Article details: 2018-0013	
Title	Validation of 5 key colonoscopy-related data elements from Ontario health administrative databases compared to the clinical record: a cross-sectional study
Authors	Jill Tinmouth MD PhD, Rinku Sutradhar PhD, Ning Liu MSc, Nancy N. Baxter MD PhD, Lawrence Paszat MD MS, Linda Rabeneck MD MPH
Reviewer 1	Balthasar Hug MD
Institution	Luzerner Kantonsspital, Internal Medicine, Lucerne, Switzerland
General comments (author response in bold)	Comments to the Author Review "Validation of 5 key colonoscopy-related data elements from Ontario health administrative databases compared to the clinical record: A cross-sectional study" CMAJ Open 2018-0013 by Tinmouth et al. General: The authors describe how they validate five data elements regarding colonscopy using administrative databases in Ontario. The study is well written and the topic timely. There are some minor issues that need to be addressed as outlined below. Methods: - Definitions, P.4, line 10: The authors mention "..these definitions…", but there were no definitions of e.g. "colonoscopy case" etc. The definitions follow on p. 5, last paragraph. But on p. 4 the reader has no clue yet of how "colonoscopy case" is defined. So either refer the authors on p. 4 that definitions will follow on p. 5 or they move the definitions to p. 4 where these expressions first appear. Definitions are crucial for this study e.g. the authors mention that "There were 14 alternative definitions for colonoscopy case…" on line 51, p.5. Definitions decide how strong the selection bias for the study will be. **Response: Please see answer to editor point #6. The 14 definitions of colonoscopy case are technical, consisting of different combinations of OHIP codes. We felt it prudent only to limit the listing of the specific codes as these could overwhelm the reader, therefore they are listed only once in Figure 2. We have added information in brackets in the methods and Table 1 to help direct the reader to this Figure where appropriate.** - Random selection, P. 4, lines 33-36: "We randomly selected 23 hospital and 5 non-hospital facilities in Ontario to participate in the study. The hospital sites were selected in a stratified fashion…" This two sentences are confusing: first, how did the authors "randomly" select the hospitals? What was their technique for that? Second, "randomly" and "selected in a stratified manner…" exclude each other. The authors should be clear about how they chose their study hospitals and express this in a clearcut manner. **Response: Please see answer to editor point #7.** - Medical Record Abstraction, p. 4, lines 49-54: the authors selected data "from April 1 2008 to March 31 2009…" in a manuscript submitted in 2018. It would help the reader to see an explanatory sentence about why the authors chose to use these data and none e.g. from 2009 to 2017 and discuss that in the limitations section of the paper. Also, more recent data might be more interesting or there might be a selection bias using this one year or there might have changes taken place in the Canadian Health Care System that might affect the results of this study. **Response: Please see answer to editor point #4. Text has been revised in the limitations section and a new appendix created. It is beyond the scope of this study for us to abstract charts from more recent years.** - Sampling strategy, last line p. 4-first line p. 5: "..using a stratified sampling strategy previously used by others 22." It would help the reader to have this sampling strategy explained in a sentence or two without having to download the paper ref #22 and read all that by himself. **Response: Added to methods (p. 5).** - "E codes", p. 5, line 54: the authors mention "E codes" probably meaning ascending colon, descending, transverse right and left? This should be explained in a sentence to the international reader who will not have a clue about the Canadian colonoscopy coding system. **Response: This is explained in the methods on p. 6 in the section "Administrative data definitions and reference standards for colonoscopy data elements"** Discussion - Limitations section, second to last paragraph on p. 9: The authors emphasize their "rigorous sampling strategy" (p. 9 line 43) although we have not heard anything about that, how the authors did that etc. except the flow chart in Figure 1, see above. They mention random selection several times regarding the clinics and the patient charts, but nowhere do they

	explain how they did it. Did they use a computer program for this step? **Response: We have provided a description of the sampling strategy for facilities (p. 4) and charts (p.5). The sampling was performed using SAS (also added to text).** - Also, as a reader I wonder what might have influenced the study next to selection bias: e.g. the years of training of the gastroenterologist? This is a point the authors definitely should discuss in depth in this manuscript. The word "training" just appears in ref. #7 by Kaminsky et al (2016). **Response: As we sampled charts randomly from 28 randomly selected facilities, we the endoscopists whose reports were abstracted would have also been randomly selected. Therefore, we would not anticipate any systematic biases related to the endoscopists who were included.** - P. 9, last line: "The methods described here are reproducible…" this is not something for the authors to mention; other authors have to analyze if these data are reproducible. These authors have not proven that their data is reproducible and this was not a goal of the study. This sentence has to get cut out. **Response: Revised as suggested.** Figures - The print in Figures 2 and 4 are too small **Response: Revised as suggested.**
Reviewer 2	Stuart Nicholls PhD
Institution	Children's Hospital of Eastern Ontario (CHEO), Research Institute, Ottawa, Ont.
General comments (author response in bold)	Comments to the Author Validation of 5 key colonoscopy-related data elements from Ontario health administrative databases compared to the clinical record: A cross-sectional study In the present study the authors seek to validate procedural code information relating to colonoscopy. The need for validation of health administrative data is an important preliminary step to future studies that make use of this data. Overall point – the STROBE guidelines are superceded by the RECORD guidelines that were developed as a STROBE extension (and is cited in the paper) for studies using routinely collected data. As such, RECORD is the appropriate reporting guideline. **Response: RECORD is the guideline that should be used for observational studies using routinely collected data, this study is a validation study of data elements derived from routinely collected data. Our study type would be used to meet the criterion listed in RECORD item 6.2. Many of the requirements for RECORD are not applicable to our study as we are not using routinely collected data to measure an association between an exposure and an outcome. We have completed the RECORD guideline as best as possible.** Methods 1. How were outpatient visits for gastroscopy, flexible sigmoidoscopy and colonoscopy identified at the hospitals? Were ICD codes, keyword searching, or billing used? If OHIP billing was used, to what extent would that be artificially inflating the accuracy of the algorithm? E.g. if the OHIP billing codes is used to find the initial cohort, how accurate is that initial tranche and does it make it more likely that an algorithm built on OHIP codes is going to be highly accurate? Are the authors confident that this initial search captured all possible colonoscopy cases? **Response: In order to identify colonoscopies done at hospitals, we used the OHIP code Z555A (alone or without other codes), which is the base code that common to all colonoscopies billed with OHIP. There is >96% overlap between OHIP and CIHI codes for colonoscopy. Less than 4% of charts found using OHIP cannot be found in CIHI and vice-versa, therefore, neither data source is likely "perfect". Any effect on the accuracy of the algorithm is likely to be small. There were additional important reasons for us to use OHIP codes to identify colonoscopies. As 25% of colonoscopies were done in non-hospital clinics, use of OHIP codes allowed us to identify colonoscopies done in these settings. This approach also allowed us to capture flexible sigmoidoscopies billed using colonoscopy codes in our sample, an important source of potential misclassification. Finally, the structure of the OHIP codes for colonoscopies was key to our stratified sampling procedure, which sampled by the extent of the colon visualized, allowing for the inclusion of rare events. The structure of the CIHI codes for colonoscopy would not permit this approach.** 2. While it may be inferred from the context, it would be useful to know which specific CIHI databases were being used **Response: As colonoscopies can be performed in outpatients or inpatients, we used CIHI's Discharge Abstract Database (inpatient) and Same Day Surgery database (outpatient). This information has been added to the methods (p.4).**

3. Given the authors also had linked CIHI data, I was surprised to see reliance on OHIP and presence or absence of a CIHI record. Was consideration given to also utilizing procedure codes within CIHI data to improve algorithm performance or as sole data source given the potential to then produce hospital level analyses in data held only by CIHI?
Response: When identifying colonoscopies and their characteristics in Ontario, in general, OHIP has several advantages over CIHI. Because of these advantages, OHIP is the most commonly used data source for research and quality improvement; hence, we felt it was the most relevant source to validate when measuring colonoscopies. Specifically, at least one quarter of colonoscopies in Ontario are performed in non-hospital facilities, which do not submit data to CIHI. As a result, OHIP is a better data source for population-based studies, as it will not miss colonoscopies performed outside of hospitals. In addition, the OHIP codes provide a greater level of detail than the CIHI codes – for example, the OHIP polypectomy codes are more descriptive (provides some information on size of polyp, method of removal and number) and the structure of the OHIP codes for colonoscopy (a base Z code + the E codes for each segment of colon reached) allows for the measurement of complete vs incomplete colonoscopies, which is an important quality indicator.

4. As per Quan et al 2004 and to a degree De Coster et al 2008, which the authors cite, some more descriptive data on the comparison between chart and administrative data would be useful. If the authors could include – as a supplementary table perhaps – some descriptive information on the presence of key items in both charts and admin data it would help with contextualizing the sensitivity, specificity, PPV and NPV.
Response: For this study, when all combinations of administrative data definitions and reference standards are considered, there were distinct 27 comparisons, each generating a 2x2 table with unweighted data and a 2x2 table with unweighted data. Therefore, there are 54 2x2 tables in total, which we felt would be too overwhelming for the reader. If there is a desire to include these tables in the supplementary materials, they can be provided.

5. I remain unclear what the implication is for sigmoidoscopy billed as colonoscopy, perhaps the authors can elaborate the discussion here. For if sigmoidoscopy is included within a case definition of completeness of colonoscopy, won't this generate false positives?
Response: As indicated in Table 1, procedures billed as colonoscopies (ie according to the administrative data) but were intended as flexible sigmoidoscopies (ie according to the chart review) were included in the cohort (not in the administrative data definition of 'colonoscopy completeness' as suggested by the reviewer). We included these procedures in the cohort for the purposes of measuring the test characteristics of the completeness data element as they contribute to false negatives (according to the billing data, appear to be incomplete colonoscopies but are really flexible sigmoidoscopies billed using colonoscopy codes).

6. I am unclear as to what the reference standard was for the colonoscopy. Was it colonscopy performed, or performed or intended? The tables also indicate a number of excluded cases. Please can you provide a complete definition of the reference standard cases.
Response: The reference standard for colonoscopy case ("colonoscopy intended or performed)" and the source (endoscopist procedure note in the medical chart) is described in Table 1. We have reworded it in the Table to try to make it more clear as well as adding similar text to the methods (p6.). The variation in cohort sizes is explained in the answer to editor point #15.

7. The authors mention 2 reference standards for anesthesiologist-assistance and for polypectomy. This fact suggests that there is no standard. Perhaps, again as per Quan et al 2004, the authors might also consider reporting Kappa agreement as a more 'neutral' description of agreement?
Response: While we understand that it may be confusing to have 2 reference standards, we chose to report in this way for polypectomy and anesthesiologist- assistance as in practice, these data elements are often considered surrogates for adenomas and the use of propofol respectively. As such, we report the test characteristics for 2 reference standards: 1) what they should actually measure (was a polyp documented by the endoscopist and did an anesthesiologist attend the colonoscopy) and 2) for the surrogates (histology confirmation of adenoma, including advanced adenoma, or sessile serrated adenoma/polyp and use of propofol). As such, we believe that the test characteristics reported are most appropriate.

Quan et al. 2004. Validity of Procedure Codes in International Classification of Diseases, 9th revision, Clinical Modification Administrative Data. Medical Care; 42, 801-809

Results
8. As per the suggestions of methodology, a more descriptive overview of the cases, and agreement between clinical and administrative data, and algorithms would be useful.

	Response: A description of the cases is provided in Table 2. The algorithms are described in Table 1 with further detail in terms of the precise codes used in Figures 2/Table 3 as well as in the supplemental tables. The issues with providing the descriptive data on the agreement between clinical and administrative data are described above in item #4. 9. I would suggest that cases that could not be linked to administrative data are not described as cases abstracted, but that the number of linked cases be a subset of all cases that were abstracted from the charts found and upon which data extraction was conducted. **Response: Only cases that were abstracted and linked to the health administrative data are described as 'abstracted' (n=1845). Subsets of these cases were used for some analyses**