Author's response to reviews

Title: A Cross-sectional Study of the Identification of Prevalent Asthma and Chronic Obstructive Pulmonary Disease among Initiators of Long-Acting beta-Agonists in Health Insurance Claims Data

Authors:

David D Dore (david_dore@brown.edu)
Najat Ziyadeh (najat.ziyadeh@optum.com)
Bin Cai (bin.cai@novartis.com)
Robin Clifford (robin.clifford@optum.com)
Heather Norman (heather.norman@optum.com)
John D Seeger (jdseeger@partners.org)

Version: 3
Date: 2 February 2014

Author's response to reviews: see over

Response to Reviewer Comments
January 2014

Reviewer: Elizabeth Wasilevich

Major Compulsory Revisions

1. The authors describe that multiple diagnoses of asthma/COPD were rare, thereby limiting their ability to look at alternate definitions based on more than 1 diagnosis. Since many case definitions do require more than one office visit claim to qualify as asthma/COPD (in the absence of a hospitalization or ED visit), the rarity of this circumstance should be explained a bit more. Was it expected?

We now note that this phenomenon has been observed in pervious work (and is therefore was expected). This topic is addressed in the second paragraph of the Discussion, as below.

“Similarly, multiple diagnoses of asthma or COPD were rare, limiting our ability to refine the algorithms by requiring more than one diagnosis for the condition of interest. Both of these findings are consistent with previous experience with health plan data. Loughlin and colleagues found with tegaserod, a drug with a clear indication (irritable bowel syndrome), only 32% of tegaserod initiators had a claim for irritable bowel syndrome in the 6-month period preceding tegaserod initiation [10]. To address this expected feature of the data, we abstracted medical records on a subset of LABA initiators with no claims for asthma or COPD, and confirmed their actual case status. The resulting data provide an estimate of the distribution of actual asthma and COPD among those without a claims-diagnosis.”

2. The authors have not specified whether they looked at only primary diagnoses for inpatient or outpatient claims...or whether they considered all possible diagnoses.

We considered all possible diagnoses, with the aim of improving sensitivity. (Specificity was addressed through the cross-classification with covariates.) The first sentence under Identification of Asthma and COPD now reads:

“Potential cases of asthma or COPD had ICD-9 diagnosis codes for asthma or COPD in any diagnosis field on claims occurring in the 6 months prior to the index date on an inpatient or outpatient claim (a similar approach requiring 12 months of continuous enrollment yielded similar results).”

3. Since there are common claims based case definitions for asthma and COPD, the authors should describe why they chose not to use them in this study.

We added an explanation of our rationale for this choice in the first paragraph under Identification of Asthma and COPD.

“We chose to develop our own definition of asthma and COPD despite the existence of algorithms in the literature (e.g., Mapel et al [8] and Dombkowski et al [9]), because in our work we aimed to classify baseline diagnoses (indications), whereas most definitions in the literature
apply specifically to outcomes or cohort definitions. Definitions of outcomes and covariates are more robust to insensitive measures of disease than case criteria for identification of off-label use.”

4. The authors state that only those LABA prescriptions filled on the same day as an ICS were considered concomitant therapy. They justify this by stating that this strict approach increases the likelihood of concomitant therapy. If there are 2 active overlapping prescriptions, it is likely that they are taking the therapy concomitantly, even if not filled on the same day. How many persons had overlapping days supply on the filled prescription of the LABA and ICS? In other words, how many may be potentially misclassified as non-concomitant?

We now address this issue more fully in the last paragraph under Cohort Formation.

“Patients with a recent previous dispensing of an ICS were in the monotherapy group. While requiring that patients received the combination formulation of LABA/ICS or concomitant ICS on the same day as the initial LABA dispensing may result in misclassification of individuals who received ICS and LABA on different days, this more stringent definition increases the likelihood that we include only users concomitant therapy in the combination LABA/ICS group. This preference is important because combination LABA/ICS for asthma appears safer than LABA monotherapy and should not be discouraged. (Empirically, this decision was unimportant. Ninety-six percent of patients initiated on a combination product. Of the 9,965 patients classified as LABA-only initiators, 1,931 [19%] used ICS in the baseline period and were potentially misclassified. This number represents only 0.9% of the full study population, and this misclassification is of little consequence.)”

Reviewer: Lucie Blais

Major Compulsory Revisions

Most of the comments were addressed correctly; however the following 3 comments need to be addressed again:

1. The authors should discuss the impact of such low NPV (61.5%) on future studies that will be based on asthma and COPD diagnoses recorded in the database under study.

We added a discussion of this topic to the discussion on the low sensitivity in Paragraphs one and two of the discussion.

“Subsets of LABA initiators with asthma, COPD, and both conditions can be identified and differentiated using claims data, although categorization of the remaining patients is largely infeasible. Within strata of selected covariates, the claims showed better predictive ability to identify asthma only among patients with claims for asthma or claims for both asthma and COPD, and to identify COPD among patients with claims for COPD. Among patients without claims for asthma or COPD, there was no subset within which the PPV exceeded 50%, meaning the classification of patients who do not have a claim indicating asthma or COPD remains uncertain. Additionally, requiring the presence of claims for asthma or COPD resulted in a population in which nearly 25% of persons appear to not have the condition of interest.

That a substantial fraction of the LABA users did not have a claim associated with asthma or COPD is problematic in that it leaves a large fraction of patients unclassified in the data.
Similarly, multiple diagnoses of asthma or COPD were rare, limiting our ability to refine the algorithms by requiring more than one diagnosis for the condition of interest. Both of these findings are consistent with previous experience with health plan data. Loughlin and colleagues found with tegaserod, a drug with a clear indication (irritable bowel syndrome), only 32% of tegaserod initiators had a claim for irritable bowel syndrome in the 6-month period preceding tegaserod initiation [10]. To address this expected feature of the data, we abstracted medical records on a subset of LABA initiators with no claims for asthma or COPD, and confirmed their actual case status. The resulting data provide an estimate of the distribution of actual asthma and COPD among those without a claims-diagnosis. Nevertheless, a large fraction of patients remain unclassified. In studies of off-label use, this limitation means that there will be overestimation of off-label use (under the assumption that some of the unclassified patients have the “on-label” indication). In studies where there is a desire to exclude patients with asthma or COPD, it may be infeasible to completely exclude these individuals.

2. The argument made by the authors for not using regression models to identify the 3 covariates with the largest absolute difference in prevalence is not scientifically valid. Regression models are routinely used in medical research. So, either provide a valid argument or present the results of a regression analysis.

In our previous response about using regression models, we wrote, “We chose the straightforward method of analysis of looking at the differences in prevalence of covariates between confirmed and non-confirmed cases to avoid the complexity of regression modeling—in particular the additional assumptions that come with regression modeling. It is possible that regression models would have provide the same answers with additional precision, but verification of assumptions is challenging.” There is nothing invalid about these comments and the reviewer’s assertion of invalidity is not explained other than to say that, “I think that regression models would have provided a more valid answer.” It is true that regression models are used routinely in medical research—too frequently in the opinion of some. In this research, we were particularly interested interactions between claims definitions and other covariates, which we empirically tested using non-parametric methods on a small sample. The approach we used is more sensitive in identification of covariate interactions than regression models. Moreover, fitting a regression model to sparse data is treacherous as small shifts in values and a small number of outliers can greatly influence the response surface.

However, because we recognize that other readers may also wish to see results from a regression model, we conducted a regression analysis and now include a brief presentation of the results. We are hesitant to emphasize the results of the regression analysis further, because it is post-hoc, whereas the analysis included in the manuscript were described in a protocol that was developed before the study began.

We added a paragraph to the methods and results:

METHODS

“We also conducted a post-hoc regression analysis as an additional means to identify covariates whose presence may modify the accuracy of the identification of asthma or COPD. Using multinomial logistic regression models, we regressed confirmed case status (asthma only, COPD only, both, or neither, with neither as the referent) on the claims definitions for these conditions, plus the covariates with product terms for each covariate and the claims definitions. We retained an empirical approach to covariate selection by using a stepwise algorithm,
entering and retaining variables with p-values < 0.05.”

RESULTS

“The multinomial logistic regression modeling identified several variables that improved prediction of true case status through interaction with the claims definitions. The variables “symptoms involving respiratory system and other chest symptoms” (ICD-9 786.xx), the number of baseline drug dispensings, and age > 40 years improved prediction of COPD only. The number of drug dispensings also improved prediction of asthma only and both asthma and COPD relative to neither. The presence of a baseline dispensing of a beta-adrenergic drug and the diagnosis “other forms of chronic ischemic heart disease” (ICD-9 414.xx) improved prediction of true asthma only. No other variables arose as significant predictors.”

3. The statement on the sensitivity of the claims data is too speculative without any data and should be removed.

We deleted the third paragraph of the Discussion section.

“A possible contributor to the lower sensitivity of the claims measure of COPD was our choice to exclude from consideration certain diagnosis codes that might include COPD, but are mixed with other disorders. Although other studies have chosen to include ICD-9 490 (bronchitis, not specified as acute or chronic), we chose to exclude this code [6].”