Background About Shared Savings Program Design Features:
Patient Attribution, Cost Target Calculation, and Payment Calculation and Distribution

Excerpted from Draft Narratives Developed in the CT SIM Equity and Access Council
March 2015

Patient Attribution

Background

Implicit in a shared savings program is that a group of providers manages the quality and cost of care for a defined population. The twin goals of such a program are to improve efficiency (typically through methods that improve utilization management) and to improve quality (typically through more effective, consistent clinical performance and through care management and care coordination). When providers achieve these goals they are eligible for incentive payments that supplement their fee-for-service revenue. Often a provider’s ability to actually share in any savings achieved is dependent on meeting the quality targets agreed to at the outset of the contract period. The process of defining the population that a given group of providers is responsible for managing under a shared savings contract is called patient attribution. The clinical participants in the shared savings contract, which can include providers, provider groups, hospitals, and other care supplier entities, collectively agree to be responsible for the cost and quality of the patients assigned to them under the contract. We refer here to the organizations or groups of organizations that enter into shared savings contracts as Accountable Care Organizations (ACOs).

Insurance plans have developed a range of methods for attributing patients to provider organizations. Every attribution methodology involves at least three main design decisions:

1) **How** the patient is assigned to a provider (i.e. the technique or “rule” used to assign a patient)
2) **To whom** the patient is assigned (i.e. the type of provider to whom a patient can be assigned)
3) **When** during the contract period the patient is assigned

There are several techniques used to assign a patient to a provider in a shared savings program. A plurality of visits technique is used by the Centers for Medicare and Medicaid Services (CMS) in the Medicare Shared Savings Program (MSSP) (CMS, CMS Medicare Shared Savings Program Final Rule, 2011), which makes up the majority of shared savings programs in the market today (CMS, Medicare Shared Savings Program ACO Fast Facts, 2014; Gordon, 2014). This technique assigns a patient to the provider that the patient saw most frequently within a defined period of time (i.e. the year prior to the performance year or during the performance year). In patient-selected attribution patients designate their primary care provider when they enroll in their insurance plan. This technique, known as “patient attestation” is used by Blue Cross Blue Shield of Massachusetts for their Alternative Quality Contracts (Chernew, Mechanic, Landon, & Safran, 2011), among others. Insurer-selected attribution relies on the insurer to designate the patient’s primary care provider when the patient selects the insurance plan (Cromwell, 2011). A geography-based (or “population-based”) technique assigns patients to a provider based on where the patients live. This technique was used for the Medicaid patients in New Jersey in combination with a plurality of visits technique (Houston & McGinnis, 2013). The technique was intended to attribute patients who did not regularly see a physician. Attribution techniques are not necessarily mutually exclusive; in some instances using more than one can be useful, as was the case in New Jersey.
The type of provider to whom a patient can be assigned is another aspect of patient attribution. The objective is to assign patients to the providers who are predominately responsible for managing their primary care needs (Cromwell, 2011). While a primary care provider (e.g. internist, family practitioner, general pediatrician) is generally the provider type that would be the most responsible for managing the primary care needs of a patient, in practice that is not always the case. For example, patients who have chronic conditions (e.g. heart disease or diabetes) that require intensive management from a specialist will often see the specialist provider as their primary care provider. For this reason CMS, in its most recent proposed rule for MSSP, proposes changes to the current patient attribution methodology to exclude specialists in the attribution process whose services are “not likely to be indicative of primary care services” (CMS, Fact Sheets: Proposed Changes to the Medicare Shared Savings Program Regulations, 2014). Many states have followed CMS’s lead in designing their shared savings programs for Medicaid and in some cases taken it a step further. In Minnesota attributing patients to an Emergency Department (ED) was considered if that was the location of the plurality of their visits (Houston & McGinnis, 2013).

A final design consideration concerns the timing of patient assignment to a shared savings program. A patient can be assigned to a shared savings program either retrospectively or prospectively. Retrospective assignment assigns a patient to a provider at the end of the first performance year of the shared savings contract. In a retrospective model, providers do not know which patients they will be responsible for at the beginning of the shared savings contract period. Conversely, prospective assignment assigns a patient to a provider at the outset of the shared savings contract period. Prospective assignment allows providers to enter into the contract period aware of the population for whom they are managing cost and quality (see figure below).

The MSSP program currently uses retrospective assignment, but is recommending prospective assignment for some of its participating ACOs1 (CMS, Fact Sheets: Proposed Changes to the Medicare

---

1 In the 2014 CMS proposed rule a third track is proposed that will use retrospective assignment and require that the ACO take on down-side risk.
Shared Savings Program Regulations, 2014). Prospective assignment allows providers to know in advance which patients they are managing, potentially improving their ability to proactively manage toward improved outcomes and lower costs in a manner that retrospective assignment does not allow. Many physicians prefer prospective assignment. However, CMS has been historically reticent to utilize prospective assignment because of its articulated concern about associated risks of under-service: “...we agree with the comment that while providing such information may be a benefit to both the beneficiary and the ACO, concerns remain that ACOs could use it to avoid at-risk beneficiaries or to stint on care.” (CMS, CMS Medicare Shared Savings Program Final Rule, 2011). Unlike CMS, commercial insurers more commonly use prospective assignment for a range of value-based contract types, including upside-only and two-sided shared savings programs (Bailit, Christine, & Burns, 2012).
Cost Target Calculation

Background
To determine if an ACO achieves savings during a shared savings program contract period, the expected (or targeted) cost of caring for the population attributed to the ACO first needs to be defined. This is known as the cost benchmark. The determinants of the cost of a population include many factors, some of which allow for a level of predictability and others that do not. Some of the more predictable factors include: current diagnoses, age, socioeconomic status, and other social determinants of health (e.g., housing, access to transportation, etc.). Less predictable factors include: new and unexpected diagnoses, catastrophic events, and unpredictable general health trends (e.g., a bad flu season). In combination, all of these factors influence how complex and potentially costly a patient is to care for and should be considered when determining a cost benchmark. The choice of population used to set a benchmark, and the risk adjustment methodology used to adjust those costs, relate to the more predictable factors associated with cost benchmarking. The risk adjustment methodology adjusts future cost projections to account for the variation in resources required to care for different populations. The risk adjustment takes into consideration demographics and the diagnoses of the population to allow for an “apples to apples” comparison in costs between populations with different risk profiles. Additional contract features relate to the less predictable factors associated with benchmarking.

Payers generally use one of two data sources to establish a cost benchmark for a given population: historical costs or control group costs. A historic benchmark sets the expected costs of a population based on the past experience of that population. A control group benchmark uses a comparator population (e.g. all enrollees in a health plan throughout a broad regional area) to determine expected costs. Importantly, the historic benchmark inherently accounts for the clinical and cost profile of a given ACO’s population, while the control group does not. For this reason, risk adjustment is an especially important dimension of a benchmarking method that relies upon a control group.

<table>
<thead>
<tr>
<th>Cost Benchmark Method</th>
<th>Role of Risk Adjustment</th>
</tr>
</thead>
</table>
| **Historical Costs**  | • A historical cost benchmark will inherently account for risk as it is based on the actual prior care experiences of the attributed population.  
• However, adjustment can be valuable as a way to more accurately predict how future costs are likely to vary from the historical snapshot. |
| **Control Group Costs** | • Unlike the historical cost benchmark, the control benchmark is based off of a population that is not part of the shared savings program and will not inherently account for the attributed population’s level of risk.  
• Risk adjustment provides an essential method to reflect the impact of risk on the cost benchmark, providing for an “apples to apples” comparison. |

Another difference between these two methods is how accurately the benchmark reflects the utilization of a population that is desirable (i.e. represents best clinical practice). A historic benchmark utilizes the historical experience of an ACO’s population, which may or may not represent best practice, whereas a control benchmark is based on performance against market-wide medians, targets, or trends. If a historic benchmark is used and historically the population has experienced unnecessary over-utilization, the benchmark will not account for excessive and unnecessary costs that a shared savings program attempts to minimize. Over time this will be addressed as the cost benchmark is adjusted over the
subsequent years, but getting to best practice may take longer than it would if the control group methodology were to be used. Regardless of which population is used to determine the cost, risk adjustment will also be necessary. Even when a historical benchmark is used, additional factors need to be considered, such as the increased age of the population or new diagnoses. CMS currently uses the historical cost methodology for MSSP and applies a risk adjustment factor (Bailit & Christine, 2011). The CMS risk adjustment takes into account acuity of diagnoses and basic demographics such as age, but does not account for any social determinants of health. In addition, as the CMS MSSPs function today, risk is adjusted annually for patient age and decreases in patient acuity are reflected to adjust cost benchmarks downward, but CMS does not adjust the benchmark upward if there is an increase in acuity (Gaus, 2015). In the healthcare market there are additional proprietary risk adjustment methodologies used by various commercial payers (Bailit, Huges, & Burns, Shared-Savings Payment Arrangements In Health Care: Six Case Studies, 2012). However, given their proprietary nature there is not an abundance of publicly available information and it is unclear which factors are adjusted for in their methodology.

To account for the less predictable factors that affect a population’s cost of care, shared savings programs often include additional contract features to help minimize ACOs’ financial risk. Common examples of these additional contract features and examples of payers that use them are outlined in the table below:

<table>
<thead>
<tr>
<th>Additional Contract Feature</th>
<th>Payers That Utilize</th>
</tr>
</thead>
</table>
| **Truncation of High Cost Claimants** – exclusion of patients with costs above the a certain percentile (commonly 99th percentile)** | • CMS  
• Vermont Medicaid Shared Savings Program |
| **Exclusion of high cost procedures or services** – plans will exclude high cost procedures such as transplant and/or less predictable services to such as behavioral health** | • Commercial payers and CMS (Bailit, Huges, & Burns, Shared-Savings Payment Arrangements In Health Care: Six Case Studies, 2012) |
| **Enhanced Per Member Per Month (PMPM) Payment for patients with chronic conditions** – payment intended to support enhanced care management needs** | • BCBS of Michigan (Share & Mason, 2012) |
| **Socioeconomic payment adjustment factor** – enhanced payment to account for non-health factors that impact the complexity of caring for an individual** | • Providers are working with payers to develop in Oregon |
Payment Calculation and Distribution

Background
When an ACO signs a shared savings contract with an insurer, the ACO becomes eligible to earn payments that represent a share of savings it achieves on medical spending for a defined population. The ACO, in turn, may distribute the savings to the ACO’s member organizations such as provider groups or hospitals, which provide services for the ACO’s attributed population. The ACO may also retain a portion of shared savings to invest in shared infrastructure used to care for the attributed population. Provider organizations that receive shared savings payments from an ACO may in turn pass these along to providers, either in whole or in part, according to formulas they devise to reward provider performance. The shared savings paid to an ACO by a payer and in turn to the ACO’s providers is known as an incentive payment, because it is incremental to what a provider will receive on a fee for service basis. In most shared savings payment arrangements, the savings generated are split between the provider and the payer.

The process used to determine the portion of cost savings that an ACO receives is referred to as payment calculation. How the ACO distributes those savings among providers and/or provider groups, or otherwise compensates service providers within the ACO, is referred to as payment distribution.

The diagram below depicts common ways in which funds flow from a payer to providers participating in a shared savings contract.

Shared Savings Payment Terminology and Structure

[Diagram showing the flow of funds from the payer to the ACO, then to the provider groups, and finally to the providers, with various components such as portion of shared savings, quality bonus, care coordination fee (PMPM), and fee-for-service.]
Note a few key features associated with the flow of funds:

- Provider groups still earn a fee-for-service payment directly from the payer
- Providers’ compensation is generally composed of a combination of guaranteed salary, productivity payments, and a portion of shared savings and/or separate quality-based bonuses
- Quality metrics are often used to establish eligibility for some or all of an ACO’s potential shared savings, but are not generally used to pay separate quality bonuses if no savings are achieved
- Note that “quality” in this construct typically includes clinical processes and outcomes as well as other measures of provider performance such as patient satisfaction surveys

In most shared savings arrangements a Minimum Savings Rate (MSR) establishes the degree of savings an ACO must achieve in order to be eligible to earn any amount of savings. An MSR is used to ensure that ACOs only share in savings that are statistically significant and don’t result from random variation in expenditures. For example, an MSR of 1% would require that the ACO’s actual costs at the end of the performance year are at least 1% lower than the expected cost benchmark in order for the ACO to share in the savings.

The MSR set in a contract often depends on the size of the ACO (i.e. the number of lives the ACO manages) and the contract type (i.e. upside vs. two-sided risk). Random variation is less likely in a larger population and thus will usually be accompanied by a lower MSR (Bailit & Hughes, 2011). The MSR used by CMS ranges from 2.0% to 3.9% depending on the size of the beneficiary population. A 2.0% MSR is used for ACOs with greater than 60,000 patients. 3.9% is the highest MSR and is applied to ACOs with 5,000 patients, the minimum population required for participation in a CMS MSSP (CMS, 2014). Some payers believe that random variation will occur in both directions (i.e. result in savings and losses) and even out over the contract period. For this reason some payers do not utilize an MSR and others utilize a very low MSR regardless of population size. A lower MSR generally makes a shared savings contract more appealing to an ACO.

**Upside vs Two-Sided Risk:**
A shared savings contract can have only an upside or can have an upside and a downside. In an upside-only contract the ACO will have the opportunity to share in savings if actual costs are below the expected cost benchmark, but will not be at financial risk if costs are in excess of the cost benchmark. In a contract that has upside and downside risk (also known as two-sided risk) the ACO will continue to have an opportunity for savings, but will also incur a loss if spending is higher than the expected cost benchmark. The loss will occur in the form of a payment back to the payer for costs that exceed what was expected. Similar to an MSR, in a downside arrangement there is a threshold of excessive expenditures that has to be met before the ACO incurs a loss. This is known as a Minimum Loss Rate (MLR). ACO expenditures must be in excess of the MLR for the ACO to be required to owe the payer for the costs beyond what was expected. In both the MSR and the MLR the amount of savings and/or losses are capped at a maximum amount.
CMS is considering the use of a deferred reconciliation for MSRs. In this scenario the MSR would be applied over an entire contract period, so if the ACO achieves consistently small savings each year that cumulatively reach the MSR, those savings will be shared at the end of the contract period (Gaus, 2015).

Three contract design features figured in the EAC’s discussions about the potential for payment calculation to impact under-service or patient selection:

<table>
<thead>
<tr>
<th>Contract Design Feature</th>
<th>Common Design Options</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Does the contract type use downside risk, or upside risk only?</strong></td>
<td>• In a downside risk arrangement the maximum share of savings an ACO can earn is usually set at a higher percentage to make participation in a downside risk arrangement more appealing; greater risk, greater reward.</td>
</tr>
</tbody>
</table>
| **Does quality performance affect the ACO’s opportunity to earn savings?** | • The majority of programs define a quality threshold that must be met to receive any savings.  
• In a varied arrangement the percentage of savings given to the ACO correlates with quality performance. Better quality performance relative to peers or over the prior year will result in a higher percentage of savings earned by the ACO.  
• In a fixed arrangement the amount of savings shared with an ACO remains the same as long as the minimum threshold for quality performance is met. |
| **How is quality performance measured?** | • Quality performance of an ACO can be evaluated as it compares to the performance of other ACOs. This is commonly called a benchmark method.  
• Quality performance can also be measured based on an ACO’s improvement over the prior year.  
• Some arrangements use a combination of the benchmark and the improvement methods  
• The improvement method helps to engage lower performers and likely will account for any risks inherent in the ACO’s population that might make achievement of a benchmark difficult  
• A benchmark method can be useful for high performers for which demonstrating further improvement may be difficult. |

References for table: (Bailit & Hughes, 2011; CMS, 2014)
Most shared savings contracts require that an ACO meet minimum thresholds on a set of quality measures in order to be eligible to receive a portion of the savings achieved. However, other payment arrangements also exist that will provide incentive payments independent of one another; one incentive payment for the achievement of quality targets and another incentive payment for the achievement of savings (Bailit & Hughes, 2011; McGinnis, Riley, Zimmerman, & Sahni, 2013).

Once the savings are received by the ACO, the ACO is responsible for distributing the savings to the provider participants in the ACO. Most ACOs consider at least three questions when deciding how to distribute savings amongst participants:

1. Should the ACO retain a portion of the savings?
2. How should money be distributed among ACO participating organizations?
3. What factors should play a role in how savings are distributed to individual providers?

Often an ACO retains a portion of the savings to support the infrastructure needed to run the ACO. ACOs that are in a downside risk arrangement or are considering moving to a downside arrangement from an upside only arrangement will take a portion of the savings to build a reserve fund. The composition of ACOs varies greatly, from having primary care providers only to having primary care providers, specialists, hospitals, and providers of other services along the care continuum (e.g. skilled nursing facilities). How the savings are distributed between these groups usually takes into consideration the centrality of the role each provider type plays in managing the quality and cost of care for attributed patients.

Lastly, individual providers may receive a portion of the savings – but the savings may be pooled by the provider group and distributed in any number of ways. Individual providers’ shares can be based on how many attributed patients they cared for, their individual quality performance, or their individual contribution to achieving the savings. A common method for determining an individual provider’s share of the savings involves using quality performance and weighting by the number of patients attributed to that provider (The Chartis Group, 2014).
References


McGinnis, T., Riley, P., Zimmerman, M., & Sahni, N. (2013, October 24). Planning for Accountability: Emerging Medicaid ACO Models and Key Issues for States. n/a, n/a, n/a, n/a, n/a: Center for Health Care Strategies, Inc.

The Chartis Group. (2014, September n/a). Chartis Shared Savings Consideration for Fundsl Flow and Incentive Distribution. n/a. n/a, n/a, n/a: The Chartis Group.