Quantifying Productivity Growth in Health Care Using Insurance Claims and Administrative Data

John A. Romley
Abe Dunn
Dana Goldman
Neeraj Sood

March 1, 2019

We assess changes in multifactor productivity (MFP) in delivering episodes of care (including that received after initial discharge from a hospital) for elderly Medicare beneficiaries with three important conditions over 2002-2014. Across the conditions, we find that MFP declined during the 2000s and then stabilized. For heart attack, for example, MFP decreased by 15.9% over the study period. While heart-attack patients experienced better health outcomes over time, growth in the cost of care for these episodes dominated. The cost of hospital readmissions among heart-attack patients appears to have increased substantially.

1. Introduction

Multifactor productivity (MFP) growth is the ultimate source of gains in living standards, and growth appears to have slowed in the United States since the turn of the century.(Byrne, Oliner et al. 2013, Fernald 2015) One view of the current situation is that the technological progress of earlier eras is unlikely to be matched in the future, notwithstanding the ongoing information revolution and foreseeable developments.(Gordon 2016) An alternative view is that government economic statistics have systematically mismeasured MFP improvement, in fact understating it.(Feldstein 2017) Recent assessments cast some doubt on this alternative view as a convincing account of the apparent slowdown in productivity growth.(Byrne, Fernald et al. 2016, Syverson 2017)

These assessments, while informative, have not squarely addressed the issue of productivity growth in health care. This sector accounted for 17.9 percent of GDP in 2017.(Martin, Hartman et al. 2018) As health spending has grown, so have better treatments become available. Quality change is a well-known challenge for measuring prices, and the mismeasurement of health care inflation was a key concern of the Boskin Commission.(Boskin, Dulberger et al. 1998) Indeed, taking account of improved outcomes, the price of heart attack treatment has actually declined markedly over time.(Cutler, McClellan et al. 1998)

With respect to MFP, there is a longstanding hypothesis that health care and other services suffer from a "cost disease," by which a comparatively meager flow of labor-saving efficiencies drives production costs higher and higher.(Baumol and Bowen 1965, Newhouse 1992, Baumol, de Ferranti et al. 2012) The Medicare Board of Trustees has adopted this position in its long-term financial projections, through an assumption that MFP within the health-care sector will grow more slowly than MFP outside of health care.(The Board of Trustees 2018)

More starkly, the Bureau of Labor Statistics (BLS) has estimated that hospitals and nursing and residential care facilities experienced negative MFP growth from 1987 through 2006.(Harper, Khandrika et al. 2010)

The BLS measures MFP by applying a rigorous and consistent framework across industries. However, this measurement framework does not adequately reflect quality change in a health care system in which treatment outcomes are not fully embodied in transaction prices, due to factors that include information asymmetry, health insurance and administrative pricing. Another challenge in this context is that production is joint between the firm and the consumer in the sense that patients present themselves to providers for care with good, bad, or middling health. Providers who face sicker patients may use more (or fewer) resources in treatment. In a prior study, we found that U.S. hospitals substantially improved their productivity from 2002 through 2011, but only after we accounted for trends in patient severity and treatment outcomes. Improvement in patient outcomes was largely responsible.(Romley, Goldman et al. 2015)

It is critically important to understand productivity change in health care. Improved productivity could allow for cost containment, or efficiencies could be reinvested in better outcomes whose social value exceeds their resource cost (due to technological constraints on the production of health.) Yet the treatment of heart attacks and other conditions does not end with discharge from the hospital. We need to understand productivity in the treatment of complete episodes of care, including, for example, rehab, follow up doctor visits, and medications. Increasingly, public and private decision makers are assessing and incentivizing episodes of care. For example, the Centers for Medicare and Medicaid Services (CMS) recently expanded its innovation portfolio to include a Bundled Payments for Care Improvement Advanced Model (BPCI-A).(Centers for Medicare and Medicaid Services)

The complexity of health care renders it difficult to assess productivity in care delivery. At the same time, health care is a setting in which there are voluminous (albeit imperfect) data to work with. As examples, BEA's recently developed Health Care Satellite Account tracks health care spending for 261 conditions, while BLS is expanding its use of quality adjustment in price measurement.(Dunn, Rittmueller et al. 2016, Dunn, Whitmire et al. 2018, Moulton 2018) In this study, we use insurance claims and administrative data to quantify trends in the productivity of treating episodes of acute myocardial infarction (AMI), or heart attack, among elderly Americans. We also consider heart failure and pneumonia.

2. Approach

The starting point for our analysis is CMS's Inpatient Files.(Research Data Assistance Center) Our version of the Inpatient Files includes a random 20% sample of Medicare beneficiaries. As Table 1 shows, there were 29,841,183 stays at 6,353 short-term acute-care hospitals over the period 2002-2014. The Inpatient File is actually a claim-level file, and multiple claims may be associated with the same stay. While the Medicare Provider Analysis and Review File reports at the stay level, we used the Inpatient File in order to implement a complex algorithm developed by CMS for the purpose of identifying unplanned hospital readmissions.(Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (YNHHSC/CORE) 2014) The publicly available code for the CMS algorithm produces a stay-level data set by combining associated claims.

Table 1 further shows that 811,517 stays at 5,510 hospitals were for patients with a principal diagnosis of AMI. The first three digits of these diagnoses were *410*, per the

International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9).(National Center for Health Statistics)

We define episodes of AMI care as beginning with admission to a short-term acute-care hospital and ending either 90 days after discharge from the initial (i.e., "index") stay or with death, whichever came first. CMS's hospital-based bundled-payment models have almost invariably used 90-day post-discharge windows.(Centers for Medicare and Medicaid Services 2017) Because we do not have access to Medicare service utilization in 2015, we exclude episodes that started in the fourth quarter of 2014 (see Table 1.) Death dates were obtained from the research-identifiable ("RIF") version of CMS's Beneficiary Summary Files (specifically, the A/B segments that report Medicare enrollment and other beneficiary attributes.) We treat a beneficiary as having died only if her reported date was flagged as having been validated by the Social Security Administration or Railroad Retirement Board. Under our Data Use Agreement, our CMS data also include uniquely encrypted beneficiary identifiers; these IDs are used to link the Beneficiary Summary Files to the Inpatient Files.

To quantify productivity in delivering episodes of care, we estimate the following production function:

$$\ln Y_{ht} = \alpha + \mathbf{ln}\,\boldsymbol{I_{ht}}\,\boldsymbol{\beta_I} + \boldsymbol{S_{ht}}\boldsymbol{\beta_S} + \boldsymbol{O_{ht}}\boldsymbol{\beta_O} + g(t) + \epsilon_{ht},$$

in which Y_{ht} is the output of episodes initiated with an admission to index hospital h during year t, $\boldsymbol{S_{ht}}$ is severity factors for the patients in these episodes, and $\boldsymbol{O_{ht}}$ is other elements of hospital production (e.g., medical education).

Our object of interest is the function $g(t)$, a common-across-hospitals but year-specific residual between measured determinants of production and measured output. As is standard, we will interpret this residual as MFP and changes in the residual over time as productivity improvement (or decline.) As is also standard, the validity of this interpretation depends on the validity of the measurement of production determinants and output.

We measure output in each index hospital-year by the number of “high-quality” episodes, as defined below.[1] Under this framework, the health care system does not receive credit (in terms of output) for a low-quality episode, yet is still responsible for the use of scarce resources (in terms of inputs) in delivering the episode. The specification above assumes that the elasticity of substitution between quantity and favorable outcomes is -1. In a sensitivity analysis, we define output by the number of episodes and add the rates of favorable outcomes as distinct model covariates.

Our framework for incorporating quality has been called the “redefine the good” approach, in contrast with the “cost of living” approach.(Sheiner and Malinovskaya 2016) The latter was used to develop the heart attack inflation measure referenced previously.(Cutler, McClellan et al. 1998) Our framework addresses production, but does not address larger issues of allocative efficiency or social welfare.

In our prior study of MFP trends within hospitals, we defined a high-quality stay as one in which the patient survived at least 30 days beyond the date of admission, and avoided an unplanned readmission with 30 days of discharge per the CMS algorithm. Both of these outcomes correspond to quality-of-care metrics publicly reported by CMS and used in Medicare hospital reimbursement.(Centers for Medicare and Medicaid Services , Centers for Medicare and Medicaid Services , Centers for Medicare and Medicaid Services) In this study, we continue to

[1]Unless there are constant returns to scale, the exclusion of fourth-quarter episodes in 2014 will mechanically generate a decrease in output and inputs in that year, all else equal, and so bias our estimate of MFP. To deal with this, we multiply measured output and inputs in 2014 by the inverse of the proportion of episodes starting in the first three quarters of the year over the period 2002-2013.

use the 30-day readmission outcome. As Table 1 reports, 558,999 AMI stays at 5,290 hospitals were consistent with the algorithm.

One reason that some stays were inconsistent is that the algorithm excludes AMIs coded to be a subsequent episode (i.e., not a beneficiary's initial / first AMI.) Another reason is that certain AMI admissions may in fact be readmissions after an earlier AMI stay, and thus not candidates for the index admission that starts an episode.[2] Still another reason is that additional inclusion and exclusion criteria apply. For example, patients must be 65 years old or older at admission and continuously enrolled in "traditional" fee-for-service Medicare (Parts A and B) to be included, and a candidate index stay is excluded if the patient was discharged "against medical advice."[3] Age and enrollment are determined from the Beneficiary Summary Files, while type of discharge is reported in the Inpatient Files. To maximize sample size, we do not include the optional requirement of 12 months of continuous enrollment prior to the index stay.

As before, we measure quality using mortality, in this case requiring survival through the end of the episode (90-days after discharge from the index stay.) In addition, we define a high-quality episode as one in which the patient "returns to the community," rather than remaining institutionalized. Under the Improving Medicare Post-Acute Care Transformation Act of 2014, discharge to the community was adopted as an interim quality metric.(Centers for Medicare and Medicaid Services) For a high-quality episode, we require community discharge for the last within-episode claim from an institutional setting (Inpatient File or Skilled Nursing Facility File.)

[2]The version of the readmission algorithm we use requires a 30-day gap between index stays. Because our episodes last 90 days after discharge from the index stay, we modify the SAS code accordingly.
[3]For its purposes, CMS excludes candidate stays in which the patient dies before discharge. We modify the SAS code so as not to exclude these episodes.

Turning to production inputs ($\ln I_{ht}$), the comparative returns to capital, labor and other factors are not of interest here, and so we aggregate resources used into the total cost of delivering an episode, aggregating all episodes at each index hospital-year. To do so, we identify claims that overlapped with each episode, including inpatient (short-term acute-care hospitals but also long-term care hospitals and inpatient rehabilitation facilities), outpatient facilities, professional (e.g., a claim submitted by a doctor for an office visit), skilled nursing facilities (SNFs), home health, durable medical equipment, hospice and prescription drugs.[4] Where a claim did not fall entirely within the episode period, we allocate costs based on the proportion of days with overlap.

CMS claims do not directly report costs, but instead measures of resource use. For example, total charges are reported for hospital stays. To estimate costs, we use the financial reports that institutional providers participating in Medicare are required to submit to CMS.(Centers for Medicare and Medicaid Services) Hospitals, for example, report not only their actual costs, but the ratio of their charges to their costs (CCRs.) So a hospital's cost for a claim is measured by linking reported charges on the claim to the hospital's reported CCR based on Medicare provider number and then multiplying the former by the latter, as is commonly done in the literature.(Cutler and Huckman 2003) SNF cost reports include revenue-to-cost ratios, and so we use claim-reported payment for these.[5]

CCRs are sometimes unavailable, and we exclude episodes for which any CCR is missing. As Table 1 shows, this criterion excludes almost one in six episodes. We are currently investigating this issue further, in an effort to improve the quality of our analytic data. In

[4]These types of claims correspond to the Inpatient, Outpatient, Carrier, Skilled Nursing Facility, Home Health Agency, Durable Medical Equipment, Hospice and Part D Prescription Drug Event Files, respectively.
[5]Charges are not in general equal to payments in health care, due, for example, to contractual discounts off list price for commercial insurers as well as administrative pricing for Medicare and other public payers.(Reinhardt 2006)

addition, we compare patient characteristics and outcomes for episodes with and without missing CCRs in the next section.

Professional claims report Relative Value Units (RVUs), a measure of the resources required to provide a particular service.(Medicare Payment and Advisory Committee 2018) The reimbursement received by a professional is equal to the number of RVUs multiplied by a dollar-denominated “conversion factor” (CF) specified annually in CMS’s Medicare Physician Fee Schedule Final Rule, adjusted for geographic differences in the cost of care.(Medicare Payment and Advisory Committee 2018) One objective in setting the CF is to ensure that professional providers offer accessible care to beneficiaries, yet federal policy makers have intervened in the CF-setting process to postpone reductions in professional payments mandated by statute for the purpose of controlling cost growth.(Guterman 2014) We assume that the CF in 2002 equated aggregate professional revenues with aggregate costs in that year, before the interventions began. We do not include prescription drug costs (Medicare Part D was introduced in 2006.)

We wish to measure the real cost of treating AMI episodes. As an input into its reimbursement policy making, CMS constructs and reports “market basket indices” and the Medicare Economic Index (MEI). The Inpatient Hospital market basket index, for example, measures changes in the cost of providing inpatient hospital care. We use this index and those for other institutional settings to deflate nominal costs into real 2014 dollars. The MEI is used for professional payment, and measures inflation in the cost of providing professional services, less an adjustment for productivity growth in the economy at large.(2012 Medicare Economic Index Technical Advisory Panel 2012) We inflation-adjust professional costs by reversing the productivity adjustments to the MEI.

Turning to patient severity ($\boldsymbol{S_{ht}}$), a key measure comes from the Agency for Healthcare Research and Quality's Inpatient Quality Indicators (IQIs).(Agency for Healthcare Research and Quality) The IQIs were developed for the purpose of assessing the quality of care across hospitals and over time using standard patient discharge records. The IQIs include inpatient mortality for a variety of conditions, including AMI. In order to reliably assess mortality performance, teams of clinical experts developed risk adjustment models. We use the average predicted likelihood of inpatient death, derived from these models, averaged across AMI episodes initiated at an index hospital in a year. Table 1 reports that predicted mortality was not available for some episodes that are consistent with the CMS readmission algorithm. The AMI IQI excludes cases whose status as the first or subsequent heart attack was not coded, while the readmission algorithm does not. We limit our analysis to episodes with predicted mortality for the sake of clinical specificity.

An important element of the IQI risk models is the All Patients Refined Diagnosis Related Group (APR-DRG), in particular, its risk of mortality scale. While the inputs into the APR-DRGs are known (e.g., diagnosis and procedure codes), a limitation of our approach is that the logic of the APR-DRG "grouper" methodology is proprietary to 3M, and so is not transparent to end users. There is a limited-license version released by AHRQ for the purpose of implementing the IQIs. We apply version 6.0 of the IQIs, the last refinement developed for use with ICD-9 coding (CMS transitioned to ICD-10 beginning in fiscal year 2015.)

We further exploit diagnostic information in our data by measuring the proportion of episodes with different numbers of Charlson-Deyo comorbidities (such as dementia) recorded in the index inpatient record. These comorbidities were selected for their utility in predicting death within 12 months.(Charlson, Pompei et al. 1987, Quan, Sundararajan et al. 2005) For AMI we are also able to characterize the type of heart attack based on the location within the heart, using the fourth digit of the ICD-9 code. The type of heart attack relates to prognosis; for example, survival is relatively favorable for a “non-STEMI” AMI (ICD-9 of 410.7x for subendocardial infarction), at least in the near term.(Cantor, Goodman et al. 2005, Cox, Stone et al. 2006) The number of diagnoses recorded on inpatient claims increased from 10 to 25 in 2010, so we limit ourselves to the first ten.

In addition, we use the proportion of patients who were female and of various races, as reported in the Beneficiary Summary Files. These files also report the zip code in which each beneficiary resides, which we link to zip code-level data from the 2000 Census on a variety of community sociodemographic characteristics used as proxies for patient severity in prior literature (Fisher, Wennberg et al. 2003, Fisher, Wennberg et al. 2003, Romley, Jena et al. 2011, Romley, Goldman et al. 2015); examples include the poverty rate and the proportion of elderly residents with self-care limitations. As Table 1 shows, about 7,200 of 457,100 initiated at hospitals for whom *none* of the patient zip codes cannot be matched to the Census data; all other episodes can be matched. Finally, we use the proportion of discharges in each quarter, as there may be seasonality in severity and fourth-quarter discharges had to be excluded in 2014 (due to incomplete follow up.)

Turning to other elements of hospital production, we account for medical education and the provision of advanced (tertiary) care to the broader health system. Either of these activities may complement AMI care, or draw resources from it. For the former, we use indicator variables for intervals of the number of medical residents per bed specified in prior literature (Volpp, Rosen et al. 2007); this data is available from the Impact Files released annually by CMS in support of its inpatient prospective payment system.(Centers for Medicare and Medicaid

Services) For advanced care, we use indicator variables for neurological and cardiovascular procedures identified in the Dartmouth Atlas of Health Care.(Wennberg and Cooper 1996) Because there is overlap between these cardiac procedures and AMI care, a sensitivity analysis excludes this indicator variable.

Our regressions clustered standard errors at the level of the index hospital. Our base regression weighted hospital-years by the number of episodes.

3. Findings

Table 2 reports summary statistics for our analysis sample. Across 28,801 index hospital-years, the average date of the initial admission is mid-2007. The average cost per episode is $46,700 in 2014 dollars. 79.5% of elderly Medicare beneficiaries admitted to a hospital with an AMI survived at least 90 days beyond the initial discharge. The AHRQ AMI IQI predicts that 92.2% would have survived the initial hospital stay. Among 90-day survivors, 85.6% avoided an unplanned readmission within 30 days of initial discharge. Among survivors without a readmission, 84.4% were discharged home from their final institutional encounter.

In terms of severity, roughly two thirds of episodes involved a non-STEMI AMI. All episodes involved at least one Charlson-Deyo comorbidity, as AMI counts. More than 7 in 10 episodes involved additional Charlson-Deyo comorbidities. The average age of beneficiaries was 78.7 years, slightly less than half were female, and almost 9 in 10 were white. Median household incomes in beneficiaries' zip codes averaged $42,700 in the 2000 Census. In terms of index hospital characteristics, slightly more than 4 in 10 episodes took place at facilities with no medical residents, while about 3 in 20 took place at a major teaching hospital (>0.25 residents per bed.)

A simple measure of productivity is the cost of a "high-quality" episode, in which the beneficiary survives, avoids an unplanned readmission, and is discharged home. Figure 1 shows this measure over 2002-2014. The cost of a high-quality episode was $63,700 in 2002 and rose in most years to reach a maximum of $100,100 in 2014.

Figure 3 shows that the rate of high-quality episodes was 56.3% in 2002 and then climbed in most years, reaching a maximum of 61.9% in 2014. As seen in Figure 4, the rate of survival through 90 days after the index discharge rose from 77.4% in 2002 to 82.7% in 2014. The rate of avoidance of an unplanned readmission among these survivors increased from 84.7% to 87.8% over this period, as Figure 5 shows. The rate of home discharge among survivors without a readmission decreased, albeit modestly, from 85.9% in 2002 to 85.1% in 2014 (see Figure 6.)

While the rate of high-quality stays improved, Figure 6 shows that the average cost of an episode (high-quality or not) started at $35,900 in 2002, rose then dipped in 2013, then rose again to reach a high of $61,900 in 2014.

Figures 7-10 show a number of trends in patient severity. As seen in Figure 7, the average age of beneficiaries was 78.5 years old in 2002, rising to a peak of 79.0 in 2008, then fell back to 78.5 as of 2014. The number of Charlson-Deyo comorbidities on the index inpatient record grew from 2.27 to 2.61 over the period (Figure 8.) Predicted inpatient survival from the AHRQ IQI declined slightly, from 93.3% in 2002 to 92.7% in 2014 (Figure 9.) Based on our benchmark regression results, we create a composite index of patient severity, accounting for all severity-related covariates.

As Figure 10 shows, the patient severity index was 100.0 (by construction) in 2002 and climbed to a maximum of 11.5 in 2011, then declined to 102.1 as of 2014. This decline is consistent with the decrease in age seen in Figure 7. The decline implies that, near the of the analysis window, episodes involved patients whose AMI and other health status were less severe, making it less costly to deliver the improved episode outcomes seen in Figures 3 and 4. A composite index of other hospital production was stable (Figure 11.)[6]

Our regression uses indicator variables for calendar year for the productivity function $g(t)$. Figure 12 shows the cumulative change in MFP since 2002. In 2003, productivity decreased by 3.2%. It decreased another 9.2% through 2007 and was then stable through 2011. It improved in 2012, from 13.1% below the 2002 level to 7.8%. A decline in 2013 wiped out these gains, and MFP ended in 2014 at 15.9% below the 2002 level ($p < 0.001$.)

The decline through 2007 is similar to the pattern we found when examining MFP within hospitals.(Romley, Goldman et al. 2015) In that study, the trough was roughly 5% below 2002. However, based on the initial hospitalization only, MFP recovered by 2008 enough to be comparable to 2002, and it exceeded the 2002 level by almost 8% in 2011, the end of that study's analysis window.

The movements shown in Figure 12 are sizable. The standard deviation of the year-to-year changes was 3.8%. As a point of comparison, our prior study of U.S. hospital MFP over 2002-2011 had a year-to-year standard deviation of 2.9%. Another study of hospital productivity over 1981-2005 found standard deviations close to 2.0%. The longer follow up over a 90-day post-discharge episode plausibly enlarges the scope for variations.

Figure 13 helps to decompose the drivers of these trends. If output were redefined to ignore the quality of episodes (i.e., quantity only), MFP in 2014 would have been 25.6% below its 2002 level, instead of 15.9% below. We also found improvement in outcome quality to be an

[6]An increase in the other hospital production index would have implied that medical education and/or advanced care capabilities made it more difficult to deliver high-quality episodes.

important factor in our prior study of productivity within hospitals. If patient severity were ignored in the analysis, 2014 MFP would have been 18.3% lower than in 2002. Thus, severity adjustment plays a modest role, as we found in the case of heart attack previously. Finally, if other hospital production were ignored, 2014 MFP would have been 14.5% lower than in 2002, similar to our main result. In sensitivity analysis for AMI episodes, our findings are similar when advanced cardiovascular procedures are excluded from other hospital production.

Our regression analysis that confirms that the story told in Figures 1-6 persists after accounting for patient severity, and quantifies the magnitude of the MFP decline. MFP decreased not because quality-adjusted output decreased, but rather because increases in output did not keep sufficient pace with rising costs per episode. Total real costs per AMI episode grew nearly 73% over 2002-2014. Figure 14 shows that the average cost of acute inpatient care after the initial hospitalization increased from $7,400 in 2002 to $28,300 in 2014. By contrast, the cost of all other care grew from $28,500 to only $33,600.

The dramatic increase in the cost of hospital readmissions merits careful scrutiny. Based on our current findings, including the regression analysis, if the cost of post-index acute inpatient care had grown at the same rate as all other care, and if that slower growth rate did not negatively impact the quantity or quality of care, then MFP in delivering AMI episodes would have increased by 10.7% over 2002-2014, rather than decreasing by 15.9%.

Figure 15 compares MFP trends in treating AMI to the trends in treating heart failure and pneumonia episodes. In all cases, productivity declined over 2002-2014. Compared to AMI, MFP decreased less for pneumonia (9.1% below the 2002 level) but more for heart failure (28.4% below 2002.) As with AMI, the bulk of the declines for pneumonia and heart failure came early in the period. In the case of pneumonia, MFP recovered nearly a third of its losses, improving from its trough of 13.8% below the 2002 level as of 2010 to 9.1% below 2002 as of 2014.

Finally, we address the issue of missing cost-to-charge ratios (CCRs.) Figure 16 reports the rate of high-quality stays, according to whether a CCR is missing for any claims in an episode. The rate of high quality stays is significantly lower for episodes with at least one CCR missing. Recall that these episodes were excluded from the results just presented. Our hypothesis is that an episode with more claims is more likely to have a CCR missing, and higher utilization reflects poorer health status, and so the poor outcomes seen in the figure. Clearly, the issue of missing CCRs is an important one to address further.

4. Conclusions

As policy makers and health practitioners increasingly focus on episodes of care, this study has extended our prior analyses of the MFP of hospitals in caring for patients admitted with heart attack, heart failure and pneumonia. We follow patients for ninety days after hospital discharge, incorporating all the care received from health care professionals and institutions and assessing patient survival, readmission and community discharge over the course of an episode.

We faced challenges in conducting an analysis as comprehensive as this. The main challenge is that cost data was not available for all providers. With significant effort, we were able to estimate costs for *all* of the institutional care received in 84% of episodes. For the other episodes, patient outcomes were systematically worse, raising a concern about the representativeness of our findings. Going forward, the causes of missing cost data can and should be investigated and addressed to the extent possible. A pragmatic approach worth exploring would assess the likely importance of the missing data. For example, missing cost data on a single claim whose reimbursement is small in comparison to other claims would seem to be a not terribly worrisome case.

Another challenge is the measurement of the quality of care that patients receive. We have focused here on three aspects of quality that are clearly important to patients, practitioners and policy makers. Functional status is also important, but not widely available for large patient populations. Perhaps this issue could be addressed by drawing on disparate sources of information; the advantage of the current approach is its internal consistency. In terms of community discharge, our measure speaks only to the last institutional Medicare provider seen. A patient could have been discharge home but later admitted during the episode window to a custodial nursing home not covered by the Medicare program, and so not observed in the current analysis. Future research should explore methods that we previously developed for use with the Minimum Data Set, which was unavailable to us at the time of this study.(Buntin, Colla et al. 2010)

Measuring patient severity is a perennial challenge in analyses of health care. During the period studied, CMS adopted a new characterization of diagnoses for purposes of reimbursement, and changes such as these can induce responses in diagnostic behavior by providers that create spurious trends in measured severity. In the case of AMI episodes, severity adjustment based on diagnoses, demographics and community context did not materially affect our productivity findings. Nevertheless, this issue remains important, and worth exploring further. Just after the study period ended, CMS transitioned from ICD-9 diagnoses to the ICD-10 classification. To extend the present analysis later in time, ICD-9 codes will have to be cross walked to ICD-10, and there would be some risk of spurious trends.

Despite these challenges, this study has produced some striking findings. In particular, for all three conditions studied, measured MFP decreased from 10% to 30% over the 12-year window. For AMI episodes, patient outcomes improved but were dominated by rising costs. Our findings underscore concerns about productivity improvement and cost control in health care. In our prior study that included hospital costs only, productivity increased between 2002 and 2011. This contrast with the present findings underscores the importance of taking a more comprehensive view.

In our hospital-only study, there was in initial decrease in MFP followed by a larger increase for the cardiac conditions. In the current study, MFP decreased early on for the most part and then largely stabilized, but did not recover. For AMI episodes, hospital readmissions appear to have driven the increase in episode costs. It is noteworthy that CMS instituted penalties for “excess” readmissions beginning in 2013. Understanding the impact of public policy and market conditions on MFP is an important and ambitious direction for future research.

References

2012 Medicare Economic Index Technical Advisory Panel (2012). Report to the HHS Secretary: Review of the Medicare Economic Index.

Agency for Healthcare Research and Quality. "Inpatient Quality Indicators Overview." from https://www.qualityindicators.ahrq.gov/modules/iqi_resources.aspx.

Baumol, W. J. and W. G. Bowen (1965). "On the Performing Arts: The Anatomy of Their Economic Problems." The American Economic Review **55**(1/2): 495-502.

Baumol, W. J., D. de Ferranti, M. Malach, M. Pablos, xe, A. ndez, H. Tabish and L. G. Wu (2012). The Cost Disease: Why Computers Get Cheaper and Health Care Doesn't, Yale University Press.

Boskin, M. J., E. L. Dulberger, R. J. Gordon, Z. Griliches and D. W. Jorgenson (1998). "Consumer Prices, the Consumer Price Index, and the Cost of Living." Journal of Economic Perspectives **12**(1): 3-26.

Buntin, M. B., C. H. Colla, P. Deb, N. Sood, J. Escarce, xe and J (2010). "Medicare Spending and Outcomes After Postacute Care for Stroke and Hip Fracture." Medical Care **48**(9): 776-784.

Byrne, D. M., J. G. Fernald and M. B. Reinsdorf (2016). "Does the United States Have a Productivity Slowdown or a Measurement Problem?" Brookings Papers on Economic Activity: 109-157.

Byrne, D. M., S. D. Oliner and D. E. Sichel (2013). "Is the Information Technology Revolution Over?" International Productivity Monitor **33**: 20-36.

Cantor, W. J., S. G. Goodman, C. P. Cannon, S. A. Murphy, A. Charlesworth, E. Braunwauld and A. Langer (2005). "Early cardiac catheterization is associated with lower mortality only among high-risk patients with ST- and non–ST-elevation acute coronary syndromes: Observations from the OPUS-TIMI 16 trial." American Heart Journal **149**(2): 275-283.

Centers for Medicare and Medicaid Services. "BPCI Advanced." from https://innovation.cms.gov/initiatives/bpci-advanced.

Centers for Medicare and Medicaid Services. "Historical Impact Files for FY 1994 through Present." from https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Historical-Impact-Files-for-FY-1994-through-Present.html.

Centers for Medicare and Medicaid Services "Hospital Compare."

Centers for Medicare and Medicaid Services. "Hospital Cost Report Public Use File." from https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Cost-Report/HospitalCostPUF.html.

Centers for Medicare and Medicaid Services. "Hospital Readmissions Reduction Program." from https://www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientpps/readmissions-reduction-program.html.

Centers for Medicare and Medicaid Services. "The Hospital Value-Based Purchasing (VBP) Program." from https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/hvbp/hospital-value-based-purchasing.html.

Centers for Medicare and Medicaid Services. "IMPACT Act of 2014 Data Standardization & Cross Setting Measures." from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/IMPACT-Act-of-2014/IMPACT-Act-of-2014-Data-Standardization-and-Cross-Setting-Measures.html.

Centers for Medicare and Medicaid Services (2017).
CMS Bundled Payments for Care Improvement Initiative Models 2-4: Year 3
Evaluation & Monitoring Annual Report.

Charlson, M. E., P. Pompei, K. L. Ales and C. R. MacKenzie (1987). "A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation." Journal of Chronic Diseases **40**(5): 373-383.

Cox, D. A., G. W. Stone, C. L. Grines, T. Stuckey, P. J. Zimetbaum, J. E. Tcheng, M. Turco, E. Garcia, G. Guagliumi, R. S. Iwaoka, R. Mehran, W. W. O'Neill, A. J. Lansky and J. J. Griffin (2006). "Comparative Early and Late Outcomes After Primary Percutaneous Coronary Intervention in ST-Segment Elevation and Non–ST-Segment Elevation Acute Myocardial Infarction (from the CADILLAC Trial)." The American Journal of Cardiology **98**(3): 331-337.

Cutler, D. M. and R. S. Huckman (2003). "Technological development and medical productivity: the diffusion of angioplasty in New York state." Journal of Health Economics **22**(2): 187-217.

Cutler, D. M., M. McClellan, J. P. Newhouse and D. Remler (1998). "Are Medical Prices Declining? Evidence from Heart Attack Treatments." The Quarterly Journal of Economics **113**(4): 991-1024.

Dunn, A., L. Rittmueller and B. Whitmire (2016). "Health Care Spending Slowdown From 2000 To 2010 Was Driven By Lower Growth In Cost Per Case, According To A New Data Source." Health Affairs **35**(1): 132-140.

Dunn, A., B. Whitmire, A. Batch, L. Fernando and L. Rittmueller (2018). "High Spending Growth Rates For Key Diseases In 2000–14 Were Driven By Technology And Demographic Factors." Health Affairs **37**(6): 915-924.

Feldstein, M. (2017). "Underestimating the Real Growth of GDP, Personal Income, and Productivity." Journal of Economic Perspectives **31**(2): 145-164.

Fernald, J. G. (2015). "Productivity and Potential Output before, during, and after the Great Recession." NBER Macroeconomics Annual **29**(1): 1-51.

Fisher, E. S., D. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas and E. L. Pinder (2003). "The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care." Ann Intern Med **138**(4): 273-287.

Fisher, E. S., D. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas and E. L. Pinder (2003). "The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care." Ann Intern Med **138**(4): 288-298.

Gordon, R. J. (2016). The rise and fall of American growth : the U.S. standard of living since the Civil War. Princeton, Princeton University Press.

Guterman, S. (2014). "The "Doc Fix" — Another Missed Opportunity." New England Journal of Medicine **370**(24): 2261-2263.

Harper, M., B. Khandrika, R. Kinoshita and S. Rosenthal (2010). "Nonmanufacturing industry contributions to multifactor productivity, 1987–2006." Monthly Labor Review.

Martin, A. B., M. Hartman, B. Washington, A. Catlin and T. N. H. E. A. Team (2018). "National Health Care Spending In 2017: Growth Slows To Post–Great Recession Rates; Share Of GDP Stabilizes." Health Affairs **38**(1): 10.1377/hlthaff.2018.05085.

Medicare Payment and Advisory Committee (2018). Physician and other health professional payment system.

Moulton, B. (2018). The Measurement of Output, Prices, and Productivity: What's Changed Since the Boskin Commission?, Brookings Institution.

National Center for Health Statistics. "International Classification of Diseases,Ninth Revision, Clinical Modification (ICD-9-CM)." from https://www.cdc.gov/nchs/icd/icd9cm.htm.

Newhouse, J. P. (1992). "Medical Care Costs: How Much Welfare Loss?" Journal of Economic Perspectives **6**(3): 3-21.

Quan, H., V. Sundararajan, P. Halfon, A. Fong, B. Burnand, J.-C. Luthi, L. D. Saunders, C. A. Beck, T. E. Feasby and W. A. Ghali (2005). "Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data." Medical Care **43**(11): 1130-1139.

Reinhardt, U. (2006). "The Pricing Of U.S. Hospital Services: Chaos Behind A Veil Of Secrecy." Health Affairs **25**(1): 57-69.

Research Data Assistance Center. "Inpatient (Fee-for-Service)." from https://www.resdac.org/cms-data/files/ip-ffs.

Romley, J. A., D. P. Goldman and N. Sood (2015). "US hospitals experienced substantial productivity growth during 2002-11." Health Aff (Millwood) **34**(3): 511-518.

Romley, J. A., A. B. Jena and D. P. Goldman (2011). "Hospital spending and inpatient mortality: evidence from California: an observational study." Ann Intern Med **154**(3): 160-167.

Sheiner, L. and A. Malinovskaya (2016). Measuring Productivity Growth in Healthcare: An Analysis of the Literature, Brookings Institution.

Syverson, C. (2017). "Challenges to Mismeasurement Explanations for the US Productivity Slowdown." Journal of Economic Perspectives **31**(2): 165-186.

The Board of Trustees, F. O.-A. a. S. I. a. F. D. I. T. F. (2018). The 2018 Annual Report of the Boards of Trustees of the Federal Hospital Insurance Trust Fund and the Federal Supplementary Medical Insurance Trust Fund.

Volpp, K. G., A. K. Rosen, P. R. Rosenbaum, P. S. Romano, O. Even-Shoshan, A. Canamucio, L. Bellini, T. Behringer and J. H. Silber (2007). "Mortality Among Patients in VA Hospitals in the First 2 Years Following ACGME Resident Duty Hour Reform." JAMA **298**(9): 984-992.

Wennberg, J. and M. Cooper (1996). "The Dartmouth atlas of health care." The Center for the Evaluative Clinical Sciences, Dartmouth Medical School, American Hospital Publishing: 15-20.

Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (YNHHSC/CORE) (2014). 2014 Measure Updates and Specifications Report Hospital-Wide All-Cause Unplanned Readmission – Version 3.0.

Table 1

Stays / Episodes	*Beneficiaries*	*Hospitals*	*Description*
29,841,183	7,880,612	6,353	All Medicare FFS stays in short term acute care hospitals, 2002-2014, based on random 20% sample of beneficiaries
811,517	635,380	5,510	Heart attack (acute myocardial infarction, i.e., AMI) stays
798,414	625,301	5,505	Excluding stays in fourth quarter of 2014 (incomplete follow up as index stays)
558,999	501,940	5,290	Stays / episodes meeting CMS readmission measure criteria
470,120	426,933	4,837	Excluding episodes with any missing cost-to-charge ratios
457,120	415,562	4,753	Episodes meeting AHRQ IQI risk measure criteria
449,950	409,423	3,859	Excluding index hospital-years with no Census sociodemographic data available

Table 2

Variable	*Mean (SE)*
Episodes, n	449,950
Hospitals, n	3,859
Hospital-years, n	28,801
Year of admission	2007.4 (3.7)
Cost per episode (000s of 2014 dollars)	$46.7 ($69.0)
Survival of episode	79.5% (12.2%)
No unplanned readmissions (30 day) among survivors	85.6% (10.4%)
Discharge home among survivors without readmissions	84.4% (13.3%)
AHRQ predicted inpatient survival	92.2% (3.7%)
Location of heart attack: Anterolateral (410.0x)	2.1% (3.6%)
Location of heart attack: Other Anterior Wall (410.1x)	8.1% (7.5%)
Location of heart attack: Inferolateral Wall (410.2x)	1.7% (3.4%)
Location of heart attack: Inferoposterior Wall (410.3x)	1.2% (2.7%)
Location of heart attack: Other Inferior Wall (410.4x)	9.9% (8.1%)
Location of heart attack: Other Lateral Wall (410.5x)	1.2% (2.8%)
Location of heart attack: True Posterior Wall (410.6x)	0.3% (1.5%)
Location of heart attack: Sub-Endocardial (410.7x)	68.3% (16.5%)
Location of heart attack: Other Specified Sites (410.8x)	1.4% (4.4%)
Location of heart attack: Unspecified site (410.9x)	5.7% (9.0%)
No Charlson-Deyo comorbidity	0.0% (0.0%)
1 Charlson-Deyo comorbidity	27.9% (12.9%)
2 Charlson-Deyo comorbidities	32.3% (12.3%)
3 Charlson-Deyo comorbidities	20.8% (11.1%)
4 Charlson-Deyo comorbidities	11.2% (9.0%)
5+ Charlson-Deyo comorbidities	7.8% (8.2%)
Age	78.7 (3.1)
Female	48.8% (14.2%)
White	88.1% (15.4%)
African American	7.6% (12.6%)
Hispanic	1.7% (5.7%)
Other race	2.5% (6.8%)
Patient zip code characteristics	
Median household income ($000)	$42.7 ($10.1)
Social Security income ($000)	$11.3 ($0.9)
Poor	12.0% (4.9%)
Employed	94.3% (2.0%)
Less than high school education	19.9% (6.6%)
Urban	70.4% (21.7%)
Hispanic	8.7% (12.2%)
Single	41.7% (4.6%)
Elderly in an institution	5.5% (2.4%)
Non-institutionalized elderly with physical disability	29.3% (4.7%)
Mental disability	11.0% (2.9%)
Sensory disability among elderly	14.6% (2.6%)
Self-care disability	9.7% (2.5%)
Difficulty going-outside-the-home disability	20.5% (3.6%)
Index hospital characteristics	
No residents	42.8% (49.5%)
Residents per bed > 0 and ≤ 0.25	41.6% (49.3%)
Residents per bed > 0.25 and ≤ 0.6	10.7% (30.9%)
Residents per bed > 0.6	5.0% (21.7%)
Advanced neurosurgical procedures	0.0% (0.0%)
Advanced cardiovascular procedures	7.1% (7.4%)

Figure 1

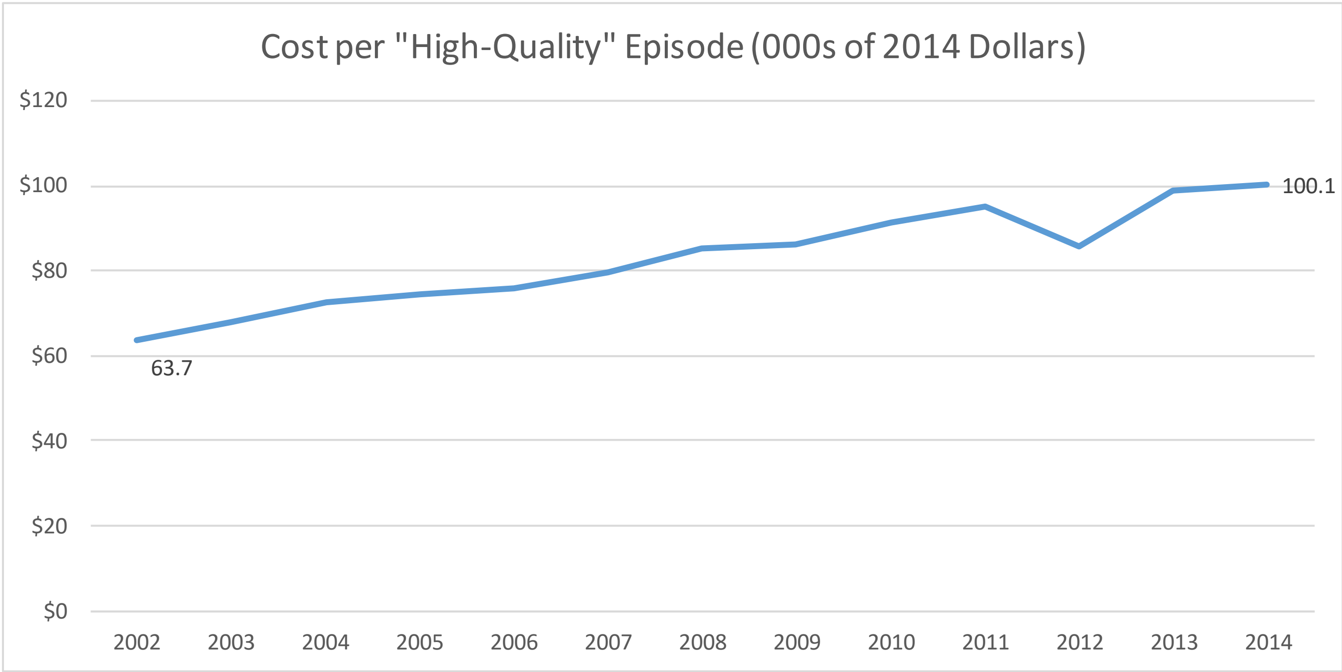
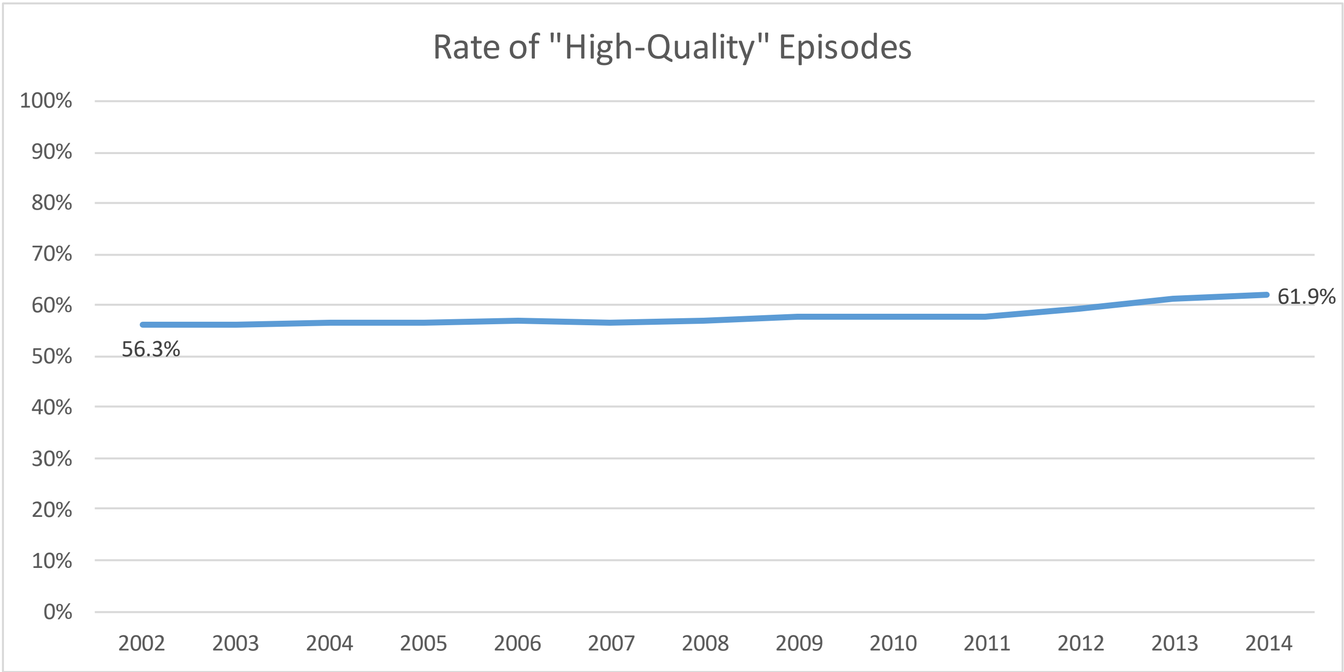

Cost per "High-Quality" Episode (000s of 2014 Dollars)
$120
$100
$80
$60
$40
$20
$0
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
63.7
100.1

Figure 2



Rate of "High-Quality" Episodes
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
56.3%
61.9%
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

Figure 3

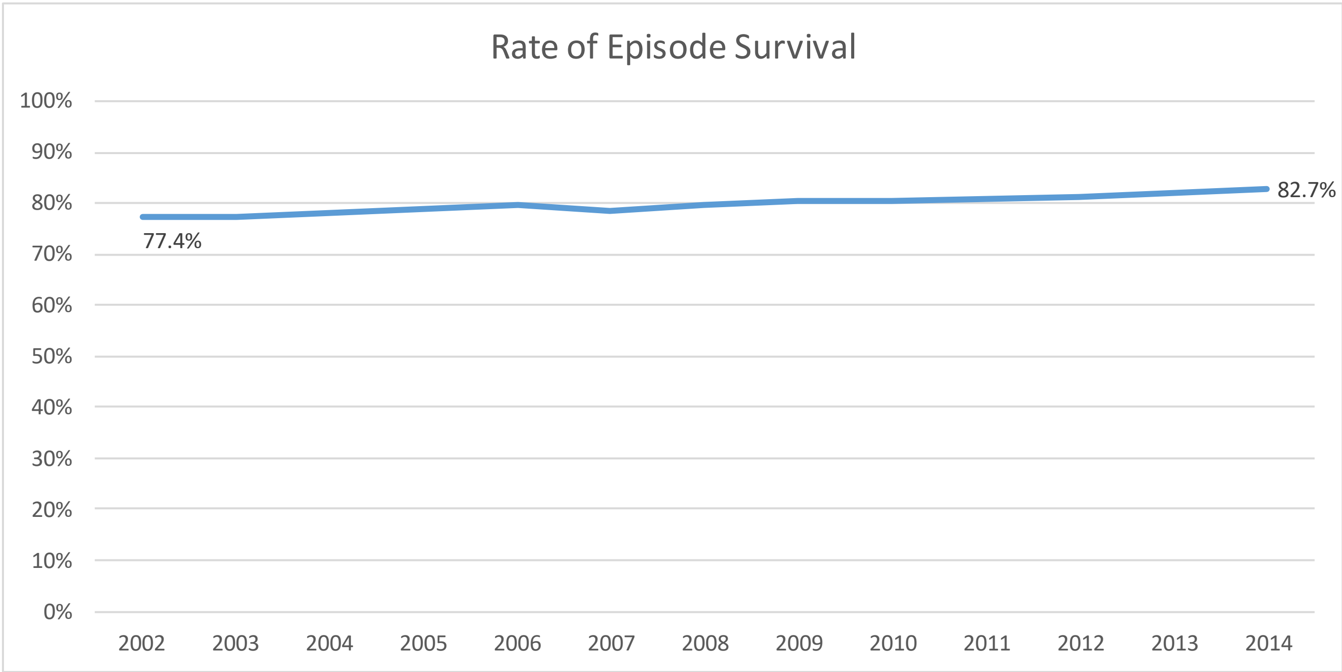
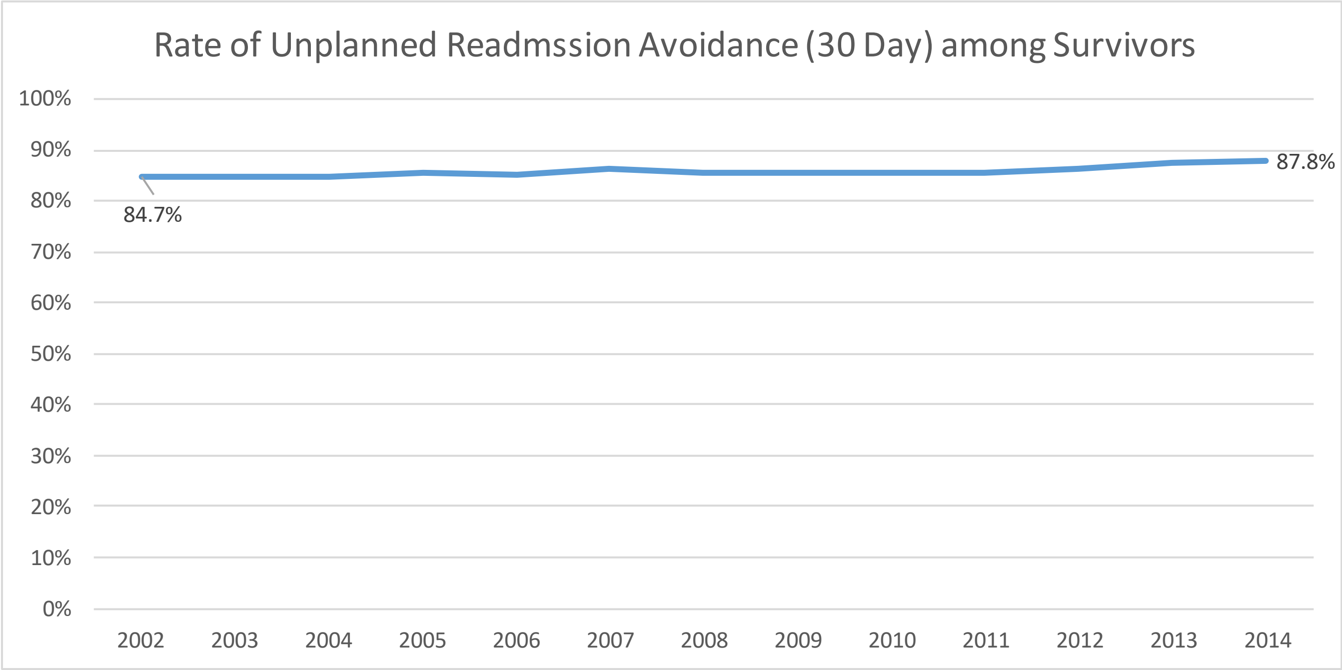

Rate of Episode Survival
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
77.4%
82.7%

Figure 4



Rate of Unplanned Readmssion Avoidance (30 Day) among Survivors
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
84.7%
87.8%
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

Figure 5

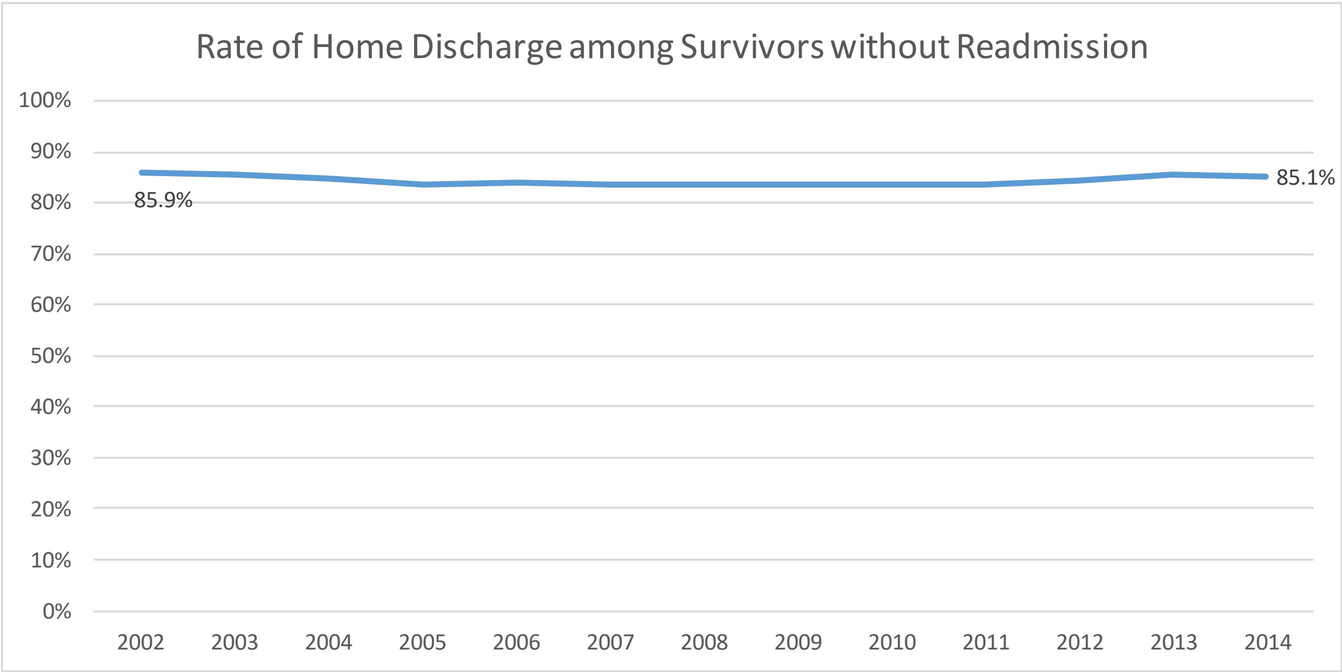

Rate of Home Discharge among Survivors without Readmission
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
85.9%
85.1%
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

Figure 6

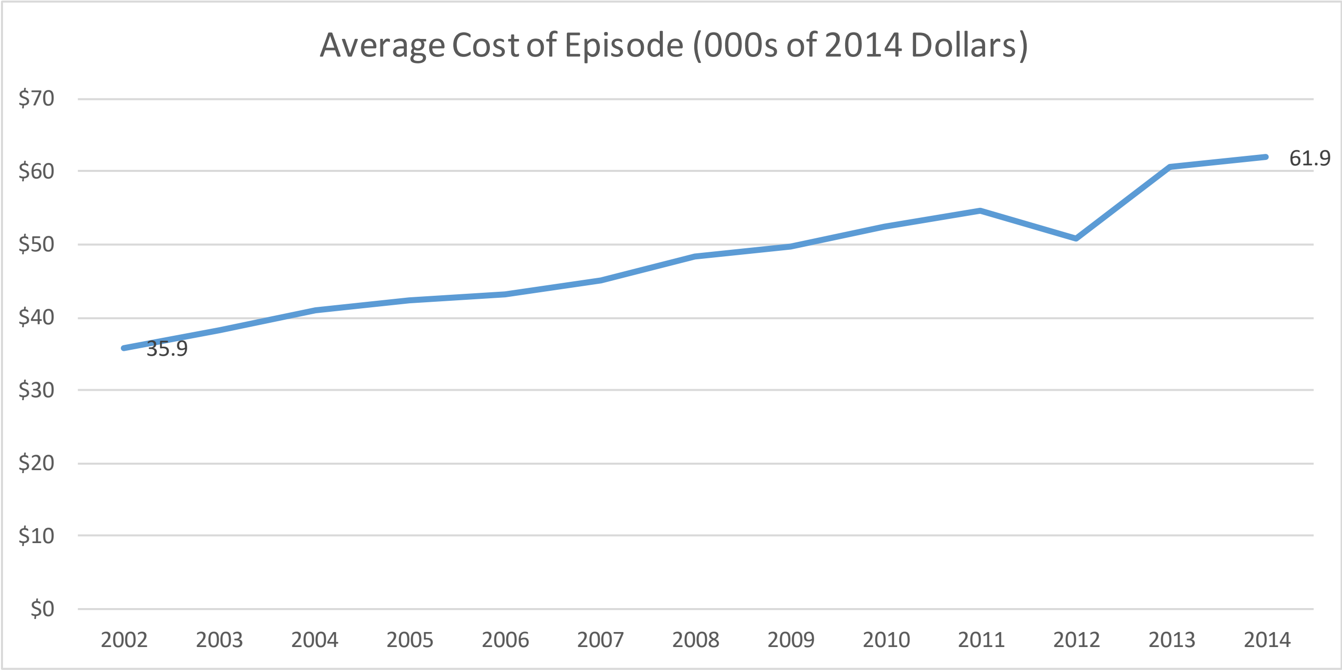
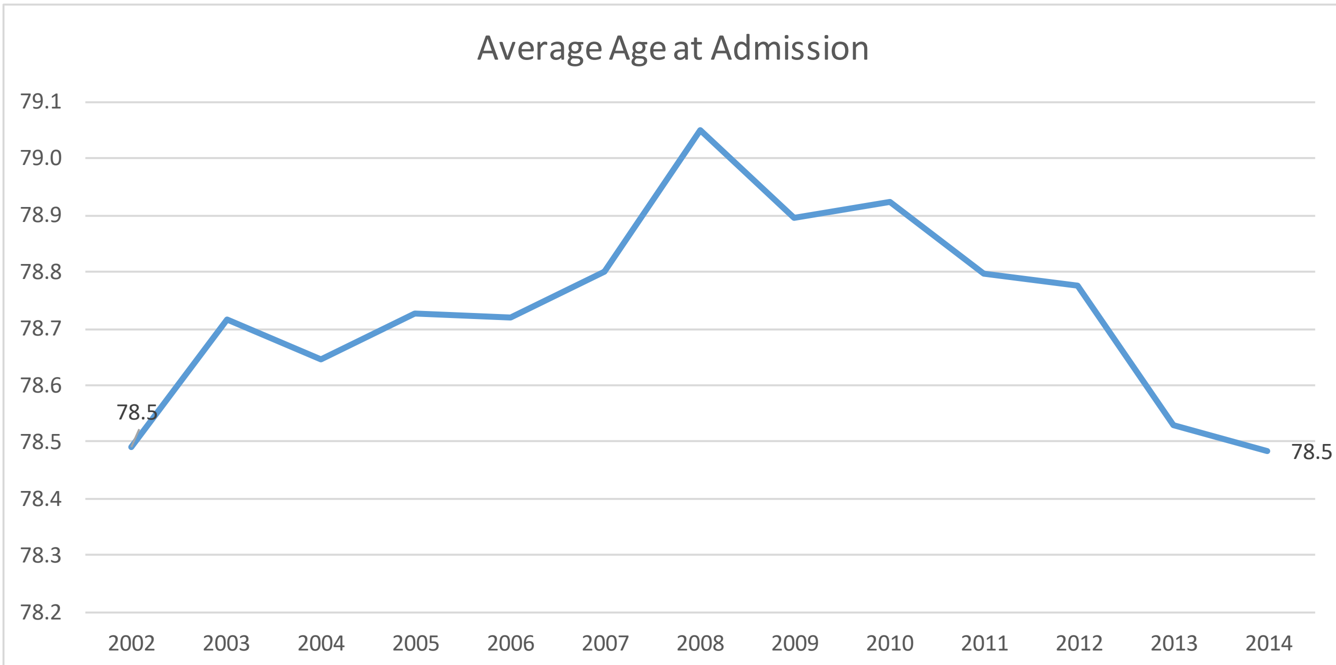
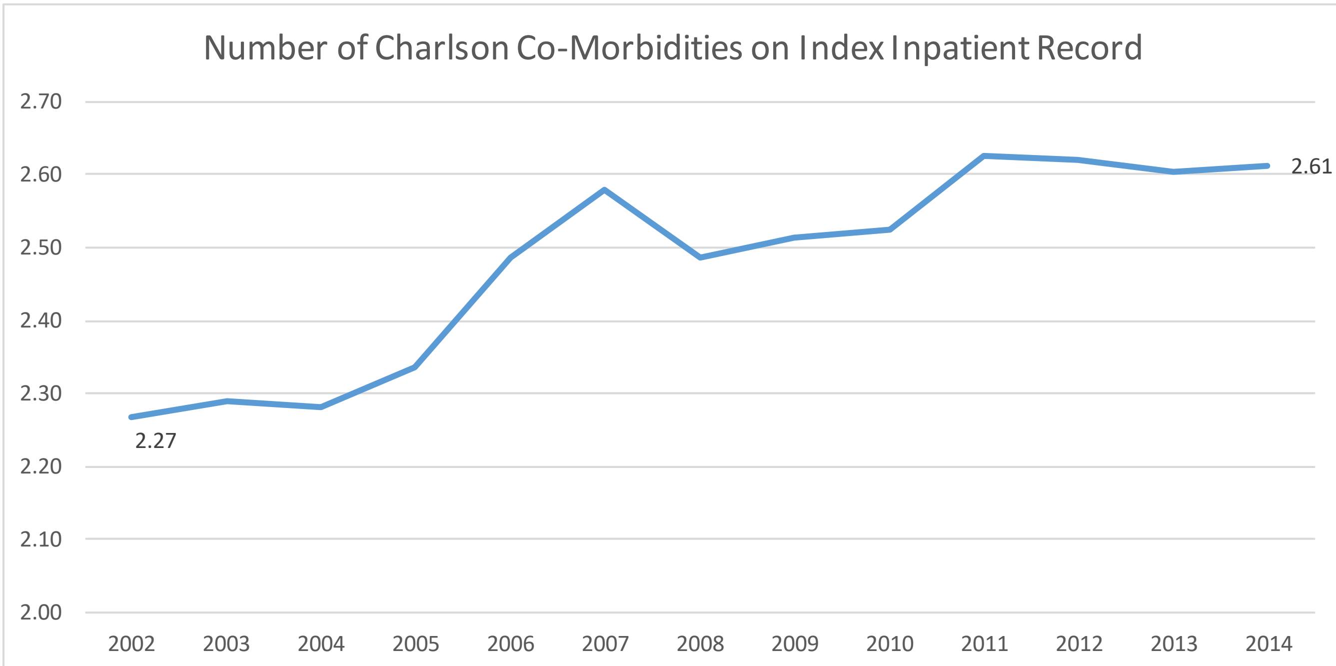

Average Cost of Episode (000s of 2014 Dollars)
$70
$60
$50
$40
$30
$20
$10
$0
61.9
35.9
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

Figure 7



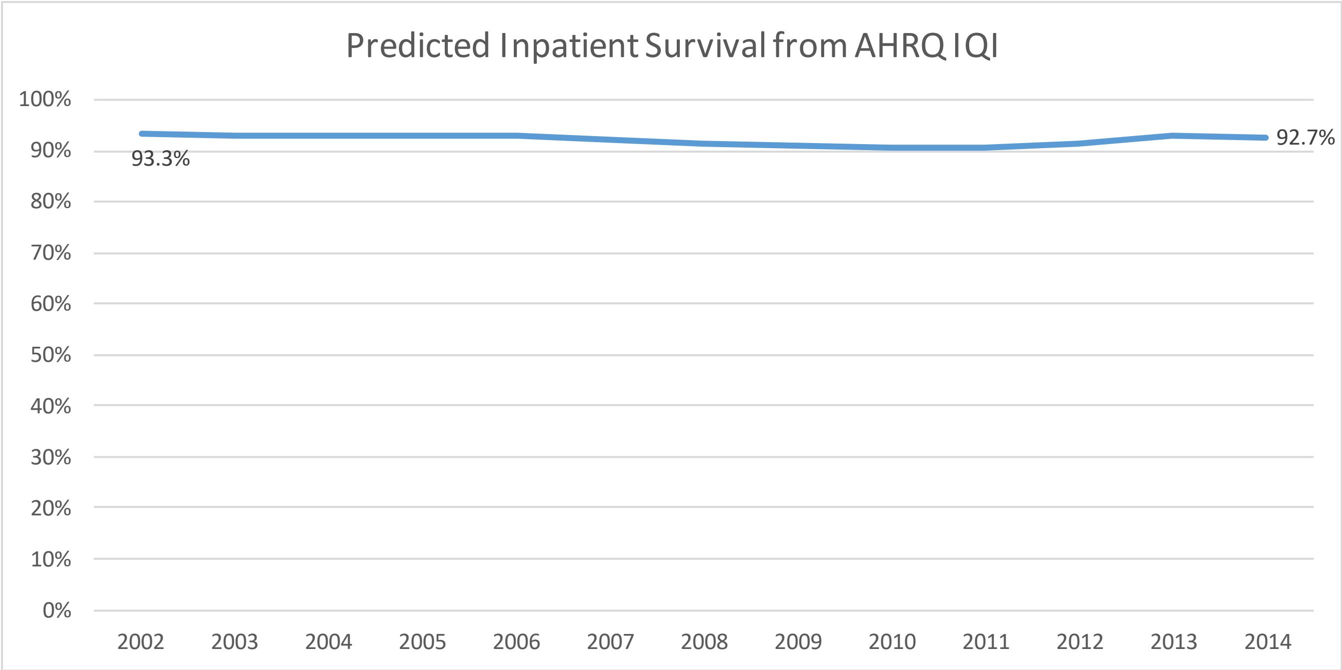
Average Age at Admission
79.1
79.0
78.9
78.8
78.7
78.6
78.5
78.4
78.3
78.2
78.5
78.5
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

Figure 8



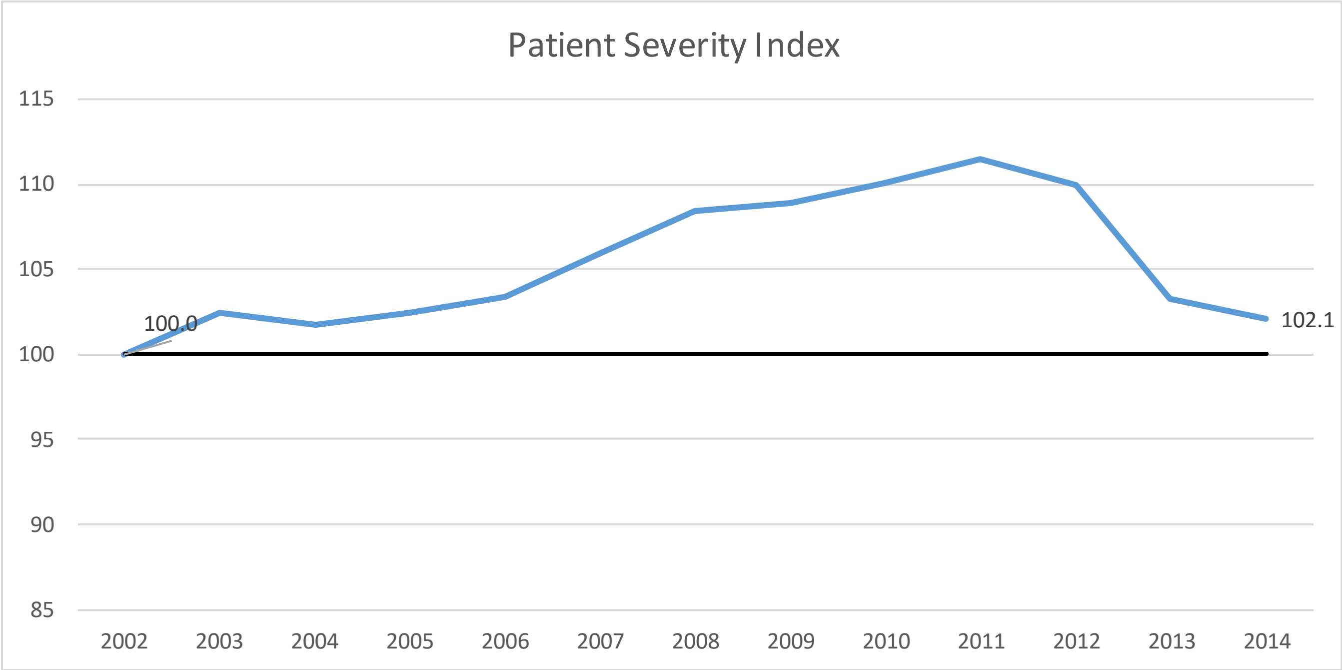
Number of Charlson Co-Morbidities on Index Inpatient Record
2.70
2.60
2.50
2.40
2.30
2.20
2.10
2.00
2.27
2.61
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

Figure 9



Predicted Inpatient Survival from AHRQ IQI
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
93.3%
92.7%
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

Figure 10



Patient Severity Index
115
110
105
100
95
90
85
100.0
102.1
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

Figure 11

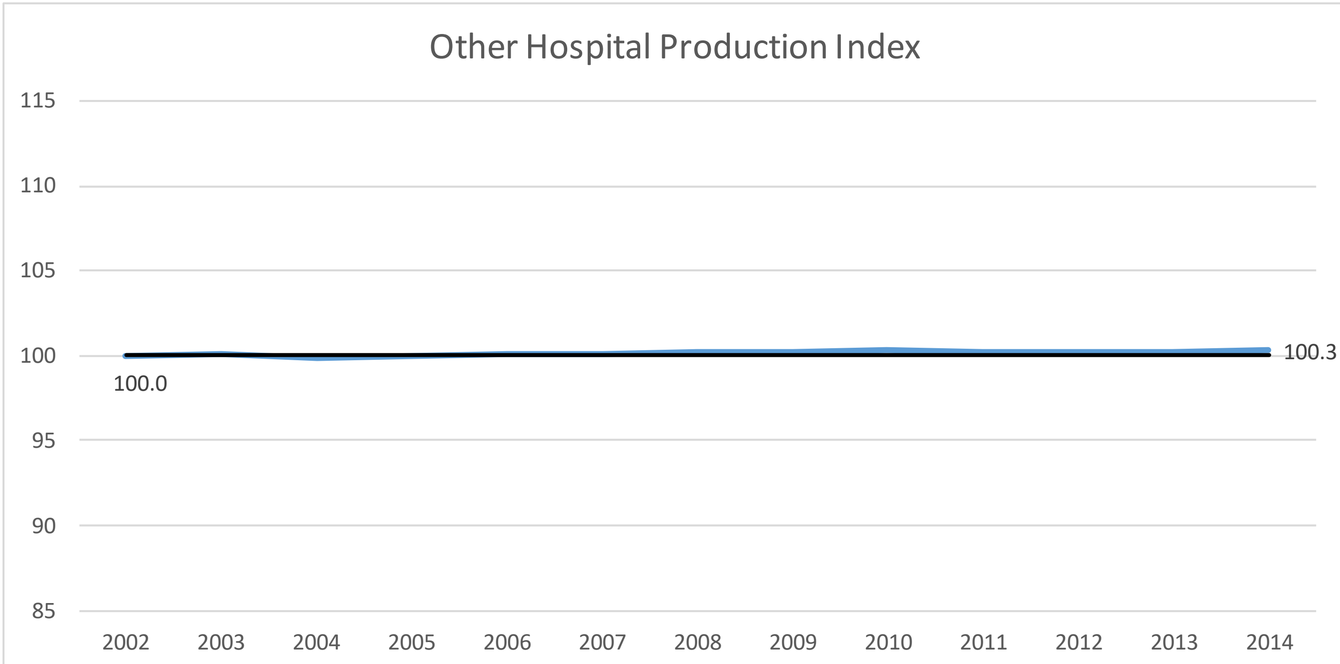

Other Hospital Production Index
115
110
105
100
95
90
85
100.0
100.3
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

Figure 12

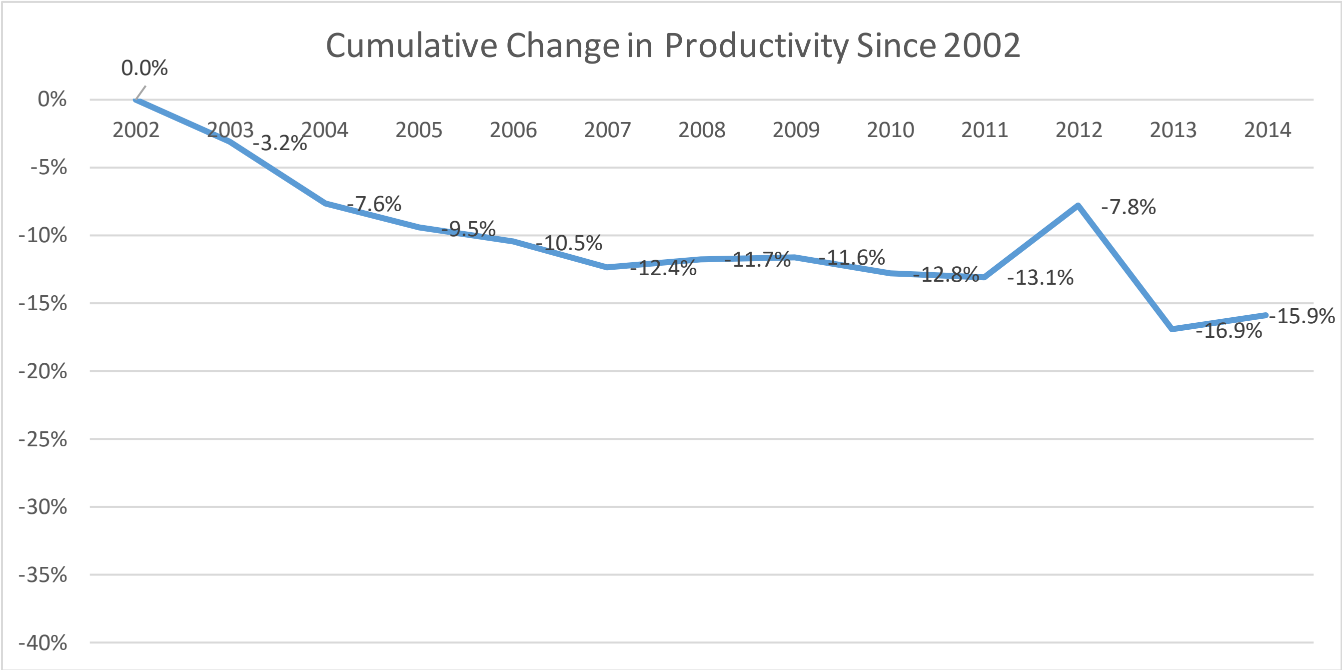

Cumulative Change in Productivity Since 2002
0%
-5%
-10%
-15%
-20%
-25%
-30%
-35%
-40%
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
0.0%
-3.2%
-7.6%
-9.5%
-10.5%
-12.4%
-11.7%
-11.6%
-12.8%
-13.1%
-7.8%
-16.9%
-15.9%

Figure 13

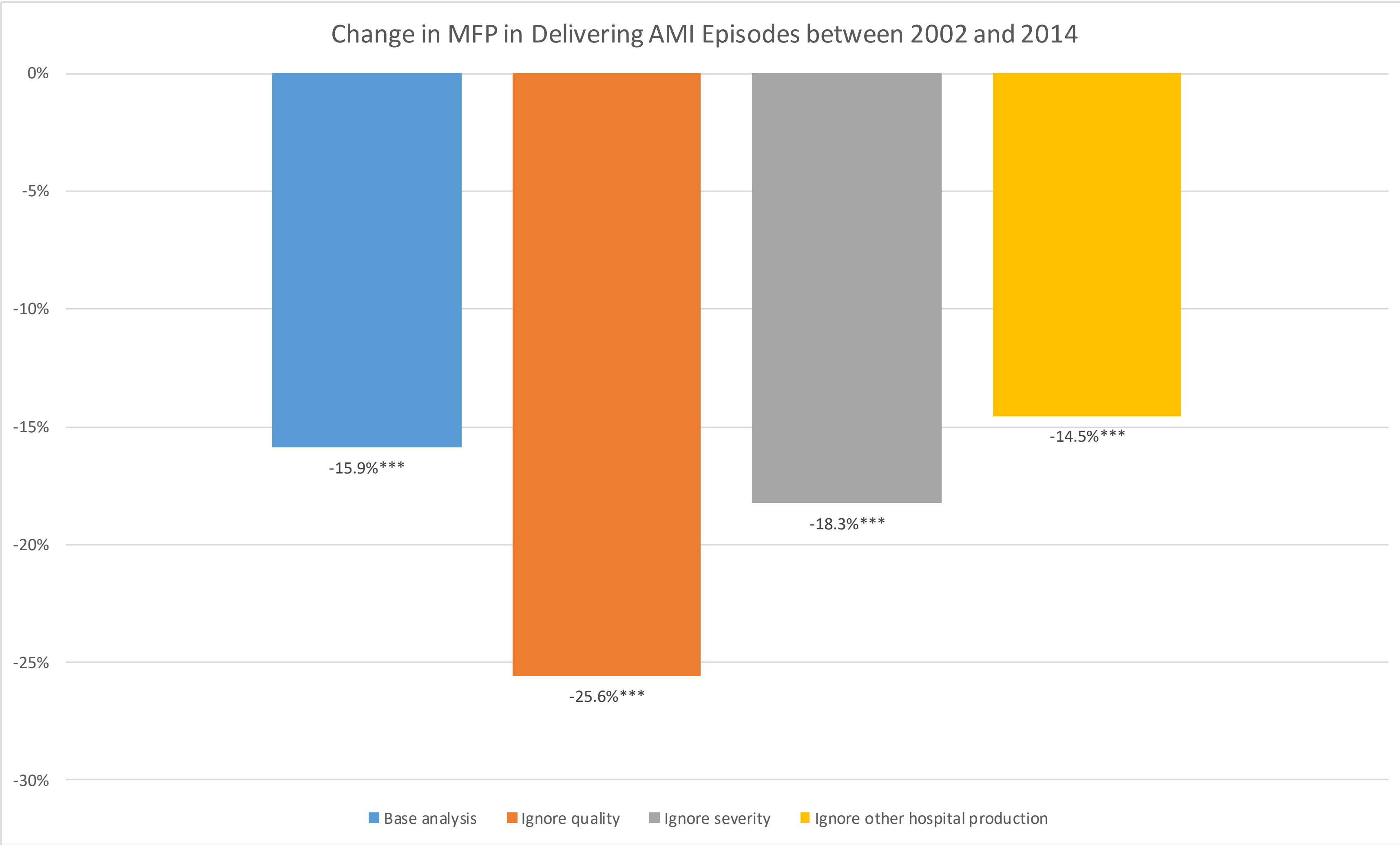

Change in MFP in Delivering AMI Episodes between 2002 and 2014
0%
-5%
-10%
-15%
-20%
-25%
-30%
-15.9%***
-25.6%***
-18.3%***
-14.5%***
Base analysis
Ignore quality
Ignore severity
Ignore other hospital production

Figure 14

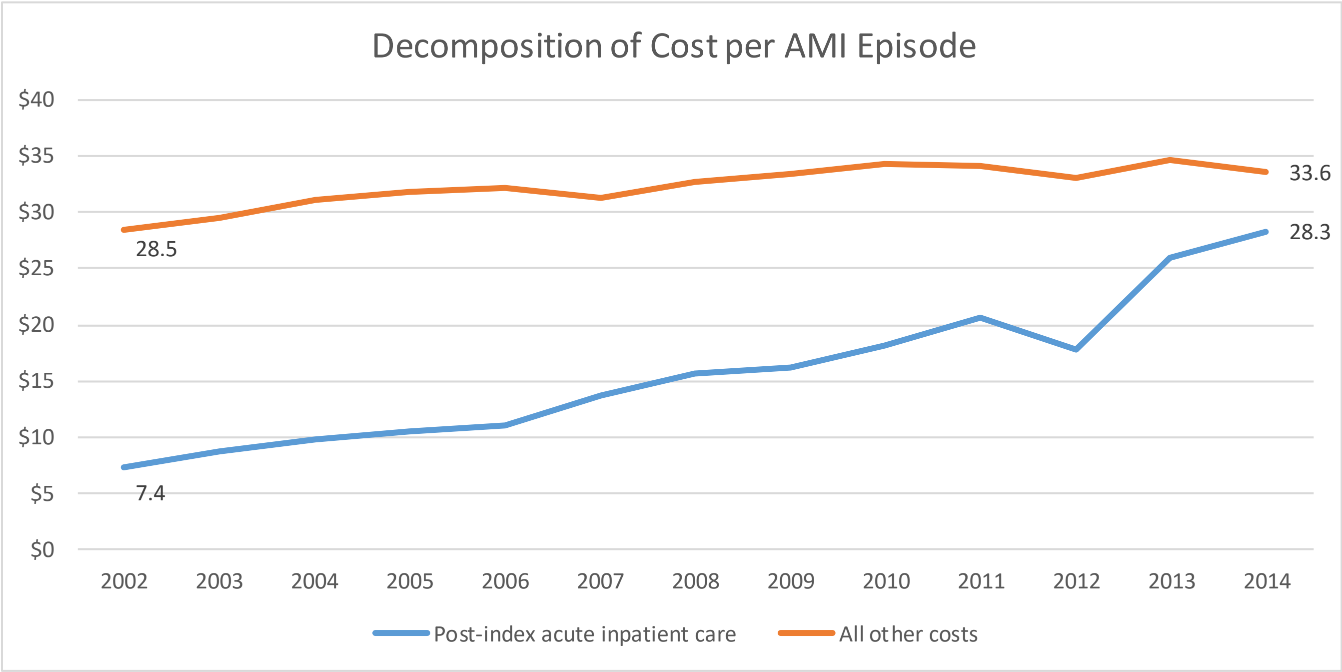

Decomposition of Cost per AMI Episode
$40
$35
$30
$25
$20
$15
$10
$5
$0
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
28.5
33.6
28.3
7.4
Post-index acute inpatient care
All other costs

Figure 15

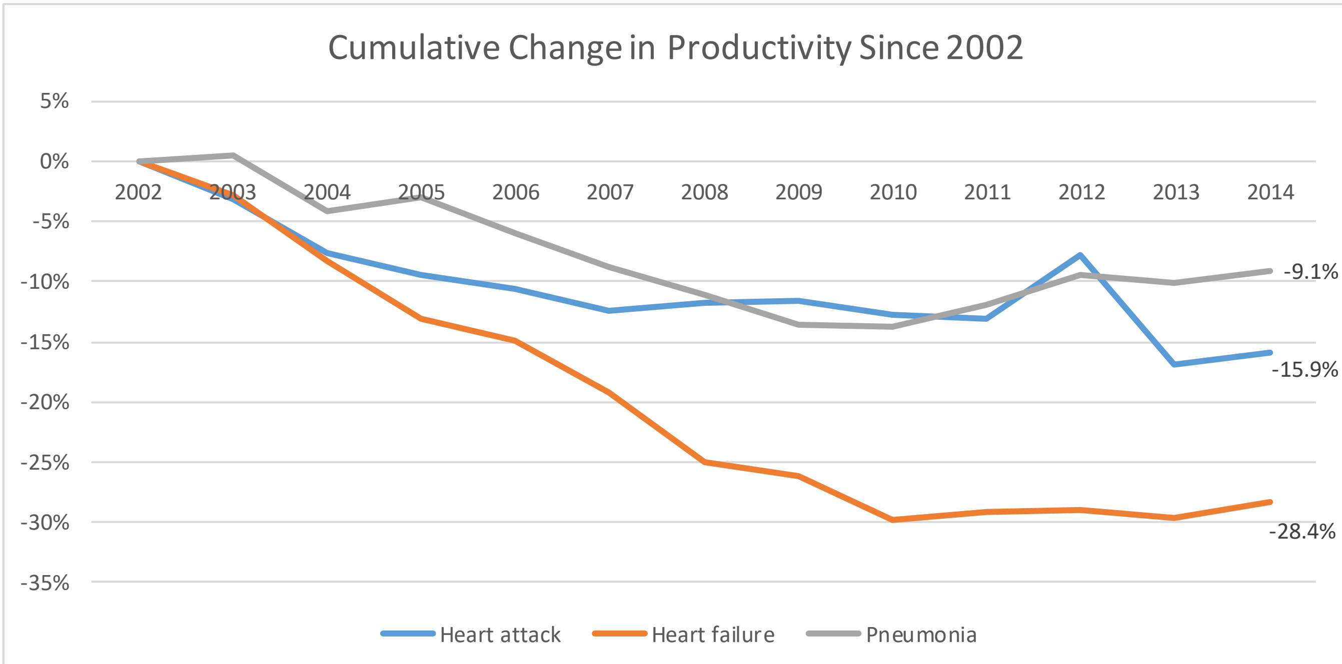
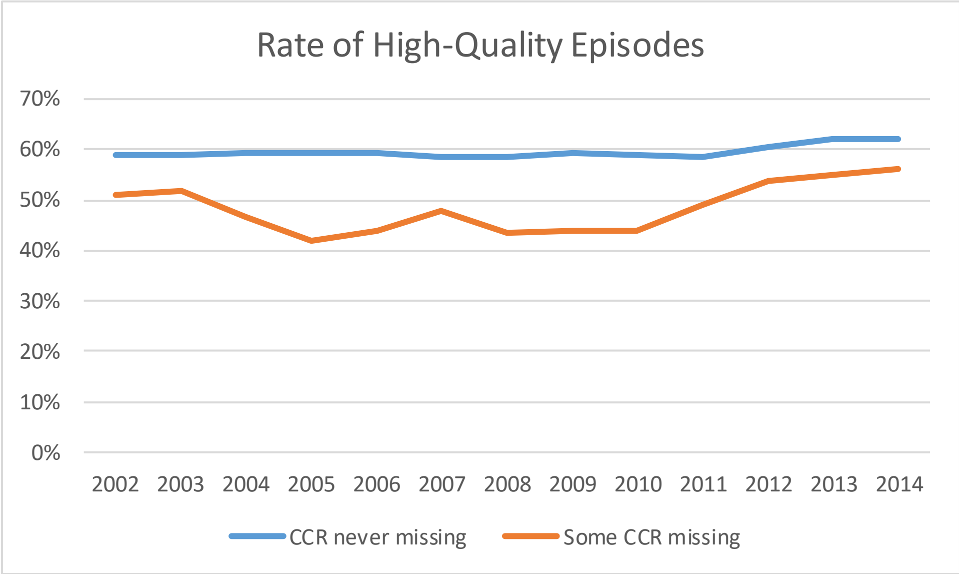

Cumulative Change in Productivity Since 2002
5%
0%
-5%
-10%
-15%
-20%
-25%
-30%
-35%
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
-9.1%
-15.9%
-28.4%
Heart attack
Heart failure
Pneumonia

Figure 16


Rate of High-Quality Episodes
70%
60%
50%
40%
30%
20%
10%
0%
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
CCR never missing
Some CCR missing