ICD-10 Coding Will Challenge Researchers- Caution and Collaboration May Reduce Measurement Error and Improve Comparability Over Time

Alexander J. Mainor, JD, MPH, **Nancy E. Morden, MD, MPH**, **Jeremy Smith, MPH**, **Stephanie Tomlin, MS**, and **Jonathan Skinner, PhD**

The Dartmouth Institute for Health Policy & Clinical Practice, Lebanon, NH 03756

Abstract

Background: The October 1, 2015 U.S. healthcare diagnosis and procedure codes update, from the 9th to 10th version of the International Classification of Diseases (ICD), abruptly changed the structure, number and diversity of codes in healthcare administrative data. Translation from ICD-9 to ICD-10 risks introducing artificial changes in claims-based measures of health and health services.

Objective: Using published ICD-9 and ICD-10 definitions and translation software, we explored discontinuity in common diagnoses to quantify measurement changes introduced by the upgrade.

Design: Using 100% Medicare inpatient data, 2012–2015, we calculated quarterly frequency of condition-specific diagnoses on hospital discharge records. Years 2012–2014 provided baseline frequencies and historic, annual fourth quarter changes. We compared these to quarter four of 2015, the first months after ICD-10 adoption, using Centers for Medicare and Medicaid Chronic Conditions Warehouse (CCW) ICD-9 and ICD-10 definitions and other commonly-used definitions sets.

Results: Discontinuities of recorded CCW-defined conditions in quarter four of 2015 varied widely. For example, compared to diagnosis appearance in 2014 quarter four, in 2015 we saw a sudden 3.2% increase in chronic lung disease and a 1.8% decrease in depression; frequency of acute myocardial infarction was stable. Using published software to translate Charlson-Deyo and Elixhauser conditions yielded discontinuities ranging from −8.9% to +10.9%.

Conclusion: ICD-9 to ICD-10 translations do not always align, producing discontinuity over time. This may compromise ICD-based measurements and risk-adjustment. To address the challenge, we propose a public resource for researchers to share discovered discontinuities introduced by ICD-10 adoption and the solutions they develop.

On October 1, 2015, the U.S. updated its healthcare code catalogue from International Classification of Disease (ICD) Version 9 to Version 10.[1] While the update promises to improve specificity of codes used, it challenges researchers and could bias health measures.

Corresponding Author: Alexander J. Mainor JD MPH, 1 Medical Center Dr. Lebanon NH 03756, Alexander.J.Mainor@dartmouth.edu, Phone (603) 653-0800, Fax (603) 653-0896.

In this brief, we illustrate the risks introduced by ICD-10 adoption and suggest a collaborative approach to addressing the problem.

Prior to October of 2015, the ICD-9 catalogue had been in use since 1979, and over time had fallen short in characterizing patient health and health care. Using ICD-9, for example, a clinician could not specify, and a researcher could not know, if a recorded surgical procedure involved the left or the right side of the body; nor could one readily distinguish care of a new condition (e.g., a new fracture) from ongoing care of a previously treated condition (e.g., follow-up care for a recently repaired fracture). The ICD update addressed this lack of detail. To do this, however, the number of available diagnosis codes increased from 14,025 in ICD-9 to 69,823 in ICD-10. The ICD-10 update also introduced combination codes (e.g., E10.21, diabetes mellitus with nephropathy), codes for the use of new technologies, and options to specify distinct settings of care.[1] So, the switch has value, but for researchers it poses a challenge, because many ICD-9 disease definitions are not easily translated into comparable ICD-10 definitions, and many ICD-10 codes have no specific ICD-9 synonyms.

Researchers routinely use ICD codes to identify and measure the health and health care of populations. The ICD code catalogue switch introduces the risk of artificial changes in these measurements. To help researchers address the challenges, ICD crosswalks (i.e. translation software creating code "equivalents") are available.[2] Unfortunately, ICD-9 and ICD-10 definitions emerging from these resources do not always align. Poor alignment of code diagnosis "equivalents" produces discontinuity in measures of disease prevalence over time; it may also compromise risk-adjustment methods that rely on ICD disease classifications.

Using analysis of Medicare data before and after the switch, we illustrate potential pitfalls of these crosswalks. We test some available translations by measuring weekly frequencies of common conditions during the transition and reveal the discontinuity of measures temporally aligned with the adoption of ICD-10 (October 1, 2015, the first day of the fourth quarter of 2015). We then suggest addressing this problem by creating a public good for all researchers, using a web-based platform, "Dataverse" for sharing ICD-9 and comparable ICD-10 definitions, rate comparisons that quantify the discontinuity in diverse datasets (to allow adjustment for comparisons over time), and the programming code used to make the comparisons. Our exploration of inpatient diagnostic code discontinuity illustrates the problem and serves as a starting point for the envisioned shared resource that would include a broad range of datasets.

Methods:

Our aim in this work was to explore discontinuities in recorded diseases associated with the adoption of ICD-10. Drawing from a full 100% sample of Medicare inpatient administrative data (MedPAR files) 2012 through 2015, we measured the weekly frequency of the appearance of conditions on hospital discharge claims using the following sources for ICD-9 to ICD-10 translations:

1. CMS Chronic Conditions Data Warehouse (CCW) ICD-9 and ICD-10 Definitions are used by CMS to create indicators for 27 common chronic conditions in the CMS research data.[3] The indicators and definitions are often

used by researchers to identify patients with specific conditions for cohort creation or risk adjustment.[4]

2. General Equivalency Mappings Software (GEMS) is an ICD-9 to ICD-10 translation product developed by CMS and the Centers for Disease Control to support ICD-10 adoption.[5] We tested this tool on ICD-9 Charlson-Deyo[6] and Elixhauser[7] disease definitions. These lists of conditions are commonly used by researchers for risk adjustment; understanding direction and magnitude of discontinuities is essential to effective biostatistical modeling.

3. Researcher-published ICD-9 and ICD-10 translations: Canada began adopting ICD-10 in 2001; their researchers have since been tackling the ICD update challenge. We explored one Canadian team's rigorous approach to ICD translation of Charlson-Deyo definitions from ICD-9 to ICD-10.[5] This team created an algorithm comparable to GEMS and used output combined with clinician judgement to inform both ICD-10 definitions and revision of previously-used ICD-9 definitions to optimize alignment over time. Exploration of discontinuities in these definitions reveals the outcome (and residual discontinuity) of extensive effort aimed at harmonizing past and present measurements.

For each week in each year, 2012–2015, we measured the number of discharge events on which each of the ICD-9 and ICD-10 condition-specific definitions listed above appeared. We also measured the weekly number of total hospital discharges. For each calendar quarter we used these weekly values to calculate the average proportion of all admissions with each condition recorded. The frequency of appearance in quarter four of 2015 (the first quarter of ICD-10 use) was compared to that of quarter four in 2014 to create a quarter four change or "delta" temporally associated with ICD-10 adoption.

To estimate how much the 2014 to 2015 fourth quarter change in diagnosis appearance might be attributable to the code catalogue update, we required comparison to baseline or secular trends in change of fourth quarter diagnosis appearance on hospital discharges. For this, we measured weekly fourth quarter diagnosis appearance in 2012 and 2013, one and two years before ICD-10 adoption. We used these weekly counts to calculate mean fourth quarter rates of diagnosis appearance for these years. These were used to calculate year-to-year fourth quarter difference or "deltas" (2012 vs. 2013 and 2013 vs. 2014). These two fourth quarter "deltas," were averaged, as a baseline rate of change. We then compared this baseline fourth quarter change to the 2014 vs. 2015 "delta" or change in appearance of each diagnosed condition, to obtain a measure of excess change beyond that expected from secular trends. All change measures are expressed in percentage terms, and graphs are used for visual portrayal.

Results

In this brief we present select results from our analysis using CCW definitions and from our use of GEMS to translate Charlson-Deyo conditions. The displayed conditions represent the range of patterns we observed in our analysis: an increase, a decrease and a stable trend. We

present full results from all analyses on our website.[9] Figure 1 portrays weekly rates of appearance of three exemplary conditions: Chronic Obstructive Pulmonary Disease (COPD), Depression, and Acute Myocardial Infarction (AMI). In the last quarter of 2015, we find a sudden and discontinuous 3.2% increase in hospital discharges associated with COPD (the largest increase), a 1.8% decrease in hospital discharges associated with depression (the largest decrease) and a relatively stable rate of AMI, compared to the same timeframe in 2014.

Figure 2 portrays the broad range of excess change in frequencies of appearance of the 27 principal CCW codes, the change disease appearance from 2014 to 2015 versus the average fourth quarter change from 2012 to 2014. Figure 3 presents these same discontinuity measures for our GEMS translation of Charlson-Deyo conditions. For this translation, we see the range of excess quarter four change ranging from –7.4% to +8.5%. Larger discontinuities were seen in the GEMS translation of some Elixhauser conditions. For example, the frequency of the appearance of psychoses increased 10.9% more in quarter four from 2014 to 2015 than it had on average in the 2012 to 2014 time frame, while cardiac arrhythmias decreased 8.9% more than average in this same period. Results were comparable (select large positive and negative changes) when we employed ICD-9 and ICD-10 definitions laboriously derived by Canadian researchers tackling this challenge with Canadian healthcare data.[5] In assessing their Charlson-Deyo translation on U.S. Medicare data, we saw similarly large discontinuities related to a distinct set of conditions, for example an 8.5 percentage point increase in the appearance of dementia and 3.3 percentage point decrease in the appearance of chronic lung disease. See our website for full analysis.[9]

Discussion

In our analysis of Medicare inpatient claims, we find substantial discontinuity in some conditions attributable to the transition from ICD-9 to ICD-10. The discontinuities, while relatively small in most cases, are substantial in others. Our findings suggest care is needed in the approach to ICD-10 translations, especially for studies spanning the transition and studies aiming to compare health and health services measures before and after the update.

In general we find no detectable, systematic difference in the clinical conditions recorded with ICD-10 codes compared to ICD-9. On our Dataverse open resource, we provide the distribution of specific disease codes appearing in the claims during the fourth quarter of 2014 (ICD-9) and the fourth quarter of 2015 (ICD-10). Two of our team clinicians examined specific code frequencies in detail and found no clinical explanations for the apparent discontinuities amenable to correction informed by clinician recategorizing some codes based on clinical insight. A rare exception to this general rule is found in asthma definitions. Using August 2017 CCW definitions, we discovered a 20% jump in frequency of asthma on hospital discharges in the fourth quarter of 2015 compared to 2014. Review of the codes by clinicians revealed three diagnoses more consistent with COPD than asthma (e.g., J441 chronic obstructive pulmonary disease with (acute) exacerbation). Removing these codes eliminated the discontinuity (see comparison on our website).[9] These codes were also removed by CCW in the November 2017 definition. This discovery reinforces the

importance of continuously accessing the latest version of definitions in use and illustrates the potential risk of code recycling.

What might explain the discontinuity? The jumps displayed in Figures 1–3 and in the full summaries available on our website,[9] likely result from a combination of (a) assumptions and tradeoffs accepted by the developers of the GEMS software, (b) fundamental structural, level of detail, and organizational differences in ICD-9 and ICD-10 that make precise cross mapping impossible, and (c) possible errors made by those entering codes on claims (clinicians and professional coders), especially during early ICD-10 adoption.

Our study has limitations. First, we use available 2015 data. This is the first year for which ICD-10 codes are in use, but the data are limited to the fourth quarter because ICD-10 was adopted on October 1, 2015. Analyses using broader time frames will be essential to understanding the impact of ICD-10 adoption on measures requiring a longer time horizon, such as annual prevalence of disease. Secondly, our exploration of CCW definition frequency was conducted at the event level not the individual level; as a consequence, we do not aim to and do not achieve replication of CCW measures. Thirdly, CCW variables included in research datasets are based on inpatient and outpatient records, so our discontinuity measures do not necessarily translate directly into changes in CCW conditions prevalence. A full 12 months of data are necessary to replicate their methods of outpatient disease measurement as many definitions require "one inpatient claim or two outpatient claims over 12 months."[3] We intend to repeat and make public similar analyses of outpatient (Part B, Carrier File) Medicare claims when we have sufficient longitudinal data (more than one calendar quarter) to ascertain cases commonly defined by the appearance of two outpatient diagnosis codes in a year.[3] Future studies will similarly examine procedure codes, subject to the same challenges and risks as diagnosis codes. We have no reason to believe discontinuities apparent in these Medicare claims will differ from other claims datasets, but that must be tested. Additionally, we document temporal association but cannot conclude causality and do not control for other possible explanations for the abrupt changes observed.

How will we address inconsistencies? In September of 2017, the CCW published an assessment of the impact of ICD-10 implementation on claims received by Medicare and on observed annual prevalence of CCW coded chronic conditions.[10] In this early exploration of the ICD upgrade, that, like ours, included just 3 months of data following ICD-10 implementation, the authors concluded that annual disease prevalence fluctuated little with the code catalogue update. While important for 2015 overall, in month-to-month analysis of inpatient data, we find appearance of some conditions changed substantially following the update. This suggests fluctuations in annual prevalence will be larger than those measured by CCW using the first three months of available data, because all months will reflect the new code and imperfect translations. Canadian researchers who tackled this problem in a different patient population (adults in hospitals in one region of Alberta, Canada),[8] drew a similar conclusion, as did a team studying the transition in the VA health care system.[11] Work examining trends in opioid-related inpatient stays also found a sharp increase in stays involving an opioid-related diagnosis associated with the transition.[12]

Researchers are likely to worry about the ICD update introducing discontinuities and thus errors in their research. In response, these teams may create their own translation algorithms or adjust available tools so measures align well enough for study-specific definitions and short-term needs. Yet these approaches may compromise our ability to compare data and results across research teams and over time, as well as adding to costs in terms of time and energy. The risk of this duplication and potential confusion motivates our proposed shared resource. To facilitate these transitions, we propose a crowd-sourced ICD-9 to ICD-10 crosswalk test and solution website by using the Dartmouth Dataverse website[9] to post our code; we will provide access for others to post valuable code (and are open to alternative approaches for posting code and analysis).

Conclusion

The U.S. clinical and health policy research community is beginning to use ICD-10 data. This requires updating and changing computer code analyzing ICD codes. While publicly available "off-the-shelf" translation programs, such as GEMS, are invaluable, measurement and management of discontinuities will be essential for reliable, comparable research. We may also be able to improve methods over time with clinically-informed fine-tuning of definitions (both ICD-9 and ICD-10). We have described some first steps towards that goal and illustrate the challenge of the ICD upgrade to research including common conditions. To achieve efficiency and leverage the collective power of diverse research groups across the U.S., we propose the creation of a publicly available website where researchers may download code – and post their own updated code, complete with discontinuity measurements. By working collaboratively, it is possible to address an otherwise overwhelming problem, and perhaps establish a foundation for additional scientific collaboration in the research community.

Acknowledgments

References:

1. CDC National Center for Health Statistics. International Classification of Diseases, (ICD-10-CM/PCS) Transition - Background 2015; https://www.cdc.gov/nchs/icd/icd10cm_pcs_background.htm. Accessed May 23, 2018.

2. Centers for Medicare and Medicaid Services. 2018 ICD-10 CM and GEMs. 2017; https://www.cms.gov/Medicare/Coding/ICD10/2018-ICD-10-CM-and-GEMs.html. Accessed May 23, 2018.

3. Chronic Conditions Data Warehouse. Condition Categories 2018; https://www.ccwdata.org/web/guest/condition-categories. Accessed May 23, 2018.

4. Drye EE, Altaf FK, Lipska KJ, et al. Defining Multiple Chronic Conditions for Quality Measurement. Medical care 2018;56(2):193–201. [PubMed: 29271820]

5. CDC National Center for Health Statistics. International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) 2017; https://www.cdc.gov/nchs/icd/icd10cm.htm. Accessed May 23, 2013.

6. Bannay A, Chaignot C, Blotiere PO, et al. The Best Use of the Charlson Comorbidity Index With Electronic Health Care Database to Predict Mortality. Medical care 2016;54(2):188–194. [PubMed: 26683778]

7. Mehta HB, Dimou F, Adhikari D, et al. Comparison of Comorbidity Scores in Predicting Surgical Outcomes. Medical care 2016;54(2):180–187. [PubMed: 26595225]

8. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Medical care 2005;43(11):1130–1139. [PubMed: 16224307]

9. http://www.dartmouthatlas.org/pages/ICD10.

10. Chronic Conditions Data Warehouse. CCW Condition Categories: Impact of Conversion from ICD-9-CM to ICD-10-CM 2017.

11. Yoon J, Chow A. Comparing chronic condition rates using ICD-9 and ICD-10 in VA patients FY2014–2016. BMC Health Services Research 2017;17:572. [PubMed: 28818082]

12. Heslin KC, Owens PL, Karaca Z, Barrett ML, Moore BJ, Elixhauser A. Trends in Opioid-related Inpatient Stays Shifted After the US Transitioned to ICD-10-CM Diagnosis Coding in 2015. Medical care 2017;55(11):918–923. [PubMed: 28930890]
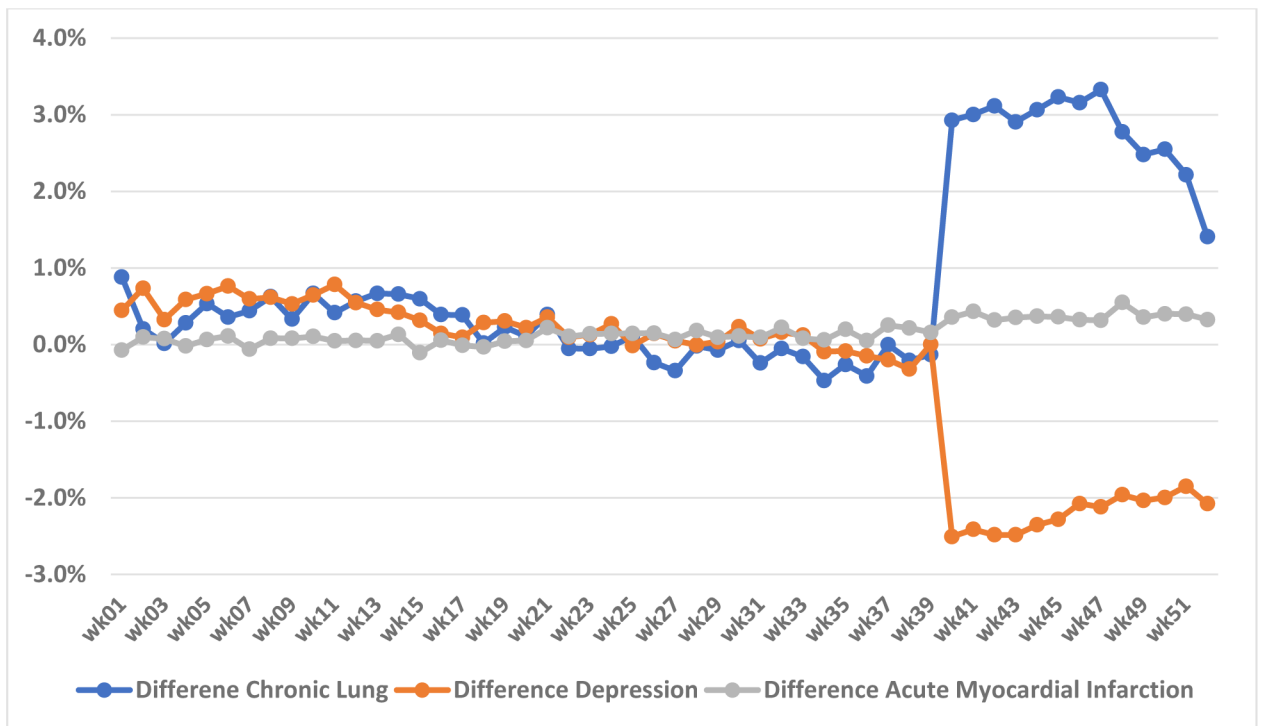
Definitions compared are from Chronic Conditions Warehouse Chronic Conditions Algorithms August 2017 version. https://www.ccwdata.org/web/guest/condition-categories Accessed September 11, 2017. Weekly frequencies.

Figure 1:
Examples of Discontinuities measured: Differences in proportion of hospital discharges associated with Chronic Obstructive Lung Disease, Depression, Acute Myocardial Infarction using Chronic Conditions Warehouse ICD-9 and ICD-10 definitions.
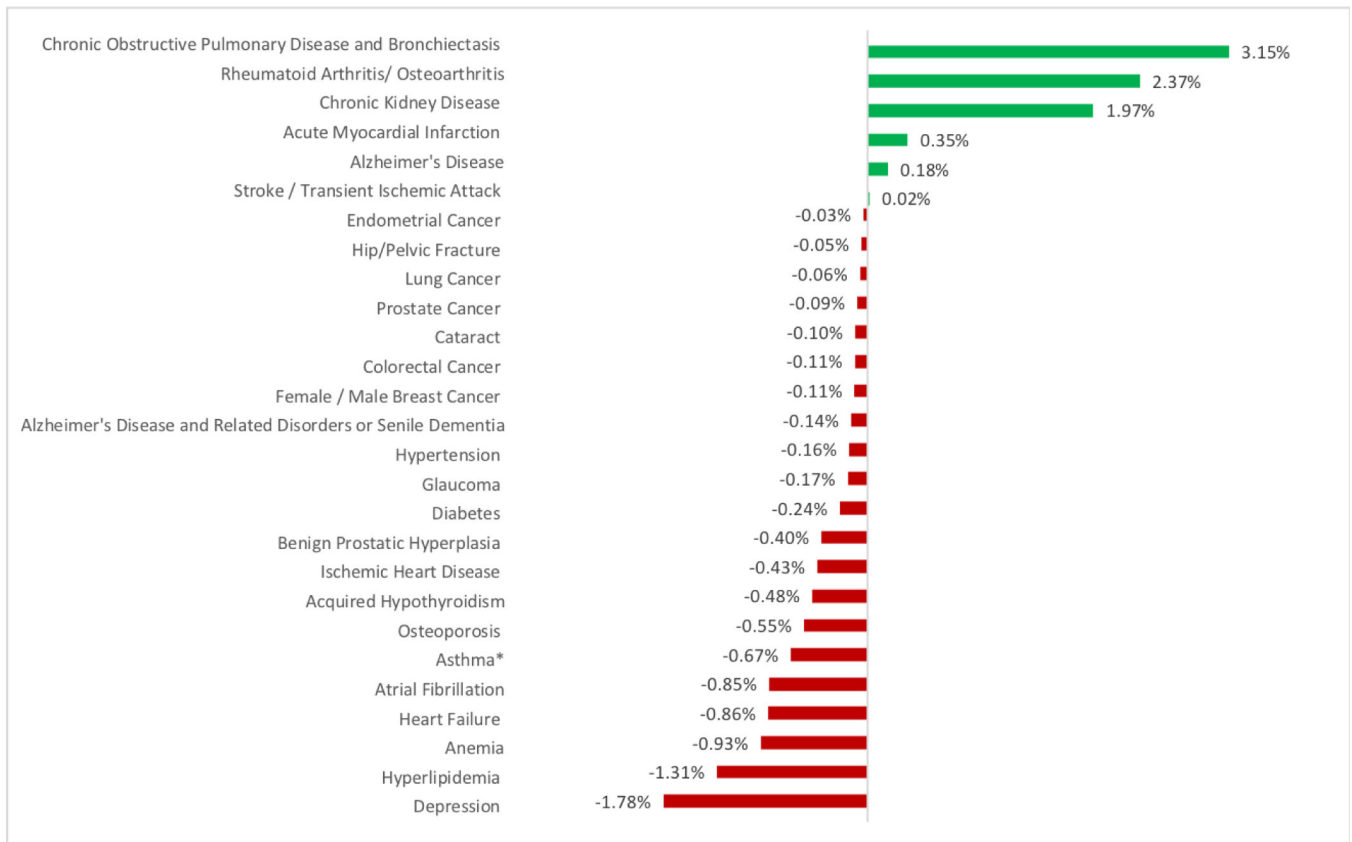
The graph displays the excess change in frequency of condition appearance in Medicare paid hospital discharges apparent in the MedPAR files. The difference between the fourth quarter mean weekly frequency in 2015 vs 2014 minus the average fourth quarter change 2012-2014.

Figure 2:
Discontinuity of Diagnosis Appearance in Medicare Hospital Discharges - Chronic Conditions Warehouse ICD-9 vs. ICD-10 Definitions
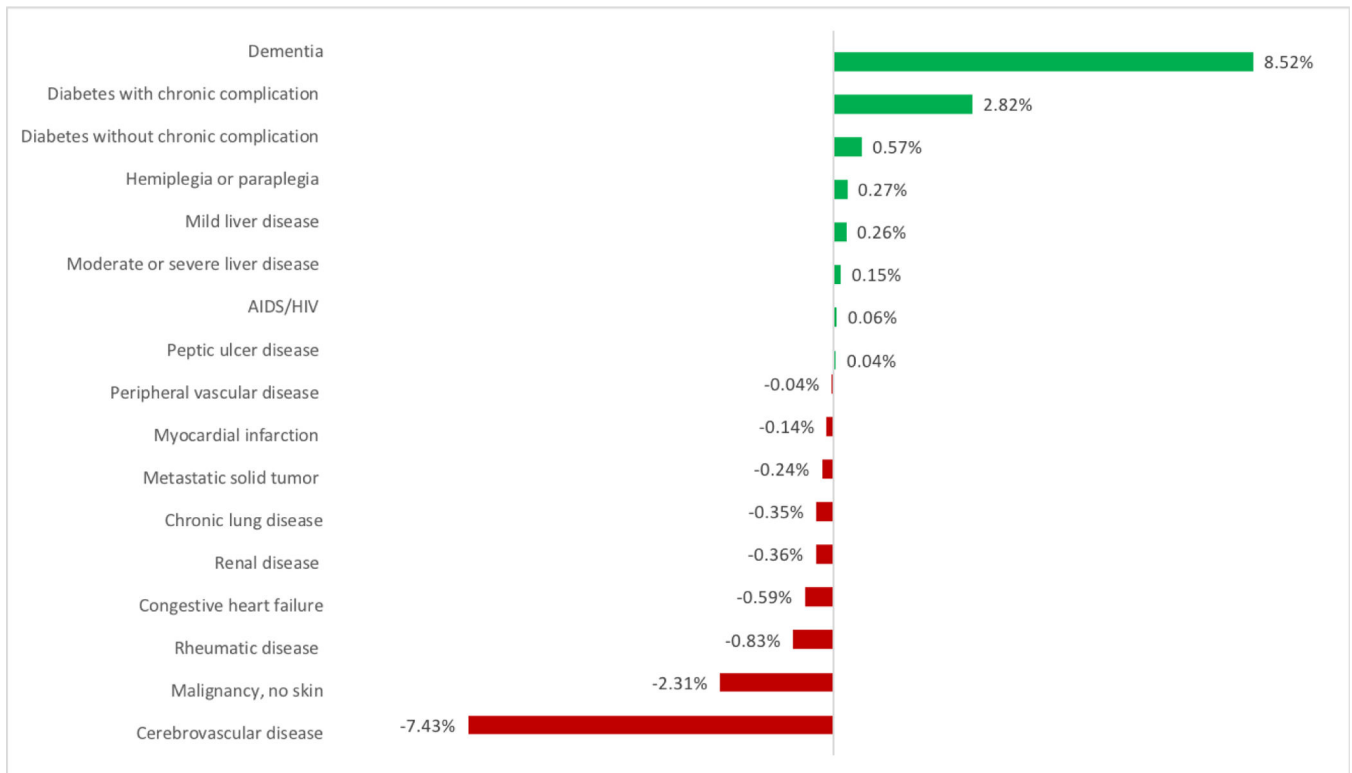
The graph displays the excess change in frequency of condition appearance in Medicare paid hospital discharges apparent in the MedPAR files. The difference between the fourth quarter mean weekly frequency in 2015 vs 2014 minus the average fourth quarter change 2012-2014.

Figure 3:
Discontinuity of Diagnosis Appearance in Medicare Hospital Discharges - General
Equivalency Mapping Software (GEMS) Translation of Charlson-Deyo ICD-9 vs. ICD-10
Definitions