Virtual Meeting Track 2: Setting Episode Budgets

Maternity Multi-Stakeholder Action Collaborative

June 15, 2017

1:00-2:00pm ET

Highlights and Key Takeaways

MAC members participated in the virtual meeting for Track 2: Setting Episode Budgets. Below are highlights and key takeaways from the interactive meeting, which included live interviews with Andy Allison, Senior Expert in the Health Care Systems and Services with McKinsey, and François de Brantes, Vice President and Director of the Center for Payment Innovation with Altarum Institute. Both discussants shared their experiences working with different payers to set episode budgets.

Methodologies for Setting an Episode Budget

Episode-based payment models set budgets to manage costs. When determining how to set an episode budget, you are really determining how much an episode should cost with the expectation that an accountable provider's performance will be assessed relative to that set budget. Under a retrospectively reconciled financial arrangement, the budget is used to determine whether the accountable provider has been successful or not.

There are various approaches for setting an episode budget. Three options discussed during the session today include:

- ***Historical Average of the Provider Group***: Administrative, this approach is relatively easy to do and it provides an incentive for the provider to be more efficient than in the past. This approach does not address variation in price across providers, though, so high price providers may continue to stay high price providers. It may also reward high-cost, less efficient providers who would otherwise have the opportunity to become more efficient, but under this approach they can reap a financial benefit without much change. Alternatively, an already efficient provider may have less opportunity to improve upon their efficiencies and therefore may not receive the financial incentives they deserve.
- ***Historical Average of a Geographic Area or Marketplace***: Administratively, this approach is relatively easy as well. It may help support reduction in price variation either immediately or in a phased in approach over time. You can phase in this model by starting the budget with a certain percentage based on provider experience and a certain percentage based on the market, and increase the market percentage over time. Under this model, high-cost, low-efficiency providers might find themselves coming out on the shorter end relative to the historic average of their own experience
- ***Building from the "Ground Up"***: Instead of using historical averages, this option predicts how much would an episode should cost and builds an ideal budget. This removes the incorporation of any strange pricing patterns or historical variations in the budget. This approach would be the most time-consuming, and would likely create financial loss among the high-cost, low efficiency providers.

This is a brief overview of some of the options and limitations. There may be ways for your organization to modify these three options and address the limitations outlined above. It's always important to analyze the options against your organization to understand who gains, who does not gain, or who gains less under any model.

Examples from the Field

Historical Average of a Provider Group
Community Health Choice, Texas

*This Medicaid health plan set budgets for two large provider groups in Houston, TX, focusing on three distinct components of a maternity episode: prenatal, labor and delivery, and newborn care. The **prenatal care** budget can be prorated based on the number of months the mother has been cared for by the health plan to account for women who may start prenatal care after the first trimester. The **delivery** budget is a blended rate looking at the historical vaginal and cesarean rates. It is important to note that the way that the blend is set may create a stronger or weaker incentive for providers to reduce the cesarean rates. Finally, for the **newborn** budget, this health plan looked at nursery levels during their pilot period to determine who should be included or excluded and found that nursery level designations were somewhat arbitrary. Community Health Choice decided to move away from nursey level designations and is instead applying a stop loss cap for their high cost outliers, regardless of nursery assignments.*

Threshold Budgets
Arkansas, Ohio, and Tennessee State Medicaid

Instead of setting one budget target, these three states used a threshold approach and established multiple targets with different financial consequences based on where the provider falls relative to the different targets or thresholds. The models used in these three states vary slightly, but are based on one common model.

To set the thresholds, the average historical cost for the entirety of the state, not for individual providers, was used. Depending on where each provider falls within the set thresholds, the provider may share in savings, receive a financial penalty, or see no change in payment from the negotiated reimbursement rate. This approach can have either the state or the health plan setting the thresholds. In Tennessee, the state determined the acceptable threshold and gain sharing limits and the contracted managed care organizations determined the commendable threshold. In Arkansas and Ohio, the state determined all thresholds.

Determining the Risk Model

There are various options for determining the risk model and sharing in savings and/or risk. A shared savings model allows the provider to share in savings if performance is within a certain threshold/budget. If the provider is outside of that threshold, there is no financial consequence. A shared risk model allows a provider to share in savings if performance is within the threshold/budget, however, there are financial consequences for being outside of that range.

There are also two full risk options. A discount arrangement has an established budget and the payer removes a certain percentage off that budget and pays the remainder to providers per episode. In this method, the payer is guaranteed some savings, and if the provider can lower the cost below that discounted rate, the provider keeps that savings. The full risk approach has a negotiated rate, presumably with some savings that the payer has built in. The provider assumes full risk if they come under budget or go over budget.

Integrating Quality into Payment

There are two general ways quality may be incorporated into maternity episodes. They are not mutually exclusive but they are often treated exclusively.

- ***Quality Performance Affects the Risk Arrangement:*** This is the most common method in current models. Minimum quality threshold must be met before the provider can share in savings, almost like a gate the provider must get through before sharing in the savings. There are also integrated models where quality is integrated into the service payment. After the provider meets the minimal acceptable quality scores, the performance within those quality scores can impact the financial gains or losses. For example, the higher the quality scores, the greater the financial benefit to the provider, or if there is a loss, the loss may be reduced.
- ***Quality Performance Financial Implications are Independent of the Risk Arrangement*:** In this model, there may be a predefined incentive amount that a provider can receive for performing well on quality measures, an open incentive opportunity. Another option is a fixed pool divided among providers. The amount a provider receives

is determined by how other providers perform. Although there are not specific examples of this method used in episode models, it has been used within other value-based payment models.

Interview: Experience with Setting Episode Budgets (Altarum Institute and McKinsey & Company)

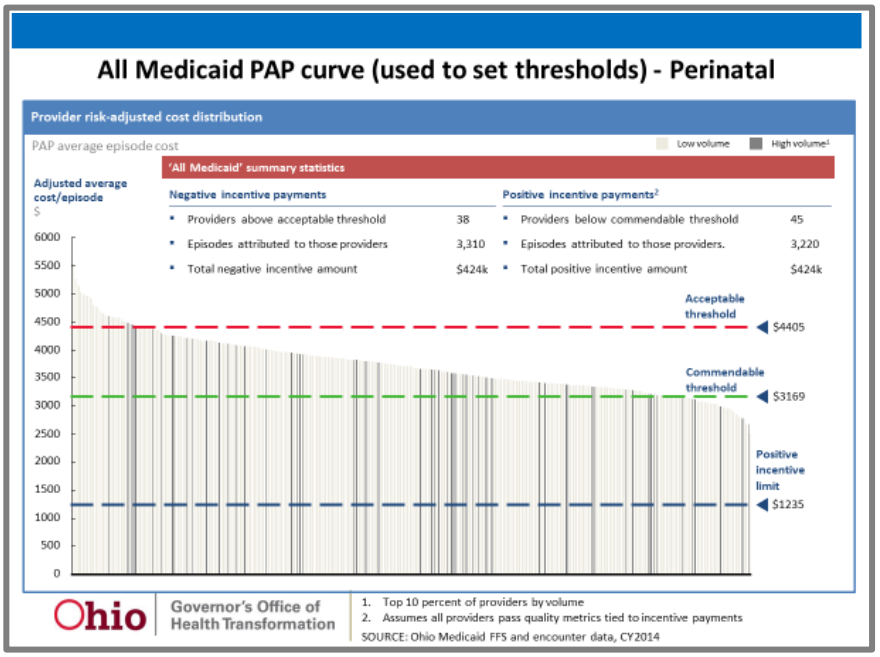
François de Brantes has vast experience creating maternity episode payment models, including the model being used by Community Health Choice in Texas today. He is also helping to design the model for New York's Delivery System Reform Incentive Payment (DSRIP) Program, among others. François was invited today explain episode budget approaches, including the one undertaken by Community Health Choice. Andy Allison helped create the episode payment models being used statewide in Arkansas, Ohio, and Tennessee. He also serves as the Medicaid Director for the state of Arkansas during the implementation of the episode model. Andy will discuss the contrasting maternity episode models used in those three states.

- **Question 1:** Community Health Choice and the New York DSRIP Program created three separate budgets for each maternity episode which were then added together. This is a different approach from having just one maternity episode budget. What is the rationale for apportioning the maternity episode into three separate budgets and what is the desired end for using this type of model?
    - [François]: Community Health Choice focused on creating a common interest of a healthy delivery and newborn for the physicians, obstetricians, neonatologists, and everyone else providing care to the mother and newborn. Having all three components together in one budget provides an incentive around the global budget as opposed to the individual pieces. By keeping each component separate, it allows you to adjust for the risk severity of the mother without the newborn being mixed in as well. There is not necessarily correlation between the newborn's underlying risk factors and potential costs of care, and the mother's underlying risk factors and potential cost of care. By keeping the budgets separate, you can create separate risk adjustment models for each component, but then sum up the components into one global budget.
- **Question 2:** Payers are often concerned with higher costs resulting from including the newborn in the bundle. Was this one of the reasons for dividing the budget into three components?
    - [François]: No, it is more about how complicated it is to appropriately adjust for the severity of the mother and the severity of the newborn when using one budget for the entirety of care. There are also difficulties with linking the mother with her newborn in claims data. In New York state, hospitals report the link between the mother's ID and the newborn's ID to the state. Although there is this separate reporting process, it is not 100% accurate. The question then becomes how to include those cases that are not appropriately linked. If using one global budget, those cases may end up being excluded from the entire episode. If the budget has a specific newborn component, a portion of the prenatal care is still relevant within the budget. An incentive may still be created around the delivery and pregnancy even if the mother/newborn link doesn't exist. Having the mother/newborn link, though, would bring all three pieces together.
- **Question 3:** How does prorating of the prenatal phase work? For example, is there a limit beyond which the prenatal period was too short, or could there be a budget for a woman who came in for one prenatal visit?
    - [François]: This varies tremendously by state because it may be based on rules around Medicaid eligibility and/or other social or demographic factors. The challenge for Community Health Choice and others was that many mothers enrolled in Medicaid on their date of delivery or had their first visit a few days prior to their delivery. In those instances, the prenatal episode doesn't count but the delivery episode does. There is no right or wrong answer to this question of proration. Each organization should analyze their own data and determine if there are certain patterns. In New York and Texas, proration works well. The gains and losses on the maternity budget come mainly from the delivery care. For example, in a blended rate between cesarean and vaginal deliveries, the bulk of the cost is within the cesarean deliveries. So, when looking at the impact on providers of these global budgets, there are two

sources of potential savings or losses: cesarean rates and excessive use of high level nursery designations.

- Karen Love from Community Health Choice noted that only about 4% of the deliveries in her health plan were "drop in deliveries" with no prior prenatal care.

- **Question 4:** The Arkansas, Ohio, and Tennessee models are different as they don't set one budget, but have multiple thresholds. Why did these states decide to use a multiple threshold approach?
  - [Andy]: I was hired in Arkansas as the Medicaid Director at the end of 2011 to help design and implement episodes, and achieve the adoption of the episodes and patient centered medical home model in the state. In that design process, we decided to use a retrospective episode approach where all incentives are calculated looking back at a provider's experience at the end of the year. One of the core objectives was to develop a relatively parsimonious number of variables in the model that payers could manipulate. It also had to be understandable to providers so they could conceptually have those in mind while guiding their practice.
  - Here is an illustration of using the multiple threshold approach. Each vertical line is the average cost over the course of the year. We could have said we will incentive any deviation from that median or average provider range, but we didn't do that because having an acceptable and commendable threshold is far more flexible. Another critical concern for us was convincing providers and ourselves that the vertical lines were truly representative of an average cost that could be compared from one provider to the next and that the variation depicted is truly reflective of provider performance and not differences in the women being served. There were discussions around risk adjustment and exclusions, but one advantage of the acceptable and commendable thresholds is you can add distance between ranges to accommodate for unmeasured but justified variation in performance across providers. Risk adjustment and exclusions alone are not pinpoint perfect, but having distance between those the ranges creates an allowance for statistical variation. That is how we ended up with an acceptable and commendable threshold, and the gain sharing limit protects against incentives to skimp.



  - There are two extremes to setting thresholds. Pick a percentile you can defend against the criticism of clinical variation and performance, or vary the thresholds episode by episode based on clinical data. In the slide above, the percentage of providers that receive an incentive or have a negative incentive is quite small. To set the thresholds and keep those percentages smaller, Ohio chose a percentile that they applied across all episodes. Ohio decided to define excessive use across all episodes in a common way using the top 10 percent of providers. From my experience as the Medicaid Director of Arkansas, I wanted to be able to convince providers that in the case of every single episode, the acceptable threshold can be defended on clinical terms. I was less interested in defending a particular percentile.

Arkansas' clinical advisors looked at examples of what someone at the 95th or 85th percentiles cost, what their practice looked like, and the variations across it.

- **Question 5:** How do you adjust thresholds or actual performance targets against patient demographics and against risk? To what extent did Arkansas and other states risk adjust the performance of the providers when assessing them against the threshold?
    - [Andy]: The goal is for providers to respond to the incentives, and for providers have the picture in mind of where they are relative to the commendable threshold. It is also important to convince providers, ourselves, and payers that the vertical lines are truly representative of performance, which means adjusting those raw average costs into an apples-to-apples comparison. Ohio went to great lengths to make those vertical lines comparable across providers by risk adjusting and finding significant measurable factors that had clinical justification for variation that could be taken account of in the data. Ohio uses 77 total risk factors, including obesity, previous cesarean deliveries, STIs, and anemia. In most cases, you may end up with 20-30 factors that are statistically significant which can be used in the risk adjustments to move the vertical lines. There are exclusions built in as well, such as mothers under 12 years of age or selected clinical factors. You can also decide to exclude outliers such as the highest cost or most intense cases. All of this can be balanced against how many cases it takes to qualify and receive the incentive, and the greater that number, the less risk adjustment you might have to do.
- **Question 6:** What data is required to implement maternity episodes and set a budget (e.g. only historical claims or ongoing real time data analysis)? Does this require access to enhanced analytic or software functionality?
    - [François]: There are multiple ways for health plans, state agencies, and employers, to start designing an episode and decide how they want to start their implementation. Have access to a good store of historical data from which to create your severity adjustment models and establishing the baseline for the physicians and facilities is great. If you don't have access to this data, it does not mean you can't start somewhere else. You can use averages and hold providers harmless so everyone can learn from the experience as they move forward.
    - In both Texas and New York, there were existing, large data sets with good sample sizes sufficient to really understand some of the reasons for variability in cost of care. We used our own software to establish budgets based on those analyses, however, for pregnancy, delivery, and newborn episodes, you don't need highly sophisticated. Even without the help of sophisticated software you can create a good budget. If you're going to include both the mother and the newborn, you do need at least one key piece of information which is the mother/newborn link which is the piece of data that helps you know this is the newborn of a specific mother. This may require a secondary data source which you bring in with your primary claims data source to link the events together. For the maternity bundle though, I don't think you need to have sophisticated software because the sum total of the claims isn't overwhelming.
- **Question 7:** Can you share what your experience has been with risk models and what your recommendations might be to those interested in beginning to pursue a maternity episode model which includes risk models?
    - [François]: If you are a Medicaid plan or agency, the heterogeneity of the population can create very significant differences in the underlying risk of the population. For example, one of the facilities in Texas ended up seeing a fair number of mothers who were incarcerated while pregnant, and that is a different risk profile than a young mother between 18-20 years of age who is having her first child and enrolls in Medicaid. In the first case, you probably will not have access to enough historical claims data to refine the risk models, so it is important to build into the contracts the ability to look at these cases on the back end and make exclusion decisions.

  - There are various approaches to measuring patient risk and a lot of those approaches depends on the patient population. If you're an employer, the risk profile of the mother is significantly better and far more homogenous than what you're going to see in a Medicaid plan.

- **Question 8:** How do you incorporate quality into the financial model for the maternity episode to impact shared savings or mitigate losses?
  - [Andy]: For Arkansas, Ohio, and Tennessee, incentives are applied to quality measures by requiring providers to meet a minimum threshold or set of qualifications to receive gain sharing. When considering which quality metrics to use, states and payers should:
    - 1) Determine which quality targets are weighed most heavily by providers (e.g., reducing cesarean deliveries, improving the number of prenatal care visits, and good birth outcomes)
    - 2) Determine which metrics are easily measured using claims data
  - After selecting the potential quality metric list, set the gain sharing and loss bars by considering the aspirations of the payers and providers in the state/region. If you want a meaningful number of providers to be eligible for gain sharing in the first year, start with a relatively achievable target for each of the chosen quality metrics. Also, consider the understanding of and investment in value-based payment models that providers in your region already have. Balance the complexity against the precision of the incentives.
- **Question 9 [From MAC Participant Jill Allman]:** Are there mechanisms in any of these models for Midwifery model care with enhanced prenatal care, longer and more personal prenatal visits with more education integrated in? Midwifery model prenatal care can significantly improve outcomes and save dollars on better outcomes, but it is more expensive due to longer visits on the front end.
  - [Andy]: I am not aware of one, but all of this could be incorporated into a model by adjusting the price used in the calculation of each providers average cost. You could also measure midwifery care in the quality targets to encourage use. These models are flexible enough to incent the advantages of midwifery care. It is a clinical judgement from payers as to whether it is an advantage for them to do so.
  - [François]: There is nothing that currently discourages the inclusion of midwifery care. You may need to start from the ground up to include midwifery model care. Pricing an episode which includes mostly midwives might cause the cost of the episode to be substantially lower which many providers might not be willing to accept. Using current historical averages, if medical/obstetric groups want to trade off some of the OB care with midwifery care, then they may have an overall lower cost of care in the end which would increase their gain sharing. From that perspective, any substitution of resources that yields a better outcome at a lower cost is explicitly encouraged in these kinds of payment models.
  - [François]: Our analysis shows that the single biggest determinant so far for the total cost of global bundles is the newborn's birthweight. If on a relatively systematic basis you can have a midwife following a patient through gestation with a good rapport and relationship at a significantly lower cost than you would have with multiple OB visits, and you end up with a nice healthy newborn through a vaginal delivery, you could do very well in these payment arrangements.
- **Question 10**: If you were to give advice to a commercial insurer, Medicaid plan or state Medicaid agency on financial model development for maternity episodes, what is most important for them to know?
  - [Andy]**:** When introducing the perinatal episode model in Arkansas, the dominant response from obstetricians was that their population of patients was different. We then heard that same response from all other providers in every episode we introduced. So, I think the number one principle is to demonstrate in the design and adoption of the process that you can account for any variation and create an apples-to-apples comparison. That is fairly achievable, especially in Medicaid, because the

volume is so high. Therefore, convince providers that it's not about the statistical variation but about deciding on what you want to achieve, what are the clinical outcomes, and what do you really want to incentivize.

- [François]**:** Population health experts point to two important outcomes from pregnancy and birth: 1) a vaginal delivery over a cesarean delivery due to the long-term impact, and 2) a normal birthweight baby. Whatever model you design, keep those two outcomes as your focus. Not that other outcomes don't matter, but those are the two outcomes that matter most for the mother and for the newborn.

Appendix: Interactive Poll Results

Setting the Episode Budget Poll Questions

Setting the Patient Population Poll Question	Participant Response	Total Percent
What approach is your organization currently using or considering for maternity episode budget setting?	Provider past experience	38%
	Regional past experience	25%
	Cost-based build up	38%
	Other	0%
What type of risk model is most feasible for your organization to pursue?	Shared savings	75%
	Shared risk	25%
	Discount arrangement	0%
	Full risk	0%

Feedback on the Value of this Meeting

Question	Answer	Total Percent
Please let us know how you would rate the value of this meeting?	Very Valuable	67%
	Valuable	22%
	Somewhat Valuable	0%
	Not Valuable	11%