Author’s response to reviews

Title: Chart Validation of an Algorithm for Identifying Hereditary Progressive Muscular Dystrophy in Healthcare Claims

Authors:

Xiaoxue Chen (xchen@healthcore.com)
Abiy Agiro (aagiro@healthcore.com)
Ann Martin (ann@parentprojectmd.org)
Ann Lucas (annmlucas@yahoo.com)
Kevin Haynes (khaynes@healthcore.com)

Version: 1 Date: 26 Apr 2019

Author’s response to reviews:

Dear Editor,

Thank you for the thoughtful review of our manuscript, Chart Validation of an Algorithm for Identifying Hereditary Progressive Muscular Dystrophy in Healthcare Claims. We have carefully considered the reviewers’ comments and have revised the manuscript accordingly. Our responses to each of the reviewers’ comments are included in the text as tracked changes as well as in the letter below.

General Comments

1. Please cite Table 3 in the main text of the manuscript.
Response: We cited Table 3 on page 9, line 20 as follows: “The PPV of the case-identifying algorithm for MD was 95% (95% CI 88% – 98%, Table 3).”

2. In the “Funding” statement, please declare the role of the funding body in the design of the study, and collection, analysis, and interpretation of data and in writing the manuscript.
Response: We declared the role of the funding body in the following sentence: “Patient-Centered Outcomes Research Institute (PCORI) did not play a role in the design of the study; in the collection, analysis, or interpretation of the data; or in the writing of the manuscript.” (Please see pg 15)

3. Please add the “Additional files” section after the Tables, describing each supplementary file included in the submission.
Response: We listed the supplementary files after the tables: Appendix Table 1. Code Lists Used to Define Baseline Comorbidity Burden; Appendix Table 2. Characteristics of Patients Whose Medical Charts Were Obtained Versus Those Whose Medical Charts Were Unobtainable (please see pg 23).

4. Please upload Appendix Tables as Supplementary files.
Response: We uploaded the appendix tables as supplementary files.

Reviewer 1
1. I would however be cautious with how the authors define disease prevalence in the different areas and in particular cardiac disease prevalence. I do not disagree how the author words it but would recommend that there should be some comments that depends on how one defines the specific disease (ie cardiac, respiratory, bone-health and endocrine-related conditions). For example for cardiac disease prevalence if you use ECG you would have a high rate of disease prevalence based on relative tachycardia and other non-specific ECG findings or if you use systolic function by shortening fraction you would not demonstrate cardiac disease until later. The first shows a high prevalence of disease early using a very sensitive test such ECG but loses specificity. On the other hand using ECHO shortening fraction which is first a blunt tool and second global function decline by shortening fraction occurs later and you underestimate prevalence. I would ask that the authors in the limitations at least address the tools to define disease may alter prevalence and is important in understanding the tool one uses clinically as well as in trials.
Response: Thank you for the thoughtful comment. This study primarily focused on the validation of an algorithm to identify patients with muscular dystrophy (MD) used by Duchenne Patient Powered Research Network to identify and recruit members. Assessing the variation of disease prevalence estimates by diagnostic instruments such as ECG is beyond the scope of the study, but we agree with the reviewer that characterization of the clinical conditions among the population with MD will vary by the tools used to define the conditions and over the life course of MD. In our study, we used claim-based definition of disease based on Pediatric Complex Chronic Conditions Classification System version 2 (Feudtner et al., BMC Pediatr. 2014;14:199) as in the Soslow et al study, which potentially helps improve the consistency of complex condition assessment across studies.
To reviewer’s question, a claims-based definition generally does not include information on the diagnostic instruments used to define diseases; instead, it reflects the diagnosis providers made in clinical practice. We added the following sentences in the limitations section to explain this limitation “…our study characterized the comorbidity burdens of patients with MD using claims data. Prevalence may vary by the type of diagnostic instruments used to define disease; however, a claims-based definition generally does not include such information…” (Please see pg 13, lines 5-9)

2. Similar to (1) above, can the authors further detail how a patient is given a disease (ie cardiac, respiratory, bone-health or endocrine issues)? Is it based on ICD coding? It would again be important to address some limitations for how each patient is given a disease. If the authors used specific disease code, medication, specific health care physician who saw the patient.
Response: We appreciate the chance to clarify this point. The diseases were defined based on ICD diagnosis coding. We used CPT/HCPCS/GPI codes when appropriate. Cardiac and respiratory diseases were defined based on Pediatric Complex Chronic Conditions Classification System version 2 (Feudtner et al, BMC Pediatr. 2014;14:199). The definitions of bone health and endocrine issues were developed after consultation with internal clinical experts using the input of Duchenne Patient Powered Research Network. Details of bone health and endocrine issue definition were attached as Appendix
Table 1.
There are few limitations associated with claim-based definitions. As we noted above, a claims-based definition generally does not include information on the diagnostic instruments used to define diseases; instead, it reflects the diagnosis providers made in clinical practice. Secondly, a claim-based diagnosis may be subject to claims coding omissions or errors, which we have acknowledged in our limitations section. (Please see pg 13, lines 6-9)

Reviewer 2.
1. The methods - examining only records of people identified as positive by the algorithm - means the study cannot examine sensitivity or NPV, therefore it is not known whether the algorithm is missing a large percentage of true cases. This is very important for studies using claims data to recruit patients, or using such data for observational research. I appreciate this would be difficult to design but I wonder if the authors could also have assessed those with only one visit coded with MD, or a relevant "control" population, to get an indication whether only absolutely certain cases were identified and others may be missed?
Response: There are a few challenges around validation of claim algorithms to identify rare diseases. For rare diseases, NPV and specificity would virtually always be 99+% simply due to the rarity of the disease. Sensitivity is difficult to achieve for claim algorithms and is a limitation of our study. The rationale of our algorithm requiring two office visits coded with MD was to avoid false positives without losing too many cases. We agree that it is important to consider both PPV and sensitivity. Therefore, we estimated sensitivity among those with at least one office visit coded with an MD code. We found that two office visits with MD code gave 64% sensitivity, and it gave 48% sensitivity with the additional age requirement. Therefore, this case-finding algorithm will retain a fairly large sample with good PPV.
We added the following sentence in the limitations sentence section to clarify: “….Despite the challenges, we obtained a rough estimate of sensitivity for our case-finding algorithm among those with at least one office visits with a MD code (assuming it has 100% sensitivity). The case-finding algorithm achieved had 48% sensitivity, which helped to retain a fairly large number of patients…..” (Please see pg 12, line 23 and page 13, lines 1-3)

2. The problem above is compounded in that records were only examined for less than 25% of patients identified by the algorithm. How do we now interpret the quality of the algorithm? Does this mean that we can only have confidence in being true cases of MD for less than a quarter of patients identified by the algorithm or do the authors feel the algorithm is accurate for all patients it identifies? How do the remaining patients (~400) differ from those examined? If patients do not have adequate information in their records to be assessed, what does this indicate about the quality of their records and likelihood of having MD, or the quality of their data for research?
Response: We thank the reviewers for this question. We believe the quality of our algorithm was mostly unaffected by the fact that less than 25% of the patients’ records were examined. Our algorithm identified 580 patients in our claims environment. We focused on 510 patients (88%) who had PCP or neurologist visits to allow access to more complete patient histories to ascertain the cases. Of these 510 patients, 204 patients (40%) were available for chart validation due to health plan operational requirements (eg, current active health plan members who are not enrolled in Administrative Service Only plans). Such operational requirements would not compromise the quality of the algorithm. In the body of the manuscript, we corrected the explanation of why the patient count dropped from 510 to 204. The initial description was inaccurate, and we apologize for causing any confusions earlier. (Please see pg 9, lines 5-8)
Among the 204 patients for whom we were able to request medical charts, charts were obtained for 109 patients. We assessed the differences in the characteristics of patients whose medical charts were obtained versus those whose medical charts were unobtainable (data in Appendix Table 2). There were no significant differences except in the rates of metabolic disease and impaired growth issues, which led us to believe the unobtainable charts were mostly likely random.

3. The design as it stands also removes any element of blinding by the reviewers. They know the patients were all identified by the algorithm and therefore likely had MD.
Response: We agree this would be a limitation of the study and cannot be addressed easily. We added the following sentence in the limitations section to acknowledge this limitation: “The chart validation study sampled cases that were likely MD according to the algorithm. As the chart reviewers were not blinded to the study design, bias may have been introduced to move the PPV upward.” (Please see pg 13, lines 3-5)

4. There is no detail on how the algorithm was developed, and whether others were attempted and fared less well. The algorithm itself is a little unclear (p.7). Is it simply relevant ICD code, male, and aged less than 18 at diagnosis or does it also include specialist?
Response: The algorithm was developed by our collaborating partner – clinical experts from Duchenne registry – under the PCORI award to facilitate patient recruitment to patient-centered research networks (PPRN). Our study aim and award focus is to validate the algorithm used by the Duchenne PPRN to query administrative claims data.
The algorithm included two office visits coded with MD, male gender, and age younger than 18 years at first diagnosis. Specialist visit was not part of the algorithm and was a necessary requirement to identify which clinician was likely to have the most complete set of medical records of interest (given the fragmented nature of the US healthcare system, it is important to target the specialty office most likely to have the needed patient records). We made some edits in the paper to clarify.

5. It would be interesting to know the level of agreement by reviewers on the first 10 cases which were jointly reviewed.
Response: The reviewers agreed on 9 of the first 10 cases. For the one inconclusive case, further clinical review was carried out by the clinical experts and consensus was reached after discussion. Additional educational materials were provided to reviewer nurses and the full chart review proceeded as planned.