Measurement of Overuse of Healthcare Resources: A White Paper for the Maryland Health Services Cost Review Commission (HSCRC)

Jodi Segal, Najlla Nassery, Hsien-Yen Chang, Eva Chang, Kitty Chan, John Bridges

Abstract

Background: Healthcare quality is frequently described with composite measures representing the overall performance. Despite growing attention to overuse in healthcare, there is little experience with composite measures of overuse. Overuse can be defined as use of a service in the absence of a clear medical basis for its use, or when the risk of harm exceeds its likely benefit.

Objective: Our goal was to create a composite measure of overuse useable with claims data. We therefore aimed to identify a set of potentially overused medical procedures (indicators), operationalize these to be measureable with administrative claims, aggregate these indicators into a single indicator of overuse, and test whether the index is associated with higher costs and worse clinical outcomes, which would be the proof of principle that the measure is detecting regional overuse.

Design: Observational study using 5% of Medicare claims from 2008 (Parts A and B)

Setting: Older patients in the U.S. receiving healthcare services in hospitals or outpatient settings

Methods: Via an environmental scan, we identified published descriptions of overused procedures. We assessed each procedure's feasibility for measurement with claims data and developed algorithms for occurrences of procedures in patients unlikely to benefit. We calculated summary statistics to illustrate variance in use across hospital referral regions (HRRs) and selected 20 for inclusion in the Johns Hopkins Index of Overuse (JHOI). We started with an initial analysis of each of the 20 indicators (denote by P). As such, we regressed P_{ijk} as $P_{ijk} = \beta X_i + \Phi_{jk} + \varepsilon_{ijk}$ where X_i is a vector of patient specific factors, Φ_{jk} is a set of regional fixed effects for j procedures across the k HRRs. From this model, the Φ_{jk} is a risk-adjusted measure indicating if region k overuses procedure j more or less than expected. The $JHOI_k$ was estimated as an average of the Φ_{jk}. Thus, in this version of the JHOI each indictor contributes equally to the index. With Spearman's correlation, we assessed the correlation of this regional measure, presumptively a measure of systematic overuse, with regional measures of total costs, risk-adjusted mortality, 30-day mortality, and total mortality.

Results: 613 procedures were identified as overused. 20 had abundant frequency and variance to be possible measures of systematic overuse, including 13 diagnostic tests, 2 tests for screening, 1 for monitoring, and 4 therapeutic procedures. Usage varied markedly across HRRs. Among 1,451,142 beneficiaries, 14% had at least one overuse event. The index was positively correlated with 30-day inpatient mortality (r=0.27, p=<0.0001), and positively correlated with total costs (r=0.39, p<0.0001). It was similarly positively correlated with inpatient days (0.22, p<0.0001) and intensive care unit days (r=0.32, p<0.0001), but not with total mortality.

Conclusions: We identified a set of overused procedures that demonstrate significant variance in their usage. This study provides proof of principle that systematic overuse exists and is measurable. We need to identify the best combination of indicators for the index, which may require inclusion of procedures prevalent in younger people, and to validate the index when applied to panel data.

Table of Contents

I. Background

The Health Services Cost Review Commission (HSCRC) had developed a new all-payer model that will control costs on a per capita basis, for both inpatient and outpatient hospital costs, while requiring important care and health improvements. In order to assure successful implementation of this modernized all-payer system, the HSCRC has established Work Groups to provide recommendations on technical implementation issues and has requested white papers to inform the Work Groups.

This white paper addresses *Potentially Avoidable Volume.* The HSCRC seeks input on what types of services could be considered potentially avoidable and what types of adjustments may be required. To this end, we describe here our recent work on overuse of healthcare resources including the novel methodologies that we have developed to identify regions that may be overusing resources. While our work to date has measured overuse at the level of the Hospital Referral Region (HRR), we anticipate that these methodologies may be appropriate at a health system or even hospital level, although this remains to be tested.

In this paper we describe the conceptual framework for our work on overuse including our working definitions, and describe, briefly, the results of our literature review about published measures of overuse. We also describe the novel method that we developed to create an index of overuse that uses bellwether procedures (or indicators). We describe the method for identifying and operationalizing these indicators and how we aggregated them to create the Johns Hopkins Overuse Index (JHOI). We demonstrate that the JHOI is correlated with costs and is not correlated with measures of benefit at the level of the HRR. We submit that this supports that there is systematic overuse of resources in the regions with high values on the JHOI.

II. Conceptual Work

Healthcare expenditures are exceptionally high in the United States [1,2], yet health outcomes lag behind other developed nations.[3-5] This disparity has led many to assume that healthcare services are overused,[6-12] a phenomenon impacted by a "perfect storm" of patient, provider, and institutional determinants.[13] This unsustainable level of spending was a driving force for the passage of the Patient Protection and Affordable Care Act of 2010. [Baker, 2008] The overuse of healthcare can cause physical and financial harm to patients,[8,14-17] and have system-wide consequences.[18] As such, the reduction of overuse would contribute to achieving the "triple aim" across the healthcare system,[19] as it provides a mechanism to impact the cost and quality of the healthcare system.[20,21]

While many efforts have been targeted at promoting appropriate use of healthcare services,[22] these efforts have focused historically on reducing underuse rather than on overuse reduction.[23-25] Furthermore, many efforts targeted on the overuse of procedures have been clouded in controversy.[25-27] Few overuse indicators are routinely measured,[28] but the identification of diverse measures of overuse have been the focus of several recent initiatives [8,11,29,30] The effect of these broad initiatives, and recent structural reforms on the health system, will be difficult to measure unless

we can define overuse as a system-wide phenomenon and develop methods to measure it broadly across a health system.

While overuse has been defined by use of practices in an *absence of evidence*,[31] it may be best thought of as the provision of care *in circumstances where the potential for harm exceeds the potential for benefit*.[32] Overuse research parallels work on the wasteful [33,34], inappropriate [35,36] or inefficient [37,38] use of health resources.
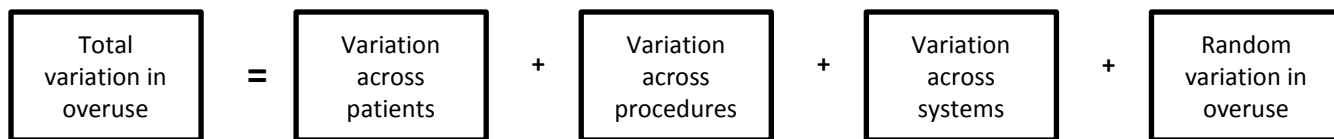
There are multiple barriers to studying overuse. First, investigating overuse requires identifying the patients for whom a procedure is inappropriate, and this requires assessment of a subjective tradeoff between benefits and harms. This subjectivity, coupled with uncertainty, implies that individual measures of overuse will be subject to noise. Second, research on overuse is often confused with work on geographic variation,[27,39] despite the wealth of empirical evidence demonstrating their differences.[40-43] Third, if overuse is systemic in the US, it might be difficult to see it varying across institutions.[44-45] Departing from these previous approaches, we suggest that it may prove more prudent to view overuse as a pervasive phenomenon within a healthcare system that affects the overuse of multiple procedures. In offering this perspective, we provide a novel way to define and measure *systematic overuse* across a health system.

Defining and Measuring Systematic Overuse
Although systematic overuse may not be directly observable, we believe that it can be identified via its consequences. Specifically, we propose that pervasive overuse would impact a range of procedures, either in specific clinical areas or across the entire system. If this is the case, then it should be measurable with a portfolio of procedures. Such "bellwether" procedures may be relatively insignificant (i.e., individually they may not be costly or subject patients to harm), but they can be seen as symptoms of a more serious and widespread problem. By combining multiple procedures, not only would the measure become more generalizable, but it would also be less subject to measurement error. Similar to a stock-market portfolio, grouping multiple indicators of overuse – each of which will be measured with error – would decrease overall volatility of the measure. Developing a portfolio of potentially overused procedures has been aided by a growing literature on overused procedures (albeit with varying definitions of what constitutes overuse).[7-12,46]

Given that certain errors may persist in such an aggregate measure, it may be necessary to statistically control for differences across the procedures and differences in case mix across the health systems being examined. As seen in Figure 1, we decompose the total variation in the overuse of the procedures into four specific types of variation: patient, procedure, health system and random. Multilevel statistical models can isolate the health system-specific component, hence minimizing the error and potential bias in the measure of systematic overuse at the health system level.

Figure 1: Source of the total variation in overuse

Total variation in overuse	=	Variation across patients	+	Variation across procedures	+	Variation across systems	+	Random variation in overuse

If this portfolio adequately reflects systematic overuse, then two consequences should be detectable – these procedures should be associated with higher expenditures and poorer health outcomes. Use of some of procedures, individually, may be associated with net cost savings; however, the procedures collectively should result in a net cost to the health system if we are proposing that the portfolio indicates systematic overuse (i.e., the direct costs of these procedures should not be offset by cost savings elsewhere in the system). Likewise, systematic *overuse* of services should not be associated with health benefits (e.g., improvement in health status, fewer adverse events, or greater life expectancy).

   While identifying and measuring systematic overuse is difficult, it should prove valuable. In contrast to approaches that have been tailored to specific procedures, conditions or clinical specialties, a global measure is more likely to inform health systems and policy makers of the serious *structural or cultural problems* inherent across entire health systems. Furthermore, policies to curtail *systematic* overuse are likely to have the broadest benefits in terms of costs, quality and outcomes.

III. Literature Review

In approaching our literature review about measures of overuse, we recognized that health service overuse can be broadly described by the following categories: (1) inappropriate for a specified clinical indication, (2) inappropriate for a clinical indication in a specific population, (3) excessive service intensity or sophistication given expected clinical benefit, and (4) excessive frequency of service given expected clinical benefit. These distinctions are important since overuse is often determined by the clinical profile (efficacy and effectiveness, benefits, risks and harms) associated with the service given a particular clinical context or scenario. For example, a service can be appropriate for high risk patients, but when given to low risk patients who are not likely to benefit, this would constitute overuse.

   The expected risks and harms to the patient is an important aspect for classifying overuse. Using advanced imaging or costly new medications can be appropriate in the absence of less expensive but equally effective alternatives. However, when less expensive alternatives are available, use of more expensive services or treatments with marginal clinical benefits can be considered overuse. Similarly, while monitoring clinical progress is an important part of good patient care, repeating tests too frequently when the probability of observing clinically important change is low can also be considered overuse.

   These categories are broadly consistent with the types of overuse identified at the Agency for Healthcare Quality and Research (AHRQ) 2009 conference on *Developing a Framework and Research Agenda for Overuse and Appropriateness Measures* including clinically harmful overuse, overuse that is

not cost-effective and overuse of appropriate care. (47) Our typology extended the AHRQ framework by incorporating clinical indication, patient population, service characteristics and temporality that more specifically describes the nature of the overuse in order to facilitate future measure development and operationalization.

We focused our environmental scan, first, on major repositories of quality of care measures. [47-61] These included the National Quality Measures Clearinghouse (NQMC), National Quality Form (NQF) Endorsed® Standards, and current and past measures under Centers for Medicare and Medicaid Services' (CMS) Physician Quality Reporting System (PQRS) program. We also searched for measures developed by organizations such as the American Medical Association's Physician Consortium for Practice Improvement (PCPI), and National Committee for Quality Assurance (NCQA). We searched using Google and Google Scholar to identify other measure sources and initiatives, including a listing of "widely adopted interventions found to be harmful or ineffective" from the Institute of Medicine and measure opportunities from the National Priority Partnership and Blue Cross Blue Shield of Massachusetts under its Alternative Quality Contract. Measure opportunities were identified from efforts such as the American College of Physicians' High Value, Cost-Conscious Care Initiative and Choosing Wisely in which medical societies partnered with the American Board of Internal Medicine (ABIM) to develop evidence-based lists of the top five most commonly used tests or procedures within their respective specialties whose use should be questioned or discussed. Finally, we complemented our search with a review of Pubmed, the ISI Web of Science and reference lists from relevant published literature.

In describing the state of the field in overuse measurement, we aimed to identify unique concepts of health service overuse. Measurement concepts, namely, general descriptions of an overused service that are not well defined in terms of clinical context, are defined in this work as "measurement opportunities."

Measures

We identified 160 measures of health service overuse; only 37 of these were measures that were completely specified, meaning that they provided definitions of denominators, numerators, and exclusions. Most overuse measures target situations where a service was directed for a specific indication or population in which the service was considered inappropriate or of low clinical value ($n = $ 110). Imaging services linked to overuse represented the largest service category within both the fully specified measures ($n = 18$, 49%) and those that we consider measurement opportunities ($n = 34$, 28%). (**Table 1**)

A number of measurement concepts were identified by more than one organization. However, we observed differences in specification or lack of specificity in these concepts. For example, a number of organizations developed a measure about imaging for low back pain. A review of these measures revealed differences in the definition of the population, the time interval

relevant for determining a case of overuse, and the specificity of the imaging test being evaluated and different inclusion/exclusion criteria.

Table 1. Number of Specified Measures and Measurement Opportunities on Overuse: Identified by Type of Overuse and Clinical Service.

	Total	Inappropriate indication	Inappropriate population	Unnecessary service intensity	Unnecessary service frequency
Screening	43	14	14	2	13
Specified	2	1	1	0	0
Opportunities	41	13	13	2	13
Diagnostic	58	24	4	17	13
Specified	16	5	4	6	1
Opportunities	42	19	0	11	12
Monitoring	34	9	0	5	20
Specified	6	1	0	0	5
Opportunities	28	8	0	5	15
Therapeutic	57	41	4	12	0
Specified	14	11	0	3	0
Opportunities	43	30	4	9	0
Overall	192[a]	88	22	36	46
Specified	38	18	5	9	6
Opportunities	154	70	17	27	40

a. Specified measures ($n = 37$); Measurement opportunities ($n = 123$); Total measures ($n = 160$); Measures may be included in more than one category (e.g., screening and therapeutic; screening and monitoring). Therefore, the sum across categories ($N = 192$) is higher than the total number of measures ($N = 160$).

Data demands for measuring health care overuse can be significant. While a number of measures can be implemented using only administrative claims data, other measures require clinical or medication information that can be challenging to obtain, including data necessary to determine "low-risk groups." Fortunately, the growing availability of data from electronic health records and other health information technology will create opportunities for implementing measures that require detailed clinical information.

IV. Identifying Indicators
Building upon the literature review described above, we sought to identify a set of procedures that we could operationalize using claims data to identify episodes of overuse. Our expectation was that this

set of procedures might then be aggregated into an index for measurement of overuse, although other uses are encouraged.

We conceptualized this project as analogous to quality of care measurement. In the field of quality measurement, indicators are not meant to assess the care given to a particular patient or even by an individual physician. Rather, they are commonly used to identify healthcare systems, defined by geography or organization, in which resources may be inadequate for good outcomes.[62] As in the quality field, we aimed to identify procedures that may be useful indicators, or tracers, of overuse in healthcare systems or indicators of regions requiring attention.[63-65] We propose that these indicators *in aggregate* may reveal systematic patterns of overuse within regions or reveal changes over time. Alternatively, these indicators may be valuable individually in quality improvement efforts.

Identifying Candidate Procedures

As described above, we reviewed publications describing medical procedures considered to be overused by these organizations. We extracted lists of procedures deemed as overused, which might be indicator procedures or bellwethers of a systematic pattern of overuse within a region. We did not review the process by which these procedures were designated as overused; it was sufficient that an organization categorized them as such. We classified the procedures as to whether they could be operationalized for measurement with only administrative claims data. For example, we did not further operationalize overused procedures that would require knowledge of symptomatology that is not available in claims data. We also excluded measures requiring information about a pharmacy dispensed medication. Finally, given that we used Medicare data, the procedure needed to be relevant to the care of older people.

Data

We acquired a 5-percent sample of fee-for-service patients insured by Medicare in 2008. We used data from the MedPar, Carrier, and Outpatient files. We required that all included individuals had 12 months of complete enrollment in Medicare Part A and Medicare Part B or death during 2008, were over the age of 65 years, and were never enrolled in a Medicare Health Maintenance Organization (HMO) during that year. With an appropriate cross walk, we linked zip codes in the Medicare data to HRRs as defined by Dartmouth Atlas. [66]

Generating Algorithms to Operationalize the Indicators in Medicare Claims

For each indicator procedure, we defined a set of individuals in whom use of a particular procedure is likely to be an overuse event given that it is not clinically indicated. We then identified use of the procedure in those individuals. Depending on the procedure, the set of at risk individuals might have been defined by age, gender, diagnoses or receipt of other services. For most of the indicator procedures that we operationalized, the set of individuals had a given diagnosis (such as sinusitis or low back pain). We used existing algorithms when possible. For some of the indicator procedures, an individual could contribute more than once if he/she could have been subjected to the intervention

multiple times during the year. One clinician (J.S.) generated all of the algorithms which were then reviewed by a second clinician (N.N). Coding software was used to identify relevant codes. (FlashCode, Medical Coding & Compliance Solutions, LLC (MCCS); Turlock, CA) (Appendix 1)

Describing Regional Use

For each indicator, we counted the occurrences of the procedure or test in the HRR among the population at risk in the HRR. We report this as a count per 1000 Medicare beneficiaries. We summarized the scores across the HRRs for each indicator by calculating a mean, median, variance, and interquartile ranges, and visually examined the plots of these as well as higher order moments. We estimated the costs that might be attributable to occurrences of these indicators. (Appendix 2)

Reducing Indicators

We discarded those indicators with little variance (which typically occurred when most HRRs had zero events) as being less informative to an index and of less value in quality improvement interventions. We also discarded indicators with very low counts of events across the HRRs combined (typically <800 events). Finally, we segregated the three indicators that describe the use of cancer screening tests (PAP test, prostate specific antigen testing, colonoscopy) as these would be overly-influential in an index given their high prevalence.

Indicator Procedures

The measures and overuse procedures were generated by a diverse group of professional societies, quality improvement organizations, and researchers as described above but broadened to include guidelines from professional societies as well.

We identified 613 indicators of overuse. The procedures spanned clinical categories and included radiological procedures, cardiac imaging, invasive diagnostic procedures, and therapeutics. Ninety-four of the 146 procedures relevant to older people required knowledge of symptoms or medications and were therefore excluded. Others were deemed too rare to develop (e.g. hyperbaric oxygen for multiple sclerosis) and others were thought to be more correctly classified as utilization measures (rather than overuse measures) and were set aside for use in validation analyses.

We identified 52 procedures that could be specified with claims. After calculating their use frequencies in the HRRs and the distributions of those frequencies, we discarded an additional 32. These 32 were either exceptionally frequent (e.g. PAP smear over age 65), invariant, or where was was very little use in most HRRs (e.g. MRI in patients with lung cancer, and screening for abdominal aortic aneurysms in women). The retained 20 procedures included 13 diagnostic tests, 2 tests used for screening, 1 for monitoring, and 4 therapeutic procedures. The clinical areas included 4 relevant to primary care practice, 3 relevant to otolaryngology, 3 to radiology, 2 to cardiology, and one each to neurology, emergency practice, allergy, oncology, end of life care, urology, physical therapy, and surgery.

Use of the Indictor Procedures

The 2008 Medicare sample included information about 551,028 men and 900,114 women after removal of 8,391 individuals with zip codes which could not be matched to an HRR. Eighty-eight percent of the enrollees were white, 7.3% were Black, and the remainder was of other races. Twenty-one percent were between 65 and 69, 23% were between 70 and 74, 20% were between 75 and 79, 17% were between 80 and 84, 12 % were between 85 and 89, and 7% were over 90. As required, beneficiaries had a full 12 months of coverage in 2008 unless they died that year (5.4%). Thirty-nine percent resided in the South, 19% in the Northeast, 25% in the Midwest, and 17% in the West.

By design, the included indicators had rates of use that varied across HRRs. (Table and Figure 2) The absolute counts of events varied; there were fairly infrequently used procedures such as MRI in mild traumatic brain injury (mean 2.6 usages per 1000 beneficiaries in the HRR), to highly prevalent procedures such as MRI for low back pain evaluation (mean of 395 usages per 1000 beneficiaries in the HRR). Some of the usage followed a near normal distribution across the HRRs, such as measurement of digoxin levels in patients with congestive heart failure and emergency department usage at end of life. Most had substantial right skewing with long, thin right tails of the distribution which were the HRRs which were exceptional outliers in their usage. An example of this is the use of sinus CT for uncomplicated rhinosinusitis, where the highest using HRR uses this in 80 of 1000 beneficiaries where the mean usage is 14 per 1000. Similarly, for hysterectomy for management of benign disease, the highest HRR used the procedure in 42 of 1000 beneficiaries where the mean across HRRs is 2.6 per 1000.

At an individual level, overuse was prevalent. Among 1,451,142 Medicare beneficiaries, 14% had experienced at least one overuse event, and this rate varied substantially across the HRRs (8.4% to 27%). Individuals were subjected to between zero and seven procedures in 2008 that can be considered overused procedures.

Table 2. Mean and Median Counts of Indicators of Overuse Across Hospital Referral Regions (N=306 regions)

Potentially Overused Procedure	Mean	Median	Interquartile Range		References
	Per 1000	Per 1000	Per 1000		
Stress echocardiography in symptomatic or ischemic equivalent acute chest pain *	33	22.7	2.7	45.5	ACCF/ASE/ACEP/AHA/ ASNC/SCAI/ SCCT/SCMR 2008 [67]
Abdomen CT use of contrast material*	222	187	133	288	QualityNet [61]
Thorax CT Use of Contrast Material*	64.9	47.5	26.8	79.7	QualityNet [61]

MRI Lumbar Spine for Low Back Pain*	395	395	356	441	QualityNet [61]
Sinus CT or antibiotics for uncomplicated acute rhinosinusitis*	14	12.4	6.9	19.1	Choosing Wisely [49]
Diagnostic tests, such as immunoglobulin testing, in the evaluation of allergy*	4.5	3.7	1.7	5.8	Choosing Wisely [49]
Follow-up tumor marker studies In asymptomatic women with previously treated breast cancer	732	672	464	984	Qaseem A [69]
Preoperative chest radiography in the absence of a clinical suspicion for intrathoracic pathology*	219	213	168	259	Qaseem A [69]
Screening for asymptomatic carotid artery stenosis in the general adult population	12.5	11	8.4	14.9	U.S. Preventative Services Task Force [68]
Traction for low back pain	115	106	77.8	137	Institute of Medicine[54]
PET, CT, and radionuclide bone scan in individuals with low risk prostate cancer	20.7	19.4	9.3	28.4	Choosing Wisely [49]
Serological tests for helicobacter pylori	9.3	8.2	5.5	11.2	National Health Service, UK [57]
Appropriate Head CT Imaging in Adults with Mild Traumatic Brain Injury*	2.6	2.3	1.1	3.8	National Quality Forum [58]
EEG monitoring in individuals presenting with syncope*	22.8	21.2	13.3	29.4	National Health Service, UK [57]
Routine monitoring of digoxin in patients with congestive heart failure	3.6	3.5	2.8	4.2	National Health Service, UK [57]
More than one emergency department visita in last 30 days of life	146	147	121	168	National Quality Forum [58]
Nasal endoscopy for sinusitis diagnosis*	34.9	24.8	12.1	45.8	The Alternative Quality Contract[51]
Fiberoptic laryngoscopy for sinusitis diagnosis*	7	5.2	2.2	9.5	The Alternative Quality Contract [51]
Hysterectomy for benign disease	2.8	2.7	1.5	3.9	National Guidelines Clearinghouse [70]

Laminectomy or spinal fusion	2.2	1.9	1.2	2.9	National Guidelines Clearinghouse [70]

*count is per 1000 eligible episodes rather than per 1000 beneficiaries, ACS=acute coronary syndrome CAD=coronary artery disease, CT=computed tomography, EEG=electroencephalography, MRI=magnetic resonance imaging, PET=positron emission tomography

V. Creating the Index

The basic structure of our empirical model is similar to those used in the assessment of healthcare quality.[71-73] We started with an initial analysis of each of the 20 indicators within the index (denote by P). Here we studied the jth procedure, for the ith beneficiary in region k, as P_{ijk}, which is equal to 1 if there was overuse for that indicator procedure and 0 if otherwise. We only included patients who were eligible for the jth procedure in this analysis. This could be decomposed into that which was attributed to the patient, that which was attributed to the region and a residual.

As such we regressed P_{ijk} as

$$P_{ijk} = \beta X_i + \Phi_{jk} + \varepsilon_{ijk} \quad [1]$$

where X_i is a vector of patient specific factors, Φ_{jk} is a set of regional fixed effects for j procedures across the k regions. The patient specific factors included gender, age variables (age and age squared), race (white, black, Asian, Hispanic, native American and others) and a claims-based morbidity indicator derived from the Johns Hopkins Adjusted Clinical Group (ACG) system (count of ADGs: 0 - 32). The ACG system assigns all ICD-9-CM codes to one of 32 diagnostic clusters (ADGs) based on five clinical dimensions: duration, severity, diagnostic certainty, etiology, and specialty care involvement. Each ADG is a grouping of diagnosis codes similar in terms of severity and likelihood of persistence of the health condition treated over a relevant period of time, typically one year. ADGs are not mutually exclusive and individuals can have multiple ADGs (up to 32). There were 306 regional indicators representing 306 HRRs. From this model, the Φ_{jk} is considered as a risk-adjusted measure indicating if region k overuses procedure j more or less than expected. This was repeated for all 20 of the procedures. These estimated Φ_{jk} become the building blocks of the Johns Hopkins Overuse Index (JHOI) for region k, with the $JHOI_k$ estimated as an average of the Φ_{jk}. Thus, in this version of the JHOI each indictor contributes equally to the index.

Correlations with Costs

As anticipated, the index was positively correlated with total costs (r=0.39, p<0.0001). (**Figure 1**)
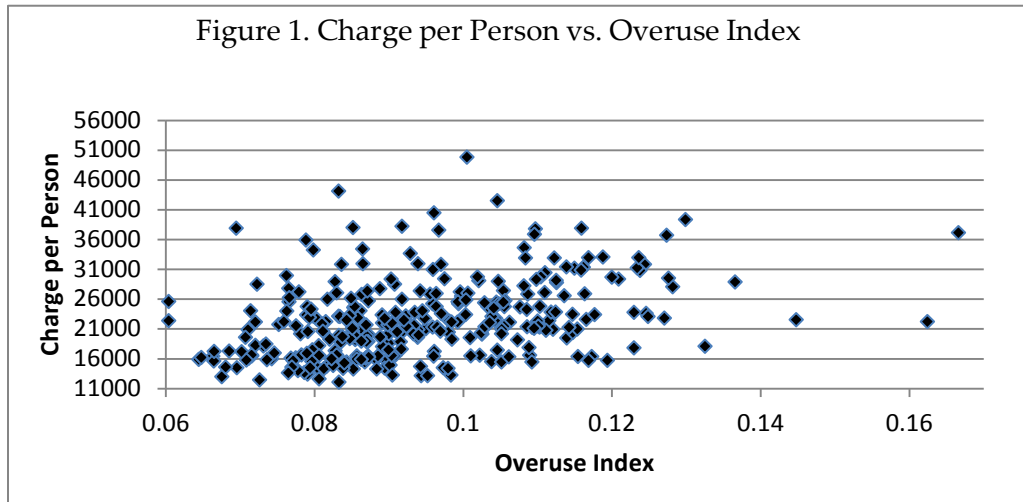
Figure 1. Charge per Person vs. Overuse Index

Legend: Overuse Index generated using Medicare Parts A and B, 2008 for each of 306 Health Referral Regions; charge is inpatient and outpatient charges

Correlation with Clinical Outcomes

Consistent with our hypotheses, the index was also positively correlated with 30-day mortality after a hospitalization (r=0.27, p=<0.0001). (Figure 2) It was similarly positively correlated with inpatient days (0.22, p<0.0001) and intensive care unit days (r=0.32, p<0.0001), although not with total mortality.
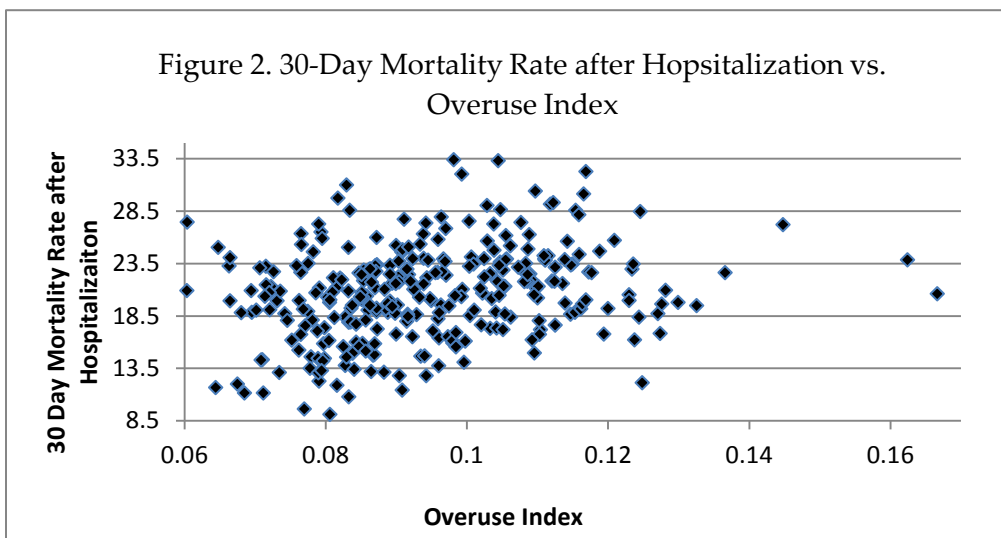
Figure 2. 30-Day Mortality Rate after Hopsitalization vs. Overuse Index

Legend: Overuse Index generated using Medicare Parts A and B, 2008 for each of 306 Health Referral Regions

Figure 3. Inpatient Days per 1,000 Persons vs. Overuse Index

Legend: Overuse Index generated using Medicare Parts A and B, 2008 for each of 306 Health Referral Regions

Figure 4. Intensive Care Unit Days per 1,000 Persons vs. Overuse Index

Legend: Overuse Index generated using Medicare Parts A and B, 2008 for each of 306 Health Referral Regions; ICU=intensive care unit

VI. Discussion
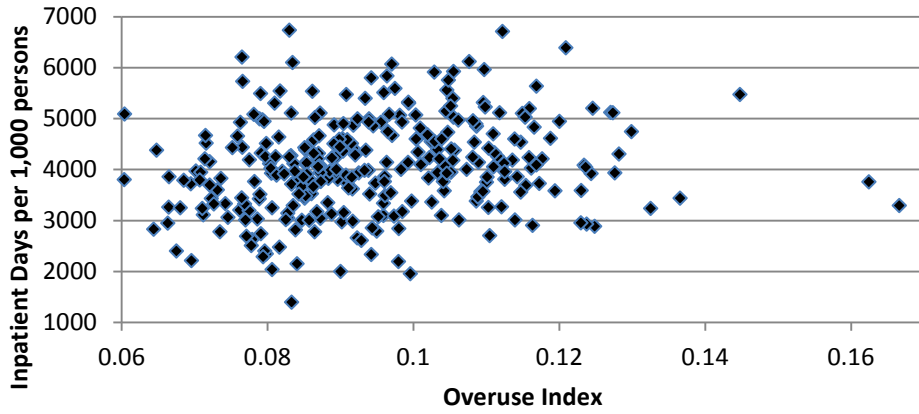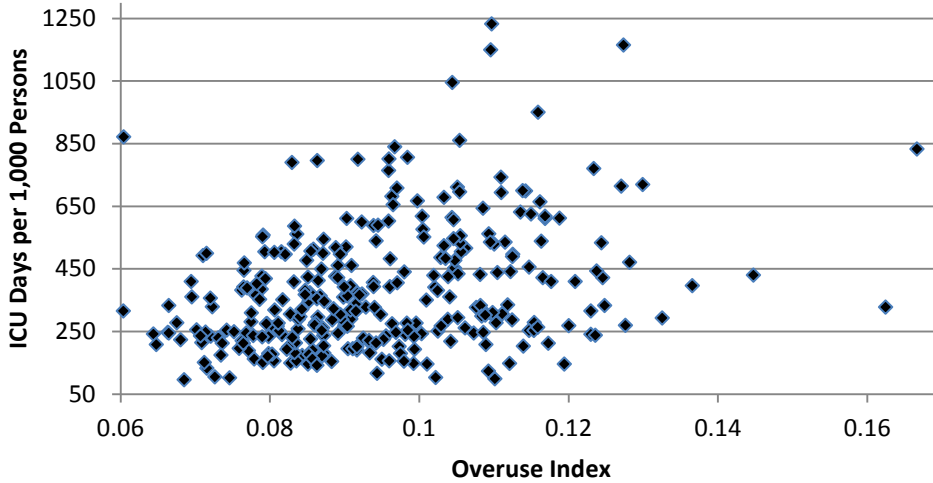
These studies demonstrate that usage of procedures and tests that may be considered overuse events can be identified within claims data. This may allow the creation of an index of overuse, as a parallel to quality of care indexes. Alternatively, these individual indicators may be valuable for measurement across regions or as targets for interventions.

We suggest that these overuse indicators identify resource use that is distinct from utilization. We anticipate that a region may be a low utilizer of health care resources but a high overuser, or the converse, yet this remains to be proven. Traditional measures of small area variation, or geographic variation, assess only variation in utilization *without attention to whether the variation is driven by underuse or overuse of services.* These 20 procedures are a useful initial set for development and testing of an index of overuse, as they are procedures across clinical specialties and include diagnostic as well as therapeutic procedures.

We anticipate that not all will agree with each of the indicators included. However, it is precisely because these indicators were identified by other bodies, through consensus processes, that we felt supported in considering them for inclusion in an index. We do not believe that these are "never procedures" –these procedures are sometimes indicated even in the populations that comprise the subpopulations of interest – the patients in whom these procedures are commonly overused. Determining the appropriateness of these procedures for individual cases was not intended. Assessment of overuse at a regional level or state level will allow for investigation into factors that drive health care overuse. As we proceed, we will attend to the recommendations in the recent IOM report that cautions against a focus on regional variation in spending,[74] although this does not necessarily mean that regional variation in overuse is to be overlooked. We were limited to including only procedures relevant to older people, given our data. These, however, may prove to be sufficient for identifying overusing regions.

The indicators might, in time, be expanded to include pharmacy claims, or expanded for use in electronic medical records data which would allow even a richer set of indicators for inclusion. Others may choose to use the specifications of individual indicators as measures for quality improvement interventions that target a reduction in use of specific procedures. If these are to be used this way, there may need to be additional refinement of the algorithms to improve their specificity.

We are optimistic that we are on our way to defining a new measure that may have very broad application in time. The ability to identify with a measure those regions or health systems that are overusing and therefore potentially harming individuals and populations, and exacerbating the national health care spending problem, is vital. Ultimately, this will allow further study of the determinants of overuse and the development of interventions to fix it for the benefit of patients.

VII. Recommended Next Steps for Utility to Health Services Cost Review Commission

How can measuring systematic overuse have an impact?

Although the term overuse has existed in a medical context for over 400 years, it remains an elusive concept. Moving the focus from single episodes of overuse to the broader concept of systematic overuse may be useful for researchers and policy makers and make health systems not only aware of this systemic problem, but responsible for addressing overuse across their entire organization.

While the proposed measurement of systematic overuse is an indirect approach (i.e., using a portfolio of potentially overuse procedures as a proxy for the underlying phenomenon), this approach has two significant benefits. First, it would make gaming the system more difficult for health systems (i.e., relative to the use of individual overuse quality metrics). Second, this approach would address the structural and system-wide determinants of overuse.

Measuring systematic overuse would likewise be valuable in exploring the multitude of patient, provider, and institutional factors that impact overuse.[13] Patient level factors such as preferences [75,76], health literacy [77], medical/psychological conditions [78,79] and wealth [80,81] are likely to have a broad effect on overuse, as are differences in providers skills [82,83], financial interests [84] and other cultural factors [82,85-87]. Tort litigation has impacted utilization at both the provider and institutional level and has led to the reactionary practice of defensive medicine.[88] The rapid development of new technologies has expanded medical options and has likewise increased resource use.[89] Utilization is largely influenced by the healthcare market, that is the supply of services [90], and the United States' fee-for-service and third-party payer systems have not incentivized patients, providers, or institutions to be conscious of expenditures.

References

1. Centers for Disease Control. Health expenditure. Available at http://www.cdc.gov/nchs/fastats/hexpense.htm. Accessed October 5, 2011.

2. Centers for Medicare and Medicaid Services. National Health Expenditure Projections 2009-2019. Available at https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/downloads/NHEProjections2009to2019.pdf Accessed October 5, 2011.

3. World Health Organization. Ranking of the World's Health Systems. Available at http://www.photius.com/rankings/healthranks.html. Accessed October 5, 2011.

4. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care. Ann Intern Med. 2003;138:273-87.

5. Starfield, B. Is US Health Really the Best in the World? JAMA. 2000;284(4):483-485.

6. Reinhardt EU, Hussey PS, Anderson GF. US health care spending in an international context. Health Aff. 2004;23:10-25.

7. Korenstein D, Falk R, Howell EA, Bishop T, Keyhani S. Overuse of health care services in the United States: an understudied problem. Arch Intern Med. 2012;171-8.

8. Grady D, Redberg RF. Less is more: how less health care can result in better health. Arch Intern Med. 2010 May 10;170:749-50

9. Goodwin JS, Singh A, Reddy N, Riall TS, Kuo YF. Overuse of screening colonoscopy in the Medicare population. Arch Intern Med. 2011 Aug 8;171:1335-43.

10. Ackerman S, Gonzales R. The context of antibiotic overuse. Ann Intern Med. 2012 Aug 7;157:211-2.

11. Cassel CK, Guest JA. Choosing wisely: helping physicians and patient make smart decisions about their care. JAMA. 2012;307:1801-2.

12. Good Stewardship Working Group. The "top 5" lists in primary care: meeting the responsibility of professionalism. Arch Intern Med. 2011 Aug 8;171:1385-90.

13. Emanuel EJ, Fuchs VR. The perfect storm of overutilization. JAMA. 2008;299(23):2789-91.

14. Warren JL, Klabunde CN, Mariotto AB, Meekins A, Topor M, Brown ML, Ransohoff DF. Adverse events after outpatient colonoscopy in the Medicare population. Ann Intern Med. 2009;150:849-57, W152.

15. Zhang Y, Lee BY, Donohue JM. Ambulatory antibiotic use and prescription drug coverage in older adults. Arch Intern Med. 2010;170:1308-14.

16. Frank SM, Savage WJ, Rothschild JA, Rivers RJ, Ness PM, Paul SL, Ulatowski JA. Variability in blood and blood component utilization as assessed by an anesthesia information management system. Anesthesiology. 2012;117:99-106.

17. Amin AP, Spertus JA, Cohen DJ, et al. Use of Drug-Eluting Stents as a Function of Predicted Benefit: Clinical and Economic Implications of Current Practice. Arch Intern Med. 2012;172:1145-52.

18. Berwick DM, Hackbarth AD. Eliminating Waste in US Health Care. JAMA. 2012;307(14):1513-6.

19. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. Health Aff. 2008;27:759-69.

20. Orszag PR, Ellis P. Addressing rising health care costs--a view from the Congressional Budget Office. N Engl J Med. 2007;357:1885-7.

21. Obama B. Modern health care for all Americans. *NEJM* 2008;359:1537-41.

22. Agency for Healthcare Research and Quality. Improving Health Care Quality. Fact Sheet. AHRQ Publication No. 02-P032, September 2002. Available at http://www.ahrq.gov/news/qualfact.htm. Accessed March 3, 2012.

23. Keyhani S, Siu AL. The underuse of overuse research. Health Serv Res. 2008;43:1923-30.

24. Chassin MR, Galvin RW. The urgent need to improve health care quality. Institute of Medicine National Roundtable on Health Care Quality. JAMA. 1998;280:1000-5.

25. Squiers LB, Holden DJ, Dolina SE, Kim AE, Bann CM, Renaud JM. The public's response to the U.S. Preventive Services Task Force's 2009 recommendations on mammography screening. Am J Prev Med. 2011 May;40:497-504.

26. Iglehart JK. The political fight over comparative effectiveness research. Health Aff. 2010;29:1757-60.

27. Fisher ES, Wennberg JE. Health care quality, geographic variations, and the challenge of supply-sensitive care. Perspect Biol Med. 2003;46:69-79.

28. National Committee for Quality Assurance. Healthcare Effectiveness Data and Information Set (HEDIS): HEDIS 2011 summary table of measures, product lines and changes. Available at http://www.ncqa.org/Portals/0/HEDISQM/HEDIS%202011/HEDIS%202011%20Measures.pdf. Accessed July 13, 2011.

29. Jaslow R. Doctors unveil "Choosing Wisely" campaign to cut unnecessary medical tests. CBS News. 4 April 2012. Accessed at www.cbsnews.com/8301-504763_162-57409204-10391704/doctors-unveil-choosing-wisely-campaign-to-cut-unnecessary-medical-tests. Accessed September 12, 2012.

30. New England Health Institute. Compendium on Overuse. Available at http://www.nehi.net/uploads/publications_other/How_Many_More_Studies_Will_It_Take_Overus e.pdf. Accessed March 3, 2012.

31. Institute of Medicine. Crossing the Quality Chasm: A New Health System for the Twenty-first Century. Washington, DC: National Academy Press; 2001.

32. Agency for Healthcare Research and Quality. Glossary: underuse, overuse, misuse. Available at http://webmm.ahrq.gov/popup_glossary.aspx?name=underuseoverusemisuse. Accessed June 8, 2012.

33. Brooks, RH. The role of physicians in controlling medical care costs and reducing waste. JAMA. 2011;306:650-1.

34. Fuchs VR. Eliminate "waste" in healthcare. JAMA. 2009;302:2481-2.

35. Sanmartin C, Murphy K, Choptain N, et al. Appropriateness of healthcare interventions: concepts and scoping of the published literature. Int J Technol Assess Health Care. 2008;24(3):342-9.

36. Lavis JN, Anderson GM. Appropriateness in health care delivery: definitions, measurement and policy implications. CMAJ. 1996;154(3):321-8.

37. AQA. AQA Principles of 'Efficiency' Measures. Accessed on October 16, 2012. Available at http://www.aqaalliance.org/files/PrinciplesofEfficiencyMeasurementApril2006.doc.

38. Romley JA, Hussey PS, de Vries H, Wang MC, Shekelle PG, McGlynn EA. Efficiency and its measurement: what practitioners need to know. Am J Manag Care. 2009;15(11):842-5.

39. Wennberg J, Gittelsohn A. Small Area Variations in Health Care Delivery: A Population-Based Health Information System can Guide Planning and Regulatory Decision-Making. Science 1973;182:1102-8.

40. Chassin MR, Kosecoff J, Park RE, Winslow CM, Kahn KL, Merrick NJ, Keesey J, Fink A, Solomon DH, Brook RH. Does inappropriate use explain geographic variations in health care service? A study of three procedures. JAMA. 1987;258(18):2533-7.

41. Leape LL, Park RE, Solomon DH, Chassin MR, Kosecoff J, Brook RH. Does inappropriate use explain small-area varations in the use of health care services? JAMA. 1990;263(5):669-72.

42. Guadagnoli E, Landrum MB, Normand SL, Ayanian JZ, Garg P, Hauptman PJ, Ryan TJ, McNeil BJ. Impact of underuse, overuse, and discretionary use on geographic variation in the use of coronary angiography after acute myocardial infarction. Med Care. 2001;39(5):446-58.

43. Keyhani S, Falk R, Bishop T, Howell E, Korenstein D. The relationship between geographic variations and overuse of healthcare services: a systematic review. Med Care. 2012;50(3):257-61

44. Keyhani S, Falk R, Howell EA, Bishop T, Korenstein D. Overuse and Systems of Care: A Systematic Review. Med Care. 2013 Jun;51(6):503-508.

45. Berenson RA, Docteur E. Doing Better by Doing Less: Approaches to Tackle Overuse of Services. Timely Analysis of Immediate Health Policy Issues, January 2013. http://www.urban.org/UploadedPDF/412834-Doing-Better-by-Doing-Less.pdf Accessed June 18, 2013.

46. Chan KS, Chang E, Nassery N, Chang HY, Segal J. The State of Overuse Measurement: A Critical Review. Medical Care Research and Review Med Care Res Rev. 2013 Oct;70(5):473-96

47. National Committee for Quality Assurance. (2009, June). Developing a framework and research agenda for overuse and appropriateness measures (Agency for Healthcare Research and Quality (AHRQ) Appropriateness Small Conference). Washington DC

48. Agency for Healthcare Research and Quality. (2008, April). *Health care efficiency measures: Identification, categorization, and evaluation* (AHRQ Publication No. 08-0030). Rockville, MD

49. ABIM Foundation. *Choosing wisely: Five things physicians and patients should question: Lists*. Retrieved from http://www.choosingwisely.org/doctor-patient-lists

50. ABIM Foundation. High Value Cost Conscious Care Initiative. http://hvc.acponline.org

51. Blue Cross Blue Shield of Massachusetts. (2009, March). *The alternative quality contract*. http://www.qualityaffordability.com/pdf/alternative-quality-contract.pdf

52. Centers for Medicare and Medicaid Services. (2010). *2010 PQRI measures list*. Retrieved from http://www.cms.gov/PQRI/Downloads/2010_PQRI_MeasuresList_111309.pdf

53. Good Stewardship Working Group. (2011). The "Top 5" Lists in primary care: Meeting the responsibility of professionalism. *Archives of Internal Medicine*, *171*, 1385-1390.

54. Chan et al. Institute of Medicine. (2008). *Knowing what works in health care: A roadmap for the nation*. Washington, DC: National Academies Press.

55. Kern, L. M., Dhopeshwarkar, R., Barrón, Y., Wilcox, A., Pincus, H., & Kaushal, R. (2009). Measuring the effects of health information technology on quality of care: A novel set of proposed metrics for electronic quality reporting. *Joint Commission Journal on Quality and Patient Safety*, *35*, 359-36

56. National Committee for Quality Assurance. (2009, June). *Developing a framework and research agenda for overuse and appropriateness measures* (Agency for Healthcare Research and Quality (AHRQ) Appropriateness Small Conference). Washington DC

57. National Institute for Health and Clinical Excellence. *Search the NICE "do not do" recommendations database*. http://www.nice.org.uk/usingguidance/donotdorecommendations/

search.jsp

58. National Quality Forum. *Identification of potential 2013 e-Quality measures*. Washington, DC

59. National Quality Measures Clearinghouse. *Measure summary: Use of imaging studies for low back pain: Percentage of members with a primary diagnosis of low back pain who did not have an imaging study (plain x-ray, MRI, CT scan) within 28 days of the diagnosis*. http://www.guideline.gov/content.aspx?id=34670

60. National Priorities Partnership. (2008). *National priorities and goals: Aligning our efforts to transform America's healthcare*. Washington, DC: National Quality Forum.

61. QualityNet. *Imaging efficiency measures*. Retrieved from http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier2&cid=1228695266120

62. Earle CC, Park ER, Lai B, Weeks JC, Ayanian JZ, Block S. Identifying potential indicators of the quality of end-of-life cancer care from administrative data. *J Clin Oncol* 2003 March 15;21(6):1133-8.

63. Mathias JS, Baker DW. Developing quality measures to address overuse. *JAMA* 2013 May 8;309(18):1897-8.

64. Neuhauser D. Assessing health quality: the case for tracers. *J Health Serv Res Policy* 2004 October;9(4):246-7.

65. Mathias JS, Baker DW. Developing quality measures to address overuse. *JAMA* 2013 May 8;309(18):1897-8.

66. The Dartmouth Institute for Health Policy and Clinical Practice. The Dartmouth atlas of health care www.dartmouthatlas.org. Updated 2012. Accessed 9/30, 2012.

67. Douglas PS, Garcia MJ, Haines DE et al. ACCF/ASE/AHA/ASNC/HFSA/HRS/SCAI/SCCM/SCCT/SCMR 2011 Appropriate Use Criteria for Echocardiography. A Report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, American Society of Echocardiography, American Heart Association, American Society of Nuclear Cardiology, Heart Failure Society of America, Heart Rhythm Society, Society for Cardiovascular Angiography and Interventions, Society of Critical Care Medicine, Society of Cardiovascular Computed Tomography, Society for Cardiovascular Magnetic Resonance American College of Chest Physicians. *J Am Soc Echocardiogr* 2011 March;24(3):229-67.

68. U.S.Preventative Services Taskforce. Screening for Carotid Artery Stenosis. 2007 http://www.uspreventiveservicestaskforce.org/uspstf/uspsacas.htm

69. Qaseem A, Alguire P, Dallas P et al. Appropriate use of screening and diagnostic tests to foster high-value, cost-conscious care. *Ann Intern Med* 2012 January 17;156(2):147-9.

70. Agency for Healthcare Research and Quality. National Guideline Clearinghouse. [http://www.guideline.gov/index.aspx . Accessed March 2013](http://www.guideline.gov/index.aspx)

71. Bridges J, Dor A, Grossman M (2005) A Wolf Dressed in Sheep's Clothing: Perhaps Quality Measures are Just Unmeasured Severity, *Applied Health Economics and Health Policy*, 4(1): 55-64

72. Feldman P, **Bridges J**, Peng, T (2007) Team Structure and Adverse Events in Home Health Care, *Medical Care*, 45(6):553-561

73. Tsai A, Votruba M, **Bridges J**, Cebul R (2006) Overcoming Bias in Estimating the Volume-Outcome Relationship, *Health Services Research*, 41(1):252-264

74. IOM (Institue of Medicine). Variation in Health Care Spending: Target Decision Making, Not Geography. Washington, D.C.: The National Academies Press; 2013.

75. Lesser Y, Rabinowitz J. Elective amniocentesis in low-risk pregnancies: decision making in the era of information and uncertainty. Am J Public Health. 2001;91(4):639-41.

76. Wong MD, Asch SM, Andersen RM, Hays RD, Shapiro MF. Racial and ethnic differences in patients' preferences for initial care by specialists. A J Med. 2004;116(9):613-20.

77. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann of Intern Med. 2011; 155:97-107.

78. De Miguel-Diez J, Carrasco-Garrido P, Bejas-Guiterrez J, Martin-Centeno A, Gobartt-Vazquez E, Hernandez-Barrera V, Dil de MA, Jimenez-Garcia R. Inappropriate overuse of inhaled corticosteroids for COPD patients: impact on health costs and health status. Lung. 2001;189(3):199-206.

79. Wagner JA, Pietrzak RH, Petry NM. Psychiatric disorders are associated with hospital utilization in persons with hypertension: results from the National Epidemiologic Survey on alcohol and related conditions. Soc Psychiatry Pschiatr Epidemiol. 2008;43:878-88.

80. Sogutlu A, Levenson JL, McClish DK, Rosef SD, Smith WR. Somatic symptom burden in adults with sickle cell disease predicts pain, depression, anxiety, health care utilization, and quality of life: the PiSCES project. Psychosomatics. 2011;52:272-9.

81. Prasad SM, Gu X, Lipsitz SR, Nguyen PL, Hu JC. Inappropriate utilization of radiologic imaging in men with newly diagnosed prostate cancer in the United States. Cancer. 2010;118(5):1260-7.

82. Lousuebsakul V, Knutsen SM, Singh PN, Gram IT. Is colposcopic biopsy overused among women with a cytological diagnosis of atypical squamous cells of undetermined significance (ASCUS)? J Womens Health (Larchmt). 2003;12(6):553-9.

83. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abbound PA, Rubin HR. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA. 1000;282:1458-65.

84. Sood R, Sood A, Ghosh AK. Non-evidence-based variables affecting physicians' test ordering tendencies: a systematic review. Neth J Med. 2007;65: 167-77.

85. Wilson IB, Dukes K, Greenfield S, Kaplan S, Hillman B. Patients' role in use of radiology testing for common office practice complaints. Ann of Intern Med. 2001;161:256-63.

86. Rello J, Lorente C, Bodi M, Diaz E, Ricart M, Kollef MH. Why do physicians not follow evidence-based guidelines for preventing ventilator associated pneumonia: a survey based on the opinions of an international panel of intensivists. Chest. 2002;122:656-61.

87. Yabroff KR, Klabunde CN, Yuan G, et al. Are physician's recommendations for colorectal cancer screening guideline-consistent? J Gen Intern Med. 2011;26:177-84.

88. Wigton RS, Darr CA, Corbett KK, Nickol DR, Gonzales R. How do community practitioners decide whether to prescribe antibiotics for acute respiratory tract infections? J Gen Intern Med. 2008:23;1615-20.

89. Baicker K, Fisher E, Chandra A. Malpractice liability costs and the practice of medicine in the Medicare program. Health Aff. 2007;26:841-52.

90. Dubois RW, Chawla AJ, Neslusan CA, Smith MW, Wade S. Explaining drug spending trends: does perception match reality. Health Aff. 2000;19:231-9.

Appendix 1. Algorithms for Indicators

Working Number	Indicator	Procedure	Population	Specifications
1	Stress echocardiography for detection of CAD/risk assessment in symptomatic or ischemic equivalent acute chest pain (i.e. Acute Coronary Syndrome)	Individuals with **CPT** codes as listed or HCPCS codes as listed for echocardiography Unit of observation: per beneficiary per date of service	Individuals with a code for **emergency visit*** with any of the **ICD-9** diagnoses OR individuals with a hospitalization with DRGs as listed, or primary or secondary diagnosis code during **hospitalization** for any of the **ICD-9** diagnoses Unit of observation: per beneficiary per date of service	**Stress Echocardiography** **CPT:** 93350, 93351 **HCPCS:** C8928, C8930 **ER visit** **CPT** 99281-99285 **Diagnoses** (associated with ER visits or Hospitalization) **ICD-9:** 410.xx, 411.1, 411.81, 411.89 **DRG** 281-287
10	Laminectomy and/or spinal fusion	Laminectomy or spinal fusion Unit of observation: per beneficiary	Everyone MINUS those with a clear indication (radicular symptoms* and we will be liberal with this) *symptoms clearly of herniated disc—radicular pain Unit of observation: per beneficiary per date of service	**Laminectomy** **CPT:** 22533, 22534, 22558, 22630 0275T 63005, 63012, 63017, 63030, 63035, 63042, 63047, 63200, 63267, 63272, 63173, 63185, 63190, 63191 **ICD-9 procedure:** 80.51, 81.06, 81.07, 81.08, 84.67, 84.65 **DRG:** 459, 460 **But NOT for following ICD-9 (exclusions from denominator)** **Neither** **Herniated disc (omitting cervical):** **ICD-9:** 722.1, 722.2, 722.3, 722.5, 722.6, 722.7, 722.8, 722.9 722.70, 722.72, 722.73, 722.80, 722.82, 722.83, 722.90, 722.92, 722.93 **Nor** ***Two occurrence within 30 days for following ICD-9 (any combination):*** **ICD-9:** 355.0, 355.7, 355.8, 355.9, 724.3, 724.4, 729.2
11	HYSTERCTOMY for benign disease	Any hysterectomy (not specified for malignancy treatment) Unit of observation: per beneficiary per date of service	All women MINUS those with a malignancy diagnosis Unit of observation: per beneficiary	**Hysterectomy** **ICD-9 procedure:** 68.3, 68.4, 68.5, 68.6, 68.7, 68.9 68.31, 68.39, 68.41, 68.49, 68.51, 68.59, 68.61, 68.69, 68.71, 68.79 **CPT:** 58150, 58152, 58180, 58200, 52810, 58260, 58262, 58263, 58267, 58270, 58275, 58280, 58285, 58290, 58291, 58292, 59293, 59294, 58541, 58542, 58543, 58544, 58548, 58550, 58552, 58553, 58554, 58570, 58571, 58572, 58573 **NOT (Remove from denominator):** **Malignancy** **ICD-9:** 179, 180.x, 182, 183, 184 **DRG:** 734-741, 754-756
20	Fiberoptic laryngoscopy for patients with a diagnosis of sinusitis.	Laryngoscopy WITH **ICD-9** code indicating sinusitis on the same claim Unit of observation: per beneficiary per claim	Individuals with a diagnosis of sinusitis (acute or chronic) – inpatient or outpatient Unit of observation: per beneficiary per date of service	**Fiberoptic laryngoscopy** **CPT:** 31575, 31476, 31577, 31578, 31579 **Sinusitis** **ICD-9:** 461, 461.x, 473, 473.x *linked to laryngoscopy in same CLAIM ID.*
21	Nasal endoscopy for sinusitis diagnosis	Nasal endoscopy WITH **ICD-9** code indicating sinusitis on the same claim Unit of observation: per beneficiary per claim	Individual with a diagnosis of sinusitis (acute or chronic) – inpatient or outpatient Unit of observation: per beneficiary per date of service	**Diagnostic endoscopy** **CPT:** 31231, 31233, 31235 **Sinusitis** **ICD9:** 461, 461.x, 473, 473.x *linked to laryngoscopy in same CLAIM ID.*
24	More than one emergency department visit in last 30 days of life	More than 2 visits with location code or **CPT** code indicating ED use within 30 days before death	Individuals with death during our observation period Unit of observation: per beneficiary	**ED Visit** Medicare location code 23 or **CPT:** 99281-99285 **ER use date** – Death date less than 30.

		Unit of observation: per beneficiary		
26	Routine monitoring of digoxin in patients with congestive heart failure	Any measure of digoxin with no hospitalizations or ER visits during that year. Unit of observation: per beneficiary	All patients* with CHF *will include atrial fibrillation patients as well Unit of observation: per beneficiary	**Therapeutic monitoring of digoxin** **CPT:** 80162 **ED Visit** **CPT:** 99281-99285 **CHF** **ICD-9:** 428, 428.0, 428.1, 428.2, 428.3, 428.4, 428.9 428.2x, 428.3x, 428.4x **AF & Flutter** **ICD-9:** 427.3, 427.3x
27	EEG monitoring in individuals presenting with syncope	EEG on the same claim as diagnosis of syncope or at any time during the hospitalization with a code for syncope Unit of observation: per beneficiary per date of service	Individuals with an outpatient visit with diagnosis of syncope or hospitalization for syncope Unit of observation: per beneficiary per date of service	**EEG** **ICD-9 procedure:** 89.14 (during hospitalization) **CPT:** 3650F, 95812, 95813, 95816, 95819, 95822, 95827 (outpatient) **Syncope** **ICD-9:** 780.2, 992.1 (heat), 337.01 (carotid sinus)
32	Serological tests for helicobacter pylori	Any code indicating testing for H. pylori Unit of observation: per beneficiary per date of service	Whole population Unit of observation: per beneficiary	**H. Pylori testing** **CPT:** 86677
34	MRI in individuals with traumatic brain injury	MRI on the same claim as diagnosis if outpatient or during hospitalization if inpatient Unit of observation: per beneficiary per date of service	Patients with traumatic brain injury Unit of observation: per beneficiary per date of service	**MRI (brain)** **CPT:** 70551, 70552, 70553, **Traumatic brain injury** **ICD-9:** 850, 850.x, 850.xx, 851, 851.x, 851.xx, 852, 852.x, 852.xx, 853, 853.x, 853.xx, 854, 854.0, 854.0x, 854.1, 854.1x, 959.01 Within first 2 days
36	PET, CT, and radionuclide bone scans in individuals with prostate CA	PET, CT, or radionuclide bone scan **AFTER** diagnosis Unit of observation: per beneficiary per date of service	Men with low risk for prostate CA Unit of observation: per beneficiary	**Prostate CA** **ICD-9:** 185, 233.4 **PET Scan** **CPT:** 78811, 78812, 78813, 78814, 78815, 78816 **Pelvic CT Scan** **CPT:** 72192, 72193, 72194 **Bone Scan** **CPT:** 3269F, 77074, 77075 **NOT** **Intermediate/high risk of recurrence** **CPT:** 3272F, 3273F
37	Traction for low back pain	Traction with diagnosis of low back pain Unit of observation: per beneficiary per date of service	Low back pain diagnosis Unit of observation: per beneficiary per date of service	**Traction** **CPT:** 97012, 97140 **HCPCS:** E0830 **Low back pain** **ICD-9:** 721.3, 721.90, 722.10, 722.52, 722.6, 722.93, 724.02, 724.2, 724.3, 724.5, 724.6, 724.70, 724.71, 724.79, 738.5, 739.3, 739.4, 846.0, 846.1, 846.2, 846.3, 846.8, 846.9 , 847.2
41	Screening for asymptomatic carotid artery stenosis (CAS) in the general adult population.	**CPT** 93880 or 3100F, ONLY IN outpatient setting (not ER) Unit of observation: per beneficiary per date of	All people with NO **ICD-9** codes for: 785.9, 784.2, 362.34, 435.9, 433.10, 342.90, 780.2, 781.3, 437.0 Unit of observation: per beneficiary	**Carotid imaging** **CPT :** 93880 ONLY IN outpatient setting (not ER), 3100F **ED visit** **CPT:** 99281-99285 **NOT**

				ICD-9: 785.9, 784.2, 362.34, 435.9, 433.10, 342.90, 780.2, 781.3, 437.0
43	Preoperative chest radiography in the absence of a clinical suspicion for intrathoracic pathology	71010, 71020 These codes must be in a 30 day window before the anesthesia code Unit of observation: per beneficiary per date of service	All patients who had anesthesia 00100-02101 (**CPT**) with EXCLUSION of diagnoses 466.xx, 480.xx-488.xx, 490.xx-496.xx, 500.xx-508.xx, 510.xx-519.xx Unit of observation: per beneficiary per date of service	**Anesthesia** **CPT:** 00100-02101 **Radiology** **CPT:** 71010, 71020 **NOT** **ICD-9:** 466.xx, 480.xx-488.xx, 490.xx-496.xx, 500.xx-508.xx, 510.xx-519.xx
45	Performing tumor marker studies in asymptomatic women with previously treated breast cancer	82378 (CEA), 86300 (CA 15-3) (CA 27.29) Unit of observation: per beneficiary per date of service	Breast cancer is 174.0-174.9 Unit of observation: per beneficiary	**CEA** **CPT:** 82378, **CA 15-3 (CA 27.29)** **CPT:** 86300 **Breast cancer** **ICD-9:** 174.0-174.9
46	Don't perform unproven diagnostic tests, such as immunoglobulin G (IgG) testing or an indiscriminate battery of immunoglobulin E (IgE) tests, in the evaluation of allergy.	Use of CPT 82701, 82784, 82785, 82787, 86005 on the same claim as a code for diagnoses in the denominator column Unit of observation: per beneficiary per claim	477.0, 477.1, 477.2, 477.8, 477.9, 493.0, 493.02, 493.9, 493.90, 493.92, 708.0, 995.3 Unit of observation: per beneficiary per date of service	**Diagnostic tests** **CPT:** 82701, 82784, 82785, 82787, 86005 **Allergic rhinitis** **ICD-9:** 477.0, 477.1, 477.2, 477.8, 477.9 **Asthma** **ICD-9:** 493.0, 493.02, 493.9, 493.90, 493.92 **Other allergies** **ICD-9:** 708.0, 995.3
49	MRI Lumbar Spine for Low Back Pain	MRI of the lumbar spine studies with a diagnosis of low back pain (from the denominator) without the patient having claims-based evidence of prior antecedent conservative therapy. CPT=72148, or 72149, or 72158 with no codes for 97110, 97112, 97113, 97124, 97140, 98940, 98941, 98942, 98943 in the 60 days preceding the MRI of the lumbar spine AND no codes for 99210-99205, 99211 -99215, 99241-99245, 99341-99345, 99347-99350, 99354-99357, 99385-99387, 99395-99397, 99401-99404, 99455-99456, 99499 between 28 and 60 days preceding the MRI of the lumbar spine Unit of observation: per beneficiary per date of service	MRI of the lumbar spine studies with a diagnosis of low back pain on the imaging claim. CPT=72148, or 72149, or 72158 AND ICD-9: 721.3, 721.90, 722.10, 722.52, 722.6, 722.93, 724.02, 724.2, 724.3, 724.5, 724.6, 724.70, 724.71, 724.79, 738.5, 739.3, 739.4, 846.0, 846.1, 846.2, 846.3, 846.8, 846.9 , 847.2 Excluded from the denominator - CPT codes: 22010-22865 and 22899 in 90 days preceding MRI; ICD-9 codes: 140-208, 230-234, 235-239, 304.0X, 304.1X, 304.2X, 304.4X, 305.4X, 305.5X, 305.6X, 305.7X, 344.60, 344.61, 729.2, 042-044, 279.3 in preceding 365 days; 800-839, 850-854, 860-869, 905-909, 926.11, 926.12, 929, 952, 958-959 in preceding 45 days; 324.9, 324.1 on same claim as MRI Unit of observation: per beneficiary per date of service	**MRI Lumbar Spine** **CPT:** 72148, 72149, 72158 **NOT** in 60 days preceding Lumbar Spine MRI **Therapies** **CPT:** 97110, 97112, 97113, 97124, 97140, 98940, 98941, 98942, 98943 **NOT** between 28 and 60 days preceding Lumbar Spine MRI **Evaluation and management** **CPT:** 99210-99205, 99211 -99215, 99241-99245, 99341-99345, 99347-99350, 99354-99357, 99385-99387, 99395-99397, 99401-99404, 99455-99456, 99499 **NOT** in the preceding 90 days (denominator) **Lumbar Spine Surgery** **CPT:** 22010-22865 and 22899 **Low back pain** **ICD-9 :** 721.3, 721.90, 722.10, 722.52, 722.6, 722.93, 724.02, 724.2, 724.3, 724.5, 724.6, 724.70, 724.71, 724.79, 738.5, 739.3, 739.4, 846.0, 846.1, 846.2, 846.3, 846.8, 846.9, 847.2 **NOT** in preceding 365 days **ICD-9:** 140-208, 230-234, 235-239; 304.0X, 304.1X, 304.2X, 304.4X, 305.4X, 305.5X, 305.6X, 305.7X, 344.60, 344.61, 729.2, 042-044, 279.3 NOT in preceding 45 days **Trauma** **ICD-9:** 800-839, 850-854, 860-869, 905-909, 926.11, 926.12, 929, 952, 958-959 **NOT** on the same claim **Intraspinal abscess** **ICD-9:** 324.9, 324.1
50	Thorax CT Use of Contrast Material	The number of thorax CT studies with and without contrast ("combined	The number of thorax CT studies performed (with contrast, without contrast or both with and without	**Thorax CT w/ and w/o contrast** **CPT:** 71270

		studies"). CPT 71270 Unit of observation: per beneficiary per date of service	contrast). CPT 71250, 71260, 71270 Unit of observation: per beneficiary per date of service	**Thorax CT** **CPT:** 71250, 71260, 71270
51	Abdomen CT use of contrast material	The number of Abdomen CT studies with and without contrast ("combined studies"). CPT 74170 Unit of observation: per beneficiary per date of service	The number of Abdomen CT studies performed (with contrast, without contrast or both with and without contrast). CPT 74150, 74160, 74170 Exclude from the denominator if on the same claim as CPT 74140, 74160, 74170 - ICD-9 code: 593.9, 120.0, 599.70, 599.71, 599.72, 251.2, 251.0, 250.8, 270.3, 255.9, 194.xx, 277.xx, 237.xx, 155.0, 155.1, 155.2, 157.0, 157.1,157.2,157.3, 157.4, 157.8, 157.9, 189.0, 211.5, 211.6, 211.7, 223.0 Unit of observation: per beneficiary per date of service	**Abdomen CT w/ and w/o contrast** **CPT:** 74170 **Abdomen CT** **CPT:** 74150, 74160, 74170 **NOT on the same claim as the numerator** **ICD-9:** 593.9, 120.0, 599.70, 599.71, 599.72, 251.2, 251.0, 250.8, 270.3, 255.9, 194.xx, 277.xx, 237.xx, 155.0, 155.1, 155.2, 157.0, 157.1, 157.2, 157.3, 157.4, 157.8, 157.9, 189.0, 211.5, 211.6, 211.7, 223.0
47	Don't order sinus computed tomography (CT) or indiscriminately prescribe antibiotics for uncomplicated acute rhinosinusitis.	Any occurrence of sinus CT (CPT 70486, 70487, 70488) in the 3 months preceding the diagnosis of acute sinusitis Unit of observation: per beneficiary per date of service	461.0, 461.1, 461.2, 461.3, 461.8, 461.9 AND NO code in the preceding 3 months for any of these AND NO code in the preceding 3 months for 473.0, 473.1, 473.2, 473.3, 473.8, 473.9 Unit of observation: per beneficiary per date of service	**Sinus CT** **CPT:** 70486, 70487, 70488 **Acute sinusitis** **ICD-9:** 461.0, 461.1, 2461.2, 461.3, 461.8, 461.9 **NOT in the preceding 3 months** **Acute sinusitis** **ICD-9:** 461.0, 461.1, 461.2, 461.3, 461.8, 461.9 **NOT in the preceding 3 months** **Chronic sinusitis** **ICD-9:** 473.0, 473.1, 473.2, 473.3, 473.8, 473.9

Appendix 2. Cost Estimates

We estimated the national costs attributable to the use of these procedures in patients who are unlikely to benefit. As described in the text, we had counts of the usage of these procedures in the 5% Medicare sample data. This was multiplied by 20 for a national estimate, as the 5% sample is selected to be representative of the U.S. With the FlashCode software (2011 version), we looked up the Medicare reimbursement for the procedures. (1) These can be considered to be the direct costs of the procedures as we did not include any downstream costs. Where appropriate, we included technical and professional fees associated with the procedures. We did not include costs of phlebotomy for the blood tests.   As the costs vary by state, we chose Minnesota as our reference state. We updated the costs to 2013 dollars with 3% inflation per year. For the two inpatient procedures (laminectomy and hysterectomy) we used the "fair costs" as listed in the Healthcare Bluebook for these procedures, which are likely to be much lower than usual stated costs. (2) The costs of an emergency visit came from the Medical Expenditure Payment Survey of AHRQ.(3)   We kep the usage the same as in 2008 rather than reflecting any growth in use of procedures. The estimated U.S. costs are reported as millions of dollars in the Table to 2 or 3 signficiant digits.

 (1)   FlashCode, Medical Coding & Compliance Solutions, LLC (MCCS); Turlock, CA
 (2)   http://www.healthcarebluebook.com/page_Results.aspx?id=102&dataset=MD
 (3)   Machlin, S., and Chowdhury, S. *Expenses and Characteristics of Physician Visits in Different Ambulatory Care Settings, 2008*. Statistical Brief #318. March 2011. Agency for Healthcare Research and Quality, Rockville, MD. http://www.meps.ahrq.gov/mepsweb/data_files/publications/st318/stat318.pdf