COLORADO
Department of Health Care Policy & Financing

Payment Alternatives for Colorado Kids (PACK)
Design Review Team (DRT)
Meeting Minutes
May 22, 2024

5:00 P.M. to 7:00 P.M.

1. Introductions

Suman Mathur called the meeting to order.

The following DRT participants were in attendance: Alison Keesler, Amber Griffin, Andrea Loasby, Cassie Littler, David Keller, Ealasha Vaughner, Hoke Stapp, Jane Reed, Laura Luzietti, M. Cecile Fraley, Mark Gritz, Melissa Buchholz, Mike DiTondo, Robert Haywood, Sarah Bennett, Sarrah Knause, and Toni Sarge.

Other attendees included Devin Kepler (Department of Health Care Policy and Financing [HCPF]), Katie Price (HCPF), Helen Desta-Fraser (HCPF), Nicole Nyberg (HCPF), Peter Walsh (HCPF), Breelyn Brigola (Stakeholder Engagement (SE) Team), Emily Leung (SE Team), Suman Mathur (SE Team), Puja Patel (PACK Support Team), and Samantha Block (PACK Support Team).

Emily Leung presented DRT Session #6 meeting minutes for approval, during which an action item was created to revise language for one DRT participant's meeting comment.

2. Level Setting

Emily Leung reminded DRT participants of the PACK North Star Goal.

Suman Mathur then stated that today's discussion was focused on quality target setting, specifically on the reward structure impacting incentive payments. She explained that the objectives for today's meeting were to 1) understand how a reward structure impacts payments for the PACK program; 2) review and get feedback on Commendable Threshold and Minimum Acceptable Threshold reward structure components from DRT Session 5; and 3) provide feedback on options to assess performance between the Commendable Threshold and Minimum Acceptable Threshold. She referenced payment design components introduced in DRT Meeting 6, which are primary care services (including payment for the Alternative Payment Model (APM) 2 code set) and pay-for-performance incentive payments (contingent on meeting standards for the six (6) Division of Insurance (DOI) Pediatric Measures).

3. Quality Target Setting and Reward Structure

Helen Desta-Fraser stated that HCPF's ('the Department) current uses a close the gap target setting methodology. With this methodology, Primary Care Medical Providers (PCMPs) are measured year-over-year improvement from their own historical baseline and rather than absolute thresholds. Helen explained that while close the gap is the current methodology, HCPF is considering other methodologies for new or re-designed value based payment programs, including PACK. She noted that today's discussion considers absolute performance rather than improvement (e.g., close the gap methodology).

Presentation and Discussion on Key Components of Any Reward Structure

Helen Desta-Fraser reviewed key components of a reward structure introduced in a previous DRT meeting on April 24, 2024, but refined based on stakeholder feedback, which include the following.

- **Commendable Threshold:** Maximum threshold based on reasonable attainability where all performance above is rewarded.
    - Based on stakeholder feedback on April 24, 2024, the Commendable Threshold and HCPF Goal (e.g., Stretch Goal) are now equivalent, and the intent is to align the Commendable Threshold and HCPF Goal with realistic goals that are in line with what is attainable for providers. Performance will be evaluated periodically, and the commendable thresholds may be adjusted in future phases of PACK.
- **Commendable Area:** High performers, who are above the Commendable Threshold and would be eligible for 100% of the reward.
- **Minimum Acceptable Threshold:** Based on minimum acceptable standards where all performance below is not rewarded.
- **Minimum Acceptable Area:** Low performers, who are below the Minimum Acceptable Threshold and would be eligible for 0% reward.

Dr. Peter Walsh communicated that the Department has determined to officially adopt Childhood Immunization Status Combination 10 (CIS Combo 10) as a quality measure, which is the example measure for this discussion. This combination does include the flu vaccine. Dr. Walsh stated that currently Colorado is performing at the 33rd percentile for CIS Combo 10, and that the national mean similarly reflects this challenge. He clarified that benchmarks will be relative to the measure. He also noted that state performance is higher for CIS Combo 10 versus CIS Combo 7 and encouraged DRT participants to consider where to set the thresholds in today's discussion.

Questions and feedback from DRT participants regarding CIS Combo 10 are below.

- A DRT participant asked about the data source for CIS Combo 10, whether it's from Colorado Immunization Information System (CIIS), or claims data. The participant also asked about continuous enrollment specifications.
    - The Department explained that the initial process utilizes claims data, supplemented by CIIS data, adhering to the CMS Child Core Set Technical Specifications. The Department clarified that modifying measures based on inclusion and exclusion criteria is not part of the current strategy.
    - The Department added that continuous eligibility requirements include 12 months prior to a child's second birthday, allowing for a single gap in enrollment of up to 45 days.
- A DRT participant asked whether the commendable and minimum acceptable thresholds are up for discussion for all measures, and whether they will be evaluated on a yearly basis.
    - The Department responded that all the other elements are up for discussion except for CIS Combo 10 specification.
- DRT participants shared that vaccination status is often beyond a providers' control, despite their best efforts. A DRT participant expressed concerns about the challenges in meeting vaccination rates, noting differences in state policies for vaccinations required for school enrollment and the penalties faced by practices for accepting unvaccinated patients. The possibility of introducing codes for immunizations not performed due to caregiver refusal was suggested.
- A DRT participant asked about the population included in the measure performance, noting that practices serving many newcomers as they become Medicaid-eligible could be impacted.
    - The Department responded that specifics related to who will be included is to be

determined. There will likely be attribution changes with ACC Phase III which may have considerations for PACK.

Samantha Block shared an example of the reward structure using an example measure, CIS Combo 10, with the caveat that the presented thresholds are hypothetical and for example purposes only. She showed the box plot for this measure and highlighted key components that applied to this measure example. Samantha explained that the top, or maximum, performance a provider achieved is at 66% while the bottom performer is at around 1%. For additional context, she explained that the Health First Colorado FFS Median Performance for this measure is 35%, and that the Medicaid National Average Performance is 32%. She further explained that the 25th percentile, or first quartile, of Colorado's performance is approximately 27% and that the 75th percentile, or upper quartile, of Colorado's performance is approximately 44%. Other key components are commendable and minimum acceptable areas (descriptions below):

- **Commendable Area:** 42% (Commendable Threshold) and above, meaning that a provider becomes eligible for the full performance reward if 42% or more of their attributed members who turned age 2 that year receive all the immunizations in Combination 10.
- **Minimum Acceptable Area:** 29% (Minimum Acceptable Threshold) and below, meaning that a provider does not qualify for any performance reward if 29% or less of their attributed members who turned age 2 that year receive all the immunizations in Combination 10.

Questions and feedback from DRT participants are below:

- A DRT participant queried whether performance evaluation is tied to individual providers (National Provider Identifier, NPI) or the practice as a whole (Taxpayer Identification Number, TIN)
    - The Department and other DRT participants confirmed that performance is based on the brick-and-mortar practice site, or Primary Care Medical Provider ID (PCMP ID).
- A DRT participant shared an example of practice's situation of having multiple locations under one TIN, and expressed concern about how member movement between these sites affects performance measurement.
    - The Department and the PACK Support Team clarified that a practice is rewarded for attributed members who received the service under Medicaid, regardless of the site of where the service was received.
- A DRT participant asked about the proposed commendable threshold and whether it aligned with the upper quartile or was merely for discussion. The DRT participant also pointed out the potential discrepancy in performance assessment based on patient distribution across different PCMP IDs.
    - The PACK Support Team explained that the commendable threshold (42%) happened to coincide with the upper quartile (~44%) in this case, clarifying that the commendable threshold should ideally be based on a national benchmark.

Suman Mathur facilitated a conversation surrounding Commendable and Minimum Acceptable Thresholds. Discussion questions posed to DRT participants are:

- Is there a performance level that justifies providers receiving the **full (100%)** reward?
- Is there a performance level that is inadequately low where **no reward (0%)** should be given to providers?

Questions and feedback from DRT participants are below:

- DRT participants agreed there should be a commendable threshold in which providers receive 100% of the reward.
- Several DRT participants worried that a minimum acceptable threshold may disincentivize participation within the measure or PACK program.
    - Another DRT participant said it may be acceptable to have a minimum acceptable threshold if the program is upside risk only.
- A DRT participant suggested that the absolute threshold should not exceed the 80% percentile because there is little room for improvement beyond that point.
- At multiple points during this meeting, but beginning in this section, DRT participants referenced "externalities" beyond providers' control that could negatively impact their performance on a metric, and that this could translate to financial risk if this impacts a practice's ability to access incentive dollars.

Presentation and Discussion of Rewarding Between Commendable and Acceptable Thresholds

Samantha Block stated that the next part of the discussion focuses on rewarding between Commendable and Minimum Acceptable Thresholds. The two potential options, which would be consistent across all measures, for scaling rewards between the thresholds, are **tiering** and **sliding scale**. The tiering option was presented and discussed first, followed by the sliding scale methodology presentation and discussion. A broader discussion about both reward options occurred at the end.

Tiering: Option 1 for Rewarding Between Commendable and Minimum Acceptable Thresholds

Samantha explained that in a tiering reward methodology, payments earned are tiered based on performance levels. Using the Childhood Immunization Status Combination 10 (CIS Combo 10) example measure, in which the Minimum Acceptable Threshold is 29% and the Commendable Threshold is 42%, practices performing below the minimum threshold receives 0% reward, practices in Tier 1 (29%-35% performance rate) receive 33% reward, practices in Tier 2 (36%-41% performance rate) get 67% reward, and PCMPs performing at or above 42% performance rate (Commendable Threshold) receive 100% reward. To demonstrate this distribution of how PCMPs may perform on CIS Combo 10, Samantha presented a table of performance rates organized by tiers with respective performance and shared a hypothetical example of how many awarded points PCMPs would receive based on their performance rate and tier.

The SE Team facilitated a discussion about the tiering reward option. Questions for consideration include:

- Are there other potential benefits or drawbacks that should be included?
- Are there any unintended consequences?
- Should there be a buffer that prevents year to year backsliding to a lower tier?
    - If so, how much should that buffer account for?

Tiering questions and feedback from DRT participants are below.

- DRT participants noted that an unintended consequence of the tiering reward structure is that a provider could bounce around tiers (e.g., backsliding or "winning by accident") simply due to statistical error or random chance if the denominator size is not large enough.
- DRT participants noted that practices need to receive timely data and practice support/coaching regularly in order to assure they are able to improve their performance.
    - Samantha acknowledged these points and stated that considerations for data and program sustainability (e.g., coaching and practice supports) would be discussed in a future meeting.
- A DRT participant asked if PACK is trying to provide incentives for continuous improvement or incentives to reach certain thresholds, explaining that with a tiering approach, practices may focus on reaching a specific threshold so rewarding improvement isn't necessarily there unless a practice exceeds that performance threshold and progresses to the next tier.
- A DRT participant inquired about the frequency of threshold assessments and their basis, questioning the stability of goalposts for practices striving to advance tiers and the assurance of reward for their efforts.
    - The PACK Support Team stated that the adjustment of thresholds and the challenge of moving goalposts were topics for discussion within the Department.
- While out of scope for the discussion, it was commented that practices appreciate close the gap methodology goals, as they allow for performance competition against their own previous results within the current APM structure, offering a relatable and understandable approach.

Sliding Scale: Option 2 for Rewarding Between Commendable and Minimum Acceptable Thresholds

Samantha presented the second option for rewarding between thresholds, which is the sliding scale reward methodology, in which payment earned is proportionate to achievement percentage. Using the CIS Combo 10 measure as an example with hypothetical thresholds similar to the previous example, the Minimum Acceptable Threshold is 29%, Commendable Threshold is 42%,

and the sliding scale range is between 29% to 42%, constituting a 13% difference. With a sliding scale option, practices that perform below the minimum threshold receive 0% payment and those that perform above the commendable threshold receive 100% payment.

Samantha then presented a hypothetical example of provider performance calculations for CIS Combo 10 to demonstrate how points are awarded by practice. This is achieved through taking practices' performance rates and calculating a normalized score (performance rate – minimum threshold) divided by the difference between commendable and minimum threshold, which is then multiplied by the total possible measure points.

The SE Team invited DRT participants to share their thoughts about the sliding scale reward option. Questions for consideration include:

- Are there other potential benefits or drawbacks that should be included?
- Are there any unintended consequences?
- Should there be a buffer that limits the amount that a provider can backslide?
    - If so, how much should that buffer account for?

Sliding scale questions and feedback from DRT participants are below.

- A DRT participant highlighted concerns regarding uncertainty and potential backsliding with both sliding scale and tiering reward structures, emphasizing the impact of sample size on variability.
- DRT participants agreed that assessing or predicting performance and potential reward with the sliding scale methodology may be overly complex for practices. They shared that it could be administratively burdensome for practices to track/manage data for multiple measures across various sites, especially for rural practices.
    - One DRT participant did not think the sliding scale math and methodology were complex.
    - Samantha confirmed that practices would not be held responsible for calculating their placement on the sliding scale option, as this information would be provided to practices by the Department.
- DRT participants emphasized that if payment associated with performance metrics were to account for a significant portion of a practice's revenue, it is essential to have a more predictable structure to limit a practice's financial risk.
- A DRT participant noted that accurate assessment of performance hinges on accurate attribution. Establishing trust was highlighted as crucial for encouraging practices to engage in quality improvement initiatives.

The SE Team led a broader discussion about both reward options of tiering and sliding scale. Questions asked were:

- Using Menti, do you prefer a Tiering or Sliding Scale reward method?
    - For those that answering Tiering, why?
    - For those that answered Sliding Scale, why?

Questions and feedback from DRT participants are below:

- Menti results show that most, 9/14, DRT participants prefer the tiering option, 2/14 DRT participants prefer a sliding scale option, and 3/14 DRT participants indicated they had no preference.
- Reasons DRT participants preferred the tiering methodology over sliding scale include:
    - Tiering provides a buffering capacity that reduces financial unpredictability and is seen as less complex and administratively burdensome compared to sliding scale.
    - Concerns that sliding scale could lead to significant payment fluctuations due to operational flow, making it particularly challenging for pediatrics where resources are already constrained.
    - The perceived vulnerability of sliding scale to minor changes affecting results, especially in smaller practices, which could impact financial stability.
    - The tiering approach is viewed as offering a clearer, more manageable risk level, which is crucial for pediatric practices that cannot afford significant financial risks or administrative burden.
- DRT participants responding to the Mentimeter activity who preferred the sliding scale methodology to tiering shared factors including:
    - The sliding scale offers rewards based on improvement without needing to meet a static goal, potentially providing a more dynamic and growth-oriented approach.
    - It presents an opportunity for practices to be rewarded for incremental improvements, which could motivate practices to continually strive for better performance.
    - Despite concerns about complexity and uncertainty, some see the sliding scale as offering potential for growth and improvement beyond fixed thresholds.
- DRT participants emphasized that pediatric practices are unable to risk possibly losing out on incentive dollars, which are essential to their revenue.

    - A DRT participant elaborated that currently the reason why most pediatric practices aren't able to accept a large amount of Medicaid patients is because they aren't able to financially break even. They suggested that PACK design should increase overall revenue to support practices in accept Medicaid patients.
- Throughout the DRT meeting, a few DRT participants noted at different points that they prefer the close the gap methodology to encourage practice improvement.

Presentation of Example Scorecard for Quality Payment

Samantha provided a broader view of how points may be generally tied to payment, regardless of proposed methods like tiering and sliding scale by presenting an example scorecard for quality payment that a practice may receive. In this scorecard, there are two measures, CIS Combo 10 and Well-Child Visits in the First 30 Months of Life. Samantha described that in this hypothetical example, the practice is awarded 33 points out of 100 maximum points for CIS Combo 10 and 67 points out of 100 maximum points for Well Child Visits in the first 30 months. In this scenario, the maximum eligible incentives are $50,000. To calculate the reward payout to the practice, the maximum eligible incentives ($50,000) is multiplied by the proportion of the practice's total points (33+67=100 points) to the maximum possible points (100+100=200 points). The rewarded payout to the practice in this example is $50,000 multiplied by (100/200) to get $25,000.

Discussion of Clinical Quality Measures

The SE team facilitated a discussion about the six (6) Division of Insurance (DOI) pediatric clinical quality measures (listed below) and invited DRT participants to share their thoughts about whether all measures should be equally weighted. The discussion question posted to DRT participants was:

- Should all measures have the same amount of points available? If not, why should some measures be weighted differently than others?

DOI Pediatric Clinic Quality Measures:

1. Well-Child Visits in the First 30 Months of Life (First 15 months and 15-30 months)
2. Child and Adolescent Well-Care Visits
3. Childhood Immunization Status Combination 10
4. Immunization for Adolescents Combination 2
5. Developmental Screening in the First Three Years of Life
6. Screening for Depression and Follow-Up: Ages 12 to 17

Questions and feedback from DRT participants are below:

- DRT participants discussed the importance of measure nuances, highlighting that not all measures may equally enable providers to maximize reimbursement, especially with variations in the DOI Pediatric Clinical Quality measures (listed above). Concerns were raised about measures like Well-Care Visits in the First 15 Months of Life, where early attribution issues could limit providers' ability to meet requirements.
- DRT participants suggested heavier weighting for Child and Adolescent Well-Care Visits as the measure is inclusive of other measures and touches more members but noted the importance of considering continuous enrollment issues.
- DRT participants wondered whether practices would be able to choose which measures they participate in, and noted that the population size, or denominators, for some measures may be very small given the age distribution of their patients.
    - Samantha stated future DRT discussions will focus on unique populations and considerations for practices who see small groups.

Suman Mathur invited DRT participants to provide any additional feedback and ask final clarifying questions. Questions and feedback from DRT participants are below:

- DRT participants raised concerns about practices performing in the bottom quartile potentially not receiving payments, highlighting the critical need for PACK to be financially sustainable to avoid driving practices away. It was argued that the absence of potential for financial gains in PACK could essentially pose a downside risk, underscoring the program's requirement to offer real financial improvement opportunities to maintain participation.
    - The design of the initial PACK model was discussed, noting its basis on national benchmarks and an upside-only approach to minimize downside risk. Despite these measures, there were lingering worries that insufficient opportunities for financial growth could discourage engagement with the program.
- Several suggestions were made by DRT participants for state-level interventions and support mechanisms to help address the challenges faced by practices, especially those performing below expectations. Identified opportunities included:
    - Engaging proactively with underperforming practices to understand and mitigate both financial and systemic obstacles.
    - Offering compensation for participation in the Vaccine for Children (VFC) program, alongside exploring broader state management of vaccines to reduce the financial and operational

strain on practices, thereby helping them to serve more Medicaid patients.
- Considering the integration of additional activities within PACK as incentives for practices to undertake quality improvement initiatives, recognizing the diverse challenges and opportunities that different geographic locations and practice settings present.

4. Looking Ahead

Suman Mathur provided a list of resources and reminded DRT participants about the next meeting on June 12th from 5-7pm. She added that future discussions will focus on some of the issues raised in today's discussion, including considerations like rural and practice size, performance improvement, and program sustainability. Suman then closed the meeting.