Review Comments

Thank you for the invitation to review the paper entitled "Returning to Baseline Daily Ambulation after Cardiothoracic Surgery".

The subject of early ambulation after cardiac surgery and VATS using accelerometers is highly interesting. However, there are serious concerns regarding scientific soundness of the study, visualization of results, and potential duplication of previous work.

I would like to comment as follows:
Comment 1. The statement that early ambulation might prevent delirium is highly interesting. I therefore looked into your reference 3 and that seems to be work from your group. Please describe the difference from your study on delirium (REF 3) and this work, as it seems to be identical.
https://doi.org/10.1016/j.amjsurg.2020.08.031

Reply 1: Thank you for thoughtfully reviewing our previous work that we referenced here. This prior work done by our group uses similar post-operative walking data, although key differences include primary outcomes, inclusion criteria and patient population. The primary outcome in the referenced study was percent preoperative baseline and subgroups included those who developed delirium and those who did not. The primary outcome here is return to pre-operative baseline and subgroups are based on type of surgery (cardiac and VATS). Inclusion criteria in the referenced paper were age ≥ 60 years, length of hospitalization >48 hours and more broadly included type of surgery (also included abdominal surgery). Inclusion criteria here did not limit patient age, included only patients undergoing cardiac and thoracic operations and patients were not excluded based on length of hospitalization.

Comment 2. Thank you for describing your work according to the STROBE guidelines. However, many items were marked as N/A. Please address all items in your manuscript.

Reply 2: Thank you for reviewing the STROBE Statement as relevant to our study. We have addressed nearly every item in the Statement, although there are some remaining items which are truly not applicable and therefore marked NA. An explanation of each item changed is described herein. Item 6b asks for matching criteria for applicable matched cohort studies or matched case-control studies for which ours is neither and therefore it has been marked NA. Item 9 asks for addressing potential sources of bias, which has been added (page 6, line 23 and page 7, lines 1-2) and again addressed in the limitations of our study (page 10, lines 6-6). Item 12b is addressed in page 7, lines 4-8, which describes the comparison between subgroups and statistical methods for correlation between post-operative day and number of steps. Item 12c and 12d were added (page 6, lines 2-4) as those without the full pre-operative or post-operative walking data were eliminated from analysis. Item 12e remained NA as we did not perform any sensitivity analyses. Item 13c

has been updated to reference Figure 1, which is the patient enrollment diagram, and this has been updated as addressed in other review comments.   Item 14b remains NA as there was no missing data for variables of interest.   Items 15b and 15c remain NA as this was not a case-control or cross-sectional study.   Item 16a remains NA as there were no unadjusted estimates or confounder-adjusted estimates in this study.   Items 16b remains NA as continuous variables were not categorized in this study.   Items 16c remains NA as relative risk was not an outcome in this study. Item 17 has been updated to reference results reporting for the correlation analysis and comparison of subgroups (page 8, lines 7-17).   Item 22 remains NA as there was no funding source for this study.

Comment 3. In the background section you state: "Ambulatory recovery following video assisted thoracoscopic surgery and open cardiac surgery has not been quantified." This is not true. Others have reported this as well, i.e. https://doi.org/10.3390/s21061979 (The authors should not feel the obligation to cite this work)

Reply 3: Thank you for introducing us to this work out of The Netherlands.   This is great work and we have referenced it with a brief elaboration (page 5 lines 6-8) since it does measure post-operative activity after cardiac surgery and a relevant background paper.   We changed the statement in question to return to baseline ambulation (page 3 line 2).   While still clinically useful, the referenced work does not discuss return to pre-operative ambulatory levels and quantifies walking as time spent per day instead of number of steps.   Number of steps per day is more frequently measured to quantify walking and more easily obtainable due to the ubiquity of commercially available smart-devices.   Furthermore, the Vivofit3 has been externally validated to video-review, which is the gold-standard as mentioned in page 6 lines 4-5.   Lastly, our study quantifies ambulation beyond the inpatient stay, providing a longer observation period.   Thank you for this comment and allowing us to include this reference and explanation in our manuscript.

Comment 4. 21 patients were included in a 3-year time interval. Without many exclusion criteria, there seems to be a high risk for bias. Please comment.

Reply 4: We do acknowledge the limitation of small sample size could allow for a high risk of bias, although similar to the sample size of the previously mentioned paper on a similar topic (https://doi.org/10.3390/s21061979).   To minimize the risk of bias, we compared baseline attributes of subgroups (demographics, comorbidities and pre-operative ambulation) to ensure known variables that may impact post-operative ambulation were similar (Table 1 and Table 2). Additionally, this is addressed in the limitations of our paper, although we have now elaborated (page 10, lines 5-8).

Comment 5. The abstract is unclear regarding: "There was a strong positive correlation between increasing post-operative day number and daily steps (ρ=0.972; p<0.001), which was similar for video assisted thoracoscopic surgery patients (ρ=0.781; p<0.001) and cardiac patients (ρ=0.928; p<0.001). " Please rephrase and make 1 statement per sentence.

Reply 5: Thank you for making this suggestion.   This has been reworded and split into three

sentences, one for each statement (page 3, lines 15-20).   Due to this change, small edits have been made elsewhere in the abstract without changing any meaning to adhere to the journal requirement of an abstract ≤ 350-words.

Comment 6. What was your original hypothesis? Please incorporate this at the end of the introduction.

Reply 6: Our original hypothesis was that daily number of steps would increase linearly post-operatively for both cardiac surgery and VATS patients and VATS patients would recovery to baseline sooner than cardiac surgery patients.   This has been added in a concise manner towards the end of the introduction (page 5, lines 16-18), although as the second to last sentence to adhere to journal instructions with the STROBE statement being the last sentence of the introduction.

Comment 7. Please be specific about your exclusion criteria: What are "non-ambulatory patients" (i.e. Katz Index of Independence in Activities of Daily Living categories?) and "those undergoing emergency operations" (Is this EuroSCORE II urgent, emergency, or salvage?)

Reply 7: Thank you for this suggestion.   We have updated the exclusion criteria to be less vague (page 6, lines 1-2).   We did not include patients in our study for whom pre-operative walking data could not be gathered: those who did not walk at baseline and those who did not undergo elective procedures (could not be given the device pre-operatively).

Comment 8. The Vivofit is only suitable for walking (see supplemental table S2 of https://doi.org/10.3390/s21061979. What is the limitation of only recording walking and not other activities?

Reply 8: While the authors of the referenced work chose to record stationary post-operative activities such as laying, sitting, and standing, only dynamic mobilization is within the scope of our study and the clinical differences between various stationary activities is not well documented and likely of limited clinical benefit.   We do not capture cycling with the Vivofit, although from the referenced work, this was only on average 1.4 minutes per day on post-operative day 1 and then increased minimally thereafter and subsequently decreased from post-operative day 4 as seen in their Figure 3e.   We do capture other activities that involve taking steps such as climbing stairs, speed walking, jogging, etc.   We have added a comment to our limitations addressing this on page 10, lines 13-18.

Comment 9. You state: "Baseline daily steps were recorded for 3 to 10 consecutive days pre-operatively". How does this work, patients visit the outpatient clinic first, are recruited for the study, get their accelerometer while on the waiting list, and get admitted? What happens during surgery with the accelerometer? Was it on during OR and ICU stay?

Reply 9: Thank you for bringing to our attention that this was not clear.   Participants that were recruited for the study were provided with the wristband at their pre-operative, outpatient visit. They would come in for surgery and the pre-operative data was downloaded.   The wristband was

not worn during the operating room and no data was collected on the day of surgery. While in the operating room, the device was kept with their personal belongings. The wristband was then worn again on post-operative day 1 through at least post-operative day 28 regardless of disposition (intensive care, floor or home). These methods have been updated to more clearly define this on page 6, lines 8-15,

Comment 10. When was the data downloaded? Daily? Please add this to the manuscript.

Reply 10: The data was downloaded on the day of surgery and at their post-operative visit. This has been added to the manuscript on page 6, lines 15-17.

Comment 11. Why were average steps collected and not median? Is average justified? Please show in your response, i.e. by presenting an histogram

Reply 11: Thank you for raising this point. The data is not normally distributed and is right skewed. For this reason, we have changed the data to median (IQR). This has been updated in the manuscript and the Spearman correlations were re-run with median steps/day and updated in the abstract and Table 2.

Comment 12. The statistical methods are unsuitable for your hypothesis. This should be a paired measurement, and preferable, a mixed model analysis with a time effect (days after surgery) and a group effect (VATS or cardiac surgery). I would say that every day consists information, and that information is thrown away with only looking at POD1 and POD28.

Reply 12: Thank you for this feedback. Provided the non-parametric distribution of data, we have updated comparisons on POD1 and POD28 to Wilcoxon Rank Sum Tests instead of Student's T-Test. We can appreciate the value of a linear regression here, although provided the limited subjects and great variability in number of steps/day between subjects and between individuals day to day, a regression analysis with a time effect and group effect provides a poor fit. We do agree with a greater number of subjects and smoother data, this may be a better way to look at the data. Despite this, we understand the reviewer's comment and added error bars to our values in Figures 2A-C, which represent the IQR. Additionally, we have added Figure 2D, which contains the ambulatory step recovery (Median and IQR) for both cardiac surgery and VATS patients. By combining this on the same graph, this provides a visual of the difference in step recovery during week 1, but convergence thereafter between groups as noted in minor comment 8. Additionally, we have commented on this (page 8, lines 17-18 and page 9, lines 1-3). We also kept the separate figures to allow for individual visualization of the data due to concern of overcrowding on Figure 2D.

Comment 13. Why were VATS and cardiac surgery groups compared? This does not seem to be part of the aim of the study? If a primary or secondary endpoint was the difference between VATS or open cardiac surgery, it can be OK to do this. Please add the original protocol sent for review to Colorado Multi-Institutional Review Board approval #16-1776 including approval letter as a file not-for publishing

Reply 13: VATS and cardiac surgery groups were compared as it is an interesting point in how these populations return to baseline and differences between them. As mentioned in other comments, we would expect quite the difference in recovery as a nature of the operation, although there are some interesting results in how these populations return to baseline when compared to one another. The differences between these groups are a secondary endpoint. We have included the original protocol sent for review to the Colorado Multi-Institutional Review Board (COMIRB #16-1776) and the initial approval letter. A more recent version has been included to demonstrate liberalization of inclusion criteria to >18 years. These attachments have been included in the resubmission as a file not for publication.

Comment 14. You state that an "average baseline preoperative ambulation was 5393 ± 2294 steps/day." This potentially does not have a parametric distribution, as the standard deviation is 50% of the mean. A median with 25-75 quartiles might be justified. Please comment.

Reply 14: Thank you for raising this point. This is non-parametric, and medians (IQRs) have been reported instead. This has been updated in Table 2 and in the manuscript (page 7, line 20-22).

Comment 15. In general, were baseline characteristics and outcomes checked for parametric distributions? Common baseline characteristics such as EuroSCORE I or II variables are missing. Are these the only ones registered? Consider to calculate Body Surface Area with the Du Bois formula from length and weight, as this can be important for (wearable) activity tracking.

Reply 15: All continuous baseline characteristics were checked for parametric distribution. Number of steps were not parametrically distributed and therefore reporting was updated to median and IQR from mean and standard deviation as mentioned in reply 11, 12 and 14. These were not the only baseline characteristics recorded, although agreed to be relevant for walking data by all authors. We do agree baseline characteristics such as EuroSCORE II (for cardiac subgroup), NYHA HF classification and BSA are relevant characteristics that have now been added to Table 2.

Comment 16. Figure 1: How many patients were screened etc? Consider using the CONSORT flow chart as information is now lacking from the figure

Reply 16: Thank you for this comment. We have updated to the CONSORT style flow chart. The new diagram can be seen as Figure 1.

Comment 17. Consider making boxplots per day or other visualizations per group, i.e. https://doi.org/10.21203/rs.3.rs-1975144/v1 or its supplemental figures (This is just an example for visualization options. The authors should not feel the obligation to cite this work.)

Reply 17: Thank you for this comment. To better visualize the number of steps per day for each group, we have included error bars which include the interquartile range for figures 2A-2C. This has been updated in the manuscript. We appreciate the box plots in the cited work, although our data is not a percentage of time, but an absolute number of steps taken. The updated visual more

closely demonstrates the progression in number of steps with each post-operative day along with variation by showing IQR.

Minor comments:

Comment 1. Consider using keywords not being in title and abstract, i.e. MeSH terms, to improve findability

Reply 1: Thank you for this suggestion.   We have updated the keywords to MeSH terms not used in the title and abstract (page 4, lines 10-11)

Comment 2. Why was referred to ERAS only, and not ERATS?

Reply 2: We consider ERATS as a set of specific recommendations for thoracic surgery that follows general ERAS protocols, although have now added it specifically (page 5, line 4).   We have referenced the Guidelines for Enhanced Recovery After Lung Surgery from 2019 (page 12, lines 4-6).

Comment 3. Which post-operative day 1 pain rating scale did you use? Pain numeric rating scale?

Reply 3: The numeric pain rating scale 0-10 was used for post-operative day 1 pain rating.   This has been clarified with a footnote at the bottom of Table 1.

Comment 4. Was length of hospital stay defined by the number of nights?

Reply 4: Length of stay was defined as the number of midnights for which a patient was admitted to the hospital.   This has been clarified with a footnote at the bottom of Table 1.

Comment 5. Why was POD 28 selected? Start of cardiac rehabilitation or general outpatient follow-up?

Reply 5: Post-operative Day 28 was selected as this was the endpoint for ambulation monitoring in the methods section.   In the discussion, we comment on how longer post-operative monitoring of ambulation may be clinically useful (page 10, 10-12).

Comment 6. Please check your manuscript for typo's and incorrect formatting.

Reply 6: Thank you for this suggestion.   This manuscript has been reviewed by all authors with each submission for typos and incorrect formatting.   Any formatting or typing errors that become apparent after the track changes feature has been turned off will be corrected.

Comment 7. Please comment on baseline steps/day in similar populations; is 5400 steps/day normal in this population?

Reply 7: This is novel data and the baseline daily number of steps in the population undergoing cardiac surgery or VATS has never previously been described.   There is a review about quantifying the number of steps/day (https://doi.org/10.2165/00007256-200434010-00001), which suggests that in healthy adults <5000 steps/day is a "sedentary lifestyle," 5000-7499 steps/day is "low active," 7500-9999 is "somewhat active," >10000 steps/day is "active," and >12500 steps/day is "highly active."

Comment 8. The small difference after week 1 is quite interesting and unexpected. VATS is 1) less

invasive than open surgery and 2) the VATS procedures itself you describe are of a lesser invasiveness anyway. Please comment. Did both patients receive the same postoperative ambulation support from physiotherapists?

Reply 8: Thank you for this comment, while this may have appeared to be the case earlier, we hope that adding figure 2D helps clarify, there is a substantial difference between the number of steps for VATS and cardiac surgery patients during week 1 and then this difference diminishes over the next 3 weeks.   We have added a comment explaining this on page 9, lines 1-5.

Comment 9. What are the limitations of walking-information only, i.e. missing out lower intensities of ambulation such as sitting or standing, and more demanding activities such as biking or walking the stairs?

Reply 9: Thank you for this comment.   This is a limitation of this study, which has been notably measured in above referenced work.   For this reason, we have added this as a limitation in the discussion (page 10, lines 13-18).

Comment 10. Table 1: change gender into sex

Reply 10: This has been changed in Table 1.

Further Review Comments

It was a pleasure to review a revision of the manuscript entitled "Returning to Baseline Daily Ambulation after Cardiothoracic Surgery" for publication in VATS. Thank you for the invitation.

I have carefully read the revised manuscript, and the Authors' reply to my comments (R1) including their STROBE checklist and approved institutional review board documents. The authors improved the manuscript with their rewriting, adding figures and updates tables. These new figures allowed for better interpretation of your work, and are well presented.

Original comments (Labeled by VATS as R1)

Original comments 1-3, 5, 7, 9, 12-17 have been answered satisfactorily, thank you.

Comment 4. Thank you for acknowledging the limitation of small sample size. The original comment on why only 21 patients were included is still open. Especially with the new figure 1, 344/374 patients were excluded because they did not undergo cardio-thoracic surgery. Is it true that you perform cardio-thoracic surgery on 30 patients in a 2-year interval? That is an extremely low and unlikely annual case volume. Please comment.

Reply 4:   Thank you for this comment.   While our case volume for cardiothoracic surgery is higher than 30 patients over a 2-year interval, we only screened elective cardiothoracic surgery patients (page 6, lines 1-2) as only patients who visited clinic pre-operatively could reliably have their number of steps recorded for the 3-10 days pre-operatively.   The remainder of our case volume is on an urgent basis while the patient is admitted.

Comment 6. Thank you. Make sure to also cite the STROBE recommendations paper at the end of

your introduction.

Reply 6:   Thank you for this comment, we have added this as citation number 8 (page 5, line 19).

Comment 8. This is clear now. Consider to cite the source of your statement "cannot quantify stationary activity such as laying, sitting or standing", i.e. your REF 6. It remains interesting from a scientific and clinical point of view where the rehabilitation field will move to: optimize activity characterisation, or deal with step count limitations.

Reply 8: Thank you for raising this point, we have added reference 6 following this statement, page 10, line 15.

Comment 10. The data was downloaded at day of surgery and post-operative visit. Would you in the future want to use the stepcounts in a more personalised rehabilitation path?

Reply 10: Interesting point and a direction of future work that our group is pursuing.   Provided that some devices can be monitored remotely, we certainly think that we can monitor a patient's number of steps in real time and if they are not returning to the baseline in the expected amount of time, then we could intervene by contacting the patient.

Comment 11. Thank you for assessing the (normal) distribution of endpoint parameters throughout the manuscript. This improved the scientific soundness of the work. For me, reporting median (IQR) is uncommon, difficult to understand, and leaves out important information about the skewness of the data. I suggest to report median (25th percentile - 75th percentile) instead throughout the manuscript text and tables. Finally, did you also check Table 1 characteristics, i.e. BMI, age, OR time, hospital stay etc for a normal distribution?

Reply 11: Thank you for the suggestion of reporting 25th-75th percentiles.   We have updated this in our results, page 7, lines 20-21 and Table 2.   All continuous variables in Table 1 were reviewed for normal distribution. Age and BSA were normally distributed and therefore left as is.   BMI, NYHA HF Classification, Operating room time, POD #1 pain rating and hospital length of stay were not normally distributed and therefore have been updated to median (25th-75th percentiles). Comparisons for these non-parametric variables have also been changed to Wilcoxon-rank sum tests with updated p-values, also updated in the manuscript (page 7, line 19).

Original minor comments

Original minor comments 1-6, 8, 9 have been answered satisfactorily, thank you.

Comment 7. Thank you for adding that source. I came across a systematic review very recently (DOI: 10.1186/s12966-022-01261-9) where Figure 2 seems to show high levels of variation between patient populations. Comparing to your study, your patients seem to mobilize adequately. Consider to add a comparison of that systematic review to your discussion, showcasing your good results.

Reply 7: Thank you for finding this relevant paper.   We have added a comment regarding our patient population and their adequate mobilization post-operatively in comparison to the data available in the literature using these two references (page 10, lines 8-9).

Comment 10. Thank you, make sure to also change gender into sex on page 10 line 2.

Reply 10: Thank you for raising this point, this has been corrected.