

# Measuring Healthcare Resources Using Episodes of Care

Mark C. Rattray, MD

*President and Executive Consultant, CareVariance LLC*

## INTRODUCTION

The past several years have brought intensive scrutiny to all aspects of our healthcare system's performance. While other industries have widely deployed continuous improvement cycles and adopted waste-reducing lean production methods, the U. S. healthcare sector lags far behind. The growing cost of healthcare in the U.S. in the face of increasingly documented lackluster quality and daily harmful and preventable errors have other industries asking, "*Where is the healthcare sector's commitment to continuous quality improvement?*" In addition, given the impact of resource waste on the affordability of and access to healthcare, monitoring the use and appropriateness of health care services is essential. Private and governmental payers of healthcare are demanding that health plans and healthcare providers address variation in quality and in the use of care resources in systematic ways. This means measuring what is important, identifying significant, actionable variation, and motivating improvement. Motivation includes providers' professionalism, (a component of which is the desire to improve when credible, actionable information is available), differential pay for differential performance, and increased transparency of measurement results.

Transparency of healthcare quality and cost has become a national mandate. We see this in an Executive Order of the President, Congress' directive to CMS to engage in value-based

purchasing of healthcare services, and an increasing percentage of initiatives by the nation's Governors and employers who have committed to preferentially contracting with health plans that provide differential provider performance information to their members.

A growing number of healthcare's stakeholders seek to quantify and reduce waste — healthcare resource use that yields insignificant or no improvements to health or health outcomes. General estimates of resource waste in the U.S. healthcare system range upwards from 30%. Identification and measurement of variation, including that among clinicians' practices, are essential precursors to improvement and transparency.

There is an increasing call for healthcare to implement other industries' best practices for reduction of wasted resources. Acknowledging that healthcare delivery is unique, and that unintended consequences must be carefully considered, the crisis of affordable healthcare coverage demands that we deploy a system of measurement and reporting.

This paper seeks to explain the episode of care concept and describe how episode of care measurement tools are commonly used to provide a view of the condition-specific use of health care resources.

---

---

## HEALTHCARE AS A LONGITUDINAL EVENT

Historically, variation of medical resource use has been measured in terms of comparative amounts of specific services, or what has been referred to as “isolated units of care.”<sup>1</sup> Payers have typically looked at medical resource use in actuarial terms with traditional monthly, quarterly, and annual time frames. Acting on variation in physician practices, plans usually focus on high and low outlier resource use at the level of these specific care components. An example of this type of variation reporting uses terms such as “office visits per member per month (pmpm)” or “bed days per 1000 members per year (days/1000)”. Physicians are often provided reports breaking down the frequency distribution compared to peers of the most common office visit (Evaluation and Management, or E&M) codes that are billed. The codes reflect degrees of visit complexity and hence payment. These variation reports also typically contain costs, not just counts, of resource usage. Costs may be represented in myriad categories, rolling up into a total “costs per member per month.”

Solon and colleagues pointed out the need for a patient and condition-centric approach to resource use variation, further advancing the emerging “episode of care” (EOC) analytic unit.<sup>2</sup> Other researchers, Hornbrook et al. discussed the concept in 1985, stating, “Health care differs from other commodities because it is typically provided in a series of separate but related delivery sessions.”<sup>3</sup>

In 1999, Rosen and Mayer-Oaks<sup>4</sup> elaborated on the EOC concept and compared and contrasted the tools, known as “grouper” software, in use at that time. They provided a more detailed discourse of the grouper approach to the largely clinical quality improvement audience they were addressing.

As information technology resources – specifically data storage, processor speed, and computer programming – became more affordable, more automated approaches to variation analysis became available. Payers are increasingly able to store many years of claims data in data warehouses, and enhanced analytic algorithms have been developed to provide new perspectives on comparative resource use at the

physician practice level. Supported over the last decade by improved technology, the shift from solely analyzing what Solon described as “isolated units of care” to “episodes of care” has enabled a focus on individual patients, their specific conditions being treated, and the variations in that treatment.

Today, using groupers’ advanced algorithmic techniques married with electronic data, the onset of treatment and conclusion of treatment may be identified for a particular condition. The software then aggregates treatment costs utilized from beginning to end, calculating the sum of the costs paid by the patient and third parties to treat that condition.

There are various current EOC methodologies originating from commercial software programs produced by several private companies. Entities such as the Centers for Medicare and Medicaid Services (CMS), the Medicare Payment Assessment Commission (MedPAC), the AQA, the National Quality Forum (NQF), and the National Committee for Quality Assurance (NCQA) are involved in addressing the EOC concept and/or investigating to what extent standardization is possible. *De novo* construction of a public domain EOC approach for highly prevalent conditions is also under consideration.

## OVERVIEW OF EPISODE OF CARE ANALYSIS AND REPORTING

As episode of care analysis is increasingly being deployed within the healthcare industry, it is important that clinicians and other healthcare stakeholders understand the major underlying concepts. What follows is an overview of the process, omitting extensive detail. These are the topics addressed:

1. Episode groupers
2. Measurement period
3. Obtaining data extracts
4. Choice of grouper output
5. Data preparation for aggregation and grouper processing
6. Grouper processing
7. Post-grouper information processing
8. Reporting

---

---

## 1. EPISODE GROUPERS

Episode groupers are software programs that create episodes of care from administrative electronic data. These programs sift through millions of claims for reimbursement submitted to a health care payer by health care providers and reconstruct the data into instances of specific patients receiving care for specific conditions. These tools are called groupers because they group billed services from all healthcare providers together into patient care episodes.

This requires sophisticated software, at the core of which are proprietary mapping schema wherein billing and diagnosis codes for services (CPT<sup>®5</sup>, ICD-9<sup>6</sup>, HCPCS<sup>7</sup>, NDC<sup>8</sup>, hospital revenue codes) are assigned to specific episode types (e.g., bronchitis, acute myocardial infarction, etc.). Groupers contain hierarchical logic that further guide the process of parsing the data into episodes. The software usually has the ability to track multiple and concurrent episodes as well as comorbidities.

The fundamental concept underlying episode construction is that most paid services, across time and settings, can be analyzed and “grouped” into specific clinical conditions. For example, a patient with diabetes and congestive heart failure (CHF) can have a routine office visit with lab work (for diabetes monitoring), an unscheduled office visit and a chest x-ray for shortness of breath (CHF-related), an ER visit and a 3 day hospitalization (for worsening CHF), a post discharge visit to a cardiologist (for CHF follow-up), then two office visits to a PCP (for diabetes management and medication adjustment), all over a 3 month period. An episode grouper program will mine the entire stream of claims associated with such a patient and define both an “episode” of CHF, as well as capture elements related to a concurrent episode of diabetes.

The two most commonly used episode groupers are Episode Treatment Groups<sup>9</sup> (ETGs) by Symmetry / Ingenix, a subsidiary of UnitedHealth Group, and the Medstat Episode Grouper<sup>10</sup> by Thomson Healthcare.

## 2. MEASUREMENT PERIOD

The measurement period reflects the time frame of services included in the grouper analysis. There are multiple considerations involved in the choice of the measurement period. The longer the measurement period, the greater the number of completed episodes and the subsequent statistical power of measurement. Measurement periods of 18 to 24 months are typical. Longer measurement periods are less sensitive to improvement efforts. Shorter, more recent, measurement periods provide more timely measurement but yield fewer episodes for analysis.

## 3. OBTAINING DATA EXTRACTS

Grouper analysis requires extracts of data from payers such as health plans that maintain claims billing and payment records, including the associated billing and diagnosis codes assigned. If multiple entities pay for patient services (for example, a pharmacy benefit manager might process pharmacy claims), data extracts from all payers may be necessary to include all patient services in the analysis. Such detail may not be available in all instances – pharmacy claims have occasionally been difficult to obtain, and mental health services information may be excluded for privacy reasons.

## 4. CHOICE OF GROUPEUR OUTPUT

As payers have been the principal users of groupers for physician practice measurement, the most frequent choice for grouper output is total episode cost, reflecting the total amount paid by the patient and his/her insurer (collectively called “allowed costs”) for the care within an episode. The episode costs are influenced by both the amount of resources used and the amounts paid for each of the resources. Resource-specific payment rates are established by contracts between the payer and the provider of services. Many plans currently supply their providers with reports on patient-specific episode costs as well as average episode costs, the latter of which are often compared to the overall peer average.

Increasingly, data are being aggregated across payers through certain local and regional initiatives. In addition, national accrediting bodies are asking for plan-level reporting. In such cases, costs may need to be converted to “standardized costs”

---

---

and in effect, normalized. A standard price for a specific service may be substituted for the actual price. Translation of actual price into standard prices usually occurs prior to grouping. When one uses standardized prices for input, the output reflects a measure of the quantity and types of services used per episode. Using standardized prices when measuring physician practices emphasizes the practices' overall patterns of resource use and deemphasizes variations in the costs of those resources. This approach normalizes, for example, variations in a payer's contracted rates for hospitals.

## 5. DATA PREPARATION FOR AGGREGATION AND GROUPE PROCESSING

The data to be processed by the grouper must conform to grouper input specifications. This requires that the data be "mapped" to the required grouper input data fields. This presents little challenge when data fields are standardized in content and nomenclature. Data sets that may be used as inputs to the grouper, however, can often vary in their architecture or degree of completeness. When data fields are non-standardized, careful analysis is required to create, where possible, "data crosswalks" between the data fields and the required grouper input fields. If the data are derived from multiple systems, such as the case with aggregation of data across payers, additional labor-intensive analyses and crosswalks may be required to provide reliable grouper input.

## 6. GROUPE PROCESSING

The first grouper step is the identification of the medical and pharmacy services provided to specific patients. The capability to perform this step relies on the quality of the preparation of the administrative claims and other data prior to entry into the grouper, the thoroughness of capture of all services provided to the patient, and the submission of the correct procedure and diagnosis(es) codes corresponding to the services rendered.

In the course of the measurement period, patients may have none, one, or multiple episodes of care. To create the episodes, the software looks at the data chronologically. It searches for the first service that signals the onset of a new episode (usually a service provided by a physician and involving a definitive

diagnosis, such as a physician E&M visit with a new diagnosis code). It then looks later in time for a period during which no further related services were rendered for that condition, called a "clean window" (which varies by condition).

The software then aggregates all medical and pharmacy services for that patient with diagnosis codes related to the episode-triggering code from all providers of services associated with that episode, from the onset of the episode to the last service provided before the clean window criteria have been met.

Other episode types that may be concurrent with this episode for the same patient are handled in a similar manner. Given that chronic conditions such as diabetes may be lifelong, a one-year episode length is typically used for such conditions. Also of note, grouper software may allow user-defined variations in the grouping calculations, such as how to handle those episodes with unusually low or high resultant episode costs or resource use (often referred to as outlier episodes).

The key results from the episode grouping process are: (1) the identification of the unique episode of care, if any, that each medical and pharmacy service is assigned to; and (2) a characterization of each episode's cost, or if price normalized data are used, a compilation of types and numbers of resources deployed in the treatment.

## 7. POST-GROUPE INFORMATION PROCESSING

**Attribution of episodes.** Once an episode is created, a process is used to "attribute" that episode to one or more care providers. Post-grouper logic may be applied to attribute or assign an episode to one or more clinicians (or the clinicians' practice group) responsible for the decisions related to the deployment of resources within the episode. Decisions around attribution methods are usually made by the measuring entity, not the episode grouper vendor.

Most episodes involve a single managing clinician. For episodes involving multiple managing clinicians, payers have historically assigned episodes to the physician that has provided the greatest amount of professional services within

the episode, excluding “non-managing” physician types such as pathologists, diagnostic radiologists, etc. Both the Bridges to Excellence / Leapfrog Group white paper<sup>11</sup> on provider efficiency as well as NCQA in its draft HEDIS 2007 Technical Specifications for Physician Measurement<sup>12</sup> propose that, in cases of multiple managing clinicians, a minimum “ownership” threshold be established of at least 25-30% of professional services within the episode. The methods of attribution of episodes are inextricably linked to the intended use of the episode-based information; given that the information can be used in multiple ways, one can expect multiple approaches to attribution and potential innovations in this area.

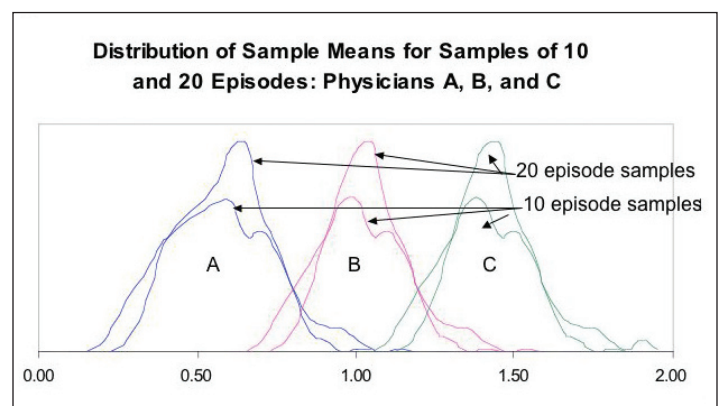
**Peer reference groups.** A significant challenge for episode of care variation measurement and reporting is defining the appropriate peer group for reference. Experience has shown that simply using the listed physician specialty in the health plan database, or even the physician’s board certified specialty, may not create a sufficiently homogenous reference group for comparison. Increasingly, physicians’ practices vary by areas of physician interest or emerging subspecialties. Examples would include orthopedics, where it is common to find physicians focusing their practices on ankles, knees, hips or spines. Other challenging areas are cardiology (generalist, interventionalist, electrophysiologist) and obstetrics/gynecology (vs. obstetrics only or gynecology only). Many group practices will have one physician performing the majority of certain procedures because of interest and skills built by case volume. Such physicians may well appear as outliers within their specialty because of skewed procedure volumes.

The underlying episode mix may assist in creating reference peer groups and identifying physicians with atypical concentrations for certain conditions. More carefully defining the specific episode types to be included for a peer group can also support more consistent measurement. However, the more peer groups created, the lower the peer group sample size, with attendant loss of differentiation power. Mechanisms for reconsideration of peer group assignment are important. More research into optimal methods for peer group assignment is desirable.

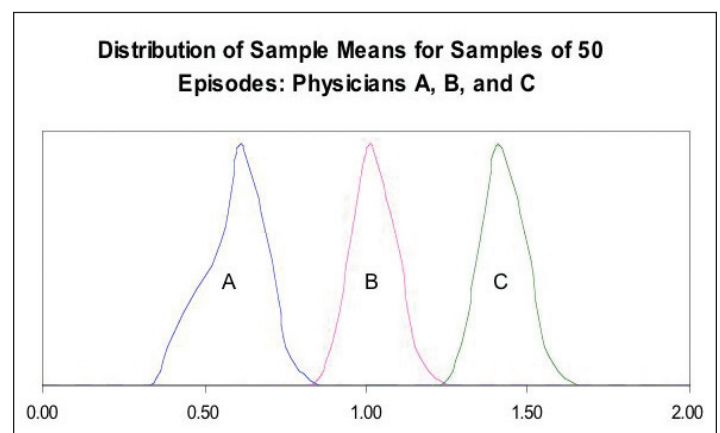
**Isolating physician performance from patient and disease factors.** The goal of episode of care assessment is to identify variations in resource use based on physician practice decisions. This requires adjustment for practice-level factors such as case-mix, patient factors such as comorbidities, and disease factors such as severity. Each of the grouper methodologies tackles adjustment to a significant degree in sophisticated ways. However, this is also an area in need of further refinement, so that treating disproportionate numbers of patients with low disease severity and thus lower resource use will not result in inappropriately favorable measurement results.

**Number of observations or “n” value.** The number of episodes (of each type and overall) has been shown to affect reliability and validity of episode-based measurement. Obtaining an adequate number of episodes for reliable reporting can be challenging.

**Figure 1a**



**Figure 1b**



Figures 1a and b are from work by Thomas<sup>13</sup> that displays the impact of increasing number of episodes on the ability to differentiate the performance of 3 physicians relative to the mean performance of the reference group (1.0 on the horizontal axis). The vertical axis values under each of the curves reflect the variation in the mean performance of each physician when analyses are performed on multiple sample sets of the size indicated. Larger episode numbers and an increasing ability to differentiate one physician's performance from another's are highly associated.

Standards for minimum numbers of episodes have yet to be adopted, but recommendations exist for minima (10,11). More research is needed to determine if different types of episodes might allow different minimum sample sizes. When episode results are used to create tiering of physicians, an evaluation of the statistical chance of misclassification into the wrong tier may be more meaningful. In such a case, the minimum number of episodes required to meet a misclassification risk statistic may vary.

Many health plans and employers have created or are involved in local and regional collaboratives to aggregate data across health plans and increase the sample size used to assess each physician or physician group. Aggregation of data across payers

as well as the inclusion of Medicare data are seen as possible remedies to current limitations on sample size.

## 8. REPORTING

Much of the post-grouper activities center around preparation of reports. Figure 2 illustrates a typical provider summary from one of the groupers.<sup>14</sup>

Understanding this type of report is important in gleaning the valuable information that groupers produce. This physician had 1,525 episodes captured by the grouper (measurement period is not stated). The top seven episode types by volume of episodes (in this case, ETGs) are depicted. The first column on the left is the description of the episode type. The number of episodes is the count of each type of episode. Usually the average amount paid to a physician for each episode type is compared to an appropriate peer average, a per-episode difference is calculated, and then the total difference calculated.

In this example, brackets indicate negative numbers and represents amounts lower than their peers' average and non-bracketed numbers indicate values higher than their peers' average. In this case, of the top 7 most frequent episode types, 6 result in lower than average costs. However, the remaining condition in the top 7, otitis media with minor surgery,

**Figure 2. Sample summary-level physician episode of care profile.**

Sample Family Practice Physician Summary Report					
Description	Number of Episodes	Average Paid	Expected Paid	Difference	Total Difference
Routine Exam	280	64	67	(3)	(840)
Chronic sinusitis, w/o surgery	245	31	39	(8)	(1,960)
<b>Otitis media, with minor surgery</b>	<b>221</b>	<b>898</b>	<b>713</b>	<b>185</b>	<b>40,885</b>
Acute bronchitis, w/o comorbidity, age 5+	82	29	47	(18)	(1,476)
Minor inflammation of skin & subcutaneous tissue	54	37	49	(12)	(648)
Asthma w/o comorbidity, age less than 18	52	66	60	6	312
Otitis media, w/o surgery	50	33	45	(12)	(600)
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
<b>Total</b>	<b>1525</b>	<b>131</b>	<b>113</b>	<b>18</b>	<b>\$27,450</b>

---

---

exceeded peers' average by \$185 per episode with a total of \$40,855 difference from peers over the measurement period studied. Further drilldown reports are necessary to determine the reason for this large variation. When all episodes are considered, this physician was \$27,450 more costly than his peers during the measurement period. The performance index is calculated as observed costs divided by the expected total episode costs accumulated over all episodes, and in this case the result was 1.16. Thus, the physician's practice is considered 16% more costly than the average of specialty peers. Most clinician practices show mixes of above and below average episode resource usage and cost performance. Groupers have value in identifying potential under use as well as overuse.

Ideally, reports are designed and tailored for the intended stakeholder. Physicians want actionable information –they want to know the drivers of variation. Employers and consumers, on the other hand, have their own needs from grouper related reporting. Some employers have asked health plans to create new “high performing” networks focused on those care providers that have overall high quality ratings while using overall fewer resources. Consumers want easy to understand summary information, and while overall performance might assist them in choosing a primary care provider, they can also benefit from identifying care providers that have demonstrated high quality and lower resource use for their specific conditions.

## **PHYSICIAN ENGAGEMENT AND LEADERSHIP**

Physician interest in episodes of care and other methods of practice variation measurement and clinical continuous quality improvement (CQI) is increasing. Large medical groups with quality improvement infrastructures have been most active in this area.

Until the methods of variation analysis and clinical CQI are understood and routinely implemented as part and parcel of contemporary medical practices, other participants in the healthcare system will deploy a variety of techniques to

drive adoption of CQI. Among those techniques are the uses of motivational incentives such as pay for performance, reduced cost sharing by patients choosing higher performing practices, exclusion of lower performing practices from particular employers' choice of physician networks, and the increasingly mandated approach of transparency of physician practice performance.

Fortunately, the past few years have seen increasing engagement of medical organization leadership in seeking system improvement and understanding of measurement techniques. Many entities have been founded or co-founded by physicians to assist in developing measures and measurement systems that help strike a balance between sensitivity to the art and science of medical practice and society's mandate to improve quality and understand and reduce system waste.

---

<sup>1</sup> Solon JA, Feeny JJ, Jones SH, et al. Delineating episodes of medical care. *AJPH* 57:3, pp 401-408, 1967

<sup>2</sup> Ibid.

<sup>3</sup> Hornbrook MC, Hurtado AV, Johnson RE. Health care episodes: Definition, measurement and use. *Medical Care Review* 42:2, 163-218, 1985

<sup>4</sup> Rosen AK, Mayer-Oakes AM. Episodes of care: theoretical frameworks versus current operational realities. *Journal on Quality Improvement* 25:3, 111-128, 1999.

<sup>5</sup> CPT (Current Procedural Terminology) is a registered trademark of the American Medical Association. <http://www.ama-assn.org/ama/pub/category/3113.html>

<sup>6</sup> <http://www.cdc.gov/nchs/about/otheract/icd9/abctcd9.htm>

<sup>7</sup> <http://www.cms.hhs.gov/MedHCPCSGenInfo/>

<sup>8</sup> <http://www.fda.gov/cder/ndc/>

<sup>9</sup> <http://www.ingenix.com/Products/Payers/CareHealthManagementPAY/EnterpriseWideDecisionSupport/EpisodeTreatmentGroups/>

<sup>10</sup> <http://www.medstat.com/Products/view?id=72>

<sup>11</sup> Rattray MC, Milstein A, Herschman R, et al. Measuring Provider Efficiency Version 1.0. [http://www.leapfroggroup.org/media/file/MeasuringProviderEfficiencyVersion1\\_12-31-2004.pdf](http://www.leapfroggroup.org/media/file/MeasuringProviderEfficiencyVersion1_12-31-2004.pdf). 2004

<sup>12</sup> HEDIS® 2007 Technical Specifications for Physician Measurement (Expanded Scope) (Draft) <http://www.ncqa.org/Programs/PublicComment/HEDISPhysician/Specifications.pdf>.

<sup>13</sup> Thomas JW. Enhancing Validity of Physicians' Economic Profiles. Presentation at Academy Health's Annual Research Meeting, Seattle, WA. June 2006. Used with permission.

<sup>14</sup> Adapted from <http://www.symmetry-health.com/products/provider.php>

The author thanks the following individuals who provided valuable assistance in the preparation of this white paper: Lewis G. Sandy, J. William Thomas, Kevin Weiss, and Timothy Zeddies.