sbäctuaries

To:	UHC File
Subject:	UHG Outcomes Evaluation Methodologies
From:	Ian Duncan PhD FSA FIA FCIA FCA CSPA MAAA
Date:	March 16th 2022
c.c.:	Tony Pistilli FSA MAAA CERA; Andrew Mackenzie FSA MAAA CERA

Santa Barbara Actuaries Inc. has reviewed a summary of methodologies used to report financial savings of certain UnitedHealthcare (UHC) medical management, payment integrity, advocacy, and out-of-network management programs to customers of UHC. We did not perform primary analysis of UHC evaluations but instead reviewed summary methodology descriptions and had the opportunity to discuss these methodologies in more detail with representatives of UHC who provided additional insights. We did not have the opportunity to conduct primary evaluation of any studies; this certification is based entirely on methodology documentation provided by UHC and discussion with UHC representatives.

In aggregate, we believe that UHC evaluates programs using transparent and robust methodologies for 100% of payment integrity, advocacy, and out-of-network management programs, and the majority of its medical management programs (representing over 90% of book-of-business reported savings). UHC considers the savings generated by individual programs to be proprietary; we have had an opportunity to evaluate UHC's reported savings numbers and document reported savings ranges for different types of programs below. Only a minor portion of the medical management savings are calculated according to methodologies that do not have a standard industry methodology, or where there is insufficient detail available to evaluate the robustness of the method. All payment integrity, advocacy, and out-of-network management savings are calculated according to robust methodologies which are either industry standard or follow industry best practices.

Overall, we believe that the majority of UHC's evaluation methodologies are consistent with industry best practices, are likely to report a fair estimate of program savings and are likely to be accepted by employers.

UHG employs a number of different methodologies for the eleven medical management and fifteen payment integrity programs evaluated. Our opinion varies by program:

Santa Barbara Actuaries Inc.
3221 Calle Mariposa
Santa Barbara CA 93105

M: 805-451-0950

Programs with transparent, robust, and industry-standard methodologies

1. Medical Management Programs
    a. Case Management
    b. Renal Management
    c. Condition Management
    d. Treatment Decision Support
2. Advocacy Program
3. Payment Integrity Programs
    a. Pre-pay Payment Integrity (all ten programs)
    b. Post Pay Payment Integrity (all eight programs)
4. Out-of-network Management

The first 4 programs account for 75 - 80% of the total medical management PMPM savings for all reported medical management programs. The latter 4 programs cover 100% of the payment integrity, advocacy, and out-of-network management program savings.

Programs that tie gaps-in-care closure to an economic effect using reasonable methodologies

5. Medical Management Programs
    a. Oncology Management
    b. Maternity Management
    c. 24/7 Access to Care
    d. Wellness Services

These 4 programs account for 15 - 20% of the total Medical Management PMPM savings for all reported programs. There is no industry-standard methodology available to evaluate programs of this type. The methodologies appear to be reasonable from the limited review SBA was able to perform.

Programs for which there is no standard industry methodology

6. Medical Management Programs
    a. Neonatal Resource Services
    b. Congenital Heart Disease
    c. Bariatric Resource Services

These programs account for a minor part of the total reported savings for all programs.

A minority of programs cannot be evaluated based on the information provided and/or we are unable to judge the reasonableness of their results. However, these programs account for a small proportion of the total savings from all programs and do not detract from the overall savings generated by UHC's programs.

In my opinion the majority of UHC's evaluation methodologies are consistent with industry best practices, are likely to report a fair estimate of program savings and are likely to be accepted by employers.

I, Ian Duncan PhD FSA FIA FCIA FCA CSPA MAAA am qualified to provide an opinion on UHG's methodologies based on my training and experience.

APPENDIX: Detailed discussion of methodologies

This memo documents the different intervention evaluation methodologies for which UHC has asked Santa Barbara Actuaries (SBA) to provide an opinion. This includes 13 medical management programs, 15 payment integrity programs, the three levels of the Advocacy program, and the Naviguard out-of-network management product. This memo will discuss these programs in that order.

Medical Management Programs Reviewed

Initially UHC identified 13 medical management programs; subsequently 2 programs were eliminated: Emergency Room Decision Support because the program is being discontinued, and Transplant Resource Services, for which UHG is revising its methodology. We did not evaluate 3 programs (Congenital Heart Disease, Bariatric Services and Neonatal Resource Services) for which limited information is available. These programs account for a minor proportion of the overall book-of-business savings. The remaining 8 programs whose methodologies we have evaluated account for 93% of UHG's reported savings.

We reviewed a summary of UHC methodologies documented in a PowerPoint presentation entitled: "Medical Cost Savings for 2020 Optum Programs."

Medical Management Program Name	Embedded / Buy-Up	Methodology
Case Management	Embedded	Participant vs. Non-participant
Renal	Embedded	Participant vs. Non-participant
Emergency Room Decision Support	Discontinuing	Participant vs. Non-participant
Condition Management	Embedded	Trend-adjusted pre-post
Treatment Condition Support	Embedded	Utilization reduction
Oncology Management	Buy-Up	Utilization reduction
Maternity	Embedded	Utilization reduction
Neonatal Resource Services	Buy-Up	Utilization reduction
Wellness Services	Buy-Up	Utilization reduction
24-7 Access to Care	Embedded	Utilization reduction
Transplant Resource Services	Excluded	Network/referral evaluation
Congenital Heart Disease	Embedded	Network/referral evaluation
Bariatric Resource Services	Buy-Up	Network/referral evaluation

UHG has emphasized that their methodologies are designed to be effective for a large book-of-business and many customer evaluations. Our review considers this trade-off between accuracy and efficiency of volume processing. Methodologies: We discuss each methodology in turn.

Payment Integrity Programs Reviewed

This memo documents the different intervention evaluation methodologies for which you have requested Santa Barbara Actuaries' (SBA) opinion. UHG identified eighteen programs in total (below). There are two methodologies that are applied to these eighteen programs corresponding to at which point in the adjudication process a claim is stopped (indicated by the "Methodology" column).

UHC employs two key evaluation methodologies, corresponding to at which point in the claim adjudication process OPI intervenes to review and potentially adjust claims.

- "Pre Pay" Programs – programs which review claims during the adjudication process, prior to claim payment. Findings resulting from the review are reflected in initial claim payment. At a high level, Pre Pay program savings are reported by calculating the difference in billed dollars before and after the adjustments made to a claim, and reporting paid dollars savings by multiplying the difference in billed dollars by a historical paid/billed ratio. Note that these claims are extracted from the adjudication process before provider contract and member benefit information is attached to the claim, making more direct calculation of allowed or paid savings impossible.
- "Post Pay" Programs – programs which review claims after claim payment. Findings resulting from the review are recovered from providers through direct payment or through offsetting future claim payments by the recovery amount. At a high level, the amount of money recovered from the provider is the savings amount reported by Post Pay programs.

OPI Program Name		Methodology
CCM (Claim Cost Management)	**Contract Audit**	Pre Pay
	PIPPA (Payment Integrity Pre-Payment Audit) **/ High Dollar**	Pre Pay
	SAM (Smart Audit Master) **Edit / Automation**	Pre Pay
	PI Automation	Pre Pay
ARO (Accounts Receivable Operations)	**AARS** (Advanced analytic Recovery Services) - Chargeable	Post Pay
	AARS (Advanced analytic Recovery Services) - Non Chargeable	Post Pay
	Recovery Services	Post Pay
	Credit Balance Recovery Services	Post Pay
	Direct Connect	Post Pay
Subrogation	**ICC** (Injury Coverage Coordination)	Pre Pay
	Subrogation	Post Pay
FWAE (Fraud, Waste, Abuse, and Error)	**Pre-Pay Professional**	Pre Pay
	Pre-Pay Facility	Pre Pay
	FCR (Focused Claim Review)	Pre Pay
	Post Pay	Post Pay
	Facility Audit	Post Pay
	IBR (Itemized Bill Review)	Pre Pay
	ACE (Advanced Communication Engine) **Reject Edits**	Pre Pay

Participant vs. Non-participant Method

This methodology applies the non-participant group as a control, measuring savings as the difference between participant and control costs PPPM. Participants for evaluation are selected based on referral, which introduces the potential for subjective bias into the evaluation. "Participant" is defined as a patient who completes one or more calls with a nurse. Because of selection bias participant and control populations are typically matched using some form of statistical matching: either directly or via propensity matching. UHG's propensity matching methodology incorporates patient conditions (via Episode Treatment Groups) and patient risk (via Episode Risk Group Score). A third matching variable is the presence of a prior admission during the previous 3 months. Outcomes are measured as the difference between median costs (truncated) post-referral for each group. We consider these methods to be robust and accepted in the industry. Two programs apply this methodology: Case Management and Renal Management.

1. Case Management
   Participants are identified with at least 1 month of pre-identification and 1 month of post-identification eligibility. The intervention population is matched to non-participants. Participants are matched on a number of variables including ETGs (Episode Treatment Groups, which assign an episode of care to the patient's treatment) and ERGs (Episode Risk Groups that assign a risk score to the patient based on condition and comorbidity). We consider the matching process and use of ETGs and ERGs to be robust.

2. Renal program
   The renal program methodology is a variant of the matched participant vs. non-participant method. Expected costs of patients with certain characteristics are calculated from the UHG book of business using a regression model. Managed patients are then compared with their expected costs from the regression model.

   An issue with clinical programs that manage progressive diseases where patients transition over time to more severe states (such as Kidney disease) is that it is difficult to measure savings from a lack of transition to a more severe state. UHG measures savings as the difference in cost between stages (for example adding $C_j - C_i,\ j > i$ where C_i, C_j are average costs of patients in different stages, i, j are CKD stages and j is more severe than i). While UHG should provide more detail around the identification of those members that do not transition, we consider that the addition of the transition prevention estimate to the underlying matching process adds an important element of sophistication to the underlying calculation.

Trend-adjusted pre- post Method

3. Condition Management:

The UHG documentation makes reference to the CCA/DMAA methodology, which is a trend-adjusted pre-post population method. The method compares the cost of a population pre-program with population cost post-program implementation. For a chronic population, pre-program costs are projected to the program period using a trend adjuster derived from the non-chronic population. This is a very well-known method which has been extensively studied and published, including in [1]. We consider this to be a robust method provided it is appropriately applied.

There are several challenges with this method:

- Inclusions and exclusions; we have reviewed the exclusions which are reasonable and consistent with industry practice.
- Because the intervention and control populations are measured over two time-periods, caution should be applied to ensure comparability between the populations over time, in turn ensuring that changes in the population risk levels over time do not affect trend.
- Re-qualification: the accumulation of identified chronic patients over time because of a feature called "once chronic, always chronic" has the possibility to distort the comparison between the populations over time. (See [2].) Because patients have chronic conditions it is assumed that once identified a patient who is still eligible for insurance will continue to be in the population. This assumption can distort the savings calculation.

Overall, however, the application of this method by UHG is reasonable.

Utilization-based Method

1. Treatment Decision Support

UHG's utilization method is a version of a method called elsewhere "pre-intent vs. post-intent" which is a well-documented and accepted methodology, particularly for surgical procedures. The principle of these methods is that when patients present for acute procedures (e.g. musculoskeletal) sufficient information is gathered about the intended procedure that this may be compared to patient behavior post-presentation; with sufficient post-presentation observation it is possible through claims data to observe whether the patient had the treatment or not. Depending on the accuracy of the information gathered this is a valid method for reporting on savings from alternative treatments. A key element of the method is the assignment of costs to the pre- and post-treatments. UHG (appropriately) takes into account both "higher shifts" (those patients that elect a more expensive procedure, post-

intervention) and "lower shifts" (those patients that elect a less expensive procedure). The calculation of savings is described as: "Savings for Lower Shifts and costs for Higher Shifts are determined by calculating the difference in annual episode treatment costs of the pre-intent and the claims result treatments (based on UnitedHealthcare claims BOB)," consistent with the application of the method by other vendors of Decision Support programs. Claims are evaluated for 6 months post decision to eliminate recidivism, an important feature of these methods to ensure that savings are not over-estimated.

A feature of the UHG methodology is measurement of the effect of "provider shifts." This measurement, while it is discussed by vendors of these programs, is difficult to implement and requires (among other things) extensive data and a well-defined and valid methodology for assessing provider quality. UHG's method begins by confirming the provider referral through examining the individual's claims; savings are calculated by assigning the difference in treatment-specific average costs between a Premium and Non-Premium Provider (based on UnitedHealthcare claims BOB). The issue here is the definition of "premium" and "non-premium." There is also a subsidiary issue of the consistency between "premium" and lower-cost, something that the method implies. UHG does not provide more detail regarding the stratification of providers; nor does the example given differentiate between savings from pre- post-intent and provider shifts. We note that savings calculated for the provider shift amount to less than 1% of savings due to decision support (which itself amounts to only 4% of total reported savings). Therefore, although the provider shift methodology raises issues it may be ignored based on materiality.

Gap Closure Methods

1. Oncology management program

   Oncology management has become a major initiative of many payers in recent years because of the incidence of cancers and the growth in new (mostly infusion) drugs. These programs are notoriously difficult to evaluate; one methodology that has gained prominence is CMS's Oncology Management Program methodology, although this is limited to 6-month cohorts from inception of chemotherapy only. As a result, unlike population based and participant-non-participant methods, there is no standard industry method for evaluating oncology programs.

   In a more general cancer population the sources of savings (other than chemotherapy) are limited to preventing side-effects, limiting wastage of expensive drugs through the (sometimes) trial-and-error nature of chemotherapy and earlier referral to hospice for patients at end-of-life. UHG lists Oncology as a Utilization methodology, although the methodology is better described as a Gap Closure method. UHG lists 10 "value-drivers" or "gaps in care" and associates savings with each of them. "Value drivers" are recorded by program nurses when the nurse identifies that a gap has been closed or patient is compliant with appropriate treatment. The source of associated savings is largely literature and expert

opinion by UHG clinicians. UHG has also obtained a review by Emory University although the latter appears to be of different conditions and not Oncology.

The savings attributed to the Oncology program amount to 7% of total reported savings. As a source of savings this makes Oncology a moderate impact program. UHG's savings methodology appropriately uses available clinical information; we recommend that the method be subjected to further testing and publication in order to gain wider industry acceptance.

2. Other utilization-based programs (Maternity; Wellness and 24-7 Access)

Maternity: the UHG savings methodology for Maternity is the same as the Oncology methodology, although the "Value Drivers" for Maternity are different. The reported savings for Maternity are similar to those for Oncology. Our assessment of this methodology is similar to that for Oncology.

Wellness: UHG assigns savings to certain behaviors based on literature, which is a reasonable and standard approach for this type of program. This approach is reasonable and standard, there being considerable literature in this area, although there are relatively few studies that link improvement in clinical outcomes to financial savings, especially in a non-controlled environment such as an employer population. UHG claims a small savings for reduction in disability and absence costs, which is also standard for many of these programs (the methodology is not reported). The difficulty with estimating disability and absence savings for any program is the lack of data integrating these costs with medical data; given its extensive data resources UHG would do the industry a major service if by publishing a rigorous study of the impact on absence and disability of some of its programs. Total savings for which UHG claims credit are about 2% of total savings so the specific methodology does not have a significant impact.

24-7 Access: UHG's methodology for this program is effectively a version of the Treatment Decision Support (pre-intent post-intent) methodology, with variations: a comparison of either actual or proposed costs prior to engagement with the nurse-line to costs after engagement. Unlike the Treatment Decision Support patients (whose conditions are likely to be well-defined, such as recommended surgery for musculoskeletal disorders) users of nurse-line services are both more variable and less well-defined. Total savings for this program are minor, so any lack of accuracy in the identification of members and conditions does not have a significant impact.

UHG increases the calculated savings with an estimate of productivity and absence effects. This is a standard procedure for employer programs; employers tend to be skeptical of the numbers because they are difficult to substantiate with hard data but understand that some estimate of savings is reasonable. We note that the estimated savings due to productivity

gains are less than 25% of the total savings for 24-7 access, which in turn represents a minor portion of total reported savings.

The largest source of savings reported for 24-7 access is telehealth visits. This is a developing area and we are not aware of a standard measurement methodology. Providers usually estimate the savings as the difference between the unit cost of telehealth visits vs. office visits. This method ignores the problem that a telehealth visit may be *duplicative* rather than *substitutive.* Studies of in-store pharmacy clinics, for example, find that some portion of patients have follow-up visits to their PCPs, presumably to confirm a diagnosis or treatment. The UHG methodology does not appear to be based on unit cost differential but instead compares pre- and post- visit cost. This does not seem correct to us, and may be related to the duplicative/ substitutive issue. A patient who has a telehealth visit does so either because of convenience or because (in the case of Covid) this is necessitated by lock-down. In these cases the visit substitutes for a PCP visit, so (presumably) the same outcome would have resulted in the case of the PCP. For 24-7 access we agree in part with counting savings due to triage and reduced time away from work. Triage, however, also needs to be considered carefully, because a high percentage of triage results in referral to an Emergency Department (to which the patient would have gone absent the 24-7 access). Counting "savings" in this case could be exaggerated; genuine savings would result from a 24-7 nurse reassuring the patient that an issue is minor and can be addressed by a PCP rather than an ER. Such an approach would require more detailed data analysis, however. This emerging area of 24-7 access currently contributes only a minor portion of overall 24-7 program savings, so concerns with the method are not material. The rapid growth in telehealth and other remote services, however, makes this an area where more study would be appropriate.

Network-based Methodologies

1. Neonatal resource: no methodology detail is provided for this program so it is difficult to comment. Savings are generated by reduction in length of stay, which may be appropriate since these stays are likely outliers and therefore subject to per diem payments. Determining an "expected" length-of-stay for a neonatal infant with which to compare the actual length of stay is a challenging problem and we would like to see more detail explaining how this is done. The high variability of neo-natal claims, frequency of NICU infants, and the lack of data to classify the infant in terms of condition and co-morbidity makes the estimation of savings an area that could benefit from a test of the savings model accuracy. The Neo-natal program accounts for 8% of all savings, making this a moderate impact program, so some test of model accuracy is probably warranted.

2. Centers of Excellence (transplant): Centers of Excellence (COEs) have expanded rapidly in recent years and incorporate several features that attract employers:

a. Referral is to a "brand-name" facility (such as Mayo or Cleveland Clinic).
b. Bundled pricing, which guarantees a flat price known in advance.
c. The benefit of "excellence" which reduces complications and re-admissions.

This methodology is a variation of the "provider shift" methodology discussed previously and is subject to the same concerns. The transplant program is, however, excluded from this evaluation because the methodology is undergoing further evaluation internally.

3. Other COE programs (Congenital Heart Disease and Bariatric Resources). The savings from these programs are small and we do not express an opinion on their methodologies.

Pre-pay Payment Integrity Methodologies

The ten OPI Pre Pay programs SBA reviewed extract claims for review during the adjudication process, before provider contract and member benefit information is attached to the claim, and prior to claim payment. At the point OPI extracts claims for review, the submitted/billed charges on the claim are the only payment information available from which to estimate savings.

Claim reviews may result in either denying an entire claim or one or more individual claim lines. In either situation, the billed dollars for the claim lines that are denied are summed. OPI reports savings on a paid basis (the costs paid by the health plan/ASO employer). Billed dollars are converted to paid dollars based on historical paid/billed ratios. The paid/billed ratios are calculated based on historical paid/billed ratios for claims that were adjusted by OPI Pre Pay programs.

Each individual product maintains slightly different methodologies to calculate the paid/billed ratios used. Some vary the calculation by provider specialty, while others use a single ratio for all specialties. It is SBA's opinion, based on industry knowledge and not a comprehensive review of UHG's data, that in addition to provider specialty, region/state would be an important variable to control for in the application of paid/billed ratios. However we believe that when this methodology is used to estimate savings for a large population on an annual basis it will produce consistent and accurate estimates.

There is one product which calculates savings based on pre and post-review paid dollars. That product is reviewing a relatively small volume of claims compared to other products, and so is able to run both the pre and post adjustment versions of the claim through an adjudication process (only the post-adjustment claim is fully adjudicated and paid). This methodology is far preferable to the billed dollar-based methodology used by most Pre Pay products, however SBA acknowledges that the billed dollar-based methodology is the only available reasonable methodology given the data limitations presented by when these claims are selected, and certifies to it's accuracy when aggregated across many claims.

There are occasions when a provider may appeal the findings of an OPI review. When these appeals are granted and OPI findings are overturned, the savings that OPI initially reported are adjusted to remove the savings from the overturned findings. This same scenario may happen in a less formal way if a provider resubmits a claim which is then paid without OPI adjustment; savings would be adjusted downward in the same way for this less formal scenario. Most products are able to track appeals and resubmissions at a claim level. However, one product is not able to do so and applies a historical average appeals/resubmission rate to savings when initially reported.

Post-pay Payment Integrity Methodologies

The eight OPI Post Pay programs SBA reviewed select claims for review after the claim has been adjudicated and paid. Any findings resulting from review will be recovered from the provider either through direct payment or through offsetting future claim payments for other claims.

Reported savings for post pay programs are the dollars recovered from providers. There may be occasions when money that is recovered is subsequently returned to a provider (i.e. in the case of a successful provider appeal). In this circumstance, reported savings will be adjusted by the amount returned. The direct nature of post pay savings means no other calculations or adjustments are needed.

Advocacy Program

The Advocacy program encompasses many member services and clinical programs that intend to optimize care through guiding members to optimal care paths and settings, addressing social determinants of health, digital engagement, and more. UHC notes that the Advocacy program "encompasses a lot of things."

There are 3 levels of Advocacy support, which vary in the level of personalized support provided by the Advocacy team. Higher support tiers, in comparison to lower support tiers, feature advocacy teams that are co-located with the employer and dedicate support to a single employer and as a result have deeper knowledge of the employer's benefit plan.

UHC does not estimate an absolute level of savings from the advocacy program because every employer is subject to some level of advocacy. Instead, UHC estimates incremental savings by comparing the outcomes of a group of employers with higher levels of advocacy to a control group consisting of employers with the base level of advocacy, controlling for the types of Medical management, Medical necessity, Payment Integrity and other UHC programs that the employers use. The methodology also removes members with outlier risk scores, high cost claims, and allowed PMPMs.

UHC attributes the difference in risk and area-adjusted Allowed PMPMs between the two higher tier study groups and the base tier control group as savings due to the higher Advocacy tiers.

This methodology is sound and industry standard. Below we will recommend a methodology to understand measurement volatility, which we believe would be especially important for the Advocacy programs due to savings be attributed using total PMPM costs. The normalization measures taken are appropriate and on a large population will produce reasonable results.

Naviguard Out-Of-Network Management

UHC has partnered with Naviguard, a vendor of out-of-network management services. Specifically, Naviguard uses a proprietary reference-based pricing methodology to outpace savings achieved by other competitive programs (shared savings, OCM, and R&C). Naviguard enforces these savings through member advocacy, negotiation, and arbitration services.

UHC estimates that Naviguard has delivered a 70.8% reduction in out-of-network billed charges in 2021, compared to a 48.6% discount for non-Naviguard out-of-network programs (derived from ASO clients who had not yet adopted Naviguard). A similar differential was seen when comparing Naviguard and non-Naviguard program spend in 2020. These savings are consistent for claims that are much higher than a reference price, and for claims that are closer to a reference price.

Recommendation

One of the challenges with any care management methodology is addressing the high variability of utilization and cost of healthcare, particularly in smaller populations. We recommend to our clients that they test any methodology on sample claims, assuming no program impact. In the absence of a program effect the expected savings measured by any methodology should be zero, although variability will result in actual measured savings being distributed around a mean of zero. If the methodology is robust the confidence interval around the expected outcome of zero should be narrow. A larger confidence interval implies uncertainty in the accuracy of measured results when a population is subject to a program, which may be addressed by changes to the measurement methodology. For more information about this proposed model test see our recently published papers ([3, 4]).

REFERENCES

1. Ian Duncan, *Managing and Measuring Healthcare Intervention Programs (2nd Edition)*. 2014, Winsted CT,: Actex Publications. 446 pages.
2. Robert Bachler FSA FCAS MAAA Ian Duncan FSA FIA FCIA MAAA and Iver Juster MD, *A Comparative Analysis Of Chronic And Nonchronic Insured Commercial Member Cost Trends.* North American Actuarial Journal, 2006. **10**(4): p. 76-89.
3. Mackenzie AJ Teppema S Wang J and Duncan I, *Payment Accuracy in Value-Based Care Contracts.* Austin Medical Sciences, 2021. **6**(3): p. 1-5.
4. Duncan I Mackenzie AJ Bonfiglio E Wrigley T and Liao X, *Shared Savings Model Risk in the MSSP Program.* North American Actuarial Journal (forthcoming), 2022.