Milliman Analysis of Total Cost of Care Benchmarks

A comparison of Blue Cross Blue Shield companies' experience for commercial large group PPO networks to medical experience benchmarks

Commissioned by the Blue Cross Blue Shield Association[1]

Adam R. Singleton, FSA, MAAA
Chris S. Tilley, FSA, MAAA
Stuart D. Rachlin, FSA, MAAA

Milliman

INTRODUCTION

Milliman was engaged by the Blue Cross Blue Shield Association ("BCBSA") to calculate total cost of care ("TCC") benchmarks for commercial large group PPO health plans using historical medical claims and enrollment data from Blue Cross Blue Shield ("BCBS") organizations and the Merative MarketScan[2] ("MarketScan") database. When "BCBS" is mentioned in this paper, it is referring to the BCBS companies. This whitepaper presents TCC benchmark results as medical allowed charge[3] per-member-per-month ("PMPM") amounts. Also, it describes the data sources and benchmarking methodology used for the analysis.[4] Key elements about the benchmark results are as follows:

- The BCBS nationwide total medical allowed PMPM (the "TCC") is 7.1% lower than the benchmark calculated by Milliman ("Benchmark").
- The 2022 claims and enrollment data for this analysis are from two independent, nationwide databases with millions of commercial enrollees and their claims.
    - Benchmark data comes from MarketScan.
    - On behalf of BCBSA, Blue Health Intelligence ("BHI") provided historical BCBS commercial data for their commercial large group PPO networks.[5]
- We filtered these datasets to include only large group commercial PPO experience for 2022 incurred claims, excluding governmental and healthcare provider (e.g., hospital systems and medical groups) employers.
- To make the datasets comparable, our methodology adjusts for demographics, health risk, and geographic mix differences, and other reasonable adjustments (e.g., an excessive mix of large claimants).

The PMPM comparison metrics between BCBS and the Benchmarks are reported at the Milliman Health Cost Guidelines ("HCG") inpatient, outpatient, professional, and other medical service category levels (excluding retail, specialty, and mail-order prescription drugs).

RESULTS

Table 1 presents the summarized 2022 nationwide and regional average allowed PMPM results for BCBS and the Benchmarks. These results are mix adjusted to the BCBS 2022 enrollment by metropolitan statistical area ("MSA"), age, and sex, as well as adjusted for medical health risk scores and disproportionate excess large claims in the two datasets.

TABLE 1: TOTAL RISK ADJUSTED MEDICAL ALLOWED PMPM FOR BCBS COMPARED TO THE MILLIMAN CALCULATED MARKETSCAN DATA BENCHMARK

Region	BCBS	Benchmark	Dollar Difference	Percent Difference
Nationwide	$416.23	$448.02	-$31.79	-7.1%
West	$408.65	$454.73	-$46.08	-10.1%
Southwest	$409.58	$442.65	-$33.07	-7.5%
Midwest	$424.12	$449.34	-$25.22	-5.6%
Northeast	$460.45	$482.59	-$22.14	-4.6%
Southeast	$392.62	$419.91	-$27.28	-6.5%

[1] Blue Cross Blue Shield Association is an association of independent, locally operated Blue Cross and Blue Shield companies. Blue Cross, Blue Shield, Blue Health Intelligence and BHI are registered trademarks of Blue Cross Blue Shield Association.
[2] MarketScan is a registered trademark of Merative.
[3] Medical allowed charge is defined as the sum of the insurer's paid amount plus member cost share (excluding balance billing) and adjustments for coordination of benefits.
[4] The Blue Cross Blue Shield Association commissioned this whitepaper. Milliman does not endorse third party solutions. The readers' experiences with this topic may vary, and they should consider differences in their own data, comparison approaches, and other variables, as well as the potential impact on their own results, which may vary from the results in this paper.
[5] Blue Health Intelligence is a healthcare data and analytics company and independent licensee of BCBSA.

Observations about the Table 1 results include:

- The nationwide BCBS PMPM of $416.23 is 7.1% lower than the Benchmark PMPM of $448.02.
- BCBS TCC is lower than the Benchmark in every region.
  - The greatest difference is in the West region at -10.1%.
  - The least difference is in the Northeast region at -4.6%
- There are significant differences in the BCBS and Benchmark PMPMs across the different regions.
  - The Northeast region tends to have the highest PMPMs and the Southeast region has the lowest PMPMs. The other two regions are similar to each other.
  - Note that these PMPMs represent each region's specific population demographic and health risk mix; therefore, one region's results are not directly comparable to another's.

Table 2 presents the 2022 nationwide results for the TCC PMPMs for BCBS and the Benchmarks by HCG service category. Like in Table 1, Table 2 shows that the BCBS TCC is 7.1% lower than the Benchmark with and without the adjustment for excess large claims.[6]

MarketScan has many data contributors which may include contributors with BCBS members and claims. Theoretically, if BCBS member results could be removed from the MarketScan data, then the BCBS TCC may be even lower relative to a benchmark that includes only non-BCBS MarketScan data.[7]

TABLE 2: NATIONWIDE BCBS AND MARKETSCAN DATA RISK ADJUSTED ALLOWED PMPM BENCHMARKS BY MAJOR SERVICE CATEGORY

Service Category	BCBS	Benchmark	% Difference
Inpatient Facility	$97.97	$93.56	4.7%
Outpatient Facility	$150.31	$173.64	-13.4%
Professional	$155.34	$167.11	-7.0%
Other Medical	$12.62	$13.71	-7.9%
Total Medical	$416.23	$448.02	-7.1%
Adj. for large claims	$416.23	$448.02	-7.1%

Other observations from the Table 2 service category level results are as follows:

- The BCBS inpatient TCC is higher than the Benchmark by 4.7%. At the same time, outpatient TCC is lower than the Benchmark by 13.4%.
- The BCBS TCC for professional and other medical services are lower than Benchmarks by 7.0% and 7.9%, respectively.

DATA SOURCES

On behalf of BCBSA, BHI provided facility and professional claims and member eligibility files for the BCBS group commercial line of business.

- These claims files have incurred dates of 1/1/2022 – 12/31/2022, processed through 12/31/2023.
- The eligibility files contain monthly member records for January through December 2022, processed through 12/31/2023.

The MarketScan data also includes facility and professional claims and comes from Milliman's internal March 2024 release of the MarketScan database. Milliman leases the MarketScan database from Merative and processes it with the Milliman HCG Grouper to organize the data into logical health service categories with Milliman's utilization counting and financial field calculations applied.

- These claims files have incurred dates of 1/1/2022 – 12/31/2022, processed through 10/31/2023.
- The eligibility files contain monthly member records for January through December 2022, processed through 10/31/2023.

For both datasets the impact of IBNR adjustments would be expected to be very small; therefore, we did not adjust for IBNR.

METHODOLOGY

The following is a description of our methodology, including our approach to normalize or adjust results for mix differences in geography, age-sex, industry, group size, and health conditions.

For each dataset, we applied several inclusions/exclusions in the eligibility/enrollment and claims data:

[6] Because both datasets contain millions of members, the adjustment for excessive proportions of large claims is negligible.

[7] For example, if BCBS makes up 25% of the volume in the MarketScan data, and assuming those BCBS members have the same nationwide average BCBS TCC of $416.23, then we can approximate the BCBS theoretical difference to a benchmark with non-BCBS MarketScan data by backing out the BCBS PMPM from the $448.02 Benchmark PMPM with a simple algebra equation and then calculate the percentage difference with the BCBS PMPM.

- Dates – include only calendar year 2022 incurred claims and membership data, processed through 12/31/2023 for the BCBSA dataset and through 10/31/2023 for the MarketScan dataset.
- Member Age – exclude all members aged 65 or older and exclude members with missing or irrational ages
- Product – include PPO products only
- Line of Business – include the commercial group line of business only.
- Sex – include only those members with an indicator of 'M' or 'F'.
- State – include only those members within the 50 U.S. states and D.C.
- MSA – exclude all members with missing member location information (e.g., no ZIP Code or a null MSA).
- Group Size – include all groups in each dataset with 51 or more subscribers for any month in 2022 to simulate large employer group health plans for large group only comparisons.

We also made several key adjustments to eliminate bias in the data from having employer groups included that tend to utilize special networks or benefit designs to steer patients to an employer's own customized reimbursement rate arrangements. These employer-specific networks are not available to most large group employers. To ensure the benchmarks are comparable for the two sets of data, we applied additional filters to the data as follows:

- Only data from "Lead Networks", as provided by BCBSA, were included in the analysis. These represent the main networks used by BCBS for large group employers. By benchmarking to these networks only, other specialized BCBS networks and custom networks are excluded. This filter is not applied to MarketScan's data.
- BCBSA communicated to Milliman that many groups in New York only use the BCBS hospital network and as a result do not have many professional claims in their historical data. Therefore, on behalf of BCBSA, BHI sent Milliman a list of group IDs to exclude to remove data for these types of groups from the analysis. Milliman excluded the members and claims for these group IDs, if the member's address was in New York, New Jersey, or Connecticut. This filter is not applied to MarketScan's data.
- We excluded health system and government employee health plans from this analysis as very large employers in these sectors often have customized provider networks that are not typical for most commercial group business and could distort the results.[8]
    - For example, health systems (e.g., hospitals) self-insured employee health plans often offer tiered networks and benefit designs which encourage their members to use their owned facilities and professional providers (and often at reduced reimbursement rates).
    - Governmental employee health plans often have unique and/or customized provider network arrangements and/or their own reimbursement rates negotiated with providers.

Both the BCBSA and MarketScan datasets are very large. Even with the above filters applied, the TCC comparisons in this paper are based on several million large group members and their claims from both datasets (i.e., it is still a highly credible study).

Milliman processed the BCBSA and MarketScan datasets with Milliman's HCG Grouper v2023.2 and Milliman's Advanced Risk Adjuster ("MARA") v4.17 software.

- HCG Grouper – sorts the medical claims data into detailed hospital, surgical, medical, and other benefit service categories using HCG definitions and it assigns and calculates logical utilization units (e.g., admits, days, visits, procedures, or services) for each service category.
- MARA – calculates and assigns a risk score for each member using the enrollment and claims history. We use the risk scores to normalize results for health risk.
    - For this analysis Milliman uses the MARA DxOPTml Concurrent model with a limitation of 11 diagnosis codes per claim record.

[8] It should be noted that we excluded these two types of employer groups by removing all employer groups with 4-digit SIC codes between 7000 and 9721. This range is broader than just health systems and government employers, and includes some other types of employers, but due to data limitations we were not able to target these employer types alone. However, the healthcare and government employer groups tend to be the largest groups in these categories. Overall results are still highly credible and these adjustments were made to both datasets to ensure the same mix of employer group types are included in the comparisons as closely as possible.

- The 11 code per record limitation is necessary because the BCBSA data file was limited to a maximum of 11 diagnosis codes on a claim record and the MarketScan data includes more than 11 per claim record.
- Without this limitation, the risk scores for the MarketScan data have an upward bias; with it applied, the two datasets' risk scores are calculated consistently.
- This model uses only demographic and medical claim history to calculate a risk score for each member based on that data alone.

To produce the nationwide and region summaries, Milliman first assigned each dataset's members and claims to a MSA to calculate MSA level results which are then summarized to enrollment weighted average region and nationwide results using the BCBS enrollment by MSA.

The BCBSA dataset reports the member location using the 3-digit ZIP code for where the member resides. Milliman mapped each 3-digit ZIP code to one of Milliman's MSA or "non-MSA" definitions in every state.

The MarketScan data is available to Milliman at the MSA and state for where the member resides. If a member does not live in an MSA, their location is defined by "non-MSA" and state.

When two or more MSAs are located mostly within a 3-digit ZIP code, we group the Benchmark data for the MSAs together and create a broader MSA definition.

We summarize the BCBSA data for all MSAs to the state, region, and nationwide levels using the BCBS membership data as the area weight for each MSA.

We re-weight the MarketScan data's membership and claims to the BCBS membership weights for each MSA (and state) and summarize it to the state, region, or nationwide levels to normalize for MSA level geographic mix differences.

Before summarizing the MSA level results to nationwide and region results, we normalize results for age-sex mix and risk score within each MSA. To accomplish this, we do the following:

- Claims data – we calculate the allowed Benchmark per member metrics at the state, MSA, age band, sex, and group size level. We multiply these metrics by the BCBS member months for these data cuts.
- Membership data – we calculate the BCBS and Benchmark's average risk score at the state, MSA, age band, sex, and group size level and assign it to each of these data groupings.
- We multiply the risk scores and claims metrics by BCBS's member months for these data groupings. We re-weight each benchmark data cut to have the same membership level as BCBS.
- This process puts the Benchmarks on the same age-sex mix as BCBS within each MSA as well as the same age-sex and MSA mix for the region and nationwide results.

As described below, we apply summarized risk scores to calculate normalized risk scores for risk adjusting our results.

First, we combine the BCBSA and MarketScan datasets, and then we summarize the MARA risk scores for each member to calculate the total of all risk scores for both datasets combined. Then we divide by the total members (for the combined datasets) to calculate a total population average risk score.

Next, we calculate normalized risk scores by dividing each member's risk score by the total population average risk score to calculate each member's risk level relative to the total population average. (This step is done for the combined BCBS and MarketScan data population after the above exclusions and after population mix adjustments.)

We divide the claims allowed PMPM metrics by the member month weighted, normalized risk score for each of the BCBS and MarketScan data populations to calculate the risk adjusted allowed PMPM metrics for BCBS and the Benchmarks.

We apply the total risk score for each member across all service categories. This is a simplified way to adjust for health condition risk differences which is appropriate for TCC but may introduce small variances for some service category specific results.

In order to consider the potential impact of extremely large claims on the allowed PMPM comparisons for each MSA and state, and for nationwide and region results, we created a large claim adjusted set of PMPM metrics to be used for final total medical allowed PMPM comparisons.

- The first step in the adjustment approach calculates the proportion of total allowed charges in excess of one million dollars from members with more than one million dollars in allowed charges during the study year in the BCBS and MarketScan data for the geography being evaluated. We refer to these proportions as the large claimant impact.
- If the difference between the BCBS and MarketScan data large claimant impacts is greater than 1%, then the total medical PMPM for the data source with the greater large claimant impact will be adjusted down to reflect a large claimant impact that is only 1% higher than other data source's large claimant impact.

- This approach reduces variation in the total medical PMPM comparison that is caused by disproportionate amounts of claims from large claimants.

CONCLUSION

When producing a total cost-of-care benchmark comparison of allowed PMPMs from two disparate databases that each rely on enrollment and claims submissions from many different data contributors, much caution should be applied when making the calculations, analyzing the data, and drawing conclusions from it.

As we have described in our Data Sources and Methodology sections, we attempted to normalize the data for many common factors that can create variances and bias in these types of comparisons. Given our approach and because both databases have claims from several million members nationwide, we believe the results presented are a reasonable representation of BCBS' commercial large group PPO, lead network, nationwide total cost-of-care relative to the Benchmark that we calculated from the MarketScan data for this comparison.

LIMITATIONS OF USE AND DATA RELIANCE

BCBSA commissioned Milliman to write this whitepaper to present the results from our engagement with BCBSA to calculate the TCC benchmarks described in this paper. The observations and conclusions reflect the work product of the authors; Milliman does not endorse any product, third-party solution, or organization. If this paper is reproduced, it should be reproduced in its entirety as sections taken out of context can be misleading.

The analyses presented in this paper are intended to be used as an illustration of the nationwide and regional comparisons of the average allowed claim costs for commercial large group PPO networks between BCBS companies and an appropriate actuarial benchmark over a specific period of time in the past. This is represented by the Benchmark we calculated using MarketScan data. Readers should use their own expert judgement and/or seek expert advice when considering the appropriate use of these results. These results are not predictions of the future and should not be used to project any future results nor should they be relied on for making any area-specific or group-specific judgements about BCBS' total cost of care relative to other insurers or third-party administrators. Any other uses of these results may also be inappropriate. These analyses should not be relied upon by third parties. Regardless, Milliman does not intend to benefit any third-party recipient of its work product.

Models used in the preparation of our analysis were applied consistent with their intended use. We have reviewed the models, including their inputs, calculations, and outputs for consistency, reasonableness, and appropriateness to the intended purpose and in compliance with generally accepted actuarial practice and relevant actuarial standards of practice (ASOP). The models, including all input, calculations, and output may not be appropriate for any other purpose. Where we relied on models developed by others, we have made a reasonable effort to understand the intended purpose, general operation, dependencies, and sensitivities of those models.

Results presented here represent best estimates from our calculations using historical experience data. Future experience will vary from our estimates for many reasons, potentially including differences in population health status, reimbursement levels, delivery systems, random variation, or other factors.

In performing our analysis, we relied on datasets and other information provided to us by BCBSA, BHI (on behalf of BCBSA), and Merative MarketScan. We also rely on other publicly available sources in our tools and models. We have not audited or verified this data, but we have reviewed it for reasonableness. If the underlying data or other listings are inaccurate or incomplete, the results of our analysis may also be inaccurate or incomplete.

QUALIFICATIONS

Guidelines issued by the American Academy of Actuaries require actuaries to include their professional qualifications in all actuarial communications. The actuaries who prepared this paper and the results communicated in it, Adam R. Singleton, FSA, MAAA, Chis S. Tilley, FSA, MAAA, and Stuart D. Rachlin, FSA, MAAA are members of the American Academy of Actuaries, and they meet the qualification standards for performing the analysis communicated in this paper.

Milliman is among the world's largest providers of actuarial, risk management, and technology solutions. Our consulting and advanced analytics capabilities encompass healthcare, property & casualty insurance, life insurance and financial services, and employee benefits. Founded in 1947, Milliman is an independent firm with offices in major cities around the globe.

milliman.com

CONTACT

Adam R. Singleton, FSA, MAAA
adam.singleton@milliman.com

Chris S. Tilley, FSA, MAAA
chris.tilley@milliman.com

Stuart D. Rachlin, FSA, MAAA
stuart.rachlin@milliman.com