MERCER

BRIEFING PAPER FOR SHARED SAVINGS FOR DUAL DEMONSTRATION

The Department of Health and Human Services is providing funding to 15 states to develop new models featuring care coordination and financial alignment to serve the disproportionately complex and costly medical needs of "dual eligibles"; (i.e., those individuals eligible for both Medicare and Medicaid). Centers for Medicare and Medicaid Services (CMS) followed up with a letter to state Medicaid agencies outlining two different methodologies for financial alignment; one using capitation and one using fee for service (FFS) reimbursement. Connecticut is pursuing a FFS financial alignment model. These models strive to improve quality and experience of care while lowering costs. There is general consensus in the State that this is an important initiative, perhaps the most important payment reform initiative the State is proposing.

One key feature of this demonstration project is that CMS will share Medicare savings with the State. As an indicator of the magnitude of expenditures on this relatively small population, in 2009, dual eligible individuals represented 10% of Connecticut's Medicaid beneficiaries and 10% of its Medicare population, but accounted for 38% of the State's Medicaid expenditures and 16% of the State's Medicare expenditures. How savings are measured and used has the potential for significant impact on the State of Connecticut.

CMS expects that states will make an upfront investment in care coordination and subsequently be eligible for a retrospective payment if a target level of certified Medicare savings result and quality thresholds are met. CMS calculates available savings as Medicare savings net of increased Federal Medicaid costs. Beneficiaries enrolled in Medicare Advantage, Medicaid managed care and Program of All-inclusive Care for the Elderly (PACE) may not participate unless they disenroll.

CMS will keep 50% of the funds saved, sending 50% to the State. Subject to CMS approval, the State has a wide range of options about what it may do with these funds. It may add them to the state general fund, use the money to improve quality systems, improve beneficiary benefits (as Massachusetts is doing) and/or share the funds with providers.

This program would operate in parallel with other shared savings initiatives; notably Accountable Care Organizations (ACOs), and Comprehensive Primary Care Initiative (CPCI), a multi-payer medical home model for which Connecticut is a finalist. This reflects the growing trend toward, and hope for, payment reform as a means to "bend the cost curve" without loss of quality.

The interest in shared savings reflects an emerging consensus that quality and cost both need to be addressed to ensure value. Specific interventions aiming to reduce waste such as duplicative screening tests, services that are not evidence based, services that are unnecessary or inappropriate, and preventable care such as avoidable hospital re-admissions are an important piece of the puzzle. Also important is creating a general climate in which providers become more mindful of cost considerations when they can do so without sacrificing quality. Rauh et al have

made the case that relatively small changes on the margins will not result in substantial savings because of high fixed costs. Changes must be systemic and costs must be part of the equation if the health care system is to be sustainable. Primary care case management models and pay for performance focused on the first half of the equation for producing value – improving quality. Adding shared savings incentives completes the mix.

Some stakeholders have expressed reservations about the possibility of the State sharing savings with providers, fearing that under certain designs providers may be motivated to withhold needed, more expensive services to accrue savings. They express particular concern because of the potential vulnerability of this population, stating that incentives related to this group should not be different from those proposed for the general population, specifically citing the person-centered medical home (PCMH) initiative which rewards quality rather than savings. While the proposed duals model would require that certain quality goals be met before savings could be shared, they regard these measures as too blunt an instrument to insure that no harm is done. They believe incentives already exist to reduce services to Medicaid because of low payment rates. They support quality related incentives, but not those tied to overall cost. This paper outlines issues around shared savings, previous experience with shared savings in other settings and recommendations for moving forward.

The duals demonstration will test two models. The first uses the existing Administrative Services Organization (ASO) to provide improved management for a large share of the dual eligibles using data analytics and intensive care management. The ASO will receive additional funding for this population, reflecting its greater complexity. However, this contractor will not have prior authorization authority over Medicare services, limiting its effectiveness in reducing unnecessary and inappropriate care. The ASO will be eligible for value based incentive payments, but will not be eligible for shared savings.

The other part of the proposed initiative involves establishing Health Neighborhoods (HN) (three to five are envisioned) and sharing savings with these provider networks in such a way that cost reductions are rewarded, contingent on providers meeting certain quality standards. That is the section under review in this paper.

Key policy question: Should savings be shared with the HN?
The first question to be decided is whether or not to share savings with HNs at all. The proposed approach is to share 25-50% of the total Medicare and Medicaid savings. In addition to the concerns expressed by advocates related to possible service disincentives, there are other reasons to be cautious:

- The possibility of rewarding random fluctuations in costs. Since there is no downside risk to providers there is no possible mirror-image state reward. This can be addressed by requiring that a certain level of savings be achieved before sharing gains (CMS has established such minimum savings requirements for its ACO program). This also points to sharing savings at a level above that of the individual provider so that there are adequate numbers of beneficiaries to ensure reasonably reliable results.

- Administrative complexity of distributing savings. The State would need to establish a protocol for sharing savings and distribute those savings directly to providers, or hire a fiduciary agent to do so, or require a formal structure of the HN to serve in a fiduciary role. Connecticut is assuming there will be an administrative lead for a HN, but its exact role is uncertain. North Carolina was required by CMS to set up an allocation structure in its 646 Medicare waiver demonstration.

  While a surmountable problem, this distribution adds a layer of cost and complexity. Depending on the approach taken it could also require negotiation with CMS, including the possibility of applying for a waiver. The most flexible waiver type, known as the "1115", is typically the most cumbersome to obtain, unless CMS puts streamlined approval processes in place. Note: the State can claim federal match on any savings shared with providers. If the State can pay for shared savings using administrative funds rather than program funds, that may provide additional flexibility. Since Connecticut gets the same federal match for both types of funds, there is no economic incentive to choose one over the other.

However, there are strong arguments in favor of shared savings with providers:

- As a practical matter, the HN model may not have adequate provider participants without a shared savings component. Providers may only participate in one shared savings initiative. Many providers will be in a position where they will be able to choose to participate in one, and only one, competing shared savings initiative. Medicare is offering shared savings incentives to ACOs, and primary care providers (PCPs) who participate in Comprehensive Primary Care Initiative. Without comparable incentives it may be hard to attract providers to participate in the HN model.

- While stakeholders have issues with potential incentives for reducing care in this program, similar incentives are likely exist to the same or even greater extent in other shared savings programs. In fact, the incentives built into the ACO and CPCI programs are stronger – with opportunities for shared savings in excess of 50% and, in the case of ACOs, there may be downside risk if savings are not achieved. The Physician Group Practice evaluation described further below found that all groups involved in this shared savings demonstration showed excellent and improved quality, whether or not they ultimately shared savings. Comparably, North Carolina found quality improvements in its first year of operations, even though there were no savings to share.

  Notably, shared savings incentives can lead providers to support *higher* levels of some services, for example, post-hospital transition services, in the interest of reducing re-admissions.

- CMS appears to be moving in the direction of having networks share savings with providers.
  - CMS advised Louisiana that it would only approve its Coordinated Care Network (CCN) Shared Savings initiative if selected networks shared savings with participating providers.

Two contractors were selected. At least one of these vendors has submitted a proposed plan, and it will limit gain sharing to PCPs serving as patient centered medical homes.
- ACO applications must include a plan for distributing gains, and possibly losses, among providers.

These arguments do not focus on the potential effectiveness of shared savings in reducing costs. Unfortunately, hard evidence as to the cost effectiveness of this approach is not in ready supply. CMS recently informed North Carolina that the first year of its shared gain demonstration program, described in the attachment, did not show net savings. Bailit Health Purchasing conducted an in depth survey of existing and planned shared savings programs and found after 32 interviews with payers and providers that most of the programs surveyed were too new to have solid results, and uncovered only one formal evaluation of a shared savings program, that of the Physician Group Practice Transition Demonstration (PGP).

The PGP experience is worth some additional scrutiny because of this evaluation and because it was a FFS model. Ten large physician practice groups were eligible for a performance bonus of up to 80% of savings if their expenditures were less than 98% of trended “Target Expenditures.” They could obtain a percentage of these savings whether or not they achieved quality targets, but a portion of savings were released proportionate to the cost-saving groups’ quality metrics. Participants faced no downside risk, except for any infrastructure investment.

A number of the evaluation findings can potentially inform Connecticut's proposal:

- Four of the 10 groups were eligible for performance payments; the two standalone groups and the two groups affiliated with academic centers, but none of the six integrated systems. The evaluators hypothesized that reducing hospital admissions may not have been a priority for these integrated systems, but academic hospitals could essentially always fill their beds.

- The evaluation found that organizational structure and previous expenditure trends were most closely associated with eligibility for performance payments. Groups whose expenditures were already trending down showed better performance.

- All the groups showed excellent quality measures and improvement in quality scores; potentially showing a willingness to invest in quality infrastructure with the hope, but no guarantee of performance payment.

- The groups used a range of interventions, making it difficult to evaluate the effectiveness of any particular strategy.

- The majority of the savings were in outpatient savings.

- There was some question of whether groups became more careful and intensive with their coding because of the risk adjustment function.

Advocates have argued that Department of Social Services (DSS) should try to encourage cost savings indirectly by giving incentive payments based on quality measures that are associated with cost reductions (i.e., reduced emergency room visits) in lieu of shared savings. However, these incentives may not be sufficient to encourage positive results. Studies of the impact of quality incentives only have yielded mixed results with respect to cost savings. Rauh et al examine the issue in their article: "The Savings Illusion – Why Clinical Quality Improvement Fails to Deliver Bottom-Line Results", while Fenton et al found that increased patient satisfaction was correlated with higher overall expenditures. When providers are focused on targeted utilization measures rather than global savings, the costs of primary care and specialist visits, home and community based services and other services that are not captured in any of the quality metrics could rise. This could occur by happenstance or through intentional cost shifting. For example, nursing facility residents tend to have fewer emergency visits than those being supported at home, potentially leading to the decision to keep someone in a nursing facility.

Additional Policy Questions

If savings are shared with providers, decisions must be made along a number of dimensions:

- **Should savings be pooled among NH? Should globally calculated savings be distributed among all participants based on quality metrics regardless of their contribution to savings?** This type of distribution could seriously skew incentives. The Dartmouth Health Atlas project, among other initiatives, has shown that over the last 20 years "glaring variations in how medical resources are distributed and used" exist across localities within a single state or region, without attendant correlation with quality. Under a pooled model, a health network that shows the greatest improvement in quality but increases cost could receive the greatest financial rewards – compensated by savings generated by others – without having improved value. Alternatively, a HN that improves quality and reduces cost could receive no reward because there were no statewide savings. This approach undermines efforts to hold networks accountable for both quality and cost; ultimately to push the system toward achieving higher value.

  In a well-publicized case study, Atul Gawande reported in the New Yorker that McAllen, Texas, at $15,000 per beneficiary, had virtually the highest Medicare costs in the country. Costs in El Paso, a comparable Texas border town with similar technology and medical personnel, were approximately half of that. On all but two of the twenty five quality metrics used by Medicare hospitals McAllen's hospitals ranked below those of El Paso. As Gawande summarized: "McAllen costs Medicare seven thousand dollars more per person each year than does the average city in America. But not, so far as one can tell, because it's delivering better health care." Under a pooled savings model, McAllen could benefit from quality improvement payments, if it managed to achieve benchmarks on identified quality measures, without making any appreciable improvements in cost or overall value.
  Additionally, providers themselves do not appear to be in favor of such arrangements. Bailit found "providers tend to argue against pooling or aggregating as it results in a perceived loss of control and motivation" and "providers appear to only support pooling across entities that have a preexisting contractual or organizational relationship".

- **Which services will be included?** Most shared savings include all covered services but some exclude some services such as outlier high cost claims, sub-capitation arrangements and out of network services. However, excluding certain services may lead to cost shifting and may reward providers even when total costs have not decreased; inadvertently increasing the overall cost of health care. Findings from the PGP evaluation support the need to have all services in the mix.

- **What populations should participate?** The Connecticut proposal is to include all "full" duals without regard to age or disability, including nursing home residents and waiver participants, high need and high cost individuals where coordinated care is critical.

  Assignment of beneficiaries to HNs can be complex and controversial. Unlike managed care, there is typically not a formal enrollment process for beneficiaries into a HN. Typically they are "attributed" to a network retrospectively based on their pattern of practitioner visits. Although savings for a beneficiary may only be counted toward one initiative, beneficiaries have total freedom of choice of providers. They may change providers or HNs at any time; there are no lock-in provisions. In fact, it should be invisible to the beneficiary that he or she is affiliated with a HN. The primary issue regarding beneficiary participation is information sharing among practitioners. While there can be privacy concerns, there are also large potential benefits to care from ensuring that all practitioners have comprehensive information about their patients. Based on lessons learned, North Carolina recommends using a quarterly attribution process, where assignment is provided retrospectively. Medicare will use attribution for its ACO models. We know of no FFS model that does not plan to use some form of attribution.

  Some stakeholders are advocating that beneficiaries be required to actively choose a HN, rather than be passively attributed to one and given the opportunity to opt-out. This can pose a significant concern. Just as there needs to be sufficient numbers of participating providers, there needs to be sufficient numbers of participating beneficiaries. Opt-in produces lower participation rates, even among those who do not object to involvement.

  Low participation can make it very difficult or impossible to get sufficient participants for the purposes of evaluating the program and may lead to reduced care management capability. A HN may not accrue sufficient resources from care management payments to cover the fixed costs associated with a robust program if participation is too low.

  Specifically, in Connecticut there are approximately 57,000 dual eligible beneficiaries who would be eligible for the demonstration. The goal is to have 5,000 to 7000 dual eligibles attributed to each of the HNs. Hartford and New Haven are the two most populous areas with approximately 16,000 dual eligibles each. Assuming no competing ACO or CPCI option, at least 30% to 40% of those eligible would have to make an active choice on the first day of implementation to make the HN statistically viable for purposes of measuring quality and costs. Experience has shown that this is extremely unlikely. For example, enrollment in Medicare Advantage plans among duals is below 15% in Connecticut. Massachusetts and NY have experienced similar enrollment rates in managed care programs. Other Connecticut

counties would require an even higher percentage of beneficiaries actively choosing to participate from the start of the initiative. This experience is not limited to the northeast. Managed care entities do not tend to use voluntary opt-in because "experience from across the country shows that recruiting on a person-by-person basis is not likely to result in sufficient enrollment to make plans actuarially sound" (Stone, Cal. Medicaid Research Inst.)

Additionally, opt-in, as opposed to passive retrospective attribution, can lead to the possibility of "cherry picking" as providers have the opportunity to persuade "easy" patients to join their practice, and to suggest that more complex and costly ones look elsewhere. Providers may avoid being held accountable for the care experience, quality or cost of these beneficiaries with more challenging needs.

For these reasons, Medicare does not allow beneficiaries to opt-out of being attributed to an ACO. Beneficiaries can opt-out of sharing clinical information with the ACO to which they are attributed, but they can only end their participation with the ACO by switching to a primary care provider who is not part of the ACO.

- **Which providers are eligible for a share of savings? Who decides how savings are allocated?** A patient's PCP, or potentially primary specialist, typically determines whether a patient is attributed to a HN, but other providers will participate in beneficiary care and in some cases may be responsible for decision making that affects some of the most expensive aspects of his or her care, and may be at risk to lose the most revenue. This distribution could affect provider choice with respect to joining the demonstration.

  Milliman has developed a briefing paper outlining some of the policy and operational issues associated with sharing savings within an ACO. Although the paper specifically addresses the issue within the ACO context, the findings relate to any network shared savings arrangement. Milliman outlines a number of considerations an ACO should take into account: simplicity and equity, encouraging quality and cost outcomes and relating a provider's incentive payment to its contribution to savings and quality. They suggest potentially looking at subdivisions of an organization, dividing savings by type of provider (e.g., hospital, PCP, clinic) and then by individual provider within that pool based on one of a number of different possible allocation formulas. Special issues arise for PCPs, specialists, low volume providers, outliers and other considerations.

Important precursors to developing a shared savings strategy are modeling outcomes and assessing risk. Important goals of such modeling are to:

- Produce savings estimates, including developing methodologies to distinguish true savings from random variation. This includes determining whether a minimum number of participants and a certain level of savings are required to receive shared savings.

- Anticipate upfront costs, both related to establishing programs and investing in preventive care that might take time to yield benefits.

- Anticipate timing of available data to calculate savings.
- Anticipate timing of potential savings and tie that to planned payments, including possible phased payments.
- Anticipate interaction with other initiatives, specifically other shared savings programs.
- Identify Connecticut specific factors that may affect outcomes.
- To the extent that quality measures are used to distribute savings, what is their statistical validity? Those that only apply to a subpopulation (i.e., diabetics) may be more difficult to validate.
- Determine whether and how to adjust for risk. Bailit found that about half of the models studied used risk adjustment.

Recommendations

In sum, the health care world is pinning much of its hopes for lower costs and improved quality on payment reform, an important and growing component of which is shared savings. This is true at both the federal and state level, and among private and public payers (many of these initiatives are outlined in the attachment). Connecticut has an important opportunity to use such reform to help coordinate and improve care for some of the most complex and costly beneficiaries of both the Medicare and Medicaid programs. We believe a carefully constructed HN program based on the following principles is our best option for obtaining this goal:

- Adequate participation of providers and beneficiaries to ensure a fair trial of the HN demonstration.
- Providers who help achieve savings should get to share in those savings, but *only* if they meet benchmarks for care experience and quality, to help promote accountability for overall value.
- Savings should be based on local rather than global performance, to improve provider motivation and engagement and to ensure a direct link to local performance.
- Savings should not be tracked to individual HN providers because there are too few beneficiaries to assure that results are not due to random variation.
- Achieve administrative simplicity where possible.
- Achieve transparency and predictability in the shared savings distribution process to the extent possible.

- Providers and services should be broadly included to preclude cost shifting and ensure local accountability for value.

- All participants need to acknowledge that shared savings are a relatively new venture, that all parties will be learning during implementation and that flexibility should be maintained.

This translates into the following recommendations with respect to Connecticut:

- Provide upfront funding for HNs to help fund infrastructure and additional supports.

- Share savings to give providers an additional incentive to organize and participate.

- Share savings based on HN specific cost performance above a minimum threshold.

- Only share savings with organizations that have met quality standards.

- Allow each HN to decide how to share savings with its member providers based on their scores on quality measures and subject to DSS approvals.

- Consider no shared savings first year dependent on availability of data and findings related to net savings.

- Track complaints and grievances and quality measures throughout the life of the demonstration to ensure no diminution of quality. Follow-up with individual providers as indicated.

- Provide data on a consumer-focused website to show how health networks are performing on quality metrics.

- Consider periodic audits to review whether providers are following accepted standards of care.

- Attribute beneficiaries quarterly to HNs but allow them to opt out with respect to sharing their data.

- Require that each HN maintain a minimum threshold of participating beneficiaries. If the number of beneficiaries falls below this threshold for a sustained period of time, its participation would be terminated.

- Revisit decisions made in designing this demonstration to the extent allowed by CMS.

Attachment
- Federal and State shared savings initiatives

ATTACHMENT – SHARED SAVINGS INITIATIVES

Medicare's interest in shared savings is reflected in the number of such projects it is supporting:

- **CPCI**, in which Connecticut may participate. This is a public-private multi-payer demonstration in which Medicare intends to share savings achieved in the covered geographic market with primary care practices, based on quality, utilization and practice size on a risk adjusted basis. It requires participating private payers to take a comparable approach. While Medicaid is specifically precluded from sharing savings, it hopes to gain CMS approval for a complementary savings-related incentive pool.

- **Medicare Shared Savings Program – ACO**. ACOs must have at least 5,000 beneficiaries and contract with CMS for three years. Beneficiaries are attributed to ACOs, not enrolled and are free to go to out-of-ACO providers. ACOs may opt for upside risk only or two-sided risk that promises higher potential gain. After accruing savings above a minimum savings requirement (2-3.9%, depending on ACO size) as compared to a projected target figure, savings are shared with ACOs. After the first yea, quality measures will affect the amount of savings shared.

- **Pioneer ACOs**, an alternative ACO model for early adopters; no Pioneer ACOs will serve Connecticut.

- **Independence at Home**. Eligible providers will provide home-based primary care to, and coordinate care for, chronically ill beneficiaries for a three-year period. Practices that meet specified quality measures and create Medicare savings can receive incentive payments after meeting a minimum savings requirement. Beneficiary Participation is voluntary.

- **Physician Group Practice Transition Demonstration**

A variety of states have undertaken their own initiatives, many of which are described on this NASHP web page. A non-comprehensive list follows:

- **North Carolina** "646" waiver (while "646" waiver authority was not specifically created to support payment reform, the North Carolina demonstration uses it for that purpose) 8 of 14 North Carolina nonprofit CCN are participating. These networks provide nurse care management, contract with PCPs and coordinate with other providers. No savings were deemed achieved by CMS, but had there been any North Carolina CCN was planning to use them for operational needs. After the first year, North Carolina CCN planned to share savings with practices but details of the formula are not finalized. There was improvement in 17 of 18 quality measures. North Carolina does have a minimum savings requirement with CMS.

- **Louisiana** Medicaid Coordinated Care Network (pending CMS approval). Louisiana is moving from a strictly FFS system and encouraging coordinated care. One model it is using is an Enhanced Primary Care Case Management plan using a per member per month (PMPM) care management fee coupled with opportunities for providers to share cost savings.

- **Massachusetts**. Payers in the medical home initiative will share savings with participating practices if they generate savings relative to a control group but the methodology is not yet available.

- **Maryland**. All participating health home practices can get incentive payments of 30-50% of savings generated by the practice if they meet performance criteria, some of which reflect use of evidence-based practices and utilization reduction. Private payers and Medicaid are participating.

- **Ohio** duals project will include quality withholds, at least 50% of which must be passed on to providers whose care lead to increases in the quality indicators.

- **Pennsylvania**. Medical home practices can share savings based on key quality and cost metrics. As the PMPM amounts decrease from year one to year three, practices can receive greater shared savings.

- **West Virginia**. Providers and payers participating in the Medical Home Performance Incentive Pilot share 5% of net savings equally.

- **Washington**. Providers in Washington State's PCMH Multipayer Reimbursement Pilot receive a PMPM that decreases after an initial period. If practices do not save at least the amount of PMPM payments, they receive reduced payments in the future.

BIBLIOGRAPHY

The Cost Conundrum
New Yorker, Annals of Medicine, Gawande, June 1, 2009

Buy-In, Large Pool, Are Keys to Success
Californiahealthline.org, California Medicaid Research Institute, UC-San Francisco, Stone January 2012

The Savings Illusion – Why Clinical Quality Improvement Fails to Deliver Bottom-Line Results
N Engl J Med; Rauh, Wadsworth, Weeks and Weinstein, December 29, 2011

Key Design Elements of Shared – Savings Payment Arrangements
Bailit and Hughes, supported by the Commonwealth Fund, August 2011

The Cost of Satisfaction; A National Study of Patient Satisfaction, Health Care Utilization, Expenditures, and Mortality
Arch Intern Med., Fenton, Jerant, Bertakis and Franks, March 12, 2012

Financial Models to Support State Efforts to Integrate Care for Medicare-Medicaid Enrollees
CMS State Medicaid Director Letter, July 8, 2011

ACO Gain/Loss Sharing
Insight at Milliman.com, Berry and Kipp, January 26, 2012

Aligning Reimbursement & Purchasing
National Academy for State Health Policy, NASHP.org Medical Homes Page, 2012

Report to Congress: Physician Group Practice Demonstration Evaluation Report
Kathleen Sebelius, Secretary DHHS, 2009