Article details: 2022-0049
Title: Low-value preoperative cardiac testing before low-risk surgical procedures: a population-based cohort study
Authors: Siavash Zare-Zadeh BHSc, Braden J. Manns MD MSc, Derek S. Chew MD MSc, Tyrone G. Harrison MD, Flora Au MA, Amity E. Quinn PhD

Reviewer 1: Dr. Finlay McAlister
Institution: University of Alberta
General comments (author response in bold)

I think the authors should be encouraged to spend some ink in the Discussion section pointing out that the use of preop cardiac testing was low even before the Choosing Wisely recommendations came out and actually increased thereafter - my bias is that often Choosing Wisely chooses not just "low hanging fruit" but fruit that is already lying on the ground and I think this study confirms that that's the case with preop cardiac testing before low risk surgery. Interestingly, the use of preop ECGs declined over time even though Choosing Wisely didn't recommend against those and the fact that patients with more comorbidities were more likely to be tested reassures me that clinicians are generally doing a reasonable job (despite the small area variations across the province).
We appreciate this point. We have added a line to the discussion to highlight the temporal findings. (Page 8: “The use of preoperative testing was low during the entire study period, even in the two years preceding the 2014 CWC recommendations.”)

It's likely on their cutting room floor, but can the authors provide any insight into how frequent preop cardiac testing was in for "RCRI high risk pts"?
This is a very interesting question. We did not define outcomes for the high risk patients that we excluded, so are unable to provide this detail. We think this would be an interesting avenue of future study to pursue.

In the abstract first line they mention MPI and it isn't til later in the text they explain that means Myocardial Perfusion Imaging so I'd take that acronym out of the abstract.
We have corrected this. (Page 2: “e.g., exercise stress testing, echocardiogram, and myocardial perfusion imaging (MPI)”)

On page 8 they discuss differences between zone 1, zone 2, etc and to readers outside of Alberta (and even those of us inside Alberta) it would be more useful to clarify where those zones were - is zone 1 rural and south, etc?
We are only comfortable including urban and rural descriptors for zones. We have added these additional details where possible.

I think they should add a sentence to the limitations section acknowledging that their estimates of RCRI may be underestimates if clinicians did not list some diagnoses during prior visits (since you only have to list one code to be paid for an outpatient visit in Alberta)
Thank you for bringing this to our attention. We have expanded our first limitation regarding the limitations of the RCRI to include this point. (Page 9: “Further, we may have underestimated risk because of limitations of administrative data, such as incomplete diagnoses reported by physicians in claims or inconsistencies in provincial billing coding.”)

Also, although the CCS recommends using the RCRI, many clinicians do additional biomarkers (like BNP or high sens trop) to further risk stratify some pre-op patients and decide on further testing - there is a Cochrane review demonstrating substantial improvements in c stats if using such biomarkers with the RCRI and it would be worth acknowledging that in the limitations section too. The citation is:
Vernooij LM, van Klei WA, Moons KG, Takada T, van Waes J, Damen JA. The comparative and added prognostic value of biomarkers to the Revised Cardiac Risk Index for preoperative prediction of major adverse cardiac events and all-cause mortality in patients who undergo noncardiac surgery. Cochrane Database Syst Rev. 2021 Dec 21;12(12):CD013139.
Thank you for bringing this review to our attention. We have acknowledged this as a limitation with respect to use of the RCRI alone for perioperative risk stratification. We agree that this would be an interesting avenue to explore, as preoperative biomarkers were only recently suggested for mainstream use in high-risk non-cardiac surgical patients by the CCS, the proportion of low-risk patients in Alberta over our cohort period that had these biomarkers measured is likely very low. Additionally, one might justify that they should not be used in our study population of low-risk patients. (Page 9: "The RCRI was not originally designed for use with administrative data and does not consider clinical findings or laboratory biomarkers which may warrant a preoperative cardiac test (i.e., abnormal heart sounds or symptoms, pre-operative NT-proBNP) (26)")

Reviewer 2: Prof. Lynne Moore
Institution: Université Laval
General comments (author response in bold)

The message in the Conclusion is a bit confusing. How do you conciliate low frequency with significant variation and increasing time trends? Given the low % of low-value testing and the fact that many of them may have been justified (admin data and the RCRI do not enable perfect accuracy to capture low-value care), would monitoring or interventions targeting reductions really be warranted?
Thank you for raising this issue. Despite the low utilization of testing in this cohort, there do appear to be opportunities to further reduce utilization (e.g., high rates of testing among older adults). Considering the utilization was low prior to CWC recommendations, perhaps CWC or any novel interventions targeting reductions are not warranted. However, if older people are being tested due to misperceptions of physicians and this is negatively impacting patient experience, interventions may be warranted. Further analysis of impact on patient experience and health care system needs would be required. (Page 9: "Although overall testing was low, the high rates of testing among low-risk older adults and variation by zone indicate ongoing monitoring of these testing metrics and their association with patient experiences may be warranted.")

Please state how the cut-off of 1 used for the RCRI for low-risk patients was selected. A reference would help.
This is the established threshold to define low-risk. A reference was added. (Page 4: "Patients with an RCRI score <1 were considered low-risk (19).")

Treatment of missing data is inappropriate but given the low % of missing data, this is unlikely to have affected results. Could be mentioned in the Limitations section.

We have added a line in the methods section regarding the small % of missing data. (Page 5: “Less than 1% of patient characteristics data was missing.”)

Figure 2 – are rates over time adjusted or unadjusted? Please specify in title. Also, I believe these are incidence proportions (%) rather than rates.
We have clarified Figure 2 to address the points you have raised. The figure presented unadjusted number of imaging tests associated with a surgery, received within 6 months prior to a surgery. (Figure 2)

Furthermore, in the Poison model, were authors were estimating incidence proportions (dichotomous outcome low-value test or not) rather than incidence rates? If the former, were robust variance estimates used? Please clarify.
We have clarified our outcome (number of tests) and modelling approach. We did not estimate proportions. We used the Poisson model to model the number of preop tests. The description of the specific modelling technique was excluded based on editorial suggestions. (Page 2: “We excluded high-risk patients using the Revised Cardiac Risk Index (RCRI, score ≥1 considered high-risk) and modeled patient and temporal factors associated with the number of tests.”)

Need to add discussion on changes over time relative to the introduction of CWC recommendations (2014). CWC recommendations appear to have had no impact – the only decrease was observed for ECG which is not in CWC recommendations. Were there any interventions to implement CWC recommendations for these tests in Alberta?
Thank you for this suggestion. Because we did not specifically test the impact of the CWC recommendations, we have noted the trend over time in the first paragraph of the discussion, including an observation that there was low preoperative testing prior to CWC recommendations.
There were not additional interventions in Alberta during the study period. (Page 8: “The use of preoperative testing was low during the entire study period, even in the two years preceding the 2014 CWC recommendations.”)

Need to add discussion of ECG use as it is the only test with a high frequency. What could explain the decrease? Should this be included in CWC recommendations? What would be the rationale?
This is an interesting point. We have already noted that novel interventions to reduce ECGs are likely not warranted. A more fulsome exploration of ECG changes over time is not within the scope of this paper. We have noted this as an area for future research. (Page 9: Future studies may further evaluate the factors driving variation in practice, including exploring drivers in reductions of ECGs and increases in echocardiography over time.)

Need more information on how costs were calculated. Was an activity-based costing method used? Please provide unit costs.
We have clarified the cost methods and provided a citation to the Alberta fee schedule. We have also corrected a supplementary table to include procedure codes for the imaging tests. Because we used the specific amount paid in the claims data, we decided not to report fees for each test in the methods section as there may be differences in amount paid due to modifiers or other billing rules. (Page 5: “The cost of cardiovascular testing was estimated by adding the amount paid reported in claims data by Alberta Health to testing facilities for all imaging procedures

identified in this study. The fees and rules governing billing for imaging tests are reported in the Alberta Health Insurance Plan Schedule of Medical Benefits.")

It would be interesting to highlight modifiable determinants in the Discussion.
Our interpretation is that regional variation and patient age are likely the most modifiable factors. We have addressed this in the discussion. (For example, Page 8: "Patient age is another driver of testing, with middle and old age individuals more likely to receive advanced preoperative cardiac testing (Figure 4). This variation may be driven by clinician misperception of risk related to age and frailty.")

Table 2: I think 'advanced (+/- ECG)' should read 'advanced (- ECG)'. Are 'advanced (+/- ECG)' and 'no test' mutually exclusive categories whose sum should equal 'overall'? Table needs to be self-explanatory. Need to indicate %(n) somewhere – n(%) is more widely used.
Thank you for pointing these errors in Table 1. We have made the changes you requested. (Table 1)

Were any sensitivity analyses performed? For example, changing the window of exposure for testing.
In this revision, we conducted a sensitivity analysis limiting the time of the assumed pre-operative specialist visit to 3 months before surgery.
The sensitivity analysis indicates patients with preoperative visits to FFS physicians (vs salaried) and cardiologists (vs internists) were more likely to receive preoperative testing. However, it is important to note that we are not able to identify who ordered the imaging tests, so it is possible this is related to referral patterns from primary care rather than specialist physician decision making. (We revised the methods and results to include the sensitivity test. Page 5: "We conducted a sensitivity analysis limiting the time of preoperative visits to the 90 days before surgery." Page 7: "When we limited this to preoperative visits to specialists in the 3 months prior to surgery, 98,895 patients had a preoperative visit to a specialist and those seeing FFS physicians (versus salary) and cardiologists (versus internal medicine) were significantly more likely to receive preoperative cardiac testing (RR payment model 1.1, 95% CI 1.0-1.1; p <0.001, RR specialty 1.2, 95% CI 1.1-1.3; p <0.001).")

Authors adjusted for within-patient clustering. Did they adjust for clustering of patients within hospitals?
Thank you for raising this point. Surgeon or surgery delivery site would have been interesting random effects to include in our statistical analysis. We have added this as a limitation. (Page 9: "Third, we did not consider contextual factors such as surgeon or local hospital practice patterns that may contribute to imaging utilization patterns or downstream medical resource use (i.e., repeat testing, consultation to review imaging, or delays in surgical booking date).")

'geographic variation within Alberta suggests geographic variation in testing rates across Canada'. I don't think you can extrapolate this from study results.
Thank you for raising this point. We have removed this line from the discussion.

In limitations, discuss other potential confounding variables in the temporal trends analyses that were not available in the data set. Also, could changes in coding over time (outcome or confounders) have biased these comparisons? Could differences in data capture used for RCRI over regions explain observed inter-region differences?

We have expanded the limitations to address potential data issues that might influence RCRI scoring. (Page 9: “Further, we may have underestimated risk because of limitations of administrative data, such as incomplete diagnoses reported by physicians in claims or inconsistencies in provincial billing coding.”)