

<b>Article details: 2017-0054</b>	
Title	Evaluation of health links on health services utilization in the Central Ontario health region: a propensity-matched difference-in-difference
Authors	Luke Mondor MSc, Kevin Walker MSc, Yu Qing Bai MSc, Walter P. Wodchis PhD
Reviewer 1	Dr. Patricia Li
Institution	Montreal Children's Hospital, McGill University Health Centre, Pediatric Consultation Centre, Montréal, Que.
General comments (author response in bold)	<p>The authors present interesting results from a preliminary but important evaluation of a system-wide model of care for high-cost/multimorbidity in Ontario. They provide a good rationale and context for the study, provide a clear explanation of the methods and analyses (which appear to be well-suited to the objectives), and display the results clearly. The interpretation is well-written, explains the possible mechanisms for the observed findings, and identifies important limitations.</p> <p>I have no major comments; just a few minor notes:</p> <p><b>We would like to thank Dr. Li for her positive review and commentary.</b></p> <p>(1) Line 90-91: suggest providing a reference for primary care model affiliation for readers who are not familiar with Ontario primary care enrolment. <b>We agree with Dr. Li's suggestion and have included two references for readers who are not familiar with primary care enrolment in Ontario (references 15 and 16).</b></p> <p>(2) Line 94: oncology and dialysis visits – are these also billed by the specialists (and therefore double-counted/potentially collinear)? Line 95: same as above – are mental health inpatient episodes accounted for in the billings by specialists? (or were the specialist billings for these visits?) <b>Although some double counting is possible, it should in no way affect our propensity matching procedure: Pearson correlation coefficients for specialist billings with oncology visits, dialysis visits and mental health episodes were 0.225, 0.214 and 0.134, respectively, and collinearity is not evident between specialist billings and these variables and should not impact our findings.</b></p> <p>(3) Line 102-103: I defer to the opinion/preference of the journal editor, but I am not sure that the c-statistic is important to demonstrate model performance (see 2009; 28(25): 3083–3107). <b>Thank you for this suggestion (including reference); we agree and have removed the sentence from our revised manuscript.</b></p> <p>(4) Line 170: should it be Central LHINs' (add apostrophe) <b>Thank you for this edit; we have revised the document accordingly.</b></p> <p>(5) There are many acronyms throughout, which makes it a bit difficult for the reader not versed in this healthcare setting but I imagine they could with the word count restrictions <b>In response to the Editor's and Dr. Li's recommendation, we have removed many of the acronyms used in this manuscript. The remaining abbreviations remain (specific to statistical analyses or to Ontario's healthcare system), and are included in a 'List of Abbreviations' at the end of the paper. All prior acronyms for ICES administrative datasets have been placed in a new Appendix (Appendix 1) along with the dataset. In accordance with the RECORD guidelines, this includes the sources of data for all variables defined in the revised manuscript.</b></p>
Reviewer 2	Dr. Bruno D. Rivierin
Institution	McGill University, Department of Epidemiology, Biostatistics and Occupational Health, Montréal, Que.
General comments (author response in bold)	<p>Main comment: The study by Mondor et al. addresses an important question – whether the initial implementation of Health Links (HL), providing care coordination for high users in Ontario, helped in improving health service utilization (e.g. reducing yearly rates of days in hospital, readmissions, and improving 7-day primary care follow-up). The study follows state-of-the-art methods for the evaluation of health systems and includes various checks for the validity of results. Overall, results presented show a worse performance HL enrollees relative to non-enrollees on many metrics (e.g. hospitalizations, ED visits and number of days in acute care (my interpretation from DID estimates in Table 2). The Interpretation section is well-written and provides a clear understanding of potential biases and/or processes at play that could explain the relative performance of HL so that future research can use this information.</p> <p>In my comments below I raise a few concerns to help authors further improve the clarity and depth of this manuscript: 1) The intervention should be described in more details; 2) The different objectives should be specified in a more transparent fashion (or modified accordingly), and the reader should be able to see the results directly to each objective; 3) I believe the DID results should be emphasized over the trends within each group; 4) Accordingly, I believe the Interpretation section should reflect a deeper understanding of potential biases and/or processes at play that could explain the relative performance of HL so that future research can use this information.</p> <p><b>We thank Dr. Rivierin for his thorough review. In our revised paper, we have rephrased our stated objectives to be clearer and more specific in context regarding the Health Links intervention. We feel that these were the source of many of the critiques made.</b></p> <p>(1) INTRODUCTION: L29-36: I am not very familiar with the Health Links in Ontario; based on the manuscript, it seems to consist only of a Care Coordination Program (CCP), but is there anything else? Are there any other elements of this intervention that may contribute to improving health service utilization? Please provide more information on what the CCP consists of. <b>We acknowledge that not all readers of the journal will be familiar with Health Links and have expanded the Introduction to include a detailed description of the program – including referrals into the program, care coordination processes (including additional services provided), and a coordinated care plan itself) as well as overall objectives of the program. Please see Introduction (lines 29-44) and Response to Reviewer Comments. With a "low-rules" implementation, the exact specification of the program delivered across Ontario is relatively uncertain.</b></p> <p>(2) L42-44: It's not clear what the difference between the two objectives are: 1) the effect of HL on health service utilization; and 2) how health service utilization patterns compare...Perhaps be more specific here so that the reader understands the distinction between the two. Do you mean to compare DID estimates and in 2) highlight the trends within each group? If so, make this clear and present/discuss results in the same fashion in the Results section. Also, the Result section does not seem to be linked directly to the objectives; this should be addressed to improve clarity. <b>In response to Dr. Rivierin's concerns, we have rephrased the stated objectives to read: "...to determine 1) whether enrolment in Health Links is associated with differences in healthcare utilization among Health Link enrollees after (vs. prior to) program enrolment, and 2) difference in utilization patterns among enrolled patients compares to trends among similarly complex patients who are not enrolled (lines 45-49). Therefore, the first objective highlights the difference (pre-post) within the enrollee group, the second objective highlights the difference-in-difference estimates (pre-post difference among enrollees vs. pre-post difference among matched comparators). We have revised the Results and Interpretation sections accordingly to follow these two objectives (i.e., first presenting the results for the measured outcomes post vs pre-index among enrollees, then presenting difference-in-difference findings).</b></p> <p>(3) METHODS: The methods appear to be appropriate to ensure valid estimates; this section sufficiently covers the important elements of the study design and approach. A few minor points: I found Appendices 1 and 2 useful; exclusions for certain outcomes are not presented in the main body, only in the Appendices (e.g., death, hospitalization for mental health, etc.) – I think this is important information that should be somewhere in the main text. *Note, in the revised manuscript, Appendices 1 and 2 are now 2 and 3, respectively.</p>

Thank you for this suggestion. However, we are unable to expand on each individual outcome in the body of the manuscript detail while staying under the journal's word count restrictions. As such, we refer readers to this Appendix for complete details used in the study.

(4) Is there a reason why you didn't look at the number of deaths? Could this be a potential outcome to look at? Or perhaps to just give people died during follow-up across both groups.

**1-year post index mortality is currently presented in the Results section (lines 143-144): 26.5% for enrollees and 24.9% for standardized difference = 0.037. We used these data to provide additional support that the 2 groups (cases and matched comparators) are similar beyond baseline covariates.**

**We did not include mortality as a measurable outcome for DID analysis. As stated in our manuscript, we sought to measure utilization measures "related to hospital care (acute and emergency services) and physician follow-up because these indicators identified as key performance markers for Health Links and are measurable with available administrative data (8,18)" (lines 143-144). We added a MOHLTC Health Links reference in the revised manuscript to support the appropriateness of the selected outcome. Although not an exhaustive list, others measures are possible and deemed important (for example, referral time for home care); however, for brevity and interpretation, we measured and interpreted only those related to hospital care.**

(5) Why only look at 7-day follow-up? Is this a requirement of the CCP? Please clarify if so.

**Ontario has largely adopted a 7-day definition for physician follow-up post-discharge as a standard of care and topic for our primary care Quality Improvement Plan indicator specifications (see link 1 below, although we did not restrict our sample to select conditions). Other institutions (CIHI, for example, see link 2): and researchers (for example, references 3 and 4 to name a few) have used 1-week to define early physician follow-up post-discharge. Therefore, a 7-day measure is consistent with existing literature.**

**1. [http://www.health.gov.on.ca/en/pro/programs/ris/docs/patients\\_with\\_primary\\_care\\_visit\\_within\\_7\\_days\\_of\\_discharge\\_qips](http://www.health.gov.on.ca/en/pro/programs/ris/docs/patients_with_primary_care_visit_within_7_days_of_discharge_qips)**

**2. [https://secure.cihi.ca/free\\_products/Physician-Follow-Up-Study-mar2015\\_EN.pdf](https://secure.cihi.ca/free_products/Physician-Follow-Up-Study-mar2015_EN.pdf)**

**3. Hernandez AF, et al. Relationship between early physician follow-up and 30-day readmission among Medicare beneficiaries with heart failure. JAMA. 2010 May 5;303(17):1716-22**

**4. Hess CN, et al. Association of early physician follow-up and 30-day readmission after non-ST-segment-elevation myocardial infarction in older patients. Circulation. 2013 Sep 10;128(11):1206-13**

(6) RESULTS: The result section is relatively well written. It should be made clear in this section that results are linked to the stated objectives above. Also, I find DID estimates are underemphasized – in my opinion, these are the most valid estimates to consider and should be presented. Post- are also informative, but I do not believe that a statistical test should be made on them; instead, I would show the figures presented to the reader gets a better sense of how the trends evolved over time across groups; and present the DID estimates as a table on its own. E

**Upon review, we agree with Dr. Riverin's assessment and have restructured the Results section as follows:**

**Paragraph 1. Baseline characteristics of the sample,**

**Paragraph 2. Robustness checks for matching and comparisons of matched vs. unmatched enrollees,**

**Paragraph 3. Pre-post differences among enrollees (i.e., results for Objective 1)**

**Paragraph 4. Pre-post differences among enrollees vs. pre-post differences among comparators – the difference-in-differences approach as well as parallel trends assessment to validate the DID approach (i.e., results for Objective 2).**

**Equal weight is given to each objective.**

**Moreover, we agree that the trend plots are informative to the reader to understand how the rates evolved over the follow-up period (generally increasing trajectories leading up to index); however, due to the overall size of the plots and because these do not show findings for objectives 1 or 2, we have kept the figures in the Appendices.**

(7) INTERPRETATION: L170-174: I do not find the first paragraph of the Interpretation reflects the main results; consider revising to highlight the first.

**We have revised the first paragraph of the Interpretation (summary of main finding) to better reflect the revised study objectives and the revised Results section:**

**"We found that hospital-related utilization patterns were comparable after (vs. before) enrolment for the initial patients in the Central LHIN's 3 'early adopter' Health Links, except for average days in acute care, which increased. In contrast, trends among comparators from the same jurisdiction (selected from health administrative data and matched on socio-demographics, comorbidities, and health system utilization) decreased for inpatient stays, ED visits and acute care days, relative to enrollees" (lines 159-164)**

(8) L175-179: Can poor definition of this program explain a poorer performance? I think this can explain no difference, not worse results. I think of other processes by which non-HL patients may get better care coordination (other than potential biases?) I suggest separating utilization from care coordination could explain results and potential biases in the evaluation of... [?]

**In our revised paper, we have rephrased this sentence to refer to these findings as "non-significant pre-post differences among enrollees" (lines 165-166) rather than "negative findings". As such, our interpretation aligns with Dr. Riverin's assessment, that poor care coordination may explain no difference in utilization patterns for the enrollee group.**

**In terms of other processes, we have added interpretation to this finding (pre-post differences among enrollees) in that hospital use and days in acute care may be driven in part by mortality (26.5% among enrollees in post-index period), as previous research has shown. Hospital use (and days in acute care, in particular) is amplified at the end of life (lines 174-175).**

(9) L179-182: I think in this sentence lies the main explanation as to why the results are negative – i.e. selection bias remains despite several adjustments. This should be highlighted to bring awareness to the research and policy community. I suggest re-organizing this section to bring forth this evidence. **This in itself is not evidence of selection bias, but rather highlights that the Central LHIN's 3 early adopter Health Links did not select the most complex cases to the program. These patients would meet the Health Links 'target population' criterion, although would not be representative of the distribution of patient complexity.**

**However, the fact that we could not find an appropriate match from the general 'target' population selected from the sample of enrollees is evidence of selection bias, which we have acknowledged in our manuscript, and have added context to in the manuscript. The (9%) sample that went unmatched were medically more complex and also greater users of healthcare across sectors (see Appendix 1). Therefore, we expect that we have underestimated the true pre-index mean for each outcome, and also may not have detected reductions among enrollees (lines 203-206) because the majority (2/3) of high-users are no longer high-users in subsequent follow-up (regression towards the mean, see reference 19). However, we would not anticipate any implications for our difference-in-differences (objective 2) as we would expect that on the whole the additional matched comparators would exhibit comparable patterns as their matched enrollee counterparts.**

(10) L186-187: This sentence is confusing when you consider that the 'improvement' in follow-up pre-/post- is the same in both group; so is there an improvement? I would revise that sentence taking into consideration

	<p><b>Here, we refer to the finding that &lt;40% of the discharges among enrollees did not receive timely physician follow-up which care coordination, acknowledging that this proportion did not significantly change pre- vs.-post index. We therefore are so because Health Links provide all the tools necessary for improved care coordination, this is an indicator subject to benchmark a target for improvements by Health Links staff and policymakers. We have revised the sentence accordingly (lines 175-177) the comparator group are inferred or should be made.</b></p> <p>(11) L188-196: This passage seems to be associated again to the selection of into HL based on unmeasured factors? I suggest regrouping section that may explain why results observed show a poorer performance.  <b>As recommended by the Reviewer, we have re-grouped the Interpretation (specifically, Explanation of main findings) to better stated objectives (i.e., first interpreting the findings of measured outcomes post vs pre-index among enrollees, then interpreting difference findings). These changes should address the Reviewer's comment.</b></p> <p>(12) Links to other published literature lacks depth and is not very convincing given that HL is not fully described in the introduction.  <b>We acknowledge that not all readers will be familiar with Health Links and have expanded the Introduction section to include description of the program – including referrals into the program, care coordination processes (including additional detail on care plan itself) as well as overall objectives of the program. Please see Introduction (lines 29-44) and Response to Editors, which links the reviewer is specifically referring to otherwise.</b></p> <p>(13) L209-212: Do patients in other HL differ from the ones studied that limit generalizability? How? And comment on how they may be reader a better idea of how much of limitation that is.  <b>As stated in our Interpretation, evaluation of the preliminary stages of the Health Links initiative within other jurisdictions are forthcoming (lines 193-195). We do not have the data to compare Health Link populations, nor are they publicly available. The flexible nature of Health Links in referring in patients from a wide target population definition and given the flexibility to provide coordinated care specific to regional contexts, we would expect some variability in the enrollee populations and Health Links in improving outcomes. As such, we acknowledge the generalizability of our findings to other health regions and anticipate evaluations from other jurisdictions as an important direction for ongoing and future work.</b></p> <p>(14) L212-213: Isn't this evidence of selection bias? Consider re-organizing to support your explanation of negative findings (mentioned in 13).  <b>Correct. We have appropriately acknowledged this as a limitation of our study, which is evidence of selection bias that limits generalizability of our findings. In the revision, we have added context around how this should affect our effect estimates.</b></p> <p>(15) L214: I am not sure that this qualifies as a population-based sample given that you do not capture all enrollees into the HL program, statement.  <b>This sentence has been removed from the revised manuscript.</b></p>
<b>Reviewer 3</b>	Dr. Daniel M Kobewka
<b>Institution</b>	Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Ont.
General comments (author response in bold)	<p>Interesting study. Well done with appropriate methods.  <b>We would like to thank Dr. Kobewka for his positive review and commentary.</b></p> <p>(1) It is interesting that health care use dropped in the control group but not in the enrollee group. You suggest that this is regression to the mean. The enrollee group did not have decreased usage after Health links enrollment suggests that the enrollees are systematically different from the control group. Propensity score matching using all available data. You discuss the possible unmeasured confounders but don't discuss how this is a weakness. There may indeed be a difference in the enrollee group that you could not detect because of an imperfect control. Please discuss this or explain why it is not important.  <b>In our revised manuscript we have discussed additional unmeasured confounders between the enrollee and comparator group that may have contributed to our observed findings. This now includes (but is not limited to): "differential patterns observed amongst control and enrollee due to other unmeasured factors such as availability of home support networks, social determinants of health beyond income, and unmet healthcare needs. As such, residual confounding is probable, contributing to the significant difference-in-differences estimate (lines 181-182)."</b>  <b>At the reviewer's request, we have also added context regarding how matching bias may impact our findings and interpretation (lines 206-207). We note that by excluding the most complex enrollees (9% lost due to inability to find a match) we therefore underestimate the utilization of enrollees in the pre-index period and potentially also under-estimate the full effect of Health Links on the health of the most complex patients.</b></p> <p>(2) Also P 10 Line 29: You state that "The quasi-experimental propensity-matched design was robust" I agree that your methods are as robust as the absence of an RCT but not perfect as the differences between control and enrollee suggest.  <b>This sentence has been removed from the revised manuscript.</b></p> <p>(3) Minor comments: P3 Line 21: Multiple uses of the word "better" are redundant. P 3 Line 23: typo "complex chronic and social needs"  <b>We have edited this sentence to now read: "In response, Ontario's Ministry of Health and Long-Term Care (MOHLTC) launched an ambitious strategy aimed to better provide coordinated, community-based healthcare for patients with complex health and social needs (lines 29-31)."</b></p>