



SUPPLEMENTAL REFERENCE

**Summary of the advanced statistical approach used
in the GIC CPI Initiative evaluation
of physician quality of care**

VERSION: 1.1
DATE: DECEMBER 15, 2009
CLASS.: PROPRIETARY, FOR CLIENT USE

Resolution Health, Inc. Proprietary Statement: This material constitutes proprietary and trade secret information of Resolution Health and shall not be disclosed to any third party, nor used by the recipient except under the terms and conditions prescribed by Resolution Health. This material is not to be copied or reproduced in any form, using any medium, without the prior written authorization of Resolution Health.

TABLE OF CONTENTS

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

1 Purpose

This document describes the statistical approach Resolution Health developed to support the Group Insurance Commission (GIC) Clinical Performance Improvement (CPI) Initiative to evaluate the quality of care delivered by physicians. Our statistical approach not only takes into account factors that could influence physicians' quality score but also provides insight into the likelihood that a physician's quality score may be higher or lower than what we observe. We also explain the decision rules that were set by the GIC for placing a physician into a quality tier based on the results of our statistical model.

2 Introduction

During the first four years of the GIC CPI Initiative, Resolution Health 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 during 2006, 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 ($[10 \times 2] + [10 \times 3] = 50$). 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 2006 would have been $40/50 = 0.80$, or 80%.

The scoring methodology described above produces a point estimate of a physician's quality of care. However, there remains some probability that a more accurate description of the true quality of care could be higher or lower than the point estimate. Some experts have 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 to modify our scoring methodology to address this statistical uncertainty regarding point estimates of physician quality of care.

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 study are typically reported as a "point estimate" with a 95% confidence interval around the point estimate to indicate the magnitude of uncertainty around the observed point estimate.

With respect to the CPI Initiative, there are many factors that may contribute to uncertainty in the quality-of-care scores that the GIC CPI Initiative 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 a health plan whose data is not collected by the GIC CPI Initiative.
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 NCQA. 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 their physicians. Therefore, the patient's circumstances may influence their behavior despite the physician's recommendations.

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 physicians' patients, each of whom has a particular likelihood of complying with his/her physician's recommendations (the “patient effect”), and
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” model¹ to produce a physician-specific probability distribution around a point estimate of quality of care. Multi-level models were developed more than forty years ago² 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.

Just as there are several ways to take account statistically 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 CPI Initiative decision rules to place a physician in a quality tier

The individual health plans use the Resolution Health results to categorize physicians into one of three quality designations (A, B, or C) based on the results of the current year’s analysis. The quality tiers are assigned by comparing physicians’ adjusted quality scores to their peers (“grading on the curve”). Performance on measures in past analyses are not currently considered.

Primary care physicians (internal medicine, family practice) and physicians in seven specialties (cardiology, endocrinology, rheumatology, OB-GYN, pulmonary medicine, pediatrics, 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 75% 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 75% 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 their particular specialty.

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 75% or greater chance that the physician truly belonged to the highest tier or the lowest tier in order to assign that physician to the highest tier or the lowest tier, respectively.

After physicians have been placed into quality-of-care tiers, the plans then incorporate the physicians' relative cost-efficiency to determine their final tier assignments.

6 References

- ¹ Diggle, PJ, Heagerty, PJ, Liang K-Y, Zeger SL. 2002. *Analysis of Longitudinal Data*. Oxford, Oxford University Press.
- ² Rao, CR. The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika* 1965;52:447–458.
- ³ Christiansen CL, Morris CN. Improving the statistical approach to health care provider profiling. *Ann Intern Med* 1997;127:764–8.
- ⁴ Shahian DM, Torchiana DF, Shemin RJ, Rawn JD, Normand S-L. Massachusetts cardiac surgery report card: implications of statistical methodology. *Ann Thorac Surg* 2005;80:2106–13.
- ⁵ Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician “report cards” for assessing the costs and quality of care of a chronic disease. *JAMA* 1999;281:2098–2105.