. Recommendation for sample size is that a minimum of 100 (and preferably 200) events (i.e. deaths) should be included in the study so that model performance and in particular calibration can be adequately assessed [2, 3]; a value much higher than the observed 36 deaths in the Kieser study.

The authors correctly assert that the widely used Hosmer-Lemeshow test is problematic for assessing calibration and should be avoided; it assesses neither direction nor magnitude of calibration. The recent TRIPOD Statement for reporting risk scores cautions against its use with preference for calibration plots [4, 5]. However, the calibration plot of Kieser et al. is also of limited usefulness (ignoring the annoyance that the two axes are not on the same scale; the y-axis is squashed), grouped by predicted risk also suffers from limitations including groups with no events and deciding how many groups. In the study by Kieser et al., we can observe that 4 out of the 10 groups have no deaths, thereby making the interpretation of their calibration plot somewhat difficult. A calibration plot should indicate with a predicted risk of X% how many patients died (which for a well-calibrated model should be close to X observed deaths); this information is not presented in or inferable from their figure. Recommendations are that loess-smoothed calibration plots (preferably with confidence intervals) should be presented so that calibration can be examined across the range of predicted values [6]. The calibration plot can then be supplemented with estimates of the calibration slope and intercept (as calculated by Kieser et al.). A final comment is related to the temporal analysis, as previously noted given the very small number of deaths (median of 4 per time period between 2003 and 2014) and an analysis that does not actually investigate model performance, very little can be concluded whether the calibration ‘evolved’ over time. Given these concerns, and others, including unclear handling of the large amount of missing data for ejection fraction, or whether a small single surgeon case series is at all interesting beyond the surgeon himself, the conclusions have limited utility and should be interpreted with a large ‘pinch of salt’.
REFERENCES


Reply to Collins and Le Manach

Teresa M. Kieserca,*, M. Sarah Rosea and Stuart Headb

* Corresponding author. Centre for Statistics in Medicine, Botnar Research Centre, University of Oxford, Windmill Road, Oxford OX3 7LD, UK. Tel: +44-1865-223460; e-mail: gary.collins@csm.ox.ac.uk (G.S. Collins).

doi:10.1093/ ejcts/ezw262
Advance Access publication 28 July 2016

Keywords: Coronary surgery • EuroSCORE

We thank Collins and Le Manach for their insightful comments on our paper [1, 2] and also for drawing our attention to their article [1] on sample sizes for external validation of prognostic models, which unfortunately was unavailable when our data were analysed. However, their article examined the effect of sample size for proportional hazards models and we used a logistic regression model (with previously published papers [3, 4] more applicable to the assessment of appropriate sample size). Our sample size, though small, is not as exaggerated as their examples (one study with 8 cases and one with 1 case). Although we reported c-statistics with confidence intervals and the P-value for the difference between them for consistency with previously published papers, we stated upfront that this has severe limitations and that huge sample sizes would be needed to detect clinically relevant differences between two c-statistics considered to be in the ‘excellent’ range.

When calculating the sample size for external validation, it is necessary to choose one or two statistics believed most important. We chose calibration, since both calibration-in-the-large and the calibration-coefficient were statistically significant, the power of our study is not an issue, but we acknowledge that bias of these estimates may be. Even though it does not directly apply to our logistic regression model, we did use their simulation study for a sample size of ~37. There was no difference in the coverage rates of confidence intervals between a sample of 37 and one of 100 (or even 200). Also bias in the calibration slope is huge when the number of events is <10, but <2.5% when the sample size is 37 and ~1.6% when 100.

Regarding our calibration plot, the scale of the axes was chosen to avoid uninformative white space in the figure; and to avoid confusion, we added a green diagonal line to illustrate the line of equality on which should lie perfect predictions. We believe that the risk of % of how many patients died is clearly presented in the legend of said table and the table itself.

We read with great interest the article by Tomić et al. [1]. It highlights an issue of considerable importance in the era of new antiplatelet agents, widely used in the treatment of acute coronary syndrome [2,3].

In the surgical scenario, the adequate timing of the double antiplatelet therapy (DAPT) withdrawal is crucial for preventing postoperative bleeding. As stated by recent guidelines, the non-aspirin platelet inhibitor, and

Coronary surgery in the contest of new and old antiplatelet therapies: is it only a matter of suspension timing?

Filippo Prestipinoa,*, Antonio Nenna, Cristiano Spadaccib and Masimo Cheloa

* Department of Cardiovascular Surgery, Universita Campus Bio-Medico di Roma, Rome, Italy
b Department of Cardiothoracic Surgery, Golden Jubilee National Hospital, Clydebank, Glasgow, UK

Received 15 June 2016; accepted 14 July 2016

Keywords: Antiplatelet drugs • Bleeding • Blood transfusion • Cardiac surgery • Coronary artery bypass grafting

Our calibration plots differ only from others in that both are presented on the same graph to illustrate difference. We agree that a less-smoothed calibration plot with 95% CI would have been ideal but not appropriate in this study due to small sample sizes. Operative mortality rate for all coronary artery bypass graft procedures is low (<5%); therefore, any prognostic model for operative mortality will necessarily have zero deaths in the smallest risk groups. Risk score validation of low-risk groups is equally important as for high-risk groups. Our analysis showed that calibration was strongest in the low-risk groups but in the highest risk groups was underestimated by logistic EuroSCORE and overestimated by EuroSCORE II. Surgeons can therefore be confident in either score for low-risk patients.

The missing ejection fraction data are unfortunate but are currently being updated by chart review. So far, most had an ejection fraction of >50% which would not alter the EuroSCORE values nor results of our study.

Finally, validation of risk models is generally accepted best assessed in the settings in which they will be used [5, 6].