



Understanding Bill Based and Aggregate Outcomes Scoring

Presented By:

Theodore Blatt, M.D., F.A.A.O.S.

Martin L. Lee, PhD, CStat, CSci



February, 2020

Understanding Outcomes Scoring

Harbor Health Systems has been the leader in benchmarking Physician performance in workers' compensation for more than a decade. Through Harbor's score-carding and benchmarking systems payers and self-insured employers gain insight into who the top-performing Physicians are, in order to utilize these top doctors to achieve better outcomes for injured workers – faster recovery at a lower cost.

This white paper summarizes the methodology by which Harbor scores these Physicians, as well as identifies issues within workers' compensation that impact the benchmarking process.

About the Authors

Theodore Blatt, M.D., FAAOS

Medical Director & Clinical Research Director

Dr. Blatt is serving as the Medical Director and Clinical Research Director for Harbor Health Systems.



Dr. Blatt was in the private practice of orthopedic surgery in Torrance California, from 1978 until 2002. He then spent twelve years as the National Medical Director of Anthem Workers' Compensation, with responsibilities for the recruitment, training and supervision

of a large multispecialty physician advisor panel and case management nurses, as well as the management of the provider network. He worked closely with clients to develop value-added programs for their medical management and lectured in seminars relating to the managed care approach for workers' compensation.

Dr. Blatt received his undergraduate degree from the State University of New York at Stony Brook, and his medical degree at the Chicago Medical School.

He completed his internship and residency in orthopedic surgery at Harbor General-UCLA Medical Center.

Dr. Blatt is a Fellow of the American Academy of Orthopedic Surgery and a Diplomat of the American Board of Orthopedic Surgery.

Martin L. Lee, PhD, CStat, CSci

Clinical Biostatistician

Martin Lee, Ph.D., has served as Clinical Biostatistician for Harbor Health Systems for 10 years.



Dr. Lee is responsible for analysis of provider scorecards and research initiatives. His career encompasses more than 35 years in the pharmaceutical and biotech industries.

Martin L. Lee is Adjunct Professor of Biostatistics in the Fielding School of Public Health at UCLA and Adjunct Professor of Internal Medicine, Charles R. Drew University of Science and Medicine. The author or co-author of over 200 scientific papers, he is a Fellow of the Royal Statistical Society and member of the American Statistical Association, the Society for Clinical Trials, the Biometrics Society, and the International Society of Thrombosis and Hemostasis. In addition he is a Chartered Statistician and a Chartered Scientist as designated by the Royal Statistical Society.

Dr. Lee received a B.A. degree in mathematics, and M.S. and Ph.D. degree in biostatistics from the University of California, Los Angeles.

Benchmarking Overview

The goal of Harbor Health Systems’ outcomes scoring is to identify doctors most closely associated with positive outcomes based on their performance. This scoring approach differs from pathway and compliance measures associated with benchmarking in that it focuses on the global population results rather than on

to work within the workers’ compensation regulatory framework. Over the years, Harbor has found top scoring providers tend to have strength in all three of these areas, and that difficulty in any one area can have an adverse impact on the overall experience of the injured worker.

Measurement Factor	Measurement Description
Service Allowed Amount	Sum of allowed amount for billed procedures after bill review
Service Charges Amount	Sum of billed amount for procedures prior to bill review
Care Duration	Total time over which medical services were provided
Care Volume	The number of unique service dates

a specific provider’s clinical decisions. In many cases, the clinical skills of the physician may not correlate with workers’ compensation outcomes. This is because at the larger claim level, multiple influences other than clinical decisions such as unique payer issues or claims handling may affect the result of the claim.

Unfortunately, the workers’ compensation system has many scenarios that can negatively influence the patient’s experience. Doctors who may have difficulty gaining the trust of the patient can have just as negative an impact on overall claims costs as doctors who struggle with producing good clinical results.

Outcomes Measurements

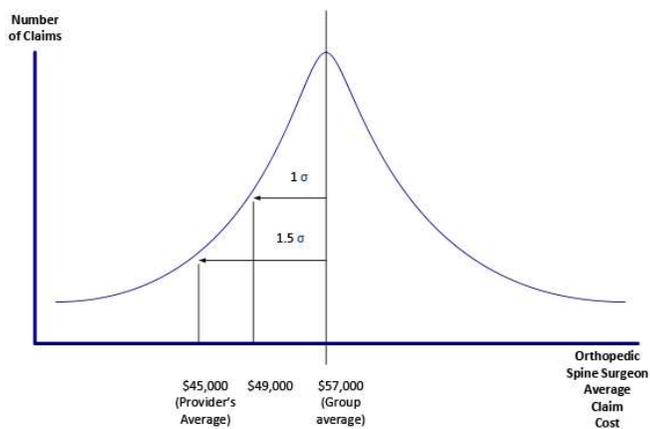
In Harbor’s models there are multiple variables reviewed in the process of developing a score depending on whether the model utilized is service allowed amount bill-based scoring, service charges bill-based scoring or aggregate scoring. Harbor has developed two (2) bill-based models for scoring: one based on post bill review “allowed amount” and the other based on pre bill review, “service charge” amount. Some of the factors are listed in the table above.

It is for these reasons that Harbor always emphasizes that this approach to scoring does not measure the quality of the Physicians. Rather it measures the Physician’s association to the overall claim outcomes.

Scoring Methodology

Scores within the benchmark focus primarily on three factors: cost, duration and frequency. For any given measurement factor, Harbor compares the average performance of the provider to the average performance of his or her peer population (specialty and, where appropriate, secondary specialty and only within their state). Scores are measured as the difference (better or worse) in that category as compared to the appropriate population mean in terms of its distance measured as the number of standard deviations.

It is important to realize that measurements are determined at the overall claim level, not at the level that a specific provider was engaged. That is because this benchmarking approach seeks to capture larger relationships that are related to clinical skills, patient engagement skills, and the overall ability



By way of example, consider the following simplified process:

1. A group of peers consisting of orthopedic spine surgeons in California are defined as a stratum. The claims selection criteria would be the California claims that include ortho-spine procedure codes. There may be dozens of providers who have billed for ortho-spine procedures against 300 claims, and the average cost of all of these claims hypothetically is \$57,000.
2. One particular physician who performs such procedures is connected to 21 claims through the billing activity and identifies him as the treating provider. The average total costs for the 21 claims linked to the provider is, for example, \$45,000.
3. Clearly this provider is averaging about \$12,000 per claim better than their peer population. If the standard deviation (σ) for the peer population was \$8,000, then this provider would be 1.5 standard deviations better than the mean.

To determine a score, all of the outcome factors considered go through a similar process and a proprietary weighting scheme is applied to each factor.

Harbor also has the capability to score out of network Physicians as well as the capability to generate scores using aggregated client data. Harbor has developed a model in which scores of Providers from each client can be aggregated and,

after a client adjustment, a single score developed for the Provider based on all of the data. This will allow for a much more robust data set to be used for the aggregate score and provide clients with scores of Providers when their own specific data would not contain sufficient claim volume to accurately generate a score. Scores are specific to specialty and jurisdiction and take into account patient complexity by looking at co-morbidities (obesity, hypertension, diabetes, mental health and combinations of each). In addition, outlier claims are statistically identified and removed prior to scoring.

Harbor is also able to accept utilization review data from the payers and the UR vendors of the payers. Distinct from the scoring models described above, Harbor has also developed a Physician scoring model based on the approval percentage of the utilization review requests submitted by that specific provider.

Sources and Use of Data

Harbor receives bill data from payers and/or vendors. Scoring is calculated based on a specific payer's set of bills. The bills could be delivered by the payer or TPA or by the Bill Review vendor. Bill data, whether pre or post bill review, is the key driver for calculating the scoring factors. The key information in the bill data is who provided services, what was done and when.

Case Mix, Statistical Significance, Adverse Selection, and Comorbidities

Workers' compensation data contains certain barriers to the typical best practices used for assuring human factors are properly accounted for in the scores. Harbor's approach to stratifying

populations is constructed to take into account case mix by using patient complexity factors.

Adverse selection (the impact of highly reputable Physicians attracting the most difficult cases) has been examined as a factor in the statistical model. As of yet most evidence of adverse selection in workers' compensation tends to correlate better with attorney impact rather than with patient selection. However using adjustment factors applied to the outcome variables based on comorbidities of the injured worker, using a proprietary process, Harbor accounts for these case-mix factors when scoring a Physician. Comorbidity factors include diabetes, obesity, mental health issues, addiction, hypertension and combinations of more than one. Based on the degree of impact noted in our analysis, an adjustment factor is applied to claim outcomes used in the scoring of a Provider to compensate for the medical complexity of the Injured Worker who would be more prone to certain complications because of these diagnoses. Both ICD-9 and ICD-10 codes are used to identify claims with these ancillary diagnoses which are in the Physician's census of claims for scoring.

A single outlier case can negatively impact a provider's score with no other pattern of practice at issue. Harbor identifies statistical outlier claims by specialty and jurisdiction and removes such claims prior to scoring.

Every applicable CPT and HCPCS code is reviewed and mapped to enhance the physician specialty listing. By doing so, we are able to better identify physician specialties by the procedure codes which are billed by that physician. We have also mapped every relevant taxonomy code to the appropriate specialty listings used in scoring. By doing so, we are able to do a more "granular" scoring of the 15 most commonly used specialties in Workers' Compensation. These specialties include Acupuncture, Chiropractic, Dental, General Surgery, Hand Surgery, Neurology, Orthopedic Surgery, Pain, Podiatry, Primary Care, Psychiatry, Spine

Surgery, Physical Medicine and Rehabilitation, Cardiology and Pulmonary Medicine.

Appeals and Educational Opportunities

Harbor has a rigorous Quality Assurance program which reviews scoring used to restructure a client network. This includes processes to ensure that the Physician is indeed identified in the correct specialty, the benchmark peer group used in the Physician's scoring is correct and that relevant outlier claims have been removed prior to scoring.

There are multiple opportunities for providers to learn more about their standing and how to maximize their performance. By understanding overall outcomes better, including the specific provider's performance relative to his/her peer group, providers can begin to think outside their practice to see how their role can impact the outcome of the claim.