Article details: 2022-0020
Title: Potentially avoidable admissions to general internal medicine at an academic teaching hospital: an observational study
Authors: Alex M. Cressman MD MSc, Ushma Purohit MD HBSc, Ellen Shadowitz, Edward Etchells MD MSc, Adina Weinerman MD MHSc, Darren Gerson MBA, Kaveh G. Shojania MD, Lynfa Stroud MD MEd, Brian M. Wong MD, Steve Shadowitz MDCM MSc

Reviewer 1: Dr. John Staples
Institution: The University of British Columbia
General comments (author response in bold)

This study seeks to identify the proportion of hospital admissions to General Internal Medicine at a single Canadian academic centre that are potentially avoidable then contrasts the characteristics of PAAs to those of non-avoidable admissions. Investigators report that 11% of admissions were potentially avoidable and that these admissions occurred among younger, less medically complicated, and less ill patients. The factors contributing to the decision to admit are plausible and diverse. Investigators state these findings might be used to inform quality improvement efforts.

Strengths:
- Clear study design
- Sizeable number of admissions screened
- Clearly written with straightforward presentation of results
- Main finding that 11% of admissions are potentially avoidable is plausible and might be a useful estimate for administrators, clinicians, and clinician-educators to know
- Table 2 highlights the diversity of factors contributing to the decision to admit and highlight the challenge of addressing PAAs at a programmatic level.
- The discussion is thoughtful and balanced

Weaknesses and uncertainties:

1. The main results present a plausible but unsurprising description of GIM admission decisions at an academic centre. No hypothesis is tested. The results thus provide a somewhat modest increment of new knowledge about admission decisions.
We agree with Dr. Staples. This is a descriptive, observational study, so we did not test specific hypotheses. We are the first to undertake a prospective, near real-time evaluation of avoidable admissions to GIM at the time of ED admission. Further, we used a detailed multi-step method for identifying PAAs and evaluated the patient, provider and system factors associated with avoidability.
The goal for this study was to develop and report on a novel method to identify PAAs more accurately at the time of ED presentation, and to identify factors that influence avoidability. While our study was focused on methods development/feasibility testing rather than testing specific hypotheses, our data may provide insights about how addressing these factors may reduce avoidable admissions.

2. Single centre results may not generalize to other sites in Ontario, in Canada, or outside of Canada.
We agree with Dr. Staples. This is a single-centre study. We have added this as a limitation to our study [Page 11, Line 5-9]. However, our method may help other institutions determine their own rate of PAAs and factors contributing to them.

3. In the setting of clinical ambiguity (perceived need for monitoring, access to social supports), inexperience (i.e., trainee typically making decision to admit), and workflow realities (at least in my institution, it is less work to admit medically and socially complicated patients in the evening than to send them home from the ED), making a conservative decision to admit and 'sort it out in the morning' seem quite reasonable. It seems appropriate for trainees to bias their independent decisions toward more conservative and safer courses of action. Perhaps the discussion should acknowledge this more fully.
We agree with Dr. Staples. We have added this point to the discussion accordingly [Page 9, Line 18-27].

4. Raising the threshold to admit might reduce the number of PAAs, but it also might increase the number of inappropriate discharges. Some acknowledgements of this trade-off and the types of research required to address it would strengthen the manuscript. Overall, it is not clear to me that 11% is too high; this design will not provide data to establish the PAA threshold that best balances patient harm and system inefficiency.
We agree with Dr. Staples. We have added this point to the discussion accordingly [Page 9, Line 18-27 and Page 10, Line 1-2]. We believe that our design allows us to estimate a rate of 11% avoidable admissions which we agree might be too low to be actionable. However, we anticipate that other institutions may have higher avoidable admission rates but do not currently have a method to assess this. Our method provides institutions with a way to measure their own avoidable admission rates and to identify potentially associated factors.

5. Table 2 should have %s. Factors should be presented in order of decreasing frequency (within each category). Correct The Title-type Capitalization to Sentence Case.
Authors use the term 'gender' (a social construct) when I believe they mean sex (a biological categorization more often captured in EMRs). e.g., p9
We agree with Dr. Staples. We have now made these changes to Table 2.

Reviewer 2: Dr. SK Peasah
Institution: Mercer University
General comments (author response in bold)

1. Authors sought to develop a methodology to reduce avoidable admissions from emergency departments, test it, and characterize contributing factors (patient, provider, and system). I believe it is a laudable objective and timely, given the impact of the pandemic on hospital-bed capacity. I have some few comments for consideration.

2. Can you define the expertise of the research clinical team- are they all staff physicians from the research site?
We agree with Dr. Peasah. We have now clarified the expertise of the members of the research clinical team [Page 4, Line 3-5] and that all research team members are from the research site.

3. If staff physicians have access to patient conditions and records, wouldn't that introduce bias during review, since it will be from hindsight and likely involve their team?

We agree with Dr. Peasah that hindsight bias is an important consideration. However, we mitigated this source of bias by removing research team member(s) from discussion if they had been the most responsible physician for the candidate PAA. We have further clarified these points [Page 5, Line 20-24].

4. Page 7 line 12 you mentioned that your model is based on prospective surveillance studies, but you have no references for that statement.
We agree with the reviewer that more information would be helpful. The Avoidability Scale was a 4-point Likert scale, based on the Likert Scales used in the Canadian Adverse Events Study and Trigger Tool Methodologies and modified for our purposes (Ref 19 and 33).
We piloted the Avoidability Scale in our initial research group meetings to ensure that all faculty agreed with the approach prior to beginning formal data collection.

5. You used Likert scale to characterize PAA into 4 groups then you collapsed it into 2 groups - could you also show the proportions in each group?
We agree with Dr. Peasah. We have provided this data here and can include in the manuscript results if deemed appropriate.

	Number of Occurrences
Avoidability Score = 1	**25**
Avoidabiity Score = 2	**25**
Avoidabiity Score = 3	**43**
Avoidabiity Score = 4	**24**

6. I am curious to know how consensus was reached by the research team after discussing the case, was it by vote?
Our research team case review process occurred as outlined below.
1) AMC presented the information for case review
2) There were questions and case discussion among the research team related to case details.
3) When there were no further questions, there was a vote using the avoidability scale (1-4)
4) In event of disagreement, which occurred infrequently, we went with majority for determination of avoidability (Yes/No)
5) For cases with PAA ≥ 3, we discussed the relevant factors as in Appendix 1/Table 2.
We have added this information to the methods and added additional detail to our Figure 1 and Figure Legend.

7. Out of the 601 cases in your cohort, only 83 (14%) came from Daytime, where you had staff physicians, and 86% from evening and night (with only residents and medical students). Of the 83 daytime cases, only 8% were avoidable cases compared to 12% from the non-day time cases. Could the difference be due to lack of staff physician oversight after daytime shifts? This should be listed as a limitation.
We agree with Dr. Peasah. We tested this hypothesis in Table 1 and outlined in results, and our analysis did not show a significant difference based on time of admission (day vs. evening/night) or weekday vs. weekend. We were surprised by this finding and do wonder if this was potentially related to inadequate statistical power.

8. I was looking at the scalability of your findings. While I acknowledge that your findings could be useful for identifying avoidable cases, it will be labor intensive, and I wonder how easily it can be implemented elsewhere. I was hoping for a tool (electronic?) for identifying such cases, testing, and sharing it, that might most likely be used by others.
We agree with Dr Peasah. There is some effort required to conduct case synthesis and reviews. At present, there is no electronic tool available currently for this process which requires significant clinical judgement. We suggest that others could practically undertake smaller reviews with similar methods to identify PAAs and contributing factors. We acknowledge that currently we do not have the ability to conduct our method in an electronic or automated way.

Reviewer 3: Dr. David Forner
General comments (author response in bold)

Thank you for the opportunity to review Potentially avoidable admissions to general internal medicine at an academic teaching hospital by Cressman and colleagues. In this prospective cohort study, the authors have developed a novel approach to identify potentially avoidable admissions (PAAs) through a multi-step review of admission lists and courses in hospital. PAAs were defined as a patient presentation that could be managed in a timely, effective, and safe manner in the ED or ambulatory-setting. This is a well-written paper addressing an important, interesting, and at times contentious, issue. Some considerations:

1. The interviews with resident physicians are described as semi-structured. Could you please provide the semi-structured interview guide, assuming one was used?
We agree with Dr. Forner that the debriefing process requires further methodological clarification. We have addressed this comment above. Please see Appendix 1 for the interview guide.

2. Similarly, how was this qualitative data analyzed? As presented, it seems to essentially be presented in a quantitative format (tallying common admission reasons).
We agree with Dr. Forner that the debriefing process requires further methodological clarification. We have addressed this comment above. Please see Appendix 1 for interview guide.
As outlined above, we did not conduct a formal qualitative analysis of interview data. We conducted brief interviews with our trainees to ensure all case data was correct and reviewed the commonly cited factors. Our manuscript has been updated accordingly.

3. Who analyzed the interview data?
We agree with Dr. Forner that the debriefing process requires further methodological clarification. We have addressed this comment above. Please see Appendix 1 for interview guide.
In line with responses to questions #1 and 2 and previous reviewer comments, this was not a formal qualitative study. The interview data was not formally analyzed apart from review of the cases, documentation and counting of the relevant factors.

4. How was consensus achieved for the avoidability score? Was any standardized method used? Can the authors comment on initial disagreement? Importantly, the proportion of disputed cases that would not have been considered further (PAA score <3) should the consensus have gone the other way.
We did not record the frequency of disputed values. There were very few cases where there were disputes about avoidability among our team, but rather most disputes were related to the extent of avoidability (e.g., 1 vs. 2 or 3 vs. 4).

5. While the authors have shown a low rate of missed PAAs within the 30 cases of non-flagged cases that they reviewed; this represents only about 5% of the total cases considered by the attending physicians. It is excellent that inter-rater reliability was assessed here, but the proportion of cases analyzed is relatively low.
We agree that a larger random sample of unflagged cases would have yielded a more precise estimate and confidence interval (0 – 17%). However, we are confident in our method given our request for staff physicians to maintain a low threshold to flag candidate PAAs to make case-finding as inclusive as possible during this initial phase.

6. The authors collected data on consult volume on both the day of admission as well as the preceding 48 hours, but this does not appear to be analyzed. This seems to be an intriguing and potentially related factor, which anecdotally may play a role in the decision to admit.
This data was analyzed using univariate analyses and was not found to be associated with differential rates of avoidable admissions. This data has been added to the Results [Page 7, Line 22-25] and Table 1.

7. Please include abbreviations in the table footer (RIW in table 1)
We agree with the reviewer. We have changed the RIW abbreviation in the Table to "Resource Intensity Weighting" Score and have included the reference to the derivation of this score through CIHI.

8. For those unfamiliar, can the authors explain the government directive of 4-hour limit for admission decision. Is this penalized in some way? Is this a trackable statistic? How does this, practically speaking, affect decision making and how is it typically considered?
We agree that this manuscript section requires additional explanation for readers outside of Ontario. This is an Ontario Ministry of Health standard that is tracked; not penalized but considered as a performance metric for ER efficiency. Practically speaking, the 4-hour limit for admission decision making influences the decision to admit as internal medicine trainees are encouraged to make rapid disposition decisions and are discouraged from holding patients for prolonged periods of monitoring in the ED.

9. Was the RIW a measure of the entire admission or the resources utilized within the ED only? That is, was it truly a predictor of PAA, or was it to show resources were wasted?
The RIW was calculated at the time of hospital discharge and used to show resources that were potentially wasted and/or could have been re-allocated to the outpatient setting. We have updated our methods accordingly [Page 6, Line 13-18].

10. Can the authors please comment on the RIW scale within the paper itself (within the methods) to give the reader a sense of the size of resource use without having to read the original paper)

We have included background on the RIW from CIHI in the Methods [Page 6, Line 13-18] and References 20-21.

11. Please add %s to the discharge destination columns in Table 1

We have addressed this comment in Table 1.