HSR Health Services Research

DOI: 10.1111/1475-6773.12498
RESEARCH ARTICLE

Guidelines for Measuring Disease Episodes: An Analysis of the Effects on the Components of Expenditure Growth

Abe Dunn, Eli Liebman, Lindsey Rittmueller, and Adam Hale Shapiro

Objective. To provide guidelines to researchers measuring health expenditures by disease and compare these methodologies' implied inflation estimates.

Data Source. A convenience sample of commercially insured individuals over the 2003 to 2007 period from Truven Health. Population weights are applied, based on age, sex, and region, to make the sample of over 4 million enrollees representative of the entire commercially insured population.

Study Design. Different methods are used to allocate medical-care expenditures to distinct condition categories. We compare the estimates of disease-price inflation by method.

Principal Findings. Across a variety of methods, the compound annual growth rate stays within the range 3.1 to 3.9 percentage points. Disease-specific inflation measures are more sensitive to the selected methodology.

Conclusion. The selected allocation method impacts aggregate inflation rates, but considering the variety of methods applied, the differences appear small. Future research is necessary to better understand these differences in other population samples and to connect disease expenditures to measures of quality.

Key Words. Health care expenditures, disease episodes, measuring expenditures of treatment

Health care is consumed by individuals to treat medical conditions, yet there are no national health care statistics that track the cost of treatment at the disease level. The need for more detailed statistics has been recognized by both academics and policy makers, who have called for the development of a Health Care Satellite Account that measures the cost of disease treatment.[1] Policy makers, consumers, and industry participants are increasingly interested in whether changes in the cost of treatment are worth the health benefit. By focusing on spending by disease rather than by service, researchers will be better able to connect expenditures for specific diseases with the associated health outcomes. Tracking expenditures at the disease level also provides a

more relevant unit of price for patients, since patients ultimately seek treatment for a condition. Moreover, recent health care reforms have led to shifts in payment structures, away from paying a fee for each service and toward bundled payments that pay for the total cost of treatment.

There is general agreement that tracking the growth in the cost of treatment is valuable, but there is little consensus regarding how a disease-based price should be defined and measured. Although several papers look at disease-price growth, they typically focus on only one disease allocation method (e.g., Aizcorbe and Nestoriak 2011; Roehrig and Rousseau 2011; Dunn et al. 2012; Bradley 2013; Dunn, Liebman, and Shapiro 2014). Those papers that do compare allocation methods typically look at the amount allocated to disease categories (e.g., MaCurdy, Kerwin, and Theobald 2009; Rosen et al. 2012), such as the allocation of expenditures to heart disease and diabetes. However, relatively little work has investigated how the different approaches for allocating disease expenditures might affect disease-price growth.[2] Understanding how these different measures might affect disease-price growth has implications for measuring inflation, productivity, and real output in the health sector.

This paper analyzes various methods for allocating expenditures to diseases using commercial claims data from MarketScan for the years 2003–2007. There are two primary aims of this paper: (1) to provide a range of estimates for disease-price inflation across different methodologies; and (2) to provide some basic guidelines for how the selected methodology may affect the measurement of disease-price inflation. There are distinct advantages in using a large data set for studying this issue. First, a larger sample is likely to be more representative of high-spending individuals that may not be present in smaller surveys (Aizcorbe et al. 2012b). Second, the larger sample will create more precise estimates, which ensures that the differences in the estimates across approaches are not driven by statistical error.[3]

We find that the different methodologies produce a range of estimates for disease-price inflation from a compound annual growth rate of 3.1–3.9 percent, with the average growth rate of about 3.5 percent across methods. As discussed in the text, the best approach may depend greatly on the particular application and the goals of the researcher. Disease-based indexes may be

---

Address correspondence to Abe Dunn, Ph.D., Bureau of Economic Analysis, 1441 L Street NW, Washington, DC 20005; e-mail: abe.dunn@bea.gov. Eli Liebman, M.A., is with the Department of Economics, Duke University, Durham, NC. Lindsey Rittmueller, M.S., is with the Bureau of Economic Analysis, Washington, DC. Adam Hale Shapiro, Ph.D., is with the Federal Reserve Bank of San Francisco, San Francisco, CA.

applied as deflators to calculate real output or to calculate health expenditures in current dollars (e.g., transforming health expenditures in 2003–2007 dollars).[4] For contracting purposes, disease-based price indexes may be useful for setting prices on bundled payments.[5] Disease-based prices may be used by insurers or employers attempting to improve the efficiency and quality of medical interventions.[6] Finally, disease-based statistics may be used to better understand spending trends in the health sector.[7]

The paper is organized as follows. The second section of this paper provides a discussion of different grouper methodologies. The third section discusses how the indexes are measured. The fourth section presents the overall results followed by a discussion of the methodologies in the fifth section. And the last section concludes.

METHODS FOR EXPENDITURE ALLOCATION

A major challenge for assigning expenditures to a specific disease is that individuals often have multiple conditions. For example, if a patient visits a doctor's office to treat both hypertension and heart disease, how should the expenditures from that visit be allocated across these two conditions? Indeed, the majority of expenditures are by patients that have many conditions; 53 percent of expenditures are allocated to those with seven or more conditions in our data.

Another challenge is determining the level of aggregation to use when allocating expenditures to mutually exclusive disease categories. One could use broadly defined disease categories, such as the Major Diagnostic Categories, which include just 25 diseases. Alternatively, one could use disaggregate categories contained in the International Classification of Diseases 9th edition (ICD-9) which includes about 13,000 disease definitions. Precisely defined disease categories are useful to account for the heterogeneity in individual conditions. However, the more granular the disease episode category, the fewer individuals will fall into each category, potentially reducing the precision of the estimates.

Following the discussion of Rosen and Cutler (2012), we focus on three general methods for allocating expenditures. The first method is an encounter-based methodology, which assigns expenditures to diseases based on the observed diagnosis at the claim line. Our second method uses episode groupers, which are software algorithms to review a patient's medical history and assign claim lines to distinct episodes. The third method, a person-based approach, uses regressions and the characteristics of the patient to statistically

divide expenditures across disease categories. Although there are three main methods analyzed in this paper, there are a variety of ways that each may be executed that may impact both the allocation across diseases and the measure of disease-price inflation. The following subsections describe each approach in more detail.

Encounter-Based Method

The encounter-based methodology uses the listed diagnosis on the claim to assign expenditures to diseases. One advantage of this approach is that it is easy to apply, since one only needs the observed diagnosis on the claim record. When applying an encounter-based method, one of the main choices is to determine the level of disease aggregation. In the MarketScan data, the diagnosis observed on the claims records are ICD-9 codes. To aggregate, we map the ICD-9 diagnosis into one of the 263 disease categories defined by the Clinical Classification Software (CCS), a free application developed by Agency for Healthcare Research and Quality and commonly applied in health research. A major advantage of this aggregation is that it groups diseases into a manageable number of clinically meaningful categories.

There are two disadvantages with the encounter-based method. First, some records have multiple diagnoses listed on a single claim. For this case, we follow one of the standard approaches used in the literature, recently applied by Starr, Dominiak, and Aizcorbe (2014), and use the first listed diagnosis on the claim, known as the primary diagnosis. The drawback here is that when there are comorbidities, the allocation of expenditures is decided entirely by the listed diagnosis, which may not be an accurate allocation for individuals with complex comorbidities. A second drawback is claims that are not assigned a diagnosis by medical care providers are not allocated to any disease category. This is often true of prescription drug claims. Claims that are not assigned to any diagnosis are removed from the analysis.

Episode-Based Method

Patients are normally given a variety of services (e.g., office visits, prescription drugs, lab tests) to treat a specific condition. Therefore, all of these services must be analyzed together to provide a meaningful unit of analysis. This is the basic concept underlying an episode-of-care (see Keeler et al. 1982; Hornbrook, Hurtado, and Johnson 1985). In contrast to the encounter-based method, the episode-grouper software uses all records for an individual,

including prescription drugs, to assign diagnoses and comorbidities using clinical knowledge of these conditions. These algorithms rely on the medical expertise incorporated into the programs. However, the episode-based approach still has drawbacks. For instance, one issue is how to identify the start and end of an episode. While chronic diseases are often analyzed over a fixed time period of a year, the determinants of start and stop times for other conditions are less clear.

This paper applies the two widely used commercial episode groupers to define disease episodes: the Medical Episode Grouper (MEG) by Truven Health Analytics and the Symmetry Episode Treatment Grouper (ETG) by Optum. Both software programs apply algorithms that examine claim lines for each individual chronologically, and each claim line is assigned to a unique episode of care. The key information used to assign episodes is the ICD-9 diagnosis on the claim (up to four diagnoses per claim line) or the national drug code (NDC) number for prescription drugs. The method for assigning start and stop times for episodes is also similar. Certain types of claim line records are determined to be "anchor" records that can initiate an episode. For example, visits to physicians are often viewed as anchor records, since physicians are viewed as qualified to assign a diagnosis to a patient. To end an episode, a certain amount of time without an associated claim record must pass, referred to as a "clean period." Although the default settings for defining anchor records and clean periods differ, both algorithms have the ability to customize these settings. Both groupers have around 500 disease categories, with ETG having 456 and MEG having 525. Both groupers provide additional disaggregation of the disease groups to account for the severity of the diseases. Neither of the grouper algorithms uses procedure codes to classify diseases or severity levels.

Many of the basic ideas underlying the groupers are similar, but the details and logic of each algorithm are distinct, leading to differences in output. One difference in these algorithms is related to severity assignment. The ETG grouper will use the demographic information to classify severity, such as age, but the MEG grouper does not. For the MEG, severity is thought of as the staging of the disease, capturing the progression of a disease toward increasing complexity: a disease stage of 1 signifies no complications, while a disease stage of 4 is death. By contrast, the ETG severity levels are based on the distribution of illness severity in the population for each disease category.

The main advantage of the episode-based method is that it is as simple to implement as the encounter-based method, but it can accommodate claims without diagnosis codes. Furthermore, it allows the researcher to use the episode-of-care definition.

However, there are some drawbacks to this method. First, grouper algorithms are proprietary and relatively costly, which may limit their usefulness for many researchers. Second, the ability of the grouper algorithms to allocate claims records to episodes depends on the claims history of the individual. For those claims without an associated diagnosis, such as prescription drug claims, the algorithm requires sufficient medical history of the individual to properly assign a claim record to a medical condition and episode. Similar to the encounter-based method, claims records not assigned are removed from the analysis.

Person-Based Method

The person-based method is based on a regression of total expenditures for an individual, including expenditures for claims that do not contain a diagnosis, on indicator variables for whether the person has a given disease in a given time period. The indicator for each disease captures the marginal effect of each disease on expenditures. The regression coefficients can then be used to allocate expenditures across diseases.

An advantage of this method is that it relies less on the diagnosis from a particular claim, and more on the statistical relationship between expenditures and disease indicators. This is particularly important for comorbidities, where certain diseases, such as diabetes, may affect the severity and treatment of other diseases, which may complicate the assignment of expenditures. Another advantage is that the person-based method allocates expenditures for those services that do not have a listed diagnosis, as is common with prescription drug claims and laboratory tests. In theory, all claims can be allocated, provided an individual has at least one diagnosis in a given year.

A main drawback for the person-based method is that it is based on an empirical model, which means the choice of specification is an important determinant of the final inflation measure. Researchers must consider a multitude of factors when deciding on a proper specification. First, one must choose the disease categories to be included in the regression (e.g., CCS, MEG, or ETG), as well as decide if disease interactions should be included to account for more complicated comorbidities, such as an interaction between diabetes and heart disease. Second, the choice of the functional form of the regression may impact expenditure allocation across diseases. When applying nonlinear models, researchers must be careful to appropriately allocate expenditures across diseases, so that they account for the patient's total dollar amount spent for the entire year, as demonstrated by Trogdon, Finkelstein, and Hoerger

(2008). Unlike the other methods, this method does not aggregate claims into specific disease categories, but instead allocates the share of total expenditures to different conditions. Therefore, there is a share of the claims that are unassigned and dropped from the analysis.

DECOMPOSITION

Expenditure per capita for disease d for time period t is denoted $C^*_{d,t}$, which is calculated by dividing total expenditures for disease d in period t by the total commercially insured population in period t. To simplify our analysis, we fix the demographics of the population, so that the changing age of the population does not affect expenditure growth. To create a measure of medical-care expenditure growth, we form the following demographically adjusted expenditure per capita index ($DECI$):

$$DECI_{d,t} = \frac{C^*_{d,t}}{C^*_{d,0}}$$

Next, we divide the $DECI$ into two-two components. One component is the treated prevalence disease index, $PREV_{d,t}$, which we define as growth in the prevalence of treated disease d, $prev_{d,t}$:

$$PREV_{d,t} = \frac{prev_{d,t}}{prev_{d,0}}$$

where $prev_{d,t}$ is simply the number of episodes treated divided by the commercially insured population.

The second component of $DECI$ is a measure of the growth in expenditure per episode, $C_{d,t}$. The value $C_{d,t}$ is calculated by dividing total expenditures of disease d by the number of episodes of disease d in period t. The medical-care expenditure index (MCE) is a measure of the growth in medical-care expenditures for the treatment of a disease episode. Denoting $C_{d,0}$ as the average expenditure per episode in the base period, the MCE for disease d is the ratio of the two measures:

$$MCE_{d,t} = \frac{C_{d,t}}{C_{d,0}}$$

Since this index controls for the health of the individual, it may be viewed as measuring the cost of treatment. Using these equations, it follows that $C^*_{d,t} = c_{d,t} \cdot prev_{d,t}$. From this, we can see that the $DECI_{d,t}$ may be

decomposed into its two components, which include the episode-based index, $MCE_{d,t}$ and the treated prevalence disease index, $PREV_{d,t}$. The multiplicative decomposition is $DECI_{d,t} = MCE_{d,t} \cdot PREV_{d,t}$.

The *DECI* will rise if there is either an increase in the $PREV_{d,t}$ or an increase in the $MCE_{d,t}$. These two components of expenditure capture distinct elements of cost growth. Changes in the treated prevalence may capture the changing health of the population or the growing awareness and treatment of a condition. The *MCE* may be viewed as the change in price for treating the disease. Assuming that the quality of the underlying treatment mix remains constant, this treatment price index may be used as a deflator to determine the change in output in the health sector. Aggregate versions of these indexes are formed using the Laspeyres index formula with 2003 as the base period.[8]

DATA

We use retrospective claims data for a sample of commercially insured enrollees from the MarketScan® Research Database from Truven Health for the years 2003 to 2007. The data are from health insurers and large employers containing medical and drug data for several million commercially insured individuals. Each observation in the data corresponds to a line item in an "explanation of benefits" of a medical claim.

We limit our sample to enrollees in noncapitated plans and include only those with drug benefits because drug purchases will not be observed for individuals without coverage. The MarketScan database tracks claims from all providers using a nationwide convenience sample of enrollees. To ensure that we observe all of an individual's expenditures for a year, we limit the sample to those enrollees that have a full year of continuous enrollment. Each enrollee has a unique identifier and includes age, sex, and region information in addition to their medical care claims information.

The claims data were processed using the Symmetry ETG from Optum and the MEG from Truven Health. Both groupers assign each claim to a particular episode group. In addition to applying groupers to the data, we also applied the mapping of the ICD-9 disease codes to the CCS codes. Demographic weights are applied to the data to make it representative of national totals and adjust for differences in age, sex, and region across years.[9]

Descriptive Statistics

The three episode-based and encounter-based disease classification systems considered here, ETG, CCS, and MEG, differ in their share of expenditures allocated to diseases. For comparison in 2007, the ETG grouper allocates 87 percent of expenditures; the MEG grouper allocates 83 percent of expenditures; and the CCS encounter-based approach allocates 84 percent of expenditures. Those claims that are not assigned to a disease episode are considered ungrouped or unallocated.

The three different methods assign each claim line to distinct diseases. Table 1 shows the top ten ETG diseases based on expenditures for 2007. The first column shows the ETG category, and the second column shows the total expenditures allocated to that ETG category. The third column shows the top four MEG categories that correspond to those ETG expenditures, ranked by MEG expenditure share. For example, the first ETG category listed, ischemic heart disease: the MEG grouper assigns 53.6 percent of those expenditures to the MEG category "Angina Pectoris, Chronic Maintenance," 19.8 percent to "Acute Myocardial Infraction," 9.2 percent to ungrouped, and 4.0 percent to "Arrhythmias." There are a few general points worth noting. First, it is clear that the MEG and ETG categories are distinct. Consequently, many of the ETG disease categories are spread across multiple, although often related, categories, which complicates any direct comparison. Second, there are many instances where the ETG grouper assigns a disease, but the MEG does not. The patterns observed in Table 1 are also observed when looking at mappings among other categories (see Tables A1 and A2 of the Appendix).

RESULTS

Table 2 lists the name of the method in the first column, followed by the aggregate 5-year growth rates for *DECI*, *PREV*, and *MCE*. The first six results apply a grouper approach, the seventh applies an encounter-based approach, and the eighth and ninth apply a person-based approach. The aggregate *DECI* is essentially a demographically adjusted expenditure per capita index. The key factor affecting the difference in the DECI is the share of expenditures allocated over time. The table shows that the *DECI* is similar across approaches, ranging from 1.22 to 1.25. The outlier is the encounter-based approach, method 7, which has a *DECI* growth of, 1.22, which is likely driven by the unallocated drug claims for this method.

Table 1: Top Ten ETG Disease Categories Mapped to MEG Classification

ETG	*Total Expenditures (Millions)*	*MEG*	*% Exp. MEG on ETG*	*Cumulative % Exp. of MEG on ETG*
Ischemic heart disease	$24,954	Angina Pectoris, Chronic Maintenance	53.6	53.6
Ischemic heart disease		Acute Myocardial Infarction	19.8	73.4
Ischemic heart disease		Ungroupable Medical Claims	9.2	82.6
Ischemic heart disease		Arrhythmias	4.0	86.6
Pregnancy, with delivery	$23,557	Delivery, Vaginal	88.3	88.3
Pregnancy, with delivery		Ungroupable Medical Claims	7.0	95.3
Pregnancy, with delivery		Ante- and Postpartum Complications	1.5	96.7
Pregnancy, with delivery		Delivery, Cesarean Section	1.0	97.8
Joint degeneration, localized—back	$19,846	Osteoarthritis, Lumbar Spine	31.2	31.2
Joint degeneration, localized—back		Intervertebral Disc Disorders: Lumbar and Lumbosacral	24.4	55.6
Joint degeneration, localized—back		Other Spinal and Back Disorders: Low Back	16.4	72.0
Joint degeneration, localized—back		Ungroupable Medical Claims	15.5	87.5
Diabetes	$19,189	Diabetes Mellitus Type 2 and Hyperglycemic States Maintenance	44.3	44.3
Diabetes		Ungroupable Medical Claims	23.3	67.5
Diabetes		Diabetes Mellitus Type 1 Maintenance	18.2	85.7
Diabetes		Diabetes Mellitus with Complications	2.7	88.4
Hypertension	$16,574	Essential Hypertension, Chronic Maintenance	60.5	60.5
Hypertension		Ungroupable Medical Claims	16.3	76.8
Hypertension		Other Respiratory Symptoms	3.3	80.1
Hypertension		Angina Pectoris, Chronic Maintenance	2.5	82.6
Routine exam	$14,765	Encounter for Preventive Health Services	84.6	84.6
Routine exam		Ungroupable Medical Claims	8.4	93.0

continued

Table 1. *Continued*

ETG	*Total Expenditures (Millions)*	*MEG*	*% Exp. MEG on ETG*	*Cumulative % Exp. of MEG on ETG*
Routine exam		Factors Influencing Health Status	3.8	96.8
Routine exam		Other Cardiovascular Symptoms	0.2	97.0
Malignant neoplasm of breast	$14,168	Neoplasm, Malignant: Breast, Female	78.3	78.3
Malignant neoplasm of breast		Ungroupable Medical Claims	13.8	92.0
Malignant neoplasm of breast		Neoplasm, Benign: Breast	1.5	93.5
Malignant neoplasm of breast		Factors Influencing Health Status	1.4	95.0
Mood disorder, depressed	$10,327	Depression	64.6	64.6
Mood disorder, depressed		Ungroupable Medical Claims	12.8	77.4
Mood disorder, depressed		Generalized Anxiety Disorder	4.3	81.8
Mood disorder, depressed		Bipolar Disorder—Major Depressive Episode	2.8	84.6
Joint degeneration, localized—neck	$9,051	Osteoarthritis, Cervical Spine	27.7	27.7
Joint degeneration, localized—neck		Intervertebral Disc Disorders: Cervical	26.9	54.5
Joint degeneration, localized—neck		Ungroupable Medical Claims	16.9	71.4
Joint degeneration, localized—neck		Other Spinal and Back Disorders: Cervical	14.4	85.8
Nonmalignant neoplasm of female genital tract	$8,898	Neoplasm, Benign: Uterus (Leiomyomas)	27.4	27.4
Nonmalignant neoplasm of female genital tract		Dysfunctional Uterine Bleeding	15.0	42.4
Nonmalignant neoplasm of female genital tract		Ungroupable Medical Claims	13.9	56.3
Nonmalignant neoplasm of female genital tract		Other Disorders of Female Genital System	13.5	69.8

Notes. Table shows the top 10 ETG disease categories based on expenditure share for 2007. Four of the corresponding MEG categories are shown in order of highest allocation to lowest allocation.

The more important difference across methods is the growth rate of the *MCE*. Method 1 applies the ETG grouper with severity adjustment, which produces relatively slow *MCE* growth of 1.14, accounting for a little

Table 2: Growth Decomposition for 2003 to 2007—Grouper-, Encounter- and Person-Based Approaches

Method	*DECI*	*PREV*	*MCE*
Grouper			
1. ETG Severity-Adjusted (Baseline)	1.250	1.103	1.139
2. ETG Not Severity-Adjusted	1.250	1.098	1.147
3. MEG Severity-Adjusted	1.240	1.074	1.163
4. MEG Not Severity-Adjusted	1.240	1.071	1.164
5. ETG Annual Episodes (Not Severity-Adjusted)	1.250	1.098	1.146
6. ETG—Episodes by MPC Class	1.250	1.077	1.165
Encounter			
7. Primary Diagnosis—CCS Disease Categories	1.224	1.028	1.178
Person—Regression Analysis			
8. ETG Severity-Adjusted	1.233	1.091	1.131
9. CCS Disease Category	1.240	1.079	1.156

Notes. The regression-based approaches of 8 and 9 apply log linear regressions. The expenditures allocated to the intercept are considered unallocated and are excluded from the analysis.
DECI, Demographically Adjusted Expenditure Per Capita Index; MCE, Medical Care Expenditure Index; PREV, Treated Prevalence Index.

over half of expenditure growth. Method 2 applies the ETG grouper without severity adjustment and shows slightly faster growth. Method 3 and 4 apply the MEG grouper, severity-adjusted and not severity-adjusted, respectively. The MEG *MCE* indexes grow slightly faster than the ETG indexes, and the MEG severity adjustments have little measurable effect on the aggregate indexes.

Methods 1 through 4 are all based on the concept of an episode, as defined by the MEG and ETG grouper, where a person may have multiple episodes in a single year. For some applications, it may be more appropriate to consider disease expenditures per person in a year or quarter, rather than expenditures per episode. For instance, Dunn et al. (2012) measure expenditure per person on a disease treatment on an annual basis. To investigate if "annualizing" episodes has an effect on the estimates, method 5 repeats method 2, but redefines an episode to correspond to a calendar year. The results are nearly identical to the results of method 2.

Next, we examine the effect of aggregation into more broadly defined disease categories. The broadly defined categories include very heterogeneous conditions (e.g., hypertension and heart disease), but they may be necessary for certain applications (e.g., small sample sizes). We broaden the 456 ETG categories by grouping them into 21 Major Practice Categories

(MPC). The results are reported in method 6. We find that relative to methods 1 and 2, the MCE growth from method 6 is slightly faster. In this case, the method 6 MCE may actually be picking up increases in treated prevalence (e.g., a person with hypertension develops heart disease). However, this impact is only a small percentage of the total growth, since the MCE for this approach grows only slightly faster than the previous alternatives.

Method 7 presents the encounter-based method, which is quite different from the other approaches, since it does not classify drug claims. Consequently, the growth rates are quite distinct from methods 1 through 6.[10] A major reason for this stark contrast is that much of the growth in treated prevalence for methods 1 through 6 is for diseases that are typically treated with prescription drugs, such as hyperlipidemia (i.e., high cholesterol), hypertension (i.e., high blood pressure), and diabetes. This highlights a potentially serious problem with this approach.

Methods 1 through 7 rely on a mapping of each claim line to a particular disease category. In contrast, the last two methods in Table 2 rely on the person-based method. Method 8 uses severity-adjusted ETG disease classification in the regression model. Although the method is quite distinct from method 1, the results are similar. One distinction is that the *DECI* is lower, due to a change in the share of expenditures allocated over time, but the *MCE* estimate does not change. Finally, method 9 also applies the person-based approach, but uses the CCS categories instead of the ETG categories. Unlike the encounter-based approach of method 7, the regression using the CCS condition categories is able to allocate drug expenditures. The *MCE* growth using CCS categories and applying the person-based approach appears similar to the other estimates.

Given the vast differences in the methods, the similarities in the MCE are striking. In particular, applying the methodologies from Table 2 (excluding method 7 that removes drug claims), the table shows MCE growth rates ranging from 1.131 (CAGR 3.1 percent) to 1.165 (CAGR 3.9 percent), which implies a difference in CAGR of about 0.8 percent across methodologies. A further exploration of additional alternatives in the Appendix shows a similarly tight range across even more person-based alternatives.

Disease Category Comparisons

While the aggregate disease-price growth patterns are similar, the patterns are not necessarily similar for specific disease categories. To explore these differences, Table 3 presents the disease-price growth estimates at the

disease category level for three different methods. Each of the methods compared in Table 3 are based on ETGs, then aggregated into broader MPCs. The first set of estimates applies the episode-based method (Table 2, method 1). The estimates show that trends for each disease category may be quite different and may not follow the general trend. The second set of results is from the person-based approach using severity-adjusted ETGs in the regression (Table 2, method 8). Generally, the results from the first and second set of estimates are positively correlated. The correlations among the indexes are all above 88 percent for the major categories (i.e., those with more than 3 percent of total spending). However, despite the positive correlation, there are many clear differences. For instance, the prevalence growth for the endocrinology condition category using the person-based method is 10 percentage points lower than the estimates using the episode-based approach.

The third column of results applies the episode-based method, but it uses the broad MPC categories (Table 2, method 6). Again, the results are positively correlated with the other two approaches, but the results are clearly different. In particular, the *MCE* price growth tends to be faster than the other two methods, following the pattern of the aggregate results. One important exception is the *MCE* growth rate for Cardiology, which actually grows slower than the other two methods.

Despite these differences, some general patterns emerge at the disease category level. Specifically, for cardiology conditions, we see that price growth and prevalence growth are quite slow. By contrast, for preventative services, we see that expenditure growth is increasing rapidly, due to both higher treated prevalence and higher prices. The growth rate for endocrinology conditions appears to be much faster than the overall trend, likely due to diseases like diabetes and high cholesterol. It is more challenging to compare methods across CCS, ETG, and MEG classifications, since the categorizations are distinct. However, an attempt at a comparison is made in the Appendix. In general, we find that differences across methods are more pronounced at the disease level, although similar qualitative patterns often emerge across methods.

GUIDELINES

Based on the estimates presented above as well as additional analysis, we have derived certain guidelines for allocating expenditures for disease-based indexes.[11]

Table 3: ETG-Based Decompositions by Major Practice Category

		1			2			3		
		Episode-Decomp.			Person-Based Decomp.			Episode-Decomp.		
		Table 2. Method (1)			Table 2. Method (8)			Table 2. Method (6)		
Description	*Expenditure Share (%)*	*DECI*	*PREV*	*MCE*	*DECI*	*PREV*	*MCE*	*DECI*	*PREV*	*MCE*
Orthopedics and rheumatology	14.0	1.300	1.119	1.165	1.240	1.079	1.151	1.280	1.054	1.215
Cardiology	12.0	1.108	1.043	1.059	1.113	1.042	1.069	1.101	1.074	1.025
Endocrinology	8.9	1.343	1.273	1.074	1.280	1.177	1.097	1.329	1.155	1.150
Gastroenterology	8.4	1.288	1.111	1.169	1.277	1.088	1.168	1.281	1.067	1.200
Preventive and administrative	6.4	1.625	1.290	1.261	1.685	1.284	1.298	1.611	1.256	1.283
Otolaryngology	5.7	1.140	1.027	1.111	1.201	1.076	1.120	1.156	1.002	1.154
Gynecology	5.5	1.215	1.006	1.203	1.160	1.022	1.132	1.223	1.007	1.214
Dermatology	4.9	1.276	1.084	1.178	1.253	1.067	1.168	1.270	1.047	1.213
Psychiatry	4.8	1.261	1.130	1.115	1.248	1.128	1.100	1.263	1.134	1.114
Neurology	4.7	1.299	1.097	1.187	1.248	1.076	1.167	1.283	1.045	1.228
Pulmonology	4.6	1.162	1.012	1.158	1.090	0.994	1.119	1.158	0.936	1.237
Obstetrics	4.5	1.259	1.076	1.171	1.243	1.072	1.161	1.262	1.043	1.210
Urology	3.1	1.224	1.118	1.112	1.190	1.090	1.107	1.220	1.046	1.166
Isolated signs and symptoms	2.9	1.115	1.002	1.112	1.368	1.168	1.171	1.334	1.168	1.142
Hematology	1.9	1.278	1.112	1.149	1.296	1.089	1.185	1.277	1.071	1.192
Hepatology	1.7	1.118	0.993	1.120	1.052	0.985	1.061	1.119	1.016	1.102
Ophthalmology	1.6	1.189	1.127	1.054	1.184	1.128	1.061	1.183	1.094	1.082
Nephrology	1.0	1.329	1.485	0.903	0.801	1.347	0.634	1.428	1.219	1.171

continued

Table 3. *Continued*

		1			*2*			*3*		
		Episode-Decomp.			*Person-Based Decomp.*			*Episode-Decomp.*		
		Table 2. Method (1)			*Table 2. Method (8)*			*Table 2. Method (6)*		
Description	*Expenditure Share (%)*	*DECI*	*PREV*	*MCE*	*DECI*	*PREV*	*MCE*	*DECI*	*PREV*	*MCE*
Neonatology	1.0	1.317	1.140	1.169	1.210	1.127	1.082	1.312	1.043	1.257
Infectious diseases	1.0	1.380	1.154	1.176	1.267	1.118	1.132	1.373	0.919	1.494
Late effects, envir. trauma and pois.	0.8	1.273	0.957	1.339	1.085	0.942	1.153	1.259	0.960	1.311
Chemical dependency	0.7	1.415	1.377	1.057	1.630	1.426	1.164	1.404	1.563	0.898
Total expenditures (in billions) 2007	$604									

Notes. Expenditures used to calculate expenditure share are calculated from the person-based decomposition, Table 2, method (8).
DECI, Demographically Adjusted Expenditure Per Capita Index; MCE, Medical Care Expenditure Index; PREV, Treated Prevalence Index.

General Guidelines

1. Given that some differences exist across methods, it is useful to check if key patterns are robust across methodologies.
2. Severity adjustment is typically desirable as it helps account for the heterogeneity across episodes or demographics, which may be important in some instances (e.g., less severe patients seeking treatment). We find that severity adjustment often has a small, but important effect on disease-price growth. Severity adjustment tends to lower the disease-price growth rate, but there are exceptions (e.g., the MEG).

Application-Specific Guidelines

1. The encounter-based method is problematic due to its naive approach to comorbidities and expenditures, and the fact that not all encounters have associated diagnosis codes assigned (e.g., drug claims). However, if the researcher has diagnosis codes on all their claims (for example, MEPS links drugs with encounters or if only using medical claims), the encounter-based approach may be preferred due to its simplicity.
2. An attractive feature of the CCS classification approach is that it is freely accessible to other researchers and results can be better compared to previous studies.
3. The person-based method is problematic if the researcher wants to allocate expenditures to particular service categories (e.g., hospitals, doctor offices, and prescription drugs). In this case, an episode-based approach may be a good alternative. However, the person-based method is an attractive alternative to the encounter-based method since it addresses the comorbidity problem and can assign expenditures that do not have diagnosis codes.
4. The CCS classification combined with the person-based approach is a good combination as it may be replicated by researchers with commonly available statistical packages. In addition, MCE estimates based on this approach fall in the middle of our range of estimates.
5. Researchers sometimes require a certain number of encounters (e.g., three or more) to classify a disease episode. The concern is that a single encounter may contain a mistaken diagnosis. We find that including diseases with only three or more encounters has minimal impact

on aggregate disease-price inflation (see method 7 in Table A3 of the appendix).

6. Commercial episode groupers may not be appropriate for all researchers or applications since they are expensive and use proprietary algorithms. Our results indicate that less expensive methodologies exist that produce similar results. For instance, annualizing claims, rather than studying episodes, appears to have little effect on MCE growth rates.
7. If groupers are used, they should be applied symmetrically across years (e.g., 1 year at a time) when analyzing time trends. This is necessary to ensure that asymmetric information about individuals in the sample do not drive trends.[12] For instance, one would expect to allocate a greater share of claims for individuals that have a long history in the data, relative to someone with a short history.

CONCLUSION

Most of the methodologies appear to give reasonably similar results. The ideal methodology will depend on the analyst's data and question. For example, commercial payers, who are either attempting to improve the quality and efficiency of medical interventions or set bundled payment rates, the episode-based approach may be preferred to the person-based approach as they allow dollars to be linked easily to the underlying services. For researchers interested in understanding spending trends, the person-based approach using CCS categories may be sufficient.

There is still a significant amount of additional research needed on the topic of disease expenditure allocation. Future research should explore these questions for other data sets, populations, and time periods. Researchers should also continue exploring new algorithms for allocating expenditures across diseases. Finally, it will be important to connect disease-price changes to quality changes. In this case, the most appropriate methodology may depend on how quality and disease expenditure information are linked.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the

Bureau of Economic Analysis, the Federal Reserve Bank of San Francisco, or the Board of Governors of the Federal Reserve System. The Bureau of Economic Analysis has sponsored work on this project. They support the work and reviewed the paper prior to distribution, but this had no substantive impact on the content of this paper.

Disclosures: None.

Disclaimers: None.

NOTES

1. See Berndt et al. (2000), Rosen and Cutler (2007), National Research Council (2010), and Aizcorbe et al. (2012a).
2. One exception is Hall and Highfill (2014), which also looks at methodologies for measuring disease prices in the Medicare population using survey data. Their data have a much smaller sample size than the claims data studied here, which raises unique issues when selecting methods. This research is highlighted in the "guidelines" section of this paper.
3. Dunn et al. (2012) show that the standard errors on inflation estimates using large claims data tend to be very small.
4. Aizcorbe and Nestoriak (2011) and Dunn, Liebman, and Shapiro (2015).
5. See Rosen et al. (2013) and http://innovation.cms.gov/initiatives/bundled-payments/.
6. See https://www.optum.com/providers/analytics/health-plan-analytics/symmetry/symmetry-episode-treatment-groups.html and http://truvenhealth.com/your-healthcare-focus/health-plan/medical-episode-grouper-health-plan.
7. Starr, Dominiak, and Aizcorbe (2014) and Roehrig et al. (2009).
8. See the appendix of this paper for additional discussion regarding the application of the different approaches.
9. Specifically, enrollees in each year are assigned weights to match the age-sex-region population of the U.S. commercially insured population in 2007.
10. Prescription drugs account for 16 percent of spending, and rise faster than average, at a rate of 5.8 percent per year.
11. The guidelines presented here are based on large commercial claims data. Additional considerations may be important for smaller sample sizes, see Hall and Highfill (2014).
12. This is demonstrated in Appendix table A7.

REFERENCES

Aizcorbe, A., and N. Nestoriak. 2011. "Changing Mix of Medical Care Services: Stylized Facts and Implications for Price Indexes." *Journal of Health Economics* 30 (3): 568–74.

Aizcorbe, A., E. Liebman, D. Cutler, and A. Rosen. 2012a. "Household Consumption Expenditures for Medical Care: An Alternate Presentation." *Survey of Current Business* 92 (June): 34–48.

Aizcorbe, A., E. Liebman, S. Pack, D. Cutler, M. Chernew, and A. Rosen. 2012b. "Measuring Health Care Costs of Individuals with Employer-Sponsored Health Insurance in the U.S.: A Comparison of Survey and Claims Data." *Statistical Journal of the IAOS* 28: 43–51.

Berndt, E., D. Cutler, R. Frank, Z. Griliches, J. Newhouse, and J. Triplett. 2000. "Medical Care Prices and Output." *Handbook of Health Economics* 3: 119–80.

Bradley, Ralph. 2013. "Feasible Methods to Estimate Disease Based Price Indexes." *Journal of Health Economics* 32 (3): 504–14.

Dunn, A., E. Liebman, and A. Shapiro. 2014. "Developing a Framework for Decomposing Medical-Care Expenditure Growth: Exploring Issues of Representativeness." In *Measuring Economic Sustainability and Progress, NBER Book Series Studies in Income and Wealth*, edited by D. Jorgenson, S. Landfeld, and P. Schreyer, pp. 545–74. Chicago, IL: University of Chicago Press.

———. 2015. "Implications of Utilization Shifts on Medical-Care Price Measurement." *Health Economics* 24: 539–57.

Dunn, A., E. Liebman, S. Pack, and A. Shapiro. 2012. "Medical Care Price Indexes for Patients with Employer-Provided Insurance: Nationally Representative Estimates from MarketScan Data." *Health Services Research* 48 (3): 1173–90.

Hall, A., and T. Highfill. 2014. "Calculating Disease-Based Medical Care Expenditure Indexes for Medicare Beneficiaries: A Comparison of Method and Data Choices. Measuring and Modeling Health Care Costs." *BEA Working Paper.* Washington, DC: Bureau of Economic Analysis.

Hornbrook, M. C., A. V. Hurtado, and R. E. Johnson. 1985. "Health Care Episodes: Definition, Measurement and Use." *Medical Care Review* 42 (2): 163–218.

Keeler, E. B., J. E. Rolph, N. Duan, J. M. Hanley, and W. G. Manning. 1982. *The Demand for Episodes of Medical Treatment: Interim Results from the Health Insurance Experiment.* The RAND Corporation, R-2829-HHS, December.

MaCurdy, T., J. Kerwin, and N. Theobald. 2009. "Need for Risk Adjustment in Adapting Episode Grouper Software to Medicare Data." *Health Care Financing Review* 30 (4): 33–46.

National Research Council, and the Panel to Advance a Research Program on the Design of National Health Accounts, Committee on National Statistics, Division of Behavioral and Social Sciences and Education. 2010. *Accounting for Health and Health Care: Approaches to Measuring the Sources and Costs of Their Improvement.* Washington, DC: National Academies Press.

Roehrig, C., and D. Rousseau. 2011. "The Growth in Cost per Case Explains Far More of US Health Spending Increases than Rising Disease Prevalence." *Health Affairs* 30 (9): 1657–63.

Roehrig, C., G. Miller, C. Lake, and J. Bryant. 2009. "National Health Spending by Medical Condition, 1996-2005." *Health Affairs* 28 (2): w358–w367.

Rosen, A., and D. Cutler. 2007. "Measuring Medical Care Productivity: A Proposal for U.S. National Health Accounts." *Survey of Current Business* 87 (June): 54–8.

———. 2012. "Challenges in Building Disease-Based National Health Accounts." *Medical Care* July: 47 (Supplement 1): S7–S13.

Rosen, A., E. Liebman, A. Aizcorbe, and D. Cutler. 2012. "Comparing Commercial Systems for Characterizing Episodes of Care." *BEA Working Paper.* Washington, DC: Bureau of Economic Analysis.

Rosen, A., A. Aizcorbe, A. Ryu, N. Nestoriak, D. Cutler, and M. Chernew. 2013. "Policy Makers Will Need a Way to Update Bundled Payments that Reflects Highly Skewed Spending Growth of Various Care Episodes." *Health Affairs* 32 (5): 944–51.

Starr, M., L. Dominiak, and A. Aizcorbe. 2014. "Decomposing Growth in Spending Finds Annual Cost of Treatment Contributed Most to Spending Growth, 1980–2006." *Health Affairs* 33 (5): 823–31.

Trogdon, J., E. Finkelstein, and T. Hoerger. 2008. "Use of Econometric Models to Estimate Expenditure Shares." *Health Services Research* 43 (4): 1442–52.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Data S1: 1. Details in Applying the Various Allocation Methods.
2. Alternative Allocation Methods.
3. Differences in Disease Definitions.