



HealthPartners®

HealthPartners Technical Documentation

# Total Cost of Care Bootstrap Reliability Analysis

*Using Optum Symmetry Episode Risk Groups (ERG)*

---

## Purpose

---

Determine the bootstrap and 90% random sample reliability of the Total Cost of Care (TCI) measure using the Optum Symmetry Episode Risk Groups (ERG).

## Table of Contents

---

Overview of Analysis	Bootstrap and 90% Random Sample Results
Overall Conclusions	TCI Consistency Over Time
Methodology	TCI Consistency Over Time Results
Bootstrap and 90% Random Sample	Definitions and Examples

## Overview of Analysis

---

Total Cost of Care (TCI) is a measure of a provider's effectiveness of managing their primary care attributed population across the care continuum. The TCI measure was applied to HealthPartners' primary care metro and regional providers as per the measure specifications and results were calculated for 2009, 2010, and 2011.

The reliability testing demonstrates the repeatability of producing the same results a high proportion of the time. To measure the reliability of the TCI measure a 90% random sample and a bootstrapping technique were employed. In these methods, reliability is measured as the mean of the variance between sampling iterations and the actual results.

In addition, the TCI measure was analyzed over time to demonstrate stability and sensitivity to provider changes or improvement initiatives.

These methods were chosen as they represent the measure intent, which is that the TCI measure represents providers' average total cost of care across their population. Since the measure is aggregated to the provider group level there is no need to quantify the variability at the member level into the evaluation.

In the 90% random sample method, the members that were attributed to a provider group were randomly sampled at the 90% membership level without replacement. This technique was employed to simulate variation within a provider group by leveraging their own population and case-mix. This method gives an indication as to the repeatability of the measure by comparing how closely the actual total cost measure is to the 90% sampled averages and simulates any potential member selection bias.

In the bootstrapping method members that were attributed to a provider group were randomly selected with replacement. This method maximizes variation around a provider group's total cost of care as each randomly selected iteration (sample populations) does not truly represent the provider's case mix of patients. This method was performed in the same fashion as above to support and validate the results found in the 90% sample method.

## Overall Conclusions

---

- The percent differences between provider Actual TCI results and both the 90% sample and bootstrap mean results are very small.
  - Ranging from -0.14% to 0.19% in the 90% sample in 2011.
  - Ranging from -0.39% to 0.37% in the bootstrap in 2011.
  - These results indicate that the TCIs for each provider group are repeatable and consistent.
- A provider's performance is relatively consistent across all three years with an average difference of 0.047.
  - These differences in provider performance over time occur because of known changes in fee schedules, collaborating provider usage and resource use saving initiatives can account for the differences.
  - Since the measure is designed to capture and reflect changes in these areas, we expect to see some explainable variability within a provider group over time.

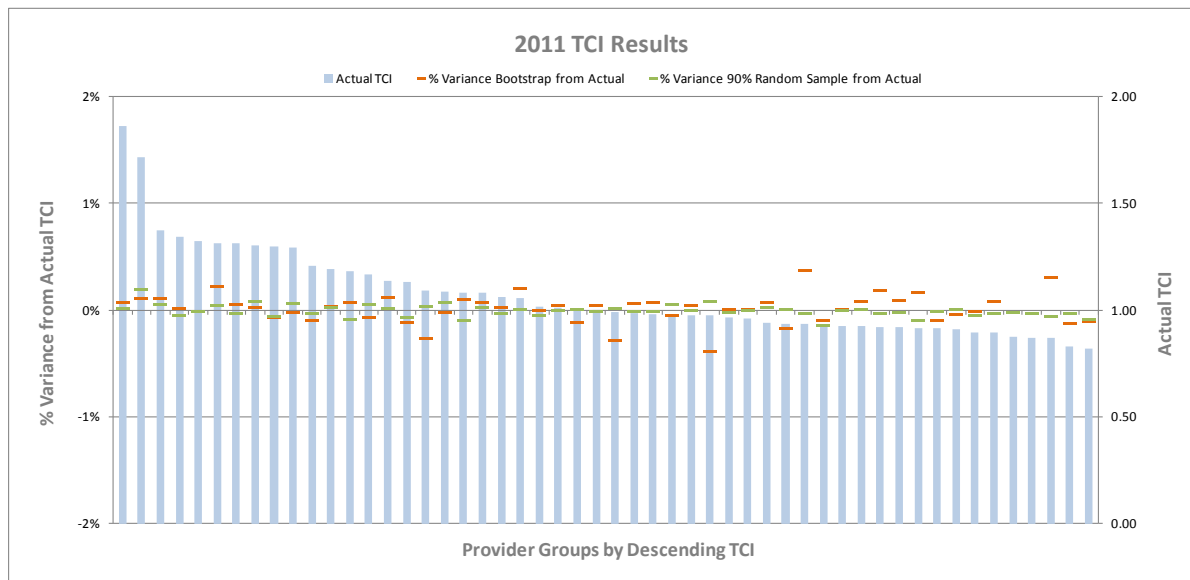
## Methodology

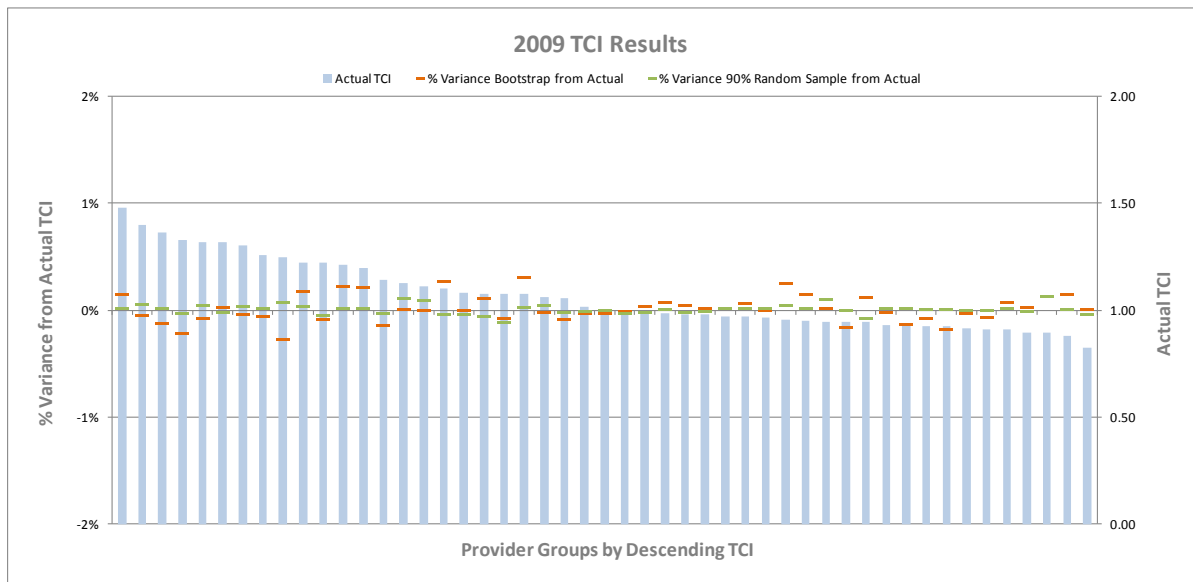
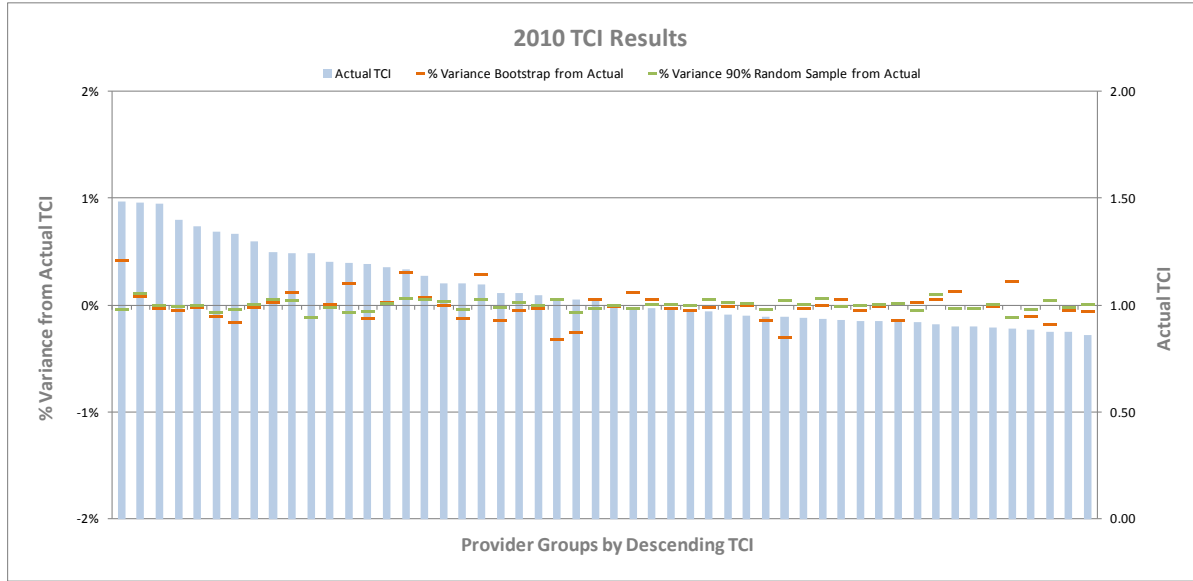
In the 90% sample method, 90% of attributed provider group members were randomly selected, without replacement. A 90% sample was used despite having the full health plan provider population, to simulate any potential member selection bias. The sampling process was performed using the SAS PROC SURVEYSELECT procedure with the Simple Random Sample (SRS) option. This method allows for each attributed member to be selected only one time until 90% of the total provider population has been reached. The 90% sampling process was repeated 500 times for each provider group and year analyzed. Attributed members' total costs were aggregated in each sample to produce 500 TCI results for each provider group for each year (see figure 1 in the definitions section for more information). Once the 500 samples were created for each provider group, the total costs of care of each sample for each provider group were compared to the metro average to produce risk adjusted indices. The Total Cost indices from each of the sampling iterations for each provider group/year were then compared to the actual TCI indices for each provider group/year and the mean variance was computed.

To perform the bootstrap, the SAS PROC SURVEYSELECT procedure with the Unrestricted Random Sample option for full replacement utilized to create a series of random samples for each provider group being measured. Full replacement means that one observation is drawn at random, recorded, and then placed back into the data pool so that it can be drawn again if randomly selected. The numbers of records sampled are drawn such that the samples created are the same size as the original number of attributed members for the provider group. In this way, it is theoretically possible (although virtually improbable) to produce a sample of size n that could consist of the same record drawn n times in a row. This was done to artificially maximize the variance within the defined populations. This sample process was performed 500 times for each year and provider group being analyzed, to produce 500 sets of risk adjusted Total Cost of Care results for each provider for each year (see figure 2 in the definitions section for more information). The Total Cost indices from each of the sampling iterations for each provider group/year were then compared to the actual TCI indices for each provider group/year and the mean variance was computed.

## Bootstrap and 90% Random Sample

The mean TCI results from the bootstrap and 90% samples compared to the actual TCI results for each provider group and year are displayed in the tables and graphs on the following pages.



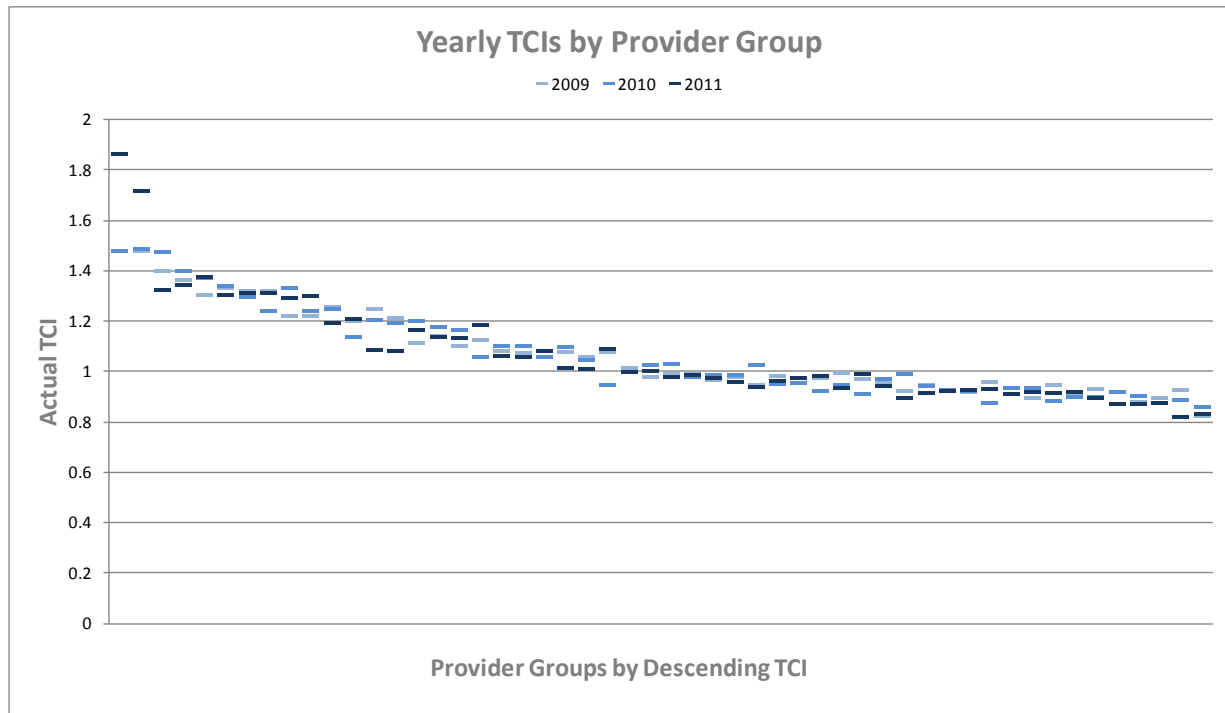


### Bootstrap and 90% Random Sample Results

- The differences between provider Actual TCI results and both the 90% sample and bootstrap mean results are very small ranging from -0.14% to 0.19% in the 90% sample to -0.39% to 0.37% in the bootstrap in 2011.
- The results indicate that the TCIs for each provider group are repeatable and consistent.
  - Ranging from -0.14% to 0.19% in the 90% sample in 2011.
  - Ranging from -0.39% to 0.37% in the bootstrap in 2011.

## TCI Consistency Over Time

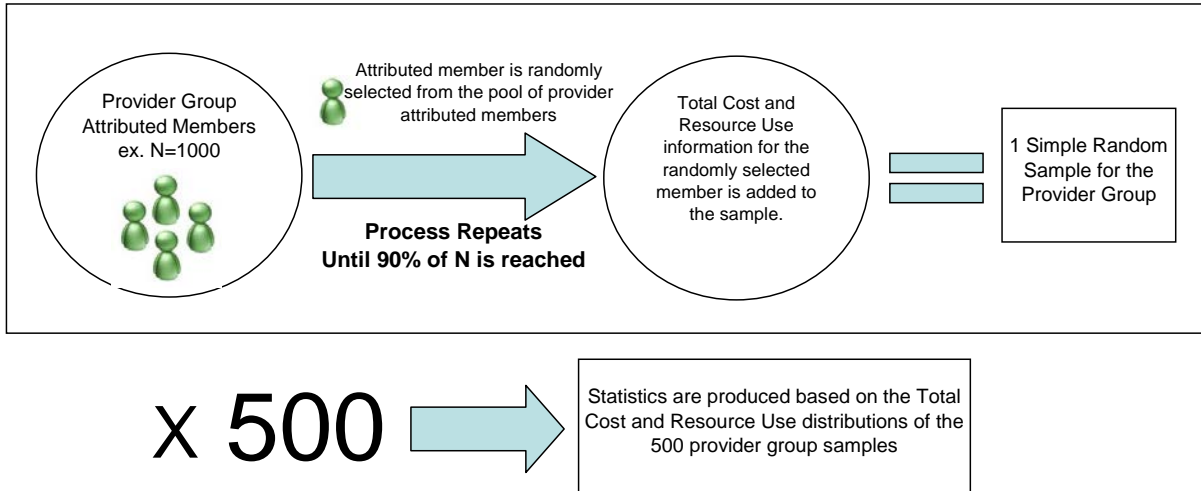
The TCI results are displayed from 2009 through 2011 for the HealthPartners Primary Care Metro and Regional Network. The measure differentiates between providers however they remain relatively consistent over time. The factors that drive variation between years within a provider are cost per unit control and resource use management.



## TCI Consistency Over Time Results

A provider's relative performance is generally consistent across all three years with an average difference of 0.047.

- These differences in provider performance over time occur because of known changes in fee schedules, collaborating provider usage and resource use saving initiatives can account for the differences.
- Since the measure is designed to capture and reflect changes in these areas, we expect to see some explainable variability within a provider group over time.
- The large differences seen in this graph between yearly TCI scores are due to changes in price, as analysis of resource use revealed a reduced yearly variation.

**Definitions and Examples**
**Figure 1: 90% Sampling – Simple Random Sample Without Replacement**

**Figure 2: Bootstrap Sampling – Unrestricted Random Sampling With Full Replacement**
