



Healthcare Cost Benchmarking for Commercial Programs

Documentation, CY2022 Data

October 10, 2024

Prepared for:

William Henderson

Principal Deputy Director
Maryland Health Services Cost Review
Commission
4160 Patterson Avenue
Baltimore, MD 21215

Submitted by:

Abt Global

6130 Executive Blvd.
Rockville, MD 20852

In Partnership with:

Milliman
Optimal Solutions Group
SAG Corporation

Contract # HSCRC-21-0274

Contents

A.	Introduction and Purpose	3
B.	Data Sources.....	3
C.	Methodological Steps.....	4
	C.1. Define Geographies	4
	C.2. Compile Characteristics for Selecting Benchmarks Communities.....	6
	C.3. Select Benchmark Regions.....	7
	C.4. Compute Metrics.....	8
	C.5. Normalize Metrics	10
	C.6. Conduct Comparisons	11
D.	Output File.....	12
E.	Data Reliance, Restrictions, Limitations	12
	Appendix A. MD APCD Exclusions and Data Cleaning	14
	Appendix B. Benefit Level Construction	20

A. Introduction and Purpose

The Maryland Health Services Cost Review Commission (HSCRC) contracted with researchers and actuaries from Abt Global and its partners, Milliman, Optimal Solutions Group, and SAG Corporation to benchmark Maryland (MD) communities to similar communities across the country to compare cost, utilization, and quality metrics for individuals covered by employer-sponsored health insurance. Used in conjunction with a similar effort for the Medicare fee-for-service (FFS) Maryland population, this information will support identifying areas to target improvement efforts, implementing incentives, and tracking performance over time.

This report documents the data sources and methodology used to construct MD regions, identify communities for benchmarking, and calculate and compare metrics for CY2022. Please also refer to the “CY 2022 Milliman Benchmark and Maryland All-Payer Claims Database Cost and Utilization Exhibits Report,” available at <https://www.milliman.com/en/insight/cy-2022-milliman-benchmark-and-maryland-all-payer-claims-database-cost-and-utilization-exhibits>.

B. Data Sources

The benchmarking methodology draws from the following data sources:

Milliman’s Consolidated Health Cost Guidelines (HCGs) Sources Database (CHSD): Milliman maintains the CHSD, a comprehensive, longitudinal, health care experience data containing detailed enrollment, medical claims, and pharmacy claims data. Milliman works with its industry partners – which include IBM MarketScan®, regional Blue Cross Blue Shield plans, and national carriers – on an ongoing basis to acquire, validate, and synthesize detailed health care data. In 2022, data for approximately 70 million commercial insured individuals were available nationwide. For the analyses presented here, the benchmark data was limited to commercially insured individuals under age 65. CHSD geographic identifiers were available at the Metropolitan Statistical Area (MSA) and state level. This was done since Medicare-covered retirees enrolled in employer-sponsored insurance could distort the per member costs if Medicare pays for the claim cost and the employer plan pays only for the remaining non-covered costs (such as drugs, deductibles and co-insurances). These data were used to identify national benchmarks and compute benchmark metrics for individuals residing in the MSAs.

Maryland All Payer Claims Database (APCD): The HSCRC provided Maryland APCD for 2022. The data contain enrollment and claims data provided by insurers for MD residents and was used to calculate MD-specific metrics. In alignment with the benchmark data obtained from the CHSD, the APCD data was limited to commercial and individual members under age 65. Other exclusions and cleaning steps are described in **Appendix A**. We used the APCD data to compute MD metrics at the ZIP-code level and aggregated to higher levels, such as MD county and Primary Service Area Plus (PSAP) hospital areas.

Primary Service Area Plus (PSAP) mappings: HSCRC provided a file on November 7, 2019 entitled “All payer and Medicare PSAP 4.2.2019.xlsx,” which assigned MD ZIP codes to MD hospitals. If a hospital served multiple ZIP-code areas, the ZIP-code areas were allocated based the share of all-payer inpatient and outpatient Equivalent Case-Mix Adjusted Discharges (ECMADs) from FY 2014 and 2015. The PSAP mappings did not change between 2018 and 2022.

Teaching costs for removal: HSCRC provided a file on December 11, 2019 entitled “Hospital file for commercial removal of GME 12-09.xlsx,” which contained estimated direct and indirect teaching costs per inpatient bed day for each MD hospital and average levels by MSA for non-Maryland MSAs. These estimates were used to remove teaching costs from APCD and CHSD inpatient metrics with the exception of residential treatment, normal newborn, and skilled nursing days. Both inpatient and outpatient GME costs were removed through the inpatient per diems. The data source for excluding teaching costs was unchanged between 2018 and 2022; costs were trended forward by 2.75% per year.

Medicare Cost Reports: We used publicly available Medicare Cost Reports from FY 2017 to obtain the percent of bed days covered by Medicare and Medicaid. Bed day statistics by payer were based on values reported by facilities and appear in Worksheet S-3 Part I of each hospital's filed Medicare Cost Report. The Medicare Cost Report data did not change between 2018 and 2022.

Publicly available geographic descriptors: We used several data sources to obtain the information needed to profile and match national benchmark regions to MD regions. These include the American Community Surveys (ACS) and Bureau of Economic Analysis (BEA). The use and timeframes of these data sources is described in greater detail below.

C. Methodological Steps

Below, we describe the methodology for identifying and comparing benchmarks for the commercially insured population. We take the following steps: (1) define the geographic area of MD and the benchmarks (2) compile characteristics used to describe the geographic areas; (3) select the benchmark regions most similar to MD regions; (4) compute the metrics to be compared between the benchmark and MD regions; (5) normalize the metrics to improve comparability between the benchmark and MD regions; and (6) compare benchmark and MD regions on cost and utilization metrics.

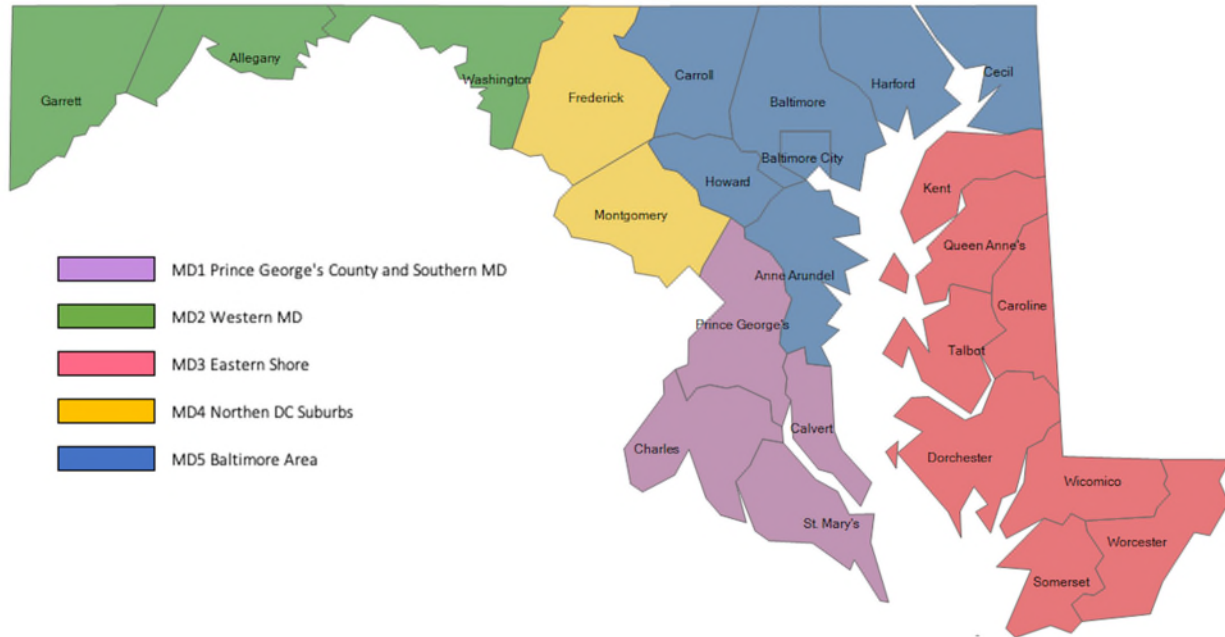
C.1. Define Geographies

The smallest geographic unit available (consistently across all data contributors) in the CHSD is the Metropolitan Statistical Area (MSA) and state. MSAs are defined by the Office of Management and Budget (OMB) and consist of clusters of counties with economic ties. The MSA generally is centered around a high population density core, such as a large city, but MSAs can also be composed of multiple smaller cities. MSAs cross state borders, but the CHSD data allow for distinguishing MSAs by state. There were 394 MSA-state areas in the U.S. in 2017 (excluding MSAs in or overlapping with MD). Geographic regions that are not considered part of any MSA are labeled as the non-MSA portion of a state (there are 50 non-MSA regions in the U.S.). We did not use any non-MSAs regions for benchmarking to MD.

To define MD areas, we delineated MSAs in MD. There are eight MSAs with all or some portion in MD, plus the non-MSA areas. We redistributed some MD counties that were geographically proximal and more economically similar, resulting in five MD regions (see **Exhibit 1**).

Changes in 2022: There were no changes in the definition of geographies between 2018 and 2022.

Exhibit 1. Maryland Counties Were Categorized into Five Regions



MD Region	FIPS County Code	County Name	Metropolitan Statistical Area
MD1 Southern MD	24037	St. Mary's County	15680
	24009	Calvert County	47894
	24017	Charles County	47894
	24033	Prince George's County	47894
MD2 Western MD	24001	Allegany County	19060
	24023	Garrett	Non-MSA
	24043	Washington County	25180
MD3 Eastern Shore	24011	Caroline	Non-MSA
	24019	Dorchester	Non-MSA
	24029	Kent	Non-MSA
	24041	Talbot	Non-MSA
	24035	Queen Anne's County	12580
	24039	Somerset County	41540
	24045	Wicomico County	41540
MD4 Northern DC Suburbs	24047	Worcester County	41540
	24021	Frederick County	43524
MD5 Baltimore Area	24031	Montgomery County	43524
	24015	Cecil	48864
	24005	Baltimore County	12580
	24025	Harford County	12580
	24510	Baltimore City	12580
	24003	Anne Arundel County	12580
	24013	Carroll County	12580
	24027	Howard County	12580

Note: Some MD counties were moved from their original MSAs when redistributing counties into the five MD regions. The recategorizations were determined in conjunction with HSCRC based on geographic and economic similarities.

C.2. Compile Characteristics for Selecting Benchmark Communities

The geographic characteristics used to describe the potential benchmark MSAs and the MD regions are listed in **Exhibit 2**. The characteristics were selected to account for existing economic and health differences across regions that would affect health care costs and utilization levels and trends of the commercially covered population. Where possible, we used similar characteristics as finalized by the Medicare benchmarking effort.

Exhibit 2. Geographic Characteristics Used to Select Benchmark MSAs

Characteristic	Data Source	Geographic Unit Used
Total population: Total U.S. population	ACS, 2013-2017 5-Year Estimates	MSA for non-MD County for MD
Population density: Total population divided obtained from the 2010 Census data divided by the land area	ACS, 2013-2017 5-Year Estimates	MSA for non-MD County for MD
Median income: median household income	ACS, 2013-2017 5-Year Estimates	MSA for non-MD County for MD
Deep poverty percentage: percentage of the population living in a household with total income below 50 percent of its poverty threshold	ACS, 2013-2017 5-Year Estimates	MSA for non-MD County for MD
Regional price parities (RPP): measure the differences in price levels across states and metropolitan areas for a given year and are expressed as a percentage of the overall national price level.	Bureau of Economic Analysis, 2016	County
HHS Platinum risk score: Risk score calculated using the publicly available HHS-HCC Platinum risk adjustment model. The HHS-HCC Platinum risk scores predict Health Insurance Exchange plan costs. The risk score is concurrent (i.e. it predicts costs for the same time period as the diagnosis collection period), based on diagnosis, and reflects a cost share rate of approximately 90%. Because the HHS-HCC Platinum risk scores are based on coefficients that represent averages for specific conditions and demographic characteristics, the risk scores generally have more explanatory power for populations than individual risk. The HHS platinum risk score model applied did not use pharmacy claims in the development of risk scores.	Milliman CHSD, 2017	MSA
Payer mix: Percent of Medicare and Medicaid bed days and excludes hospitals with unreasonable bed days counts: (1) total hospitals days less than 0, (2) days available greater than what is implied by the staffed bed count in the data, or (3) hospital and subprovider occupancy rates greater than 100% or less than 0%. Includes Medicare Advantage and managed Medicaid bed days.	Medicare Hospital Cost Reports, FY 2017	MSA for non-MD County for MD

Note: HHS = Health and Human Services; HCC = Hierarchical Condition Category; ACS = American Community Survey

The characteristics were computed at the MSA-level for each MSA nation-wide. Based on discussion with HSCRC, we assigned the RPP for MSA 47894 (Washington-Arlington-Alexandria, DC-VA-MD-WV) to MD1 (Southern MD). For the five constructed MD regions, where possible, we aggregated county-level values using total population weights. The HHS Platinum risk score was calculated from CHSD data and we assigned a score to each MD region based on mix of MSAs from which each MD county belongs, weighting to the MD region level using the MSA population counts from the CHSD. Note that this was done to select national benchmark MSAs that were similar to the MD regions. We did not use MD APCD information to select benchmark MSAs because APCD data did not become available until later in the process, after it was cleaned and analyzed; instead, we used MD data in the CHSD to compute risk scores for MD. However, below, in the discussion of the calculation and comparison of metrics, we use MD values from the APCD.

The Payer mix was calculated using the fiscal year 2017 Medicare Cost Report filing for each hospital within an MSA. Hospitals with incomplete staffed and occupied bed day information were excluded from

the calculation. The percent Medicare and Medicaid bed days was calculated by adding up the reported bed day counts for each MSA. Total occupied bed days were summarized separately from Medicare and Medicaid bed days. Traditional Medicare, Medicare Advantage, Medicaid fee-for-service and managed Medicaid bed days were included in the Medicare and Medicaid bed day count. The percent Medicare and Medicaid bed days was calculated by dividing by the Medicare and Medicaid bed days by the total occupied bed days.

Changes in 2022: There were no changes in the characteristics for selecting benchmark communities between 2018 and 2022.

C.3. Select Benchmark Regions

We created nine categories of national MSAs (excluding MD MSAs) based on percentiles of total population and population density, as shown in **Exhibit 3**. The number of categories were selected based on the frequency of MSAs that fell within each category and distribution of MD regions into categories. Due to low frequencies, MSAs falling in the lowest third of total population and the top quartile of density were moved to the second group (25-75) of total population and MSAs falling in the lowest two quartiles of density and top quartile of total population were moved to the middle category (25-75) of density. Shaded cells indicate the resulting nine categories of national MSAs for potential benchmark selection.

Exhibit 3. Categorizing National MSAs Based on Total Population and Density

	Population Density				
		<25	25-50	50-75	≥75
Total Population	<25	44	33	22	0
	25-75	56	69	62	24
	≥75	0	0	25	66

Note: Due to low frequencies, some MSAs were shifted into other categories. Shaded cells indicate the resulting nine categories of national MSAs for potential benchmark selection.

The five MD regions fell into only two of the nine categories, as shown in **Exhibit 4**. We thus limit the potential benchmark MSAs to 69 national MSAs for MD2 (Western, MD) and MD3 (Eastern Shore) and to 66 national MSAs for MD1 (Southern, MD), MD4 (Northern DC Suburbs), and MD5 (Baltimore Area).

Exhibit 4. Categorizing MD Regions Based on Total Population and Density

	Population Density				
		<25	25-50	50-75	≥75
Total Population	<25	-	-	-	-
	25-50	-	-	-	-
	50-75	-	MD2, MD3	-	-
	≥75	-	-	-	MD1, MD4, MD5

For each MD region, we applied a methodology to select comparison areas with similar characteristics by minimizing distances between the MD region and national MSAs in the assigned category across the characteristics shown in **Exhibit 2**. To do so, we calculated an Euclidean distance between the MD region and each national MSA in the category. We standardized each characteristic to have a mean of zero and standard deviation of one so that differences in scales did not influence the resulting distance scores. Thus, the Euclidean distance was based on their standard deviations rather than a mixture of units. The

distance between each MD region and national MSA was calculated by adding together the standardized differences of each of the characteristics. The formula used was:

$$D_{Nat,MD} = \sqrt{\sum_{k=1}^K \frac{(X_{Nat,k} - X_{MD,k})^2}{\sigma_k^2}}$$

Where K is the total number of characteristics; $X_{Nat,k}$ is the k^{th} characteristics for the national MSA in the category and $X_{MD,k}$ is the k^{th} characteristic for the MD region. σ_k represents the standard deviation of the k^{th} characteristic.

National MSAs with smaller distances were more similar to the MD region based on the characteristics used. For each MD region, we selected the top 20 national MSAs in terms of the lowest calculated distance. The same MSA could be (and was) selected as a benchmark for more than one region. Overall, there were 56 unique MSAs selected. The list of selected benchmark MSAs for each MD region is shown in tab “**6 Benchmark MSA descriptives**” in the attached Excel Workbook.

Changes in 2022: There were no changes in the benchmark regions between 2018 and 2022.

C.4. Compute Metrics

For the purposes of this analysis, cost refers to the allowed cost including primary insurer payments, member payments, and payments from secondary sources of coverage. The reported allowed amounts reflect CY 2022 dates of service and were adjusted for the estimated incurred but not reported (IBNR) claim costs, based on the claims completion patterns in the data. Service category was assigned using the Milliman HCG Grouper software. The HCG Grouper assigned each claim service line to a Health Cost Guidelines (HCG) service cost category based on the medical coding on the claim. In general, the following claim elements were used to assign service category:

Inpatient: Inpatient claims were identified based on the presence of room and board revenue codes and are categorized based on the diagnosis related group (DRG) assigned to the claim.

Outpatient: Outpatient claims were identified based on the presence of revenue codes on the claim or based on the provider type (e.g. ambulatory surgical center). Outpatient claims were categorized primarily based on the revenue codes and Healthcare Common Procedure Coding System (HCPCS) or Common Procedure Terminology (CPT) codes on the claim.

Professional/Other: Professional/other claims were identified based on the lack of revenue codes or based on the provider type. Physician revenue code line items billed as part of a facility claim were categorized as professional/other. Professional/other claims were categorized primarily based on the HCPCS/CPT, place of service, and provider type (e.g. home health provider) information.

The HCG Grouper categorization is continuously updated by Milliman and reflects the most current medical code sets. The HCG Grouper assigned the following traditional utilization metrics:

Admissions: Count of unique admissions for a single patient at the same facility. Interim bills and same day readmissions were combined into a single admission.

Days: Count of unique days of confinement for a single patient at a facility. Days were calculated based on the admission and discharge dates reported on the claim.

Visits: Count of unique visits for a single patient with a single provider on a unique date of service.

Procedures: Count of unique service lines for a single patient.

GlobalRVUs

Case mix was measured using the Milliman GlobalRVUs™ and Milliman Resource Based Relative Value Scale for Hospitals™ (RBRVS for Hospitals™) to assign Relative Value Units (RVUs) to inpatient, outpatient, and professional/ancillary services. Refer to “CY 2022 Milliman Benchmark and Maryland All-Payer Claims Database Cost and Utilization Exhibits Report,”¹ Attachment E for more information on GlobalRVUs. RVUs represent the relative amount of work required for each service. Dividing allowed charges by RVUs yielded a conversion factor that measured the relative case mix and severity adjusted reimbursement. An area with a higher conversion factor had higher reimbursement per unit of work than an area with a lower conversion factor.

The inpatient RVUs were assigned based on the All-Patient Refined (APR) Diagnosis-Related Group (DRG) and length of stay of each admission. The outpatient RVUs were assigned based on the HCPCS/CPT, modifier and unit coding for each service line. The professional RVUs were assigned based on the HCPCS/CPT, modifier, unit coding, and place of service for each service line.

Because the CHSD is composed of multiple contributors, we reviewed the data quality for each contributor. Some contributors were excluded based on this review. For example, we excluded contributors when the International Classification of Diseases (ICD) coding was not complete enough for DRG assignment. Additionally, we reviewed the data quality for specific service categories and if insufficient information was available to accurately assign RVUs, we imputed RVUs based on the provider type and payment information for these services. For example, we imputed RVUs when specific outpatient service categories had HCPCS/CPT code and modifiers that were not reliably populated.

Removing medical education costs

HSCRC provided the estimated teaching costs per inpatient bed day for each MD hospital. We assigned the teaching costs to the detailed APCD claims data based on the hospital identifier in the APCD. These teaching costs were assigned for all inpatient bed days except residential treatment, normal newborn, and skilled nursing days. The teaching costs were then subtracted from inpatient costs with the exception of residential treatment, normal newborn, and skilled nursing days. For the CHSD, teaching costs were assigned (and removed) based on MSA-level per day averages developed by HSCRC, rather than by hospital, because hospital-specific identifiers were not readily available in the CHSD. Both the estimated indirect medical education and graduate medical education costs were removed. Unless otherwise state, all reported inpatient costs have the teaching costs removed.

Computing metrics

Metrics were computed for each of the benchmark MSAs (using the CHSD) and averaged across the MSAs (without weights) to compute an average benchmark value per MD region. In MD, the APCD data were available at the member ZIP-code level. Metrics were computed at the ZIP-code level and aggregated up to the PSAP or the MD region, depending on the comparison (see **Section C.6**).

Additional information on the metrics used can be found in “CY 2022 Milliman Benchmark and Maryland All-Payer Claims Database Cost and Utilization Exhibits Report” available at <https://www.milliman.com/en/insight/cy-2022-milliman-benchmark-and-maryland-all-payer-claims-database-cost-and-utilization-exhibits>.

Changes in 2022: For CY 2022, we used updated CHSD data corresponding to CY 2022, but the approach for computing metrics was otherwise similar to the approach used from 2018-2021. There were no changes in the CHSD data used for the computing metrics between 2021 and 2022.

¹ Available at <https://www.milliman.com/en/insight/cy-2022-milliman-benchmark-and-maryland-all-payer-claims-database-cost-and-utilization-exhibits>

C.5. Normalize Metrics

Although benchmark MSAs were selected to be more similar to the MD regions, MSAs are fairly large and could differ from MD regions and hospital service areas (PSAPs) within the five Maryland regions. To enhance comparability within the five MD regions and to address the residual differences between the MD regions and the benchmarks, we developed a regression adjustment methodology to further control for differences in the characteristics of median income and the deep poverty percentage. We also added a measure of plan benefit levels to control for the generosity of coverage since commercial benefits can vary, affecting utilization. The adjustment factors used are shown in **Exhibit 5**.

Exhibit 5. Adjustment Factors

Factor	Data Source
Median income: median household income	ACS 2018-2022 5-Year Estimates
Deep poverty percentage: percentage of the population living in a household with total income below 50 percent of its poverty threshold	ACS 2018-2022 5-Year Estimates
Benefit levels: Based on average member cost sharing levels to account for differences in member utilization level caused by plan design. Appendix B provides additional detail on how benefit levels were calculated.	2017 Milliman CHSD and 2022 MD APCD
HHS Platinum risk score: Risk score developed by HHS to predict Health Insurance Exchange plan liability. ²	2022 Milliman CHSD and MD APCD

Note: HHS = Health and Human Services; ACS = American Community Survey

Prior to the regression calculations, we risk score and benefit-level normalized both total PMPM allowed charges and total PMPM RVUs using the HHS Platinum risk score standardized to the MD average risk score and plan benefit levels. To standardize the risk score, we divided each benchmark MSA's risk score by the average MD risk score across the state, weighted by ZIP-level membership. The average MD risk score in 2022 was 1.356536. We then divided total PMPM allowed charges and total PMPM RVUs by each region's standardized risk score and benefit-level.³

Next, we followed the steps below to further adjust the risk score and benefit-level normalized allowed charges for demographic differences between the MD and benchmark regions. No further adjustments were made to RVUs based on the regressions.

1. Regressed total risk score adjusted PMPM allowed charges on median income, the deep poverty percentage, and the benefit-level adjustment factor. The regression was conducted on only benchmark MSAs. For the 2022 data, we used the same set of demographic adjustment coefficients as calculated using the 2018 -2021 data. We continued using the coefficients calculated using the 2018 data, because the coefficients from the 2018, 2021, and 2022 data were

² The MD risk score was calculated using the APCD identifier field named "Master Patient Index (MPI)" to ensure that unique members could be identified consistently throughout each calendar year. For the purpose of matching claims data to enrollment, "concatenated payer-encrypted enrollee identification number, birth date, and gender (PIDBDGP)" was used. Please refer to the "Other Inconsistencies in Member ID" notes in Appendix A for more information.

³ The HHS platinum risk scores predict aggregate costs for all services, not costs by service category. However, we have applied the risk scores to normalize costs for each major service category (inpatient, outpatient, and professional/other) in some analyses. Caution should be used when interpreting these results since costs by service category may not vary linearly with risk score.

similar, and we felt the benefits of added stability in the model outweighed the small change in the results. **Exhibit 6** provides the regression coefficients.

Exhibit 6. Regression of Allowed Per Member Per Month Charges on Normalization Factors Using Benchmark MSAs, CY 2018

Covariates	Coefficient	P-value	R-squared	N
Median income	\$2.28 per \$1,000 increase in median income	0.000	0.2799	56
Deep poverty percent	\$10.50 per 1 percentage point increase in deep poverty	0.087		
Constant	\$168.4	0.013		

Note: We conducted a linear regression of risk score adjusted PMPM allowed charges on covariates using de-duplicated benchmark MSAs using CY 2018 data.

- Obtained an expected adjusted total PMPM allowed charges for each benchmark MSA as well as MD region from the regression results.
- Aggregated benchmark MSAs to each MD region (20 per region) using a straight, unweighted average.
- Calculated observed to expected ratios (O/E) for each benchmark MSAs and MD region.
- Calculated mean benchmark adjusted total PMPM allowed charges, unweighted across de-duplicated benchmark MSAs (56 MSAs). In calculating the demographically adjusted values for 2019, we used the average benchmark adjusted total from the 2018 data. Holding the average benchmark value constant with 2018 allows the 2022 demographic-adjusted values to be more directly comparable to the values from 2018-2021. Any differences between the demographically adjusted values in 2022 and prior years will be due to differences in the O/E ratios or to changes over time in median income or deep poverty percent (the adjustment factors included in the regression).
- Multiplied the expected ratio (#4) by the mean benchmark value (#5) to obtain adjusted values for each MD region and each benchmark MSA. This resulted in demographic-adjusted total PMPM allowed charges for each benchmark MSA and each MD region.

The MD regions and sub-regions used in Steps 2-6 could be defined as the ZIP, county, region, or Primary Service Area Plus (PSAP). ZIP-code level values were rolled up to higher levels by weighting by the number of member months. For PSAP roll-ups, HSCRC provided shares of each MD ZIP-code that was assigned to each PSAP.

Changes in 2022: For CY 2022, we used updated measures of median income and deep poverty corresponding to CY 2022, but the approach for normalizing metrics was otherwise similar to the approach used from 2018-2021. Notably, we used the same risk adjustment coefficients and average benchmark adjusted totals as used from 2018-2021.

C.6. Conduct Comparisons

We compared the demographic adjusted total PMPM allowed charges of benchmark MSAs to MD regions. For each region, we calculated the average of the 20 benchmark MSAs (without weighting). When comparing to MD PSAPs, we blended the benchmark MSAs when necessary. For example, some MD PSAPs include MD ZIP-codes in more than one MD region. When this occurred, we weighted the by the proportion of the applicable MD region overlapping with the PSAP.

Changes in 2022: For CY 2022, we conducted comparisons similarly for CY 2018, CY 2019, and CY 2021.

D. Output File

The attached workbook, entitled “Commercial_Benchmark_CY2022_Data_20241008.xlsx” contains descriptions of the MD areas, benchmark MSAs, and resulting comparisons. Each worksheet is listed and described below:

0. **Purpose and Caveats:** these are also repeated in **Section E** below.
1. **Tab Descriptions:** A summary of the contents included in each tab.
2. **MD regions:** the list of MD counties mapped to the five MD regions and the MD map.
3. **Data dictionary:** a brief description the metrics/fields in the spreadsheet.
4. **County BM summary:** MD region, county, and aggregated benchmark comparisons on characteristics and allowed PMPM charges, risk score and benefit-level adjusted allowed PMPM charges, and demographic-adjusted PMPM charges.
5. **County BM ratios:** MD region, county, and aggregated benchmark ratios on allowed PMPM charges, risk score and benefit-level adjusted allowed PMPM charges, and demographic-adjusted PMPM charges.
6. **Benchmark MSA descriptives:** characteristics of individual benchmark MSAs.
7. **County Cost Comparisons:** comparison of MD regions and counties to aggregated benchmarks on various metrics, including cost per RVU, RVUs per member per year, etc.
8. **PSAP county shares:** listing of PSAP and PSAP shares within each MD county.
9. **PSAP descriptives:** PSAP characteristics and allowed charges.
10. **PSAP BM summary:** PSAP and aggregated benchmark comparisons on characteristics and allowed PMPM charges, risk score and benefit-level adjusted allowed PMPM charges, and demographic-adjusted PMPM charges.
11. **County PSAP compare:** comparison of PSAP and county characteristics
12. **PSAP cost ratios:** PSAP and aggregated benchmark ratios on allowed PMPM charges, risk score adjusted allowed PMPM charges, and demographic-adjusted PMPM charges.
13. **PSAP cost comparisons:** comparison of PSAP to aggregated benchmarks on various metrics, including cost per RVU, RVUs per member per year, etc.

Changes in 2022: There were no changes in the format of the output delivered between 2018 and 2022.

E. Data Reliance, Restrictions, Limitations

Milliman has prepared the 2022 benchmark and Maryland APCD processed and summarized data for the use of the Maryland HSCRC. This information is intended solely for educational purposes and presents information of a general nature. It is not intended to guide or determine any specific individual situation and persons should consult qualified professionals before taking specific actions. Milliman does not intend to benefit or create a legal duty to any third party recipient of its work.

In preparation of the analysis, Milliman relied upon the accuracy of data and information gathered from or provided to us by the Centers for Medicare and Medicaid Services (CMS), HHS, and our data partners. Milliman has not audited this information, although it has been reviewed for reasonableness. If the

underlying data or information is inaccurate or incomplete, the results of our analysis may likewise be inaccurate or incomplete.

Milliman has also relied on the data and other information provided by the HSCRC for this analysis. Milliman has performed a limited review of this data and other information and checked for reasonableness and consistency. Milliman has not found material defects in the data or information used other than those described in this report, which also describes how those defects were addressed to enable this analysis to be reliably performed. If there are other material defects in the data or other information, it is possible that they would be uncovered by a detailed, systematic review and comparison of the data to search for data values that are questionable or for relationships that are materially inconsistent. Such a review was beyond the scope of this assignment.

Appendix A. MD APCD Exclusions and Data Cleaning

Maryland’s 2022 APCD was used to calculate the 2022 Maryland commercial and individual medical cost and utilization. Milliman processed the APCD eligibility and detailed claims information and calculated metrics consistent with the 2022 Milliman Benchmark Database. The data used in this analysis were limited to commercial and individual members under age 65. Milliman worked with HSCRC and Abt Global to perform the following steps to clean the APCD data:

1. **Imported APCD:** Imported APCD and processed the eligibility and detailed claims data. Milliman utilized the data dictionaries, guidance, and other information about the Maryland APCD provided by HSCRC.
2. **Included only commercial and individual experience:** Milliman included only the APCD to commercial and individual experience, using the “MARKET_SEGMENT” field. Specifically, Milliman included only the following:

MARKET_SEGMENT	Description
3	Individual Market (not sold on MHBE)
5	Private Employer Sponsored or Other Group
7	Public Employee – Other
8	Small Business Health Options Program (SHOP) not sold on MHBE
A	Student Health Plan
B	Individual Market (sold on MHBE)
C	Small Business Health Options Program (SHOP) sold on MHBE

Individual and ASO experience is excluded for the purpose of reconciliation to financial statements (see below). Non-commercial and non-individual benefit plans are excluded: Medicaid, Medicare Advantage, and Medicare Supplement. Additionally, we understand that some self-funded employers do not submit experience data to the Maryland APCD, and therefore will not be included in our summaries.

Due to inconsistent reporting of Federal Employees Health Benefits Program (FEHBP) data by the APCD contributors, we excluded FEHBP claims and membership in the CY 2022 data (MARKET_SEGMENT = “6”). In prior reports published in CY 2017 and CY 2018, FEHBP claims had been available and included for some insurers (CareFirst Blue Choice, Inc., CareFirst of Maryland, Inc., and Group Hospitalization and Medical Services, Inc.). However, starting with the CY 2019 APCD, these insurers ceased reporting FEHBP experience, so all FEHBP claims were excluded from CY 2019, CY 2021, and CY 2022 reports.

3. **Reconciliation and validation of APCD:** Milliman reviewed the APCD for reasonableness and checked the medical code values for validity and consistency with current coding standards. Milliman did not audit the APCD. To reconcile the APCD, Milliman compared the commercial enrollment and paid claim cost amounts for each APCD contributor to readily available MD financial statement reports. Milliman worked with HSCRC to exclude payers with incomplete or invalid APCD submissions.
4. **Completed the APCD for incurred but not reported (IBNR) claims:** Milliman calculated and applied completion factors for IBNR amounts. The CY 2022 APCD had three months of runout, so Milliman reviewed the completion patterns for the 2017, 2018, 2019, 2021, and 2022 APCD datasets and developed an overall CY 2022 completion factor of 0.9813 (i.e. Milliman estimated that the data was 98.13% complete). Milliman applied the completion factors to the allowed, utilization, and RVU amounts.

5. **Assign metrics:** Process the Maryland APCD data and assign analytic measures using:
 - a. The Milliman Health Cost Guidelines™ (HCG) Grouper™ to assign service category and standard utilization counts,
 - b. The Milliman GlobalRVUs™ to assign RVUs to all medical services, and
 - c. The 2017 HHS-HCC platinum risk score model to assign risk score. Additionally, benefit adjustment factors were assigned based on the observed average member cost sharing levels to adjust for the estimated impact of benefit levels on member utilization.
6. **Remove estimated teaching costs:** Remove teaching costs from the detailed APCD data using the hospital identifier in the APCD data and the HSCRC's estimate of hospital teaching costs per bed day. The HSCRC provided the estimated teaching costs per inpatient bed day for each Maryland hospital. These teaching costs were assigned for all inpatient bed days except residential treatment, normal newborn, and skilled nursing days.
7. **Develop cost-model summaries:** Tabulate results by area and service category. The area-level results include county- and PSAP-level summaries and cost models. Cost and utilization are reported by service category. Risk and benefit adjusted costs and RVUs are calculated overall for each area.

Major defects Milliman found with the APCD and their workarounds were:

1. **Incomplete data for some contributors:** As discussed above, Milliman compared the commercial membership and costs for each contributor to readily available financial reports. Milliman worked with the HSCRC to review inconsistencies and HSCRC determined the final set of payers' APCD experience to include.
2. **Comparison to NAIC statements:** Milliman compared APCD data to National Association of Insurance Commissioners (NAIC) Annual Statements for the Year 2022 and NAIC 2022 Supplemental Health Care Exhibit (SHCE) reports. The results of these comparisons are included in Attachment D-1 of Milliman's Report.⁴ The primary purpose of these comparisons is to understand the completeness of carriers' APCD data submission and to help inform the HSCRC in their discussions with carriers. As a result of the lack of agreement between these data sources, additional information was used to reconcile the APCD data used in our processes. See (3), below, for more information.
3. **CareFirst reconciliation:** CareFirst's APCD experience (including the data for CareFirst of Maryland, Inc. and Group Hospitalization & Medical Services, Inc.) does not reconcile to their National Association of Insurance Commissioners (NAIC) Annual Statement for the Year 2022 or their NAIC 2022 Supplemental Health Care Exhibit Report. For CareFirst of Maryland Inc. and Group Hospitalization and Medical Services Inc. contributors to the APCD, the total paid dollars reported in the APCD are approximately 1% lower than the paid amounts in the NAIC Annual Statement after accounting for prescription drug rebates. In contrast, reported member months in the APCD are 5% higher than the member months reported in the NAIC statement. The CareFirst organizations represent almost half of the total paid dollars included in the APCD and the results in this report are sensitive to errors contained in this data. In the past, the HSCRC worked directly with CareFirst to reconcile the APCD data to CareFirst's internal reporting, and the findings of this reconciliation process were implemented into the processing of the APCD data for this report. However, for this processing cycle, the HSCRC provided Abt and Milliman

⁴ Available at <https://www.milliman.com/en/insight/cy-2022-milliman-benchmark-and-maryland-all-payer-claims-database-cost-and-utilization-exhibits>

with summary reports of CareFirst's 2022 allowed dollars which Milliman reconciled to while developing this report's APCD results.

4. **Incomplete ICD-10-PCS (procedure code) data:** The APCD contained only procedure codes in the first position and did not contain any secondary procedure code information. This defect resulted in Milliman being unable to reliably assign APR DRG groupings and, instead, relying on the APR DRGs provided by the HSCRC with limited review. It is our understanding that the APR DRGs supplied by the HSCRC were assigned based on claims data which included secondary procedure codes. This understanding is supported by our review of the supplied data and the observed distribution of APR DRGs. See item 2-c in the below Section, Summary of Changes from Prior Reporting.
5. **Missing Subscriber IDs, maternity claims, and newborn claims and enrollment:** In prior years, where a maternity claim indicated a newborn and a subscriber ID did not include an associated newborn, Milliman added a newborn eligibility record to the enrollment dataset. Additionally, where no newborn claim was found, the maternity claim was assumed to be a combined mother and baby claim, and RVUs were added to the maternity claim to account for the utilization associated with the newborn.
 - For the 2022 reporting results, it was not possible to link newborns' enrollment or claims to the mother's maternity claim using a Subscriber ID. This difference may result in too many newborn-associated RVUs being added to the maternity claims. Moreover, because newborn enrollment records were still present in the data, but we were unable to link them with maternity or newborn claims, we did not add any newborn enrollment to the enrollment dataset.
 - Ultimately, we observed 1.20 newborns or neonates per maternity claim. Based on a benchmark rate of 1.03 babies per mother from the Health Cost Guidelines, this may represent an overstatement of newborns of approximately 16 percent. However, because of the low average service intensity of newborn claims, we estimate that this amounts to an overstatement of less than 0.02 total RVUs PMPM, or a potential RVU overstatement of 1.7 percent for inpatient services.
6. **Risk score decreases between 2021 and 2022:** Risk scores have declined between the 2021 and 2022 results. We have investigated this decline to rule out causes including the back-mapping of diagnosis codes from CY 2022 to Federal Fiscal Year (FFY) 2021. We have found that the observed risk score reductions have varied across two age cohorts:⁵
 - Individuals aged 1 to 64, inclusive, have had risk scores decline by an average of 2.8%. We have reviewed the matching of claims records to enrollment records and exposure period per member and have not identified significant changes from prior years. After comparison to the benchmark results, a review of PSAP-level changes, and discussion with HSCRC, we have concluded these reductions may be reasonable.
 - Individuals of age zero (i.e., born in 2022) have had risk scores decline by an average of 36%. Based on comparisons to prior results and an analysis of the claims data available for these members, we believe this decline is attributable to a significant decrease in the number of newborn claims reported in the APCD. Missing newborn claims results in a reduction in the HHS-HCC risk score, and we believe the reductions are attributable to this. Because of the possible sensitivity of our results to risk score changes, we have

⁵ Note that the reductions in risk scores discussed here are on a pre-exclusion basis. I.e., while we have restricted the analysis to individuals of age 64 or less, we have not excluded members enrolled in other excluded cohorts (e.g., FEHBP enrollees who are 64 years old or younger).

sensitivity tested the impact of a uniform underestimate of risk scores for newborns on a PSAP basis. This test was performed by artificially increasing all newborn risk score by a factor of 1.56 (i.e., $1 / (1 - 0.36)$). Based on this analysis, it appeared that the PSAP-level results and rankings were not heavily impacted by a uniform underestimate of newborn risk scores

Exhibit A-1 below shows the starting allowed amounts and the allowed amounts added for the IBNR amounts.

Exhibit A-1 – Summary of 2022 Maryland APCD Adjustments

Adjustment	Allowed	Percent of Total
Total	\$18,600,330,456	100.0%
Payer and line of business exclusions	\$12,766,381,342	68.6%
Commercial and Individual Subtotal	\$5,833,949,114	31.4%
+ IBNR Amount	\$111,127,137	0.6%
- Teaching Costs	\$59,388,454	0.3%
= Amounts in Cost Model	\$5,885,687,797	31.6%

SUMMARY OF CHANGES FROM PRIOR REPORTING

1. Changes between the 2021 and 2022 processing:
 - a. For the 2022 reporting process, the APCD processing was completed in collaboration with Abt Global and the SAG Corporation. Abt and SAG received claims data from HSCRC and prepared the raw data for processing by Milliman’s HCG Grouper and GlobalRVUs. SAG also provided Milliman with summary data for reconciliation and data quality review. More information on these Milliman products may be found below in the Service Category Assignment and Unit Price Normalization sections below, and Appendix F of Milliman’s Report.⁶
 - b. Per direction from HSCRC, SAG relied on the “versioned” APCD data as the starting data source for their processing. In prior years, HSCRC has directed Milliman to use the “unversioned” data source as the base data source. See more information on the “versioned” and “unversioned” data sources below.
 - c. We made annual updates to the risk normalization factor, completion factors, and teaching cost estimates.
 - d. For federal fiscal year (FFY) 2022 (i.e., the fiscal year ending September 30, 2022), CMS made significant revisions to the mapping from ICD-10-CM (diagnosis) codes to condition categories. In prior years, the mapping update had been restricted to more minor revisions and the incorporation of changes to diagnosis codes (e.g., new codes being added).⁷ Because of this change, we have applied the FFY 2021 diagnosis code to condition category mapping (i.e., the mapping from the latest version of the 2020 HHS-HCC risk score model) to our CY 2022 claims data by first mapping the claims’

⁶ Available at <https://www.milliman.com/en/insight/cy-2022-milliman-benchmark-and-maryland-all-payer-claims-database-cost-and-utilization-exhibits>

⁷ See, for example the August 2021 revisions reported here: <https://www.cms.gov/sites/default/files/2022-04/CY2021-DIY-instructions-03.31.2022-FINAL.pdf>

diagnosis codes from a CY 2022 basis to a FFY 2021 basis. We then apply our standard methodology, mapping these diagnosis codes from a FFY 2021 basis to the condition categories used in the HHS-HCC model.

2. Changes between the 2019 and 2021 processing:
 - a. We made annual updates to the risk normalization factor, completion factors, and teaching cost estimates.
 - b. Because 2021 MarketScan data was not available at the time of the benchmark data processing to support these analyses, the MarketScan data was excluded from the 2021 benchmark results. One consequence of this exclusion is a decline in the member months reported in the benchmark exhibits. In light of this data source change, Milliman confirmed the 2021 benchmark results were reasonable and consistent with prior results. No adjustment was made to the benchmarks to account for the change in the available benchmark data sources.
 - c. Milliman is now including individual market data in the benchmarks.
 - d. Like the 2019 processing, the HSCRC provided supplemental files in addition to the claims data extracts which contained corrected and additional APR DRG information. The first of these supplemental files was provided as a SAS dataset, “ip_claims_drg_review_hscrcfinal.sas7bdat,” on March 22, 2023. A second file, “hscrc_aprdrgrg_xwalk_2021_oct23.sas7bdat,” which was used to improve the APR DRG assignment rates for the subset of inpatient claims that did not have DRGs reported in the first supplemental file, was provided on November 11, 2023. After both files’ DRGs were applied, DRG availability by broad service categories appeared comparable to prior years’ analyses.
 - e. In prior years, Milliman attempted to account for apparently unreported newborns in the enrollment dataset. This year, Milliman has not made such an adjustment. (This is discussed in more detail below.)
 - f. Because of observed inconsistencies between reported allowed and the sum of paid, coordination of benefit (COB), and patient paid dollars and at the direction of the HSCRC, the patient paid amount reported in the exhibits is calculated as allowed minus paid and minus COB. In prior years, the exhibits reflected the patient paid amount reported in the claims data provided by the HSCRC.
 - g. For the APCD processing, and consistent with prior years, secondary ICD-10-CM codes (diagnosis codes) from the professional service claim dataset are excluded from the APCD risk adjustment process. Results including all diagnosis codes from the professional service claims dataset are included in the appendices.
 - h. Other minor processing changes including changes to the data mapping and cleaning process.
3. Changes between 2018 and 2019 processing
 - a. Annual updates to the risk normalization factor, completion factors, and teaching cost estimates
 - b. Federal Employees Health Benefits Program (FEHBP) data is now excluded from all APCD carriers’ results. In prior reports, this data was excluded for only Aetna Health Inc. Starting in 2019, FEHBP no longer submits claims to the APCD for any carriers and is therefore no longer available for inclusion.

- c. Milliman relied on APR DRGs contained in the APCD data provided by HSCRC, as well as a supplemental file containing corrected DRGs for some claims, provided by HSCRC on January 26, 2022. For prior reports, Milliman reassigned APR DRGs as part of our processing. The supplemental file used more detailed information contained in the HSCRC's Casemix hospital discharge dataset to re-assign DRGs where the HSCRC believes limitations on data contained in APCD resulted in suboptimal mapping.
 - d. The 2019 Maryland benchmarks were built on the “unversioned” APCD data. This is a change from 2018 which was built on the “versioned” data. The difference is in how a small percentage of overlapping claims are treated. The process of reconciliation with a major carrier conducted as part of the 2019 benchmarking showed that the “unversioned” data was a better match for the correct total received from the carrier.
 - e. Other minor processing changes including slight changes to the area information used to identify Maryland claims, changes to the data cleaning process, etc.
4. Changes between 2018 and 2017 processing
- a. Annual updates to the risk normalization factor, completion factors, and teaching cost estimates
 - b. For the purposes of risk score development, the member identifier MUID was used. For the 2017 report, PIDBDGP was used for the development of risk scores.

For additional detail on the APCD validation and exclusions applied to the APCD, see “CY 2021 Milliman Benchmark and Maryland All-Payer Claims Database Cost and Utilization Exhibits Report,” available at <https://www.milliman.com/en/insight/cy-2022-milliman-benchmark-and-maryland-all-payer-claims-database-cost-and-utilization-exhibits>.

Appendix B. Benefit Level Construction

Milliman developed plan benefit normalization factors to account for allowed cost and utilization differences caused by plan design. The allowed PMPM cost factors normalize both utilization and unit price. The RVU PMPM factors normalize utilization and service intensity, but not unit price. The plan benefit factors were developed as follows:

1. Milliman calculated coinsurance at the employer group and product level using the 2017 Milliman benchmark data. This process effectively assumed all members with the same product and employer had the same benefit levels.
2. The plan benefit factors are calculated as the ratio of the risk adjusted allowed amount for each coinsurance level to the total risk adjusted amount. These factors are calculated for both allowed dollars and RVUs and by inpatient, outpatient, and professional service categories separately.
3. We reviewed the results with the HSCRC and made one smoothing adjustment: the coinsurance 95% plan benefit factor for professional services was increased by refining the data contributors included in the calculation.

Exhibit B-1 below shows the plan benefit factors based on the 2017 CHSD benchmark data.

Exhibit B-1. Benchmark Plan Benefit Factors Combined IP, OP, and Professional Coinsurance Range Commercial, All MSAs, Limited to Under Age 65

2017 Milliman Benchmark (Nationwide)	Plan Benefit Factors					
	RVU			Allowed		
Plan Coinsurance Range	IP	OP	Prof	IP	OP	Prof
95%+	1.142	1.061	1.070	1.216	1.063	1.050
85%-95%	1.043	1.039	1.032	1.054	1.037	1.046
75%-85%	0.992	0.987	0.989	0.973	0.995	0.986
65%-75%	0.850	0.910	0.942	0.821	0.895	0.934

To apply the normalization factors, Milliman calculated the average overall coinsurance by area and interpolated between the plan benefit factors. For example (values are illustrative):

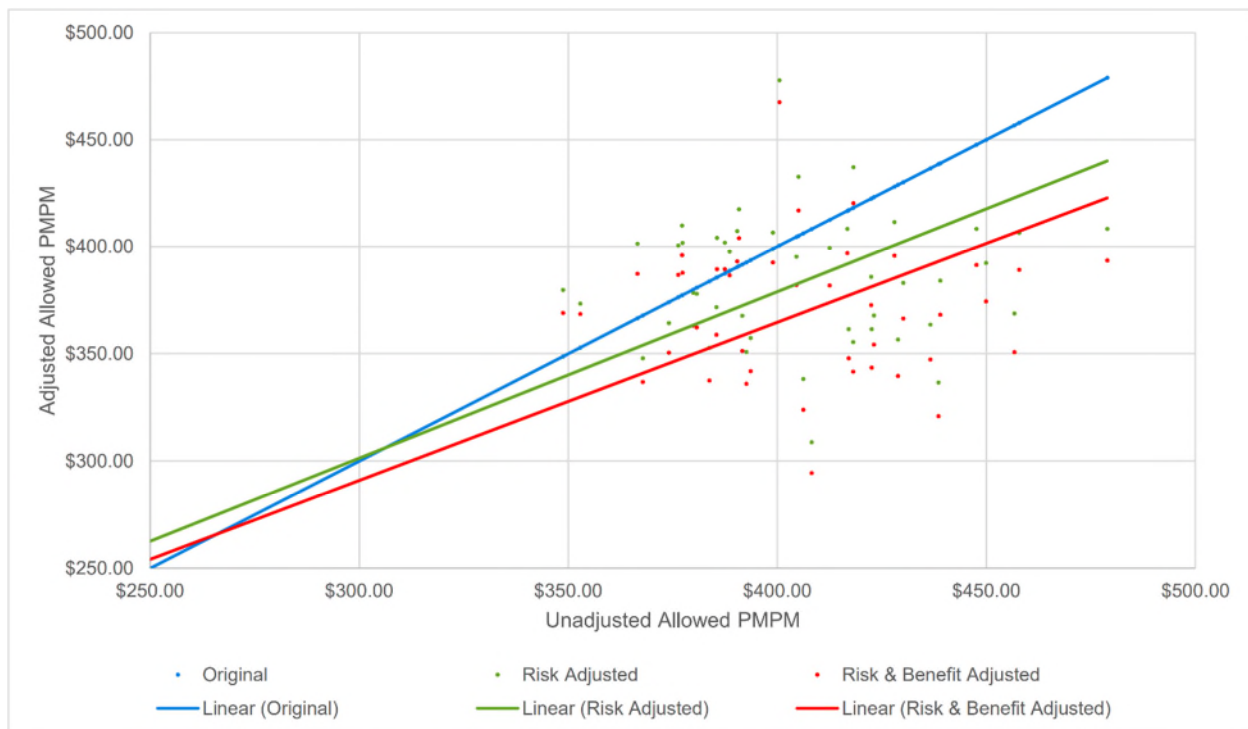
1. For a MSA, the average overall coinsurance is 84% and the risk adjusted inpatient (IP) allowed PMPM is \$100.00.
2. Milliman interpolated from mid-point to mid-point:
 - a. Midpoint of range [85% - 95%] = 90% (high bound), and
 - b. Midpoint of range [75% - 85%] = 80% (low bound).
 - c. The weight given to the factors in the [85%- 95%] range is (4/10) which was calculated as $(84\% - 80\%) / (90\% - 80\%) = [(actual\ overall\ coinsurance) - (low\ bound)] / [(high\ bound) - (low\ bound)]$.
 - d. The weight given to the factors in the [75% - 85%] range is (6/10) which was calculated as $1 - (4/10)$.
3. The plan benefit factor was $((4/10) \times (1.054)) + ((6/10) \times (0.973)) = 1.0054$.

4. Please note Milliman set the lower bound at 50% and upper bound at 100%. However, this generally did not affect the results since Milliman was looking at the overall coinsurance by MSA.
5. The plan benefit-adjusted IP allowed is calculated as (risk adjusted IP allowed PMPM) / (interpolated plan benefit factor) = $\$100 / 1.0054 = \99.46 .

Milliman calculated the plan benefit factors by area for both the CHSD benchmark results and the MD APCD. **Exhibit B-2** below shows the impact of risk and plan benefit-adjusting the allowed PMPMs on the 2022 CHSD benchmark MSAs:

1. The blue line is the original allowed PMPM,
2. The green line is the risk adjusted allowed PMPM, which has a flatter slope than the original unadjusted allowed PMPM line, and
3. The red line is the risk and plan benefit-adjusted allowed PMPM.

Exhibit B-2: SCATTERPLOT OF APCD DATA ALLOWED PMPM FOR EACH PSAP, RISK ADJUSTED ALLOWED PMPM, AND RISK & PLAN BENEFIT ADJUSTED ALLOWED PMPM



Note: Based on data from “CY 2022 Milliman Benchmark and Maryland All-Payer Claims Database Cost and Utilization Exhibits Report,” Exhibit 2.

This figure shows that after normalizing for risk score, the allowed PMPM moved towards the average allowed PMPM across all areas with limited exceptions – resulting in a more horizontal linear fit line. The results were similar for RVUs PMPM, but with the benefit normalization having a less significant effect in that case.