Article details: 2014-0116

Title: Effect of physician specialist alternative payment plans on administrative health data: a validation study

Authors: Ceara Tess Cunningham MA, Nathalie Jetté MD, Bing Li MA, Robyn Ravneet Dhanoa MN, Brenda Hemmelgarn MD PhD, Tom Noseworthy MD, Cynthia A. Beck MD, Elijah Dixon MD, Susan Samuel MD MSc, William A. Ghali MD, Carolyn DeCoster PhD RN,[3] Hude Quan MD PhD

Reviewer 1: Sanober Motiwala: University of Toronto, Dept of Health Policy, Management and Evaluation

Reviewer 1 comments and author response

1) REVIEWER #1 COMMENTS: Page 9, Lines 25-26: what proportion of the 904 physicians surveyed were APP versus FFS, i.e. was there a noteworthy difference in consent rates among the two groups? Authors note a total response rate of 35% from all groups.

CHANGED: Page 7, lines 12-13: "Of the 317 consenting physicians, 38.1% were FFS, 14.1% were part APP and 47.8% were full APP."

2) REVIEWER #1 COMMENTS: Page 13, lines 12-19: Recommendation regarding incentive programs should be clarified. In the preceding sentence, the authors note that physician payment programs do not have much of an impact on quality or quantity of claims. Are the authors proposing that incentive programs be introduced for APP physicians, since the study showed a significant difference is pre/post APP billing rates? The word quality is sometimes used in places where quantity or frequency or submission rate may be the more accurate term.

CHANGED: Page 11, lines 15-22: "These findings suggest that contrary to popular beliefs, the implementation of APP physician payment programs in Alberta may not have as much of an impact on the quantity or frequency of physician claims submitted. Since Alberta uses shadow billing incentive programs, future research is needed to determine whether incentive programs should be considered in other provinces or nationally in order to preserve the overall quality of physician claims data. Additionally, there should be continuing vigilance with respect to completeness and frequency of physician claims submitted, regardless of the payment system."

3) REVIEWER #1 COMMENTS: Page 11, lines 44-50: does the statement cited from reference number 16 pertain to Alberta or another province?

CHANGED: Page 10, lines 1-4: Added reference for billing context incentives for Alberta. Also see added reference # 17.

4) REVIEWER #1 COMMENTS, minor typos:
- On page 5, line 25, check subject-verb tense agreement between objectives and was
- On page 8, line 34, spell out the number 3
- On page 11, lines 19-20, use of "it was"
- On page 11, lines 25, there appears to be a word missing in "inaccurate data inaccurate linkage"

CHANGED:
- On page 3, line 7-10: "Given these circumstances, the goal of the project is to determine the proportion of claims submitted by APP and FFS physicians and to identify and compare the validity of information coded in physician billing claims submitted by APP and FFS physicians in Calgary, Alberta."
- On page 6, lines 10-11: "The physician claims database contains up to three diagnoses.
- On page 9, lines 15-16: "We assumed that submission rates would be close to 100% for FFS physicians, however, rates were closer to 90%."
- On page 9, lines 16-18: "Potential explanations that could account for missing billing submissions include physicians or administrative staff forgetting to submit a claim, rejected claims and inaccurate data linkage."

Reviewer 2: Eric Benchimol: The Hospital for Sick Children, Gastroenterology, Hepatology & Nutrition

Reviewer 2 comments and author response

1) REVIEWER #2 COMMENTS: Another major weakness is the exclusion of outpatient claims, which are those most often used by researchers involved in health services research. This is addressed in the Discussion. However, it was not clear that the DAD and ACCS data are inputted by the physicians themselves (or their billing/coding representatives). It was my understanding that CIHI data is entered by professional hospital coders. Therefore, the fact that the physicians are on APP should not impact completeness of data in any way. Is this also the case for ACCS? I believe the physician billing databases are the ones that should have been validated? physicians on APP may not be submitting those billings since they don't affect their income.

COMMENTS: In regards to the comments from reviewer #2, the objective of this study was not to validate the DAD or ACCS databases. They were used to determine the frequency of the medical services that were provided, which was then linked with the diagnosis code from physician claims database in order to examine whether the claims/ICD code matched the service provided. The purpose was to validate the physician claims database; where most often the claim is completed by the physician or billing clerk; however billing clerks are not professional coders as are the individuals who do the coding for hospital discharge abstract administrative data . In Alberta, physicians on an APP who do not submit a shadow billing claim may be subject to different punitive actions such a withholding of funds at the end of the year, if their shadow billing claims are not deemed to be submitted on regular basis or in comparison to their physician peers. Thus, APP physicians' income may be affected by irregular or under submission of shadow billing claims. We hope this clarifies the reviewer's questions.
CHANGED: Page 8-9, lines 23-24, lines 1-2: "In Canada, hospital and emergency discharge abstract administrative data are coded by professional health record coders. These data are not impacted by physician payment programs. However, physician claims are coded by physicians themselves or billing clerks who are not professional coders. "

2) REVIEWER #2 COMMENTS: One valuable addition to this study would be a multivariable regression model to determine predictors of completion of claims, with a resulting predictive score. For example, if the co-variables (e.g. physician specialty, years of practice, APP/FFS status, incentive Y/N, etc) could be weighted mathematically, other researchers could use this score to determine/adjust their analyses when determining completeness of codes.

COMMENTS: While the comments from reviewer #2 are important to address, the goal of this study was not to explore the predictive factors of completion of claims, we were interested in exploring the data features of the physician claims database and especially the difference of claims

submissions between FFS and APP physicians. Also due to the small number of participants there was low statistical power to assess a predictive model and be confident about the results. Our expectation was that the estimates would be imprecise if stratified analysis or predictive models were developed. We agree that this is an area that will need to be explored in future study.
CHANGED: Page 11, lines 1-4: "We did not assess predictors of non-submission due to lack of statistical power. Models predicting or screening physicians who are less likely to submit shadow claiming should be developed in the future. The model will be helpful for promoting submission of billing claims among targeted physicians.

3) REVIEWER #2 COMMENTS: Results, Paragraph 1, Page 9, lines 25: Regarding the inclusion/exclusion of physicians, the number of physicians who responded to the request but declined to participate should be included (of the 587 physicians who did not consent). Could these physicians be different in some way? For example, could the consenting physicians be more interested in research or academic contribution and therefore be more likely to submit their claims? Could a random sample of charts by physicians who did not consent be extracted to determine whether these are different?

CHANGED: Page 7, lines 7-12: Of the 587 physicians who declined to participate, a follow-up questionnaire was sent to with the aim to explore reasons for non-response. Of those who responded to the follow-up survey (n=63), 70.5% were males and 29.5% were females. Respondents for this follow-up survey were from the following specialties: Internal medicine (34.2%), general surgery (27.8%), neurology/neurosurgery (0.03%), paediatrics (18.2%) and psychiatry (16.9%).

COMMENTS: A random sample of charts from those physicians who did not consent is not possible as these physicians did not provide us access to their unique physician identifier number (i.e. PRAC ID). We would need this number and their waiver of consent to access their claims data.

4) REVIEWER #2 COMMENTS: Results, Paragraph 1, Page 9, lines 27: Why were only 182 of 317 consenting physicians included? I cannot find this plan described in the Methods. How were these physicians selected? Why were the rest excluded?

CHANGED: Page 7, lines 14-16: We were only able to link 182 out of the 317 unique PRAC ID's, this is most likely due to a wrong or inaccurate PRAC ID provided by physician in the original survey.

5) REVIEWER #2 COMMENTS: The authors also attempted to validate the accuracy of the diagnostic codes used by APP and FFS physicians, and PPV was calculated by comparing the administrative data codes to chart review of the chart. Again, if professional coders are submitting the administrative data, I'm not certain the comparison of APP and FFS physicians is of much value. In addition, it would be useful to know the most common diagnoses by these physicians, and to determine whether there were differences in the types of diagnoses seen by APP vs. FSS physicians. The PPV of codes of different diagnoses should be included in Table 3 (perhaps just the top ten).

COMMENTS: Again it is important to clarify that we were indeed validating physician claims data, not CIHI data coded by professional coders. It was the purpose of the study to examine differences between the payment plans (i.e. APP vs FFS) and claims submission. Due to small amount of physicians included in the study, exploring the diagnosis types would not be especially beneficial, due to low power and sample size. Authors CTC, HQ and NJ have recently published a paper that examines the different types of diagnosis by FFS and APP physicians. See the following paper: Mining Rich Health Data from Canadian Physician Claims: Features and Face Validity. Cunningham, CT., Cai, P., Topps, D., Svenson, LW., Jette, N., Quan, H. BMC Research Notes. 2014; 7: 682-670.

6) REVIEWER #2 COMMENTS: The above analysis (PPV of diagnostic codes) is really the PPV of the claim data reflecting the chart documentation. It does not necessarily reflect the patient actually having the disease in question. This should be included in the Discussion. However, if the chart abstractors were trained to accurate classify/identify certain diseases, this should be reported in the Methods.

CHANGED: Page 11, lines 5-7: Additionally, it is important to note that in this case the PPV does not necessarily reflect the patient actually having the disease in question. It shows the proportion of services provided by physicians was billed.

7) REVIEWER #2 COMMENTS: Table 3 ? The overall claims row demonstrates lower PPV in the APP group compared to FFS (85.8% vs. 87.0%), however each of the specialties reported showed the opposite (higher PPV in the APP group). Why is this? Are we missing some of the specialties that drive the PPV lower in the APP group?

COMMENTS: There is an OTHER group that is driving the lower PPV. We did not report this sub-specialty group such as Intensive Care Unit (ICU) physicians, hospitalists, psychiatrists and paediatricians.
CHANGED: Page 17, Table 4: "*We did not report PPV for other specialties, such as such as Intensive Care Unit (ICU) physicians, hospitalists and psychiatrists.

1) REVIEWER #2 COMMENTS Minor comments: Did the authors assess the softer claims involved with hospital admissions or ACCS visits? For example, activities such as patient/family counselling, or care team meetings can be claimed by FFS physicians in many jurisdictions. These may be much less likely to be claimed by APP physicians. I'm not certain whether these types of claims are included in the DAD and ACCS.

COMMENTS: No we did not asses any type of softer claims, these types of claims the reviewer describes are more focused on outpatient treatments and due to time and feasibility issues, we opted to assess inpatient claims. We acknowledged this limitation.

2) REVIEWER #2 COMMENTS: Introduction, Paragraph 2, Page 4, Lines 34-50: I wouldn't recommend embarking on an indictment of the inefficiencies of FFS medicine in the Introduction. Perhaps this should be saved for the Discussion, or omitted altogether.

CHANGED: Page 2, lines 14-18. Removed the following statement: "The FFS payment mechanism creates financial incentives for physicians to encourage over consumption of care, since physicians are rewarded for a high volume of services. In other words, physicians get paid more when their patients consume more care [9]."

3) REVIEWER #2 COMMENTS: Methods, Page 8, Line 18-19: I would prefer that a Kappa (with confidence intervals) replace this ?97% agreement? statement. The Kappa accounts for agreement by chance, while simple agreement does not.

COMMENTS: When the reference standard is not accepted as 'gold standard', Kappa is used. Kappa is impacted by positive and negative agreements and is hardly understood by lay-statistical readers. Thus we chose to report agreement between two reviewers. The agreement level is very high (97%). In this case, the Kappa score will also be high even if we report that. Thus we felt that the proportion of agreement between chart reviewers was easier to understand to the general reader as opposed to a Kappa score. If the editors believe the Kappa score is necessary, we can calculate and replace the proportion of agreement.

4) REVIEWER #2 COMMENTS: Discussion, Paragraph 1, Page 10, Lines 38-45: I would suggest rewording this summary sentence to reflect the fact that there was a statistically significant difference between APP and FFS physicians. For example, this could be reworded to found that Alberta APP physicians submitted statistically fewer medical service claims than FFS physicians, however the vast majority of claims were submitted in both groups.

CHANGE: Page 8, lines 20-23: "found that Alberta APP physicians submitted statistically fewer medical service claims than FFS physicians, however the vast majority of claims were submitted in both groups and that the accuracy of diagnostic coding of these claims is also high."