Summary of the Advanced Statistical Approach Used in the GIC CPII Evaluation of Physician Quality of Care
# TABLE OF CONTENTS

1. Purpose ........................................................................................................................................ 1
2. Introduction .................................................................................................................................. 1
3. Potential sources of uncertainty .................................................................................................. 2
4. Statistical approach ...................................................................................................................... 3
5. GIC CPII decision rules to place a physician in a quality tier .................................................. 4
6. References .................................................................................................................................... 5
1 Purpose

This document describes the statistical approach Resolution Health (RHI) developed to support the Group Insurance Commission (GIC) Clinical Performance Improvement Initiative (CPII) to evaluate the quality of care delivered by physicians. Our statistical approach accounts for factors that could influence a physician’s quality score and provides insight into the probability that the physician’s quality score may be higher or lower than what we observe. We also explain the decision rules set by GIC for placing a physician in a quality tier based on the results of our statistical model.

2 Introduction

During the early years of the GIC CPII, RHI scored a physician’s quality of care based solely on the percentage of times that the physician complied with clinical practice guidelines applicable to his/her patients, as observed in the claims data. For example, if Dr. Smith saw 20 patients in a year and two quality of care measures applied to 10 of these patients, and three other quality of care measures applied to each of the other 10 patients, then there would have been 50 “observations” reflecting the quality of care provided by Dr. Smith in 2006 ([10 x 2] + [10 x 3] = 50). If the care provided by Dr. Smith was consistent with the relevant clinical practice guidelines in 40 of these 50 instances, then the quality-of-care score for Dr. Smith in that year would have been 40/50 = 0.80, or 80 percent.

We have recognized that a score of this type constitutes a “point estimate” of a physician’s quality of care and that there is some probability that the true quality of care provided by the physician is actually higher or lower than what we observe. Experts also expressed concern that point estimates could be accurate yet still inadequate because, among other things, they fail to take into account patient non-compliance or the relative difficulty of complying with particular measures. To address this issue, we worked with expert biostatisticians from the Johns Hopkins School of Public Health in 2008 to modify our scoring methodology to address this statistical uncertainty regarding point estimates of physician quality of care. Starting from
year 2010, we have introduced further enhancements to the methodology to produce more stable and reliable physician quality scores.

3 Potential sources of uncertainty

Physicians deal with probability in their practices every day. For example, most tests that physicians order for their patients are not perfectly accurate. As a result, physicians recognize that there is some probability that test results for their patients may be higher or lower than their true values. Similarly, when clinical trials are reported in a journal article, the results of the studies are typically reported as a point estimate with a 95 percent confidence interval around the point estimate to indicate the magnitude of uncertainty around the observed point estimate.

With respect to the CPII, there are many factors that may contribute to uncertainty in the quality of care scores that the GIC CPII produces:

1. We are not observing the care that physicians provide to all of their patients. For example, although our database is very large, most if not all physicians have patients who are members of health plans whose data is not collected by the GIC CPII.

2. The total number of observations that are available for a physician influences the uncertainty in the quality of care estimate for that physician. Having more observations available lowers this uncertainty. Conversely, a lower number of observations increases the uncertainty.

3. The quality of care measures that are used are based on well established clinical practice guidelines that have been developed by clinical specialty societies or other highly regarded professional organizations such as the National Committee for Quality Assurance (NCQA), the Agency for Healthcare Research and Quality (AHRQ), and the National Institutes of Health (NIH). These measures, however, are limited and do not completely reflect the range of relevant types and aspects of care that a physician may provide to his/her patients.

4. Certain clinical practice guidelines may be more difficult to comply with than others. An observed difference between the performance of Dr.
Smith and Dr. Jones may reflect a difficulty in complying with some measures as compared to others, rather than true differences in the quality of care delivered by the two physicians. For example, it may be more difficult to convince a patient to undergo a chlamydia screening than to convince a patient to get a mammogram.

5. There may be differences among a physician’s patients, such as variations in the patients’ ability to pay for medications, behavioral profiles, or education levels. These variances may influence the likelihood that a given patient will comply with the care recommended by the physician. Therefore, the patients’ circumstances may dictate their behavior despite physicians’ best efforts.

4 Statistical approach

Our statistical model quantifies the uncertainty around a point estimate of the quality of care delivered by a particular physician (called the physician effect) while also attempting to control for three areas of potential bias:

1. The mix of quality measures applied to the physician’s patients (the “measure effect”).

2. The mix of a physician’s patients, each of whom has a particular likelihood of complying with the physician’s recommendations (the “patient effect”).

3. The effect of the number of observations for a particular physician available in the GIC database (the “sample size effect”).

We have employed a multi-level logistic regression model, also known as a hierarchical or mixed model\(^1\) to produce a physician-specific probability distribution around a point estimate of quality of care. Multi-level models were developed more than forty years ago\(^2\) and have been used in many areas of science. These models are now commonly used in hospital profiling\(^3\)\(^4\) and physician profiling.\(^5\)

Our model assumes that the probability of a physician complying with a particular set of quality of care measures that apply to a particular number of patients is a function of measure, patient,
sample size and physician effects, the latter of which we use to quantify the physician’s performance after adjusting for the other effects. This model allows us to make adjustments for the number of observations available and the variability among measures and patients.

Starting from year 2010, we enhanced the existing methodology to increase the reliability of physician scores. We now include physician quality results from the prior year analyses in addition to the current year results in the model. The physicians are evaluated based on analyses of data from two years rather than one year. Physician observations from the prior year and the current year are weighted equally in the model. Just as there are several ways to account statistically for the uncertainty around the point estimate (adjusted quality score) of a physician’s quality of care, there are several ways to use that information to place a physician in a particular performance tier.

5 GIC CPII decision rules to place a physician in a quality tier

GIC uses RHI results to categorize physicians in one of three quality designations (A, B or C) based on the results of the statistical model. The quality tiers are assigned by comparing physicians’ adjusted quality scores to their peers (“grading on the curve”). Primary care physicians (internal medicine, family practice) and physicians in eight specialties (cardiology, endocrinology, rheumatology, OB-GYN, pulmonary medicine, pediatrics, nephrology and otolaryngology) are assigned to quality tiers using the following decision rule:

1. To be assigned to the highest quality tier (A), a physician must have:
   a. At least 30 quality-of-care observations available
   b. A 90-percent chance or greater that the physician truly belonged in tier A, based on the observations available for the physician and the results of our statistical model.

2. Similarly, to be assigned to the lowest quality tier (C), a physician must have:
   a. At least 30 quality-of-care observations available
   b. A 90-percent chance or greater that the physician truly belonged in tier C, based on the observations available for the physician and the results of our statistical model.
3. In all other instances, a physician is assigned to the middle quality tier (B) for the particular specialty. Physicians who do not meet the minimum observation threshold are placed in a default category (*B). The minimum number of quality-of-care observations for physicians in neurology has been lowered to 10 due to the low numbers of observations for neurologists. However, there still needs to be a 90 percent- or greater chance that the physician truly belonged to the highest tier or the lowest tier to correctly assign that physician. After a physician has been placed in a quality-of-care tier (the subject of this paper), the GIC Clinical Performance Improvement Initiative then considers the physician’s cost-of-care (efficiency) tier before making final tier assignments.

6 References


