Article details: 2020-0034	
Title	Association between pre-operative frailty and outcomes among adults undergoing cardiac surgery: a prospective cohort study
Authors	Carmel Montgomery, Henry Stelfox, Colleen Norris, Darryl Rolfson, Steven Meyer, Mohamad Zibdawi, Sean Bagshaw
Reviewer 1	Luke Mondor
Institution	Institute for Clinical Evaluative Sciences, Toronto, Ont.
General comments (author response in bold)	1. Page 1. Introduction. The objective of the study is not evident from the introduction section. Please state clearly, followed by your hypotheses. **We have revised the introduction to clarify the objective of the study. PAGE: 1** 2. Page 3. Measures of Frailty. This section appears to describe frailty, as well as outcome measures (i.e., "additional data were captured on …"). Consider separate sections to improve readability. **The additional data captured refers to pre-operative data reflecting risk associated with frailty in older patients referred for cardiac surgery. The paragraph has been updated to confirm additional data was pre-operatively captured. PAGE:3** 3. Page 4. Data Sources. Did the authors utilize the longitudinal CFS scores that were captured (at 6mo and 12-mo post-surgery)? I don't see this reported elsewhere in the manuscript. **We appreciate the Reviewer's thoughtful question. The longitudinal CFS scores are the focus of an additional manuscript from this study. PAGE: 4** 4. Page 4. Main Exposure and Outcome Measures. Consider dropping frailty from this section as it has already been described previously in the methods. **Frailty has been removed in the Main Exposures and Outcome Measures section. PAGE: 4** 5. Page 5. Statistical Analysis. Many outcomes are examined; some p-values will be <0.05 purely by chance. A correction for multiple testing (for example, Bonferroni correction) should be made. **We appreciate the Reviewer's suggestion regarding Bonferroni corrections. Table 2 is simply descriptive so we have not adjusted for multiple comparisons. In Table 3 we have reported a number of clinical outcomes, adjusted consistently for age, sex, EuroSCORE II and CFS. A Bonferroni adjustment for the 3 comparisons of hospital discharge disposition has been added, with explanation in the notes at the end of Table 3. PAGE: 5, Table 2; Table 3** 6. Page 5. Statistical Analysis. Why isn't multivariable adjustment performed for all outcomes throughout? Unadjusted results will be subject to unmeasured confounding. **We appreciate the Reviewer's comment. We have reported our primary and all secondary outcomes as adjusted. We recognize there will be residual confounding present. PAGE: 5, Table 3** 7. Page 5. Statistical Analysis. Please state what variables were adjusted for

	in multivariable analysis. “Covariates of clinical significance identified a priori” is not transparent for readers. **We have revised this table and have updated Table 3 to clarify the covariates used in logistic and linear models. PAGE: 5, Table 3** 8. Page 5. Statistical Analysis. Why examine post-discharge mortality using both logistic regression and survival analysis? It seems redundant. The Cox model would suffice to meet your goal (“describe factors associated with mortality”). **We now report the Cox Proportional Hazard models only. PAGE: 5, 7, Table 3** 9. Page 6. Results. Can the authors clarify, in text and in Figure 1, how n=529 was derived? From the methods section, the 2 sites perform approx 2700 urgent/planned cardiac surgeries annually and the study period was approx 2.5 years (Nov11-Mar14). **We appreciate the Reviewer’s comment. The two surgery programs perform approximately 2,800 surgeries per year; however, we excluded several patients based on their surgery type (i.e., emergency, transplant, TAVI), some patients were missed who were admitted for urgent surgery; and we were subject to the coverage of our research coordinators capacity to enrol. PAGE: 6** 10. Page 6. Results. What is “EuroSCORE”? This is the first mention in the paper. Please state all variables in the methods section (consider supplemental materials if necessary). **We appreciate the Reviewer’s suggestion. The Methods have been updated to reflect that EuroSCORE is a cardiac surgery mortality risk score. Further details about EuroSCORE can be added to supplemental materials if requested. PAGE: 3** 11. Page 7. Results. What is the distribution of the 3- and 4-level CFS score in the population? This should be included since the authors state these findings (i.e.,multilevel CFS on mortality) as being the most impactful. **These details have been added to Table 4. PAGE: Table 4** 12. Page 9. Results. “however, the mean difference was similar (1, 95% CI, -5-8, p=0.68) among all survivors. (Table 3)” The mean difference at which time point? (baseline? 12-months?) I don’t see this result in Table 3. **Thank you for the thoughtful suggestion to improve this explanation. The mean difference refers to the mean difference in EQ-VAS for frail patients compared to non-frail patients between their baseline assessment and 12-months post-surgery. This has been clarified. PAGE: 9** 13. Page 10. Interpretation. Largely absent from the discussion is a comparison to published literature. Many studies already exist. How do the current results compare? And what gaps has this study addressed in the literature? **We thank the Reviewer for this comment. We have added relevant reference to prior published work where relevant. We submit that our study adds new knowledge on the long-term mortality at 5-years for frail compared with non-frail patients after cardiac surgery. PAGE: 7, Table 3** 14. Page 12. Limitations. A limitation is the small sample size. Only 51 persons

	were frail and many outcomes are very rare (N<5). Thus, the authors have limited degrees of freedom to enable a robust evaluation (for example, additional covariate adjustment, test possible interactions of CFS with age and sex etc). **We agree with the Reviewer. We have highlighted this in our Discussion. PAGE: 10**
Reviewer 2	Marsha Cohen
Institution	
General comments (author response in bold)	1. It would be useful to have more comments on the assembly of the prospective cohort. Presumably all patients undergoing cardiac surgery were eligible; but what proportion of patients actually participated? What proportion of patients completed the preoperative interview? **We appreciate the Reviewer's comment. While the two surgery programs perform approximately 2,800 surgeries per year, not all patients were eligible and we were limited in our capacity to enrol all potentially eligible patients. See response to Reviewer 1 above. We do not have detailed data on those patients not enrolled in this study. Figure 1 provides details of patients excluded during the screening process and thereafter. PAGE: Figure 1.** 2. The paper would be strengthened by including more information on the CFS instrument. A brief description of this instrument should be provided including its domains. Which aspects of the instrument are acquired by interviews and which by record reviews? The reader should not have to go to the reference list to find an article about the CFS; more information should be provided for in the paper. **The CFS is a validated global measure of frailty. As such, there is a measure of inherent subjectivity; however, it has shown good reliability and predictive validity in numerous ambulatory and acute care populations. We have updated the Methods section to reflect this. PAGE: 3** 3. There is no inter rater reliability testing reported for the study researchers completing the CFS who gathered the data. While the authors acknowledge this as a limitation; it still is a limitation. **We agree with the Reviewer. PAGE: 3** 4. The use of in hospital mortality as the main outcome variable should be justified. Since the number of deaths was very low, the use of multivariable analysis is restricted by the small number of outcomes and the models may be overfitted. Similarly, the number of frail patients is also very low n=51, so that the proportion who had poor post operative outcomes is also very small. This is reflected in the very wide confidence intervals around many of the estimates. In addition, there are a large number of outcomes examined. **We agree with the Reviewer. We chose hospital mortality as the primary as this is a relevant patient centered outcome following cardiac surgery.** 5. Other than to include sample size, is there a rationale for grouping valve and CABG procedures in the study? **These were combined procedures – a patient had both a valve replacement and bypass grafting. As such, this surgical procedure and intraoperative course are different than patients having isolated value procedures or isolated CABG procedures. PAGE: 6**

	6. The results of the study are not surprising. The frailer patients were also more likely to undergo more complex surgical procedures, so it is expected that their mortality and complication rates would be higher. **We appreciate the Reviewer's comment.** 7. The paper has too many tables and figures. The only figure required is the flow diagram. **We have reduced the number of tables to 4 and figures to 3 in the main manuscript and will defer to the Editor if any of these tables/figures should be moved to the Supplement.** 8. The authors might gain some insights into confounding from the paper by Sourial N et al. CMAJ 2019; 28:191. **Thank you for this suggestion.**
Reviewer 3	Greg Hirsch
Institution	Division of Cardiac Surgery, Dalhousie University, Halifax, NS
General comments (author response in bold)	1. It is unclear from the manuscript whether all eligible patients were approached, what the recruitment rate was, and whether the sample recruited was representative of the 2800 patients who undergo cardiac surgery each year at the two centres. **See responses to Reviewers 1 and 2 above. PAGE: Figure 1** 2. The approach to determination of a CFS score is vague. CFS determination varies from foot of the bed assessments, based on the CFS pictogram to an 80 item detailed inventory performed by a qualified gerontologist, and it has been demonstrated that interjudge reliability, sensitivity, and specificity of the CFS varies widely as the stringency of the process by which it is determined is relaxed. Mention is made of both interview and retrospective chart review approaches. This methodology needs to be tightened up in terms of approaches taken, with limitations acknowledged. **We have revised the Measure of Frailty section to better reflect the process for ascertainment of the CFS score. This CFS score is one validated global measure for screening for frailty that is relatively simple and shown good predictive validity. The more complex inventory and calculation of a frailty index implied by the Reviewer is another method to screen for frailty; however, that method was not used in this study. PAGE: 3** 3. It is unclear how the functional assessment of frailty (timed up and go test) was incorporated into frailty assessments. Its impact on outcomes is not clearly delineated. **Functional measures were not specifically integrated into the CFS score assigned of frailty status. Rather, this were used as ancillary measures for descriptive purposes, not to define frailty, but that may be more indicative of a frail state themselves (e.g., gait speed for example shows strong correlation with both frail states and perioperative mortality in cardiac surgery (PMID: 27438112).** 4. The output of both the logistic regression modeling are not displayed in a manner where the impact of frailty is comparable to other factors associated with frailty, for example age and major comorbidities. The details of both the logistic regression and Cox proportional modeling are incomplete.

We appreciate this Reviewer's observation. The covariates in the model have been noted in Table 3. The outputs of the Cox proportional hazard model is displayed below. We have simplified reporting in the manuscript to HR only. PAGE: Table 3

5-year mortality	Hazard Ratio	z	p	95% CI
Frailty	2.21	2.41	0.016	1.16 – 4.21
EuroSCORE	1.03	2.77	0.006	1.01 – 1.06
Age	1.03	1.93	0.053	0.99 – 1.06
Sex	2.34	2.51	0.012	1.20 – 4.55

1-year mortality	Hazard Ratio	z	p	95% CI
Frailty	4.34	2.78	0.005	1.54 – 12.19
EuroSCORE	1.06	3.70	0.000	1.03 – 1.11
Age	1.01	0.56	0.578	0.96 – 1.07
Sex	2.87	1.61	0.107	0.80 – 10.40

6-month mortality	Hazard Ratio	z	p	95% CI
Frailty	6.02	2.90	0.004	1.79 – 20.23
EuroSCORE	1.07	3.05	0.002	1.02 – 1.11
Age	1.03	0.82	0.413	0.96 – 1.10
Sex	3.05	1.40	0.163	0.63 – 14.63

Hospital mortality	Hazard Ratio	z	p	95% CI
Frailty	3.84	1.82	0.068	0.90 – 16.34
EuroSCORE	1.06	2.34	0.019	1.01 – 1.12
Age	1.10	1.95	0.051	0.99 – 1.20
Sex	2.29	1.00	0.319	0.45 – 11.63

CVICU mortality	Hazard Ratio	z	p	95% CI
Frailty	1.43	0.28	0.776	0.12 – 16.72
EuroSCORE	1.08	1.82	0.069	0.99 – 1.17
Age	1.18	1.73	0.085	0.98 – 1.42
Sex	0.97	-0.02	0.982	0.07 – 12.61

5. There are only 51 cases in the CFS > 5 group, and while significant differences in outcome were demonstrated, the widespread applicability of these finding from such a small sample is a concern.

We appreciate the Reviewer's comment. We recognize the small proportion of frail patients presents a challenge with regard to generalizability. PAGE: 13

6. This was a two hospital approach, from different regions of Alberta, with presumably different approaches in referral, pre-operative preparation and selection, and the care of patients. A hierarchical approach should have been taken.

We respectfully disagree. These are the only two cardiac referral centers in Alberta and both are administered within the same public health jurisdiction (Alberta Health Services). Both centers provide all cardiac surgical services, with the exception of transplantation, which is only offered in one center.

7. Page 5 (ln 20): Is there a particular rationale for an inclusion criteria of age ≥50? This seems to be at the lower end of other frailty papers where the age cut-off centers closer to 65. This could potentially capture a number of non-frail/low-CFS score patients?

We appreciate and agree with the Reviewer's comment. At a population level, frailty is less prevalence in ambulatory patients between 50-65 years; however, the population "ages" and accumulates deficit and diseases at variables rates. Patients undergoing major procedures may be more susceptible and prior work in acute and critical care settings have found that frailty is common in this age range and portends greater risk of complications and poor outcome (PMID: 31144259, 27922747, 27263535). This would provide sufficient rationale for lowering the age eligibility for such a descriptive study in our view.

8. It is unclear how Euroscore was incorporated into risk adjustment (see comments re: logistic regression above) and unclear if Euroscore I or II.

The EuroSCORE II was used. The logistic EuroSCORE was a covariate used in multivariable regression. PAGE: 3

9. It is unclear what procedures were performed in the frail vs. non-frail group. The published literature in this area indicates that frail and older patients face more complex procedures, which would have made risk adjustment more challenging.

We agree with the Reviewer. We submit it is conceivable that frail patients are more likely to undergo more complex procedures compared to non-frail patients, and the variation in complexity of these procedures may be difficult to completely adjust for. We chose to minimize subgroupings of surgical procedures, due to limited frailty in the cohort. We also looked at the analysis from a patient perspective where specific details of complex procedures may not be as important as identifying that baseline frailty increases their overall risk of poor outcome.

The table below outlines frequency of procedure type by frailty status.

Procedure	Overall (n=529)	CFS ≥5 (n= 51,10%)	CFS ≤4 (n= 478, 90%)
Isolated CABG	202 (38)	11 (22)	191 (40)
Isolated Valve	219 (41)	24 (47)	195 (41)
Combined CABG and Valve	91 (17)	15 (29)	76 (16)
Other (myxoma, ASD, Aorta)	17 (3)	1 (2)	16 (3)

10. Page 4 (ln 13-15)"...limited incorporation of frailty-related functional measures into cardiac surgery risk scores." There are several publications which employ risk scores taking frailty into account and may be worth referencing here.

We agree with the Reviewer. However, for clarity, we were more specifically

	referring to the frailty as a core element assessed in existing cardiac surgery risk scores or developing new scores (PMID: 22396586, 20627611, 30256921). While these studies are excellent examples of the incremental value and improved model discrimination frailty can add to existing risk scores, we are not aware of further validation or routine adoption of frailty assessment in cardiac surgery risk calculations. We would; however, advocate and support this. PAGE: 12 11. Page 5 (ln 29-30): Did the authors identify number of complex cases from neighbouring provinces/territories? How was missing data dealt with for patients where vital status was not accessible? **We did not identify patients according to their location of residence; however, we aimed not to enroll out of province (or country) patients into the study. In capturing mortality outcomes, we searched the provincial discharge abstract database to capture deaths in acute care and the provincial cardiac interventional registry (APPROACH) to capture deaths registered by the provincial vital statistics department. The remaining patients lost to follow-up represented 0.9% (n=5) of the total cohort. PAGE: 5** 12. Page 6: In a well-cited paper by Savva et al. (2013), they suggest the Timed-Up-and-Go (TUG) test is a sensitive and specific measurement where the full application of Fried's criteria is impracticable; however, they note the test is less able to discriminate the nonfrail from the prefrail or frail populations. What was the justification for using the TUG over gait speed as a measure of mobility? **We appreciate the Reviewer's comment. The TUG was a convenient and standardized test to administer as part of a battery of frailty-related screening tests prior to surgery. We did not intend for it to be a stand-alone assessment of frailty. PAGE: 3** 13. Page 7 (ln 6-7): Did research coordinators have adequate training to effectively identify duration and intensity of organ support, the occurrence of complications and adverse events? **We appreciate the Reviewer's question. Research coordinators were trained in the use of the CFS. A comprehensive case report form and data dictionary was developed and pilot tested. Duration and intensity of organ support was documented in the case report form definitions. Complications and adverse events were listed and their presence documented if found in the patient chart for the index hospital stay. PAGE: 3, 4** 14. Page 8 (ln 17-18): How did the authors deal with missing data beyond those of the CFS scores? **Cardiac surgery presents a reliable approach to non-emergent surgery, with patient histories, procedure notes, lab values, ICU and inpatient charts and discharge summaries consistently containing comprehensive information. The research coordinators were instructed to document the presence of comorbidities found on the chart. In the absence of documentation, the comorbidities were considered not present. Missing data was eliminated with this approach. PAGE: 4** 15. Page 9 (ln 9-11): Given the number of patients who reported receiving help

	at home (n=288), it would be important to know how this question was communicated to patients. Could the authors provide a clearer definition of the living arrangement question and whether they referred to particular ADLs? **This was a relatively simple discriminator. Patients were classified as independent if they lived in a private dwelling and received no outside help (except family care-giver); were classified as at home with help is they lived in a private dwelling but received any form of regular non-family caregiver assistance at home; and were classified as assisted living if they lived in a primary assisted living facility (encompasses all spectrums of support). Patients were asked specific questions about where they lived and whether the received non-family caregiver assistance.** 16. Page 9: Would prefer means (SDs), medians (IQRs), etc. be consistent in one place after decimal point throughout document. Additionally, notation throughout document is convoluted at times. **We have revised to be consistent throughout.** 17. Page 9 (ln 33-35): What is the rationale for including the out-of-province patients? **We appreciate the Reviewer's question. The intent of the study was to measure prevalence of frailty across the provincial cardiac surgery program. We did not target out of province patients for exclusion but few were recruited.** 18. Page 10 (ln 40-48): Very wide confidence intervals for adjusted ORs/Hrs may be an indication of a small sample size. How was sample size and power determined? **Sample size was determined from previous research focused on ICU admissions. Results from previous ICU studies on frailty show a prevalence of 28%-33% in the same provincial population (PMID: 31144259, 24277703). Unfortunately, the cardiac surgery subgroup presents with a lower prevalence of frailty.** 19. Page 12 (ln 3-8): Please include 6-month scores in the result section. As well, the last sentence is unclear as this data is not available in Table 3, please revise. **This has been updated. PAGE: 9** 20. Page 12 (ln 3-8): Why did the authors not include EQ-5D questionnaire scores in their results? This would provide added value for patient-centered outcomes as suggested in the Introduction. If including these scores, the Canadian valuation set for EQ-5D is available from Bansback et al. (2012) to these health states to a summary index score. **We appreciate the Reviewer's suggestion. We chose to present the EQ-VAS rather than the full set of EQ-5D results to keep within word limits for manuscript submission. We will aim to detail the comprehensive EuroQoL data in a subsequent manuscript.** 21. Page 24 – Table 1: It would be useful to include patient use of inotropes, hx of cardiac shock, endocarditis and **As our patients were recruited as outpatients in pre-operative clinics or as**

	in-patients on the hospital ward, no patients were receiving inotropic support, had cardiogenic shock or active infective endocarditis at the time of screening. 22. Page 24 - Table 1: Could the authors elaborate on the 85 patients with documented neurologic dysfunction (as defined as: Disease severely affecting ambulation or day-to-day functioning.) Could this have influenced CFS response? **We agree that the neurologic dysfunction may have influenced the assignment of the CFS and may also have genuinely reflected a worse state of frailty.** 23. Page 13 (ln 6-16): The use of 2-4 level gradients of the CFS would be valuable to see across both primary and secondary outcomes. **We agree that this would be of interest, but given the large number of analyses performed and the amount of data described in the manuscript already, we have largely restricted this to the primary outcome.** 24. Page 14 (ln 49-51): It would again be valuable to include EQ-5D scores in addition to EQ-VAS to make this determination. As well, mean scores of EQ-VAS for frail patients appear to decrease from 6-month to 12-month interval. Could the authors elaborate? **For the purposes of this manuscript we focus on reporting the change to mean EQ-VAS from baseline to 12-months. The EQ-VAS incrementally improves for non-frail patients across the three time frames for ascertainment; however, for frail patients, the EQ-VAS scores appear to not improve. An additional manuscript will explore HRQL further. PAGE: Table 3.**