For our next section for “Introduction to Current Risk Adjustment Model” and “Model Exploration Topics.” We have joining us, two additional CMS speakers.

I would like to announce Krutika Amin. Krutika is a Health Insurance Specialist at CMS, Center for Consumer Information and Insurance Oversight, CCIIO, within the Payment Policy and Financial Management Group.

In this role, she leads policy and analysis efforts relating to Advance Payments of the Premium Tax Credit, Cost-Sharing Reduction, the Premium Stabilization Program and the User Fees for the Federally-facilitated Marketplaces.

Krutika also works on policy for health insurance Exchanges created under the Affordable Care Act, ACA, and coordinates policy and regulations related to the Notice of Benefits and Payment Parameters. Before joining CCIIO, Krutika worked at Avalere Health, providing strategic advisory services to clients. Welcome, Krutika.

In addition, we have Kelly Drury. Kelly currently serves as the acting Deputy Director of the Division of Policy and Analysis and the Payment Policy and Financial Management Group at CCIIO. Since joining in May 2012, Kelly has focused mainly on Risk Adjustment policy. Prior to CCIIO, Kelly worked on home health policy in the Center for Medicare in the Financial Services at Legg Mason. Kelly has an undergraduate degree from the University of Maryland and an M.B.A. in finance and management from Johns Hopkins University. Welcome, Kelly.

Thank you. So, I’ll just get started. As you’ve heard so far, today we think the methodology is working as intended directionally, at least. We could adjust two levers to better reflect the orders of magnitude and these are the risk score calculations and the transfer formula.

In this session, I will go through the current Risk Adjustment model and how the risk score is calculated to get everyone on the same page as to what the model currently does. Then I will describe some comments we have received regarding the model. And lastly, in this session, my colleague, Kelly Drury, will discuss additional topics we have been considering to introduce the model, to improve the model, excuse me.

So, for the individual and small group markets, the HHS-developed Risk Adjustment model is used for states in which HHS is operating Risk Adjustment. This model is calibrated on the MarketScan data and just to go over a few things about the MarketScan data, this data includes employer-sponsored insurance claims data from employers and health plans. Employee’s spouses and dependents are included in the data. The MarketScan data includes data from all 50 states and D.C.

As we have described in the paper and also previously, we made the decision to use the MarketScan data because of the better fit with the individual and small group market enrollees. And so, the Risk
Adjustment model, the HHS model is concurrent. And what this means is that the current benefit data, current benefit year data are used to predict enrollee health care costs in the same year.

When we were developing the model, we decided to use concurrent model because we expected the enrollees in the market to churn both between the markets and across issuers as opposed to the more stable Medicare enrollees. And because of the population in the individual and small group markets, we will still think the concurrent model is appropriate, but Kelly will discuss that a little more, later in the session.

So, due to the inherent clinical and cost differences between the adults ages 21 and over, children ages two to 20 and infants ages zero and one, we used separate Risk Adjustment models for each of these groups. The model was developed by estimating how demographics, that is age and sex, and how diagnoses relate to health expenditures. So, the risk score calculated from each of these models is meant to be a relative measure of how costly that individual is anticipated to be to the plan.

In addition to separately modeling by age group, because of the differences in plan actuarial risk across the actuarial value level, we developed a separate plan liability model for each of the levels. That is the platinum, gold, silver, bronze and catastrophic levels. Based on these, we are setting different Risk Adjustment co-efficients for each actuarial value level category within the adult, child and infant models. The plan liability models are based on model plan benefits at each actuarial value level.

The model plan benefits include deductibles, coinsurance, and the maximum out-of-pocket limits. We plug in each enrollee’s health expenditures to those model benefits to get to the estimated plan expenditures.

So, then these simulated plan expenditures are used to develop the separate plan liability models for each actuarial value level. Because of the structure, the Plan Liability Risk Score predicts the medical expenditures that a plan is actually liable for given its actuarial value and cost-sharing structure.

Next, I will go through more detail about each of these models. So, for each of the models, the left-hand side dependent variable is the annualized simulated plan liability expenditures. The adult model right-hand side variables include age and sex demographic categories, 114 HCC diagnosis groups and 16-disease severity indicator, interactions, excuse me.

The child model includes eight age and sex categories and 119 HCC diagnosis groups. The thing to note here is that the child model does not include the disease severity interactions.

The infant model is a categorical model based on 25 categories. These categories are defined by birth maturity for age zero and one and diagnostic severity. The model is based on female, so the males do get an additive term in the infant model. We also wanted to address the potential higher utilization among individuals who are enrolled in Cost-Sharing Reduction plans and so a multiplicative adjustment to the risk score is used to account for induced utilization associated with the more generous actuarial value Cost-Sharing Reduction plans in the individual Marketplaces.
So, next I will talk about the data that goes into calculating the risk scores and how the risk score is calculated.

So, to calculate the risk scores, institutional claims to both inpatient and outpatient claims are used together with the professional claims and supplemental diagnosis files. Of course, right now, we don’t use pharmacy claims in the model, but something we will discuss later in the session as well.

Okay, so next I will go over the steps to calculate the risk scores and so for the adult model; let’s just switch over to the next slide please. So, the adult model includes enrollees 21 years and older. As I mentioned before, for the first step, the Risk Adjustment-eligible diagnosis codes are mapped onto condition categories in the model. When an enrollee has multiple diagnosis codes for related condition categories that vary in cost and severity then a hierarchy is imposed and an HCC is created. When this occurs, the HCC with the most costly and severe manifestations will count for an enrollee and other related HCCs are dropped.

To determine whether HCCs should be included in the Risk Adjustment model, in the past we considered several different statures and this included, we thought about whether the HCC represents clinically significant medical conditions with significant costs to the target population. We also considered whether there will be sufficient sample size to ensure the table results for the HCC. We considered whether excluding the HCC would limit or exclude the impact with the diagnosis, particularly subject to discretion coding. Whether the HCC identifies chronic or systematic conditions that represent insurance risk selection or risk documentation rather than random acute events was also considered in developing the HCCs. Whether the HCCs represent poor quality of care and whether the HCC is applicable to the model age group was also considered in determining whether the HCC should be included.

So, next after the HCCs are created, we group many of the HCCs into HCC groupings and run the model so that the constraints are imposed on the coefficients for each HCC in the grouping to constrain them to be equal in the hierarchy.

Next, severe illness indicators are assigned based on certain HCCs and then the severe illness indicators are interacted with some of the HCCs. So, we group these interaction terms into high and medium costs. So, we’ve constrained the interaction coefficients for the high interactions to be equal to each other and similarly for the medium cost interactions we also constrained on to be equal to each other. Then we will not have both high and medium severity indicators, the ones that we assigned the single highest severity level.

Okay. So then next for the child model, the child model includes enrollees ages two to 20 and like the adult model, it’s also additive in parameters, but the child model doesn’t include the severe illness indicator interactions that I mentioned before. So, for the child model, the Risk Adjustment eligible diagnosis codes are also converted to the condition categories, same as the adult model, and then the condition categories may be grouped into hierarchies creating an HCC. The HCCs still get grouped into HCC groupings for the child model.
One issue that we did experience with the child model that I wanted to bring up is that, and the main reason why we originally blended the coefficients in the 2016 Risk Adjustment model was because of small sample sizes particularly for some conditions in the child model and this led to some instability in the coefficients and that is why we had originally blended the data in the 2016 model. We continue to think about how to best address the small sample sizes as we go back to using one year of data, but this is something we’ll also discuss a little later. I wanted to flag that as another point to keep in mind as we go through this presentation.

Okay. So, then for the infant model, this includes infants ages zero to one. The infant’s diagnosis codes are mapped to the condition categories and HCCs like in the adult and child model. For an age zero infant, birth maturity level is assigned based on the birth weight and gestation length. The birth maturity level determines the HCCs that are used to determine an infant’s maturity category for an age zero infant.

An age one infant is given a maturity level of one. The infant’s HCCs determine the infant’s severity level from one to five. An infant with no HCCs is assigned a severity level of one. And then the age model factor for infants is determined by the single highest maturity and severity interaction.

And lastly, for the infant model, male infants get an additive in factor like I said before and this is because the infant coefficients assume female and then all the males get the additive factor.

Based on the risk scores that are calculated within each of these models, the Plan Liability Risk Score is created and this is the overall level of risk associated with each plan and PLRS is what is used in the calculation of the transfer payments. And I think the formula is actually on the next slide.

So, the risk score is defined as the total predictive relative plan liability expenditures for the enrollee based on the relevant model for the enrollee’s age group and plan metal level. An adjustment to the risk score is made for enrollees in the Cost-Sharing Reduction plan variations in the individual Marketplaces like I mentioned before.

We wanted to account for this adjustment as I mentioned before because enrollees may utilize higher health care services at a higher rate than would be the case without these Cost-Sharing Reductions.

And so, since the premiums for all Cost-Sharing Reduction plan variations are required to be the same despite increased AV level for these plans, we account for the demand associated with these CSR variant plans in the model, not the transfer formula.

The PLRS then is the average enrollee risk in the plan like I mentioned before. It's also weighted by the billable member months so that enrollees who are in the plan for longer have a higher rate.

Okay, so with that, that’s all the background we’re going to provide here for now. There’s more in the paper and there, we have published a lot more information previously. But with that background in mind, we, of how the risk score is calculated, we really want to now turn to the some of the topics we have been considering.
So based on the first year of experience with the Risk Adjustment model, we are evaluating various topics for future recalibrations and we’ll discuss a few of those in this session. Okay, so some of the topics we are considering include whether and how to adjust for partial year enrollment. We’re also considering if and how to model prescription drugs in the model. We’re also thinking about including high-risk enrollee pooling in the Risk Adjustment model. We will discuss concurrent and prospective Risk Adjustment models in this session, and lastly, we will discuss recalibration of the 2018 model. But in the session with the recalibration, we won’t be discussing the enrollee level recalibration. We will discuss that later in the day.

Okay, so I will talk about or announce these on partial year enrollment and then Kelly will speak to the other topics that are part of the session.

So, we received feedback that some issuers experienced higher than expected claims costs for partial year enrollees for the initial year of the Risk Adjustment program. We also heard from stakeholders that the methodology does not capture enrollees, but chronic conditions who may not have accumulated diagnoses in their partial year of enrollment. But on the other hand, compared to full year enrollees, partial year enrollees are less likely to have costs that exceed deductibles or the maximum out-of-pocket limits and so less cost would be incurred to the issuer. Another issue we heard is that enrollees with partial year enrollment of six months or less had higher MLRs and financial losses for issuers.

And so, having heard these comments, we asked for public comments in the 2017 Payment Notice as to how we should consider including partial year enrollment in the Risk Adjustment model. We got a bunch of different comments generally supporting the partial year enrollment adjustment in the model. Some suggested that we incorporate prescription drug claims which will help capture partial year enrollees that have a chronic condition, but do not have a provider encounter with a diagnosis that is reported. Some comments also said that we should consider member level adjustment or duration adjustment. We got a comment saying that maybe we should consider something like the Massachusetts model that includes an adjustment for partial year enrollment. We also got comments that we should make sure that any changes to the model improve the reliability and predictive power and that we don’t unnecessarily influence clinical judgement or plan behavior with such an adjustment.

Some commenters also said that they were concerned that including partial year enrollment would encourage loose enforcement of Special Enrollment Periods. And so, in all the comments we got, were generally supportive, but there was some word of caution about any methodology that we do incorporate.

In general, our belief is that in the individual and small group health plans, for whom enrollees are representative of the market and those will be risk adjusted accurately under our current methodology. However, for current Risk Adjustment, our current Risk Adjustment methodology may be inaccurate when a plan’s enrollees are significantly different from the market as a whole with respect to characteristics that are not included in the model.
And so, we analyzed the MarketScan data a few different ways to see whether we should adjust for partial year enrollment and how to adjust for it. So first, we looked at partial year enrollees in the MarketScan data and how the model was doing in predicting expenditures compared to their actual expenditures. Then we added in dummy variables in the model for enrollment duration to see how the model would do with those, and then lastly, what we’ll discuss here is our analysis when we created separate models by enrollment duration groups. In the next set of slides, I’ll go through each of these analyses and our initial thinking here and what we still want to consider with partial year enrollments.

So, like I mentioned, we first looked at the adults by their enrollment duration and the MarketScan data with the silver model. We grouped enrollees by their enrollment duration. So, we grouped them from one month of enrollment duration, two months, all the way up to 12 months. What we wanted to see is what their predicted expenditures were for each of these groups based on the 2014 adult silver model and compared the groups’ predicted expenditures to their actual expenditures. What we found was that enrollees that didn’t have a full year’s enrollment, so those that were in the one month to 11 month group enrollment, enrollment duration groups, these enrollees, their predicted expenditures with the 2014 adult silver model under predicted expenditures, so compared to their actual, what their actual expenditures were, their predicted expenditures were lower. The lower enrollment groups particularly, particularly for the one to seven months of enrollment were under predicted by the 2014 model. On the other hand, enrollees that had a full year of enrollment, a 12-month enrollment, were slightly over predicted by the 2014 silver model. And this could have been because enrollees will get HCCs or high cost acute conditions. But those with a full year’s enrollment, these high cost acute conditions or the cost for these gets spread out for the full year, whereas for those that only have a partial year’s enrollment, these expensive acute events end up being concentrated in a shorter period.

So, having seen this, the next thing we did is we included each category of enrollment duration in the 2014 silver model, as an indicator variable and we wanted to see how the model coefficients would change having added the enrollment duration indicators. We included an indicator variable separately for each month increment and the 12-month group was the reference group in the model. The way this model with these results read as that the risk score would be increased by .424 for enrollees with only one month of enrollment and that would be on top of the diagnoses, the HCCs and the demographic factors.

Results of factors themselves then reflect the higher cost associated with partial year enrollment and this is partially because the factors don’t distinguish the risk for partial year enrollees with no payment HCCs and the enrollees with payment HCCs. We also saw that the other coefficients were moving in different directions for different groups of people without a clear pattern. So, we felt that just including these indicator variables was misleading and one of the things we are considering is grouping by HCCs that are sensitive to the partial year enrollment or make a distinction somehow between payment HCCs and no payment HCCs when incorporating partial year enrollment.

So, the third analysis we did is we modeled the 2014 silver models separately by enrollment duration groups. So, we grouped them into the three groups of one-to-four months of enrollment, five-to-eight
months and nine-to-12 months of enrollment. The idea being that based on which category the enrollee is in for the length of enrollment, the coefficients from that model would be used. The thing we do like about this model is that other things being equal, separate models by enrollment duration are better for predicting accurately for each group of partial year enrollees. But on the other hand, there are tradeoffs to those. We would be making this methodology a lot more complex. So, right now with the different metal levels, the adult, child and infant models, we have 15 separate models and separating the models by enrollment duration groups would bring it to 45 separate models. And while the model would become more specific for the enrollment duration groups, we would be worried about presenting soft precision would be used in predicting costs. So, we were also concerned about the model stability, especially around HCCs with the smaller sample sizes with separating the model like this.

So, there are a few trade-offs here. We did find that when we separately modeled by enrollment duration groups, the coefficients were very different for expensive acute conditions across the duration groups. That was compared to the chronic conditions, which seemed relatively stable and we modeled it like this. The other thing I want to note here is that these results are based on a commercial data set and so they may not reflect what some commenters and issuers have experienced and reported to us. The MarketScan data, of course, is based on the employers’ plans, which may not reflect the churn that I was mentioning before with the individual and small group markets.

So, lastly as the next step, having considered these analyses because of the issues that I presented in all of these analyses that we did, we are still considering what to do about partial year enrollment. We’re considering a hybrid approach where we would take the HCCs that are the most sensitive to partial year enrollment and interact them with enrollment duration groups. We’re also continuing to think about the most appropriate way to account for the partial year enrollees and the model. We’re keeping in mind that by including the partial year enrollees, we don’t want to inadvertently discourage issuers from retaining enrollees. We also don’t want a scenario where we’re misrepresenting Risk Adjustment coefficients because of the small sample size of it.

And with that, I will pass it onto Kelly Drury to present the next few topics we have considered.

Thanks, Krutika. I’m going to start off here with drugs, which will probably be the bulk of this presentation as it was in the paper. We’ve received a lot of feedback on prescription drugs since at least the 2014 Payment Notice when we finalized the original Risk Adjustment methodology and, you know, commenters have mainly requested that we incorporate prescription drugs as predicted in the model. We’ve evaluated it for some time now, a couple of years I think we’re going on and we decided that we should proactively solicit feedback on whether or not we should incorporate prescription drugs in the model and how in the 2017 Payment Notice. And we received overwhelming support for it. Commenters generally suggested that prescription drugs be incorporated as quickly as possible while others, you know, suggested that we go through our typical rule making process as we would and wait and lay out a methodology for potentially incorporation on the 2018 Payment Notice, which is what we’re considering here.
Some other commenters also supported that we include prescription drugs in a limited manner. Others supported including every drug that’s out there in the model. Commenters also felt that using prescription drugs could increase payment accuracy and Risk Adjustment, especially when you set the proxy for a missing diagnoses for an indicator for a severity of illness. And lastly, commenters also thought that prescription drug data is better data, it’s faster, it’s more uniform across providers and their incentives and lastly, commenters shared our concerns about overprescribing and discretionary prescribing.

So, incorporating prescription drugs remains one of the biggest challenges you’ll hear about today and a lot of these concerns that I’ll talk about will influence or decision whether or not include these and how in the 2018 Payment Notice, if we do.

In evaluating a hybrid diagnosis and drug model, which is what we’re considering, as opposed to a drug-only model, there are a number of factors to consider. Hybrid models retain the corrective power of a diagnosis model and they seem to be most beneficial when adding just a few drug classes, but establishing the criteria to include these drug classes is complicated as you’ll see and requires a lot of benefit and cost analysis. So, some of the benefits that we found to using hybrid models is first and foremost for imputing diagnoses. We received a lot of feedback to this effect that this is one of the most important roles for prescription drugs and the HHS Risk Adjustment methodology and this could be due to clinician’s either failing to enter the condition on a patient’s medical record or there could be a stigma associated with certain conditions which leads providers not to record it on claims. It’s also possible that reporting systems focus on the diagnosis under current treatment and don’t require the full set of conditions associated with the patient. Also, as Krutika mentioned, there is a very real possibility that enrollees aren’t enrolled for long periods and so, if they don’t see a doctor, but they do have a prescription, this could capture that enrollee where otherwise we have not, using a diagnosis.

Another benefit would be a Severity Indicator, which we mentioned in the Payment Notice. So, while it should be, if already captured a lot of information about severity of illness, the drugs could potentially enhance that picture of a patient. So, for example, a patient could receive a first, second or third line of treatment and we could use that to indicate increasing levels of severity and drugs that are used for more serious cases or have potentially harmful side effects that you would not prescribe unless absolutely necessary.

Again, another comment in the Payment Notice was that prescription drugs have a higher quality and more timely data and we’ve received comments that prescription drug data is available more quickly than diagnosis data. It’s standardized and doesn’t vary with coding patterns and that again, it could be a more complete reflection of an enrollee’s health picture.

And lastly, a major benefit is that it would mitigate the financial disincentive to prescribe patients’ expensive medications. By compensating for a prescription drug, we would essentially incentivize plans to provide access to expensive drugs. It could also encourage change in Plan Formularies, which we see as a good thing, but this is a really challenging and careful balance that we need to consider because we want to eliminate the financial disincentives while not creating new inappropriate prescribing...
incentives. So, for the concerns that we are considering adding these, again it’s gaming and perverse incentives. We’re concerned about drug models to be susceptible to gaming and perverse incentives and changing prescribing behavior. And that is because gaming occurs when prescriptions trigger a higher payment. So, whether it’s intentionally or unintentionally when that happens, treatment decisions could be influenced by financial consideration towards drug- rather than non-drug treatment. And with less compensation from Risk Adjustment for drug utilization management, we think that plans could also loosen their effort to tightly control drug utilization.

So, in affecting the gameability of a lot of the drugs that we’ve considered, it’s not clear that many drugs are non-discretionary, and that’s because there’s a lot of disagreement around certainty of cost providers about the circumstances in which drugs should be provided. For example, for HIV AIDS, those are used prophylactically, and that, you know, we haven’t found a way to not impute a diagnosis that isn’t there in situations like that. So, even that said, we think that changes in prescribing behavior are also not necessarily a bad thing. Like I said, where we promote greater access to individuals who need treatment, we think that, that is good. We just wouldn’t want to see extreme swings based on the inclusion of drugs in the model.

Another concern is sensitivity to variations in prescription drug utilization. So, with the inclusion of drugs in the model, Risk Adjustment will reflect any factors associated with utilization. So, even if you have prescribing rates that differ by region, provider network, those would be reflected. And then plans that have more prescribing would appear sicker when maybe they aren’t necessarily sicker. So, for example, with cost-sharing differences and plans, so bronze or silver plans, they could have, they could appear to be healthier, but also not have as much access to drug utilization and have stricter drug utilization management, but not necessarily reflecting the sickness of those enrollees, and alternatively the opposite could occur with gold and platinum plans. And we also think that adding drugs to the model could increase the complexity of maintaining the model, which is already somewhat complex with our data lag and keeping up with new treatments and utilization patterns, and then the availability of outpatient drug data-only presents potential bias for severity-only models since our MarketScan data doesn’t include inpatient drug data. However, we also believe that for an imputation model, it’s very like that inpatient stays result in a diagnosis. And even for the severity model, we think that patients are very likely to refill prescription once they leave the hospital.

And lastly, multiple indications for most drugs. Few drugs are indicated for just one condition. So, some drugs can be used for high- and low- cost conditions or diagnoses that are included in our model or have been purposely excluded from our model. So, the clear one-to-one associations between those drug classes and diagnoses makes development of a hybrid model very challenging. So, in evaluating the criteria for the hybrid models and prescription drug utilization incorporation, we have five overriding criteria that we try to take into consideration. The first is clinical and face validity. The hybrid model should have clinical face validity in the relationship between the diagnosis and drugs and healthcare expenditures, as well as between drugs and their associated diagnoses.
Criterion 2 is the empirical and predictive accuracy. Drugs added to the model should increase the model’s accuracy in predicting total health expenditures, and that’s with drug and non-drug, and the predictive accuracy should be improved overall and for specific sub-populations, so, for people with different severity levels of a particular disorder. Adding drugs should also improve the completeness of the picture of an enrollee health status by identifying medical conditions that are not captured by diagnoses. So, again, this goes back to having a more complete picture of the enrollee’s health.

Criterion 3: We’d want to consider the incentives for prescription drug utilization. We’d want to incorporate drugs in a way that minimizes the incentives for inappropriate prescription drug prescribing, and we would never want drugs to be prescribed to maximize Risk Adjustment transfer. But again, we would not want to discourage very medically necessary prescribing patterns.

Criterion 4: Sensitivity to variations in prescription drug utilization. We don’t want the model to be overly sensitive to variations that aren’t due to health risks, and we’re still evaluating how exactly we can best do that, but we wouldn’t want the model to reflect factors that aren’t included in the model.

Criterion 5: Incentives for diagnosis reporting. This will come up later when we talk about the flexible model, but even with incorporating drugs in the model we would never want to discourage a full reporting of all of an enrollee’s diagnoses. So, some of this may seem repetitive, but because this is such an iterative process in decision-making, we need a set of principles and criteria for each step so that we can move forward.

And so, we’ve included here the principles for the RXC classification. Those are the prescription drug categories, and for categorizing these drugs we basically took all of the drugs from the National Drug Code, or the NDCs, and we categorized them into RXCs to determine their suitability for inclusion in the Risk Adjustment model. And we based the RXC classifications on the American Hospital Formulary Service, AHFS, and you know, we’re open to comment on whether that’s appropriate, but we used it, because it’s widely used, widely available, and most importantly, it’s frequently updated. So, after we classified all of the NDCs into RXCs based on this classification, we needed to determine which RXCs should be included in the Risk Adjustment Model. And we used a set of 10 principles, which were very similar to the principles that we laid out in the proposed 2014 Payment Notice, but of course, they’re tailored to drug categories rather than HCCs.

And the first principle is that the RXC category should be clinically meaningful. So, this essentially means that all of the codes should relate to a well-specified pharmacologic therapeutic or chemical characteristic that defines the category, and it must be sufficiently clinically specific to minimize opportunities for gaming or discretionary coding. Principle 2 is that RXCs should predict total medical and drug expenditure. So, basically, NDCs in the same RXC should be homogenous with respect to their prediction for current year cost. Principle 3 is that RXCs that will affect payments, should have adequate sample sizes to permit accurate and stable estimates of expenditures. And this is pretty self-explanatory. You know, if we have small sample sizes the data cannot reliably predict other costs of extremely rare categories.
Principle 4: In creating an individual clinical profile, hierarchies should be used to characterize the person’s illness level within each RXC where appropriate, while the effects of unrelated prescriptions accumulate. This is also similar to our HCC hierarchy in that, you know, unrelated prescriptions should increase the person’s predicted cost of care, but also similarly, we would take the most severe manifestation of a particular disease and treat it hierarchically, so that the more severe manifestations would take precedence to the less severe ones.

Principle 5: The RXCs should not reward prescription proliferation. And this is just that the RXC shouldn’t measure greater disease burden, because the same drug or closely related drugs were prescribed a lot. So, the frequency would not increase predictive cost.

Principle 6: Providers should not be penalized for prescribing additional NDCs. So, this conflicts two consequences for modeling, and the first is that no RXC should carry a negative payment weight, and that a RXC that’s higher ranked in a drug hierarchy should have a payment weight that is greater than or equal to a lower ranked RXC in the same hierarchy.

Principle 7 is similar in a way, in that the classification system should be internally consistent in how we rank, so if RXC A is greater than RXC B, and RXC B is greater than RXC C, RXC A should always be greater than RXC C, or greater meaning higher ranked.

Principle 8: The classification should assign all NDCs be an exhaustive classification. And that’s because we believe that each NDC can potentially offer some valuable information. So, by being consistent we make sure that we cover all NDCs and make sure that we assign it to the primary RXC based on how it’s administered, apply the ingredients to essentially label it, etcetera.

Principle 10 is that discretionary drug category should be excluded from payment model, which you heard. So, this is again, that we would like to prevent drugs from included in the model that are particularly susceptible to gaming or discretionary prescribing. So, in designing in those RXCs we mostly stuck to Principles 6, 7, 8 and 9, and we needed to use some clinical and statistical analysis for some of the trade-offs and then the other principles, like for example, Principle 1, which is best served by creating a large number of detailed clinical groupings that they should be clinically meaningful, is a lot of times set off by Principle 3, which is that we should have statistical validity, in those small sample sizes. So, when we need these trade-offs we use like empirical evidence, we, clinical judgement and relatedness and professional judgement on incentives, and provider responses.

So next, and after we’ve categorized all of these NDCs into RXCs and determined which RXCs can potentially be included for payment in the Risk Adjustment model, we have another set of criteria in determining which drug and diagnosis pairs we could include in the model. And there are more than 7,000 possible drug diagnoses pairs within the RXCs that we determined are eligible for payment. So, again, like these are similar criteria, but it’s for a different purpose, so we would seek to select drugs with patterns of non-discretionary prescribing. We would seek to avoid drugs where there are incentives for overprescribing. We would seek to avoid drugs where there are variations and prescribing across providers, practices and areas, and we would carefully consider the selection of high cost drugs, as these...
costs may indicate exactly the types of health risk selection that Risk Adjustment is intended to account for, but compensating these expensive drugs may also again reduce the incentive for issuers to seek greater efficiency in prescribing.

We would also seek to avoid drugs indicated for multiple diagnoses. We would seek to avoid drugs indicated for diagnoses not included in the HHS-HCC model, and we would carefully consider a selection of drugs in area exhibiting a rapid rate of technological change to exempt possible, again because it’s hard enough to keep up with the data lag for the diagnosis model. So, also trying to keep up with the new treatments coming down the pipe would prove exceptionally difficult as well as, you know, if we see new treatments being used in one year, we wouldn’t want to base future years off of a one-year anomaly.

So, in some of the considerations for these drug diagnosis pairs and how to evaluate our models and create them, we looked at empirical considerations, because initially we’d only considered using drugs as a proxy for missing diagnoses or in a, you know, severity. But then as we needed more analysis, we realized that there could be iterations to those two separate ideas. So, we initially looked at the counts and the expenditures of RXCs to get a general idea of the composition of these drug groupings. And we found that this was useful to get a sense of how common the RXC was, and how enrollee expenditures vary across these RXCs. Particularly, it was useful to see whether the RXC predicts high expenditures because of medical treatments or because of the drugs themselves. We also looked at drug diagnosis codes since each drug that enters the hybrid model must be assigned to particular diagnosis, we examined how closely related these are to each HCC. We did some statistical analysis and created rankings of the most closely associated drug diagnosis pairs, and this was used to propose pairs that may have otherwise been overlooked. We supplied regressions of drug classes and drug diagnoses interactions, and that was useful to determine which variables can add the most incremental predictive power to the model. So, we estimated a series of models, starting with the baseline diagnosis model, and then we added one additional regressor each round and each other regressor was selected on the criterion of achieving the maximum increase in the R-squared to the model.

And so doing that either, by adding them either individually or collectively, we found that there were modest improvements to the R-squared in these hybrid models. However, like I stated earlier, most of the incremental predictive power added was from a very select few diagnosis pair. We consulted with doctors and pharmacists who are familiar with risk modeling when discussing these proposed pairs to them. We also wanted their clinical expertise on treatments and protocols in evaluating these results. We also just got the potential for discretionary variable or gameable drugs, so some of the drugs were rejected based on this criteria. However, that said, that’s not to say that the drug pairs that we included in these hybrid models are completely non-discretionary. For example, I think we mentioned in the White Paper. If Risk Adjustment was to pay on anti-diabetic drugs, it’s very possible that providers would be less likely to prescribe a treatment of diet and exercise rather than the drug if there’s a financial incentive to do so.
And so besides the empirical and clinical criteria, we also considered a couple of other topics that we haven't necessarily incorporated yet, but that we seek the public input on. One is imposing model restrictions based on day of supply or number of prescriptions. We consider others’ evidence of prolonged usage to trigger a drug indication, or using a minimum day of supply to distinguish severely ill patients from those that are less severe. We didn’t include those, but again, we, we welcome your input. We also considered sub-diving or splitting the RXCs or including individual drugs. But if we were to do this, you know, the drug diagnosis relationship we more narrowly defined and have greater clinical precision, but there would be much greater complexity and smaller sample sizes, so we would need to consider that if we were to do something of this nature.

We also considered the imputation only versus imputation severity relationship. In theory, we could select some drug costs if you use only an imputing diagnosis, while other drug process could be used to impute a diagnosis and predictive severity. However, we found that most or all of the drug costs that can be used to impute a diagnosis can also give us some level of detail about severity. So, instead we only make a distinction between drug costs as used for imputation and severity versus severity only. We also consider the prophylactic use of drugs. As I mentioned earlier, with HIV anti-retrovirals, we know of no straightforward solution to falsely impute a diagnosis in these kinds of situations so we welcome feedback on that. For multiple indications of drugs, we do have some concerns that drug costs are often indicated for more than one condition. For example, we have DMARD in our imputation based models and our severity models, and these are commonly used for rheumatoid arthritis, but less commonly used for inflammatory bowel disease. So, we actually use DMARDs in these models twice; we use it as an imputation type of drug cost for rheumatoid arthritis, but where there’s a presence of inflammatory bowel disease, we use it as an indicator of severity since it’s more likely that someone with inflammatory bowel disease who is being prescribed DMARD has a very severe case of that condition.

Next we’ve provided the drug table. Unfortunately, it’s on two slides, but it has the approximately 13 drug diagnoses pairs we are initially illustrating for inclusion in the four hybrid models. The first 10 RXCs listed here are included in all four models, and then the last three are only included in the severity-only model. And included here you included the positive predictive value, which we consider to be the proportion of people with the RXC who are also observed to have the HCC. And we consider this proportion one measure of strength between the association of the RXC and the HCC. And again, as I noted before, you’ll see DMARDs on here twice; on the imputation model variance side, and then on the severity only model side.

Okay, moving on to the next one. For the imputation-only model, so this model presumes that any individual with a particular health condition should be predicted to have the same incremental costs as someone with the prescription drug-only for both the prescription drug and the diagnosis. So, this model has one level of predicted incremental expenditure field. You’ll see that each if you look in the White Paper, every predicted cost should be the same for each drug diagnosis pair regardless of how they’re identified. This is the most constrained model, the most conservative. For example, the predicted cost for someone with chronic Hepatitis in the baseline diagnosis model is around $16,000.
However, when we impute that in the imputation-only model by Hepatitis B antiviral is present, then the predicted incremental expenditures are around $25,000, regardless of how it’s identified.

So, this is a pretty large increase in the size of the predicted expenditures compared to the baseline, but as you’ll see with the other model, it’s a bit constrained. The R-squared is lower than the baseline model and the main reason is probably because it’s constrained. It’s likely because we’re constraining individuals with the diagnosis who are relatively cheap compared to individuals taking the drug who are relatively expensive. Next yeah.

So, the next model under consideration is the RX-dominant model or drug-dominant model. This is the less restrictive version of the hybrid model or imputation-only model, and estimates a predictive incremental expenditure for people who are identified by their HCC only, and then the same other predictive incremental expenditure for people identified by the RXC-only or the RXC and the HCC. And we call this drug dominant because the cost prediction is the same when drug use is present, regardless of whether there’s an underlying diagnosis. So, this model has the basic assumption that people taking the drug are assumed to be more ill, and that assumption carries through, as you can see with the chronic Hepatitis example. So, they’re predicted to have around $2,500 of expenditures with just the HCC but when the Hepatitis C antiviral is described, even without the diagnosis, that predicted incremental cost is up to $110,000, so it’s a very large gap, especially compared to the imputation only model and, you know, rightfully so if the drug is flagged it’s accurately reflecting the cost of those drugs. This does have a higher R-squared and this is likely due to the drug-dominant model predicting high expenditures for these expensive drugs and it’s not constraining those expenditures for people with the HCC as in the imputation model. In this model, for many of these pairs, there is a large difference between HCC only, predicting incremental expenditure and then the drug-only or drug and HCC predicted expenditure. And I think that’s important because it is showing that this model is demonstrating important distinctions in expenditures that are not shown in the imputation-only model.

The next model is the flexible hybrid model and unlike imputation or the drug dominant model, this actually has three different levels of predicted incremental expenditures, so that’s HCC-only expenditures, RXC-only, and then RXC and HCC. Typically in this model, you’ll see that costs increase from HCC-only through RXC-only and then RXC and HCC but we have noticed that in this model there is the possibility that someone’s predicted expenditures could actually decrease from RXC-only to RXC and HCC, and that violates one of the basic principles of Risk Adjustment modeling in that we would not want to discourage full diagnosis coding, and that occurs here for chronic Hepatitis and multiple sclerosis. So, if we were to implement this model we would first need to come up with some way of constraining these coefficients so that we don’t violate one of our most basic principles. Again, the R-squared is improved, demonstrating that this model catches some significant severity distinctions.

The last model under consideration is the severity-only hybrid model and this model again has only two levels of predicted expenditures. That is the HCC-only model; it predicts zero incremental expenditures for the drug-only and then another level for people with the drug and the diagnosis. And so this has no imputation qualities. You always need to be quoted with a diagnosis to receive credit in this model and
it’s the only one of the models to include the three drug diagnosis pairs at the bottom of the drug
diagnosis pair slide, and those are high severity diuretics for congestive heart failure; ammonia
detoxicants and cirrhosis of the liver; and DMARDS for inflammatory bowel disease, as I mentioned
earlier. Those pairs aren’t included because they don’t impute a diagnosis since they’re used for other
conditions. As I mentioned, with the inflammatory bowel disease example, the R-squared here isn’t
quite as great for this model but it’s also more constrained than the drug dominant and the flexible
models in that you still need the diagnosis in order to indicate severity with the drugs. And so, you
know, these models are very fluid. We’re changing them weekly and improving upon them and looking
at them in different ways but some of our initial evaluation of these model options are that models that
utilize drugs indicate severity have a greater degree of clinical-based validity.

Use of the drug class typically will contain some sort of information about severity so the drug-dominant
model may also have greater clinical face validity than the flexible model because it’s not necessarily
clear why the presence of a prescription drug then doesn’t affect incremental predicted cost when the
drug class imputes the diagnosis.

The models have broadly similar predictive accuracy but, of course, some constrained model has the
least predictive accuracy and that’s the imputational model. For some groups that the hybrid models are
adjusting for, for example, individuals utilizing Hep C treatments, the gains in predictive accuracy for
them are substantial in these hybrid models. But, models that add the most predictive accuracy predict
high expenditures from enrollees using expensive drug classes. And so those would be the drug-
dominant and the flexible models where the imputation severity-only models seem to create the least
strong incentives to overprescribe.

So, moving ahead, CMS would assess several factors. We need to look at the operational costs, both for
issuers and for operations; we need to continue to evaluate which drug classes, drug diagnosis pairs,
and which model should be incorporated into our Risk Adjustment model; and along those lines, you
know, we need to consider whether we want to take more of an imputation approach or focus on
severity or combine both. Regardless, we think that starting with a relatively small number of drug
classes is where we need to be.

And so the research that we’ve done for this so far has been conducted on the adult models and the
adult sample and we still have yet to evaluate what the effects will be on the child model given the small
sample size issue and whether there is a role for drug information in the infant model, so that is ongoing
and will be a major part of this decision.

So, moving on to high risk enrollee pooling and HHS Risk Adjustment, unlike the prescription drug work
which we’ve been working on for quite some time now, it’s a relatively new concept that we have
addressed in the White Paper, high risk enrollee pooling came about because traditional Risk
Adjustment doesn’t predict for high, very high cost enrollees or outliers. Currently when we include
outliers and high cost enrollees in the recalibration sample, what it does is effectively slightly inflate the
predicted cost for everyone else with that condition but without actually providing the necessary
compensation for an issuer who has an outlier to cover the cost of that enrollee. So we’re considering
recalibrating the models to set expenditures at a certain threshold, above which in practice the issuer could be liable for a percentage of those costs. We think that this would improve the predictive cost and recalibration immensely but also would provide a pool of a percentage of high cost claims that we could uniformly allocate across Risk Adjustment issuers to protect them from the high cost enrollees.

Some of the considerations that we are, that are still under consideration, next slide, is whether or not we should implement this nationally, at a state level. Even if we implement it nationally, whether it should be by market or just completely across the United States and even allocation and pooling. So, we think that a national pool probably make more sense because it will provide the best protection, especially for small issuers within a state and it will probably result in the lowest allocation adjustment to fund the pool but we still need to discuss that and we welcome feedback on what stakeholders’ preference is and as well as the percent of cost to be reimbursed from the pool above the constrained by the threshold to be determined. I think in the paper we suggested maybe a threshold of $1-million dollars but we welcome comment on that and what percent of cost should be reimbursed above that. We also have not yet determined the matter in which the pool would be allocated across all Risk Adjustment issuers, whether it’s through per member per month or percent of premium adjustment or charged to all issuers, that’s also under discussion.

So, moving on to a discussion of concurrent and perspective Risk Adjustment models. We received some feedback and questioning of why HHS uses the concurrent model versus a prospective model like Medicare does. And, so, we just thought it would be helpful to touch lightly on what they are. Concurrent models use current year information to predict current year costs. They better explain variation in current costs, which reduces unsystematic risk. And, so, our use of a concurrent model is helpful because we did not have prior year information and we think that a concurrent model supports the intent of the ACA, which is choice, competition, growth, moving in and out of plans, which is better reflected in the concurrent model than a prospective model.

A prospective model is typically favored because it tends to emphasize ongoing chronic conditions as opposed to random current year costs and it predicts on the basis of prior encounters for conditions for the coming year. However, there are significant challenges for HHS to use a prospective model. As I mentioned, we have a lack of previous year information, especially for the initial year and that, you know, people mover in and out of these plans. We want to promote that, but further our distributed data approach makes that not possible currently and we would need, you know, an enrollee identifier, which we do not have. So, a prospective model really isn’t currently operationally feasible.

And then for 2018, lastly, recalibration, we are considering an accelerated Payment Notice schedule, which in this, in that case we would not have access to 2015 MarketScan data which becomes available in December. And we saw comment on both the data lag and the Payment Notice and Congress was very supportive of reducing it, more heavily weight most current year of data and MarketScan. Finalizing our recalibration methodology in the Payment Notice and then issuing coefficients in like sub-regulatory guidance or elsewhere. And so, we are considering for 2018 recalibration given that we will not have 2015 data using only one year of data, which would be the 2014 MarketScan data. However, again, as
Krutika mentioned, we need to evaluate the small sample sizes, you know, we need to look at that in concert with all of these other proposed changes. You know, if you were going to implement separate cross year model, combined with new drug diagnosis codes and one year of data. We have a lot to look at, not in isolation as we have been and I think the interaction of all of these potential changes will determine which funds we go forward with and how they affect the model as a whole.

So, with that, I think we’ll take questions down here.

If you have any questions, we’ll collect those as the panelists are assembling at the table.

All right, before we get started with questions, I was an exceptionally poor host earlier. I want to remediate that now. First of all, I want to thank ARDX, who you see coordinating the show today. They are our stakeholder engagement contractor. They do a variety of training programs that probably all of you have participated in. It’s across all of our enterprise. They deal with the Plan Management, E&E, and Financial Management aspects of Marketplaces as well as like all of our Financial Management programs. And they’re the ones that made getting into this building today so smooth. So, thank you to ARDX.

We also have up here with us representatives of RTI and we’ll have different representatives of RTI at different sessions. RTI does all of our modeling. So, they create the models that we’ve been talking about today, do all the predictive ratios and all that so that we understand what’s going on. They’ve also done a lot of simulations as we look at suggestions on the transfer formula. So, we’ll have different folks that have done different aspects of work joining us for the Q&A sessions for technical aspects of that.

And, last but certainly not least, we have a couple of folks, I think from the Agency for Healthcare Research and Quality. Patricia Kenan I know I saw, and I’m not sure if Paul is here as well. Patricia was one of the original leaders of the Risk Adjustment development for us in CCIIO. Now, she’s over doing research side of things, all those nice charts and graphs that I was able to show you, that’s based on research they’re doing for us to tell us how Risk Adjustment works. It’s a great partnership because have folks that actually understand what we’re trying to get done, and then they’re looking at it and evaluating it and saying, “Is it working the way we want it to work?” So, thanks to all of these folks for a job well done and continue job well done in the future hopefully.

All right, so before we get into the questions, just a quick reminder that you can also submit questions by selecting the feedback button within your Simply Attend app. And our first question here is from Ashley Smithson of Blue Cross/Blue Shield Alabama. That question is: Could you model the partial year enrollment factors to be a multiplier instead of an additive factor:

Sure, so...

...interactive partial enrollment indicator. So, there’s several ways that you can look at this.

Thank you. And our next question comes from the John Berco of Covered California. We actually have two questions here. So, the first question is: To minimize gaming with the pharmacy indicator, is CMS
considering a long minimum drug usage such as the 181 days of daily dosage used by Germany and the Netherlands?

I think and I’ll defer to Greg as well; I think we have looked at certain places where length of use indicates kind of different conditions and different treatment patterns and I think that gets to some of the comments you had on the prophylactic use and such, but I don’t think we’ve come up with a firm decision on that, but I know we’ve considered it in some things and maybe Greg can talk about it.

Yeah, that’s something we had, no, it’s an issue and plan to look at it further, but we’re relatively early in our consideration of that issue and so I think that’s an important thing to raise and John, you know, if you have any particular thoughts on that, we’d like to hear them.

And this is where the public comments will be very useful, so if folks think in considering prescription drugs, that there’s something beyond just the presence of the drug that should be considered in modeling it, this is the time-period to look at that and tell us what you see and tell us what we should be considering as we work on these models.

Thank you for your responses. The next question here is: For partial year enrollee modeling, would MarketScan data be relevant? This is because it is ESI data from employees and dependents with stable 12-month employment. New hires and TERMs might be quite different from the ACA individual market partial year enrollees. Do we need to wait for actual ACA data?

Right, and this is what I was mentioning as part of my presentation that all the data we showed up there for the partial year enrollees is based on MarketScan data, which, like the commenter said, it’s not based on individual and small group market. And so later in the, later today we will talk about some of the other data considerations and potentially getting into enrollee data and using that to create partial enrollment adjustments.

But, I’ll make a couple of other points. Feedback that we’ve already gotten on the model in terms of how well it predicts, areas where it doesn’t predict as well, areas where it predicts really well are consistent with kind of what we see in our own observations. So, no model is perfect, right. When you do these Risk Adjustment models and then you slice and dice the population, you can slice it so many different ways. You will find some populations where you have a better predictive value and other populations where you do not, especially the more finely you slice it. And, this is what actuaries out there are doing and they are like, you know, “Forty year old diabetics, you seem like to be under-paying them, but, you know, you got like, you know, 50-year-olds with heart attacks and you’re over-paying them.” You can sit there and slice it many different ways, but we are finding that folks empirical results with their claims experience and people coming in and saying, “Our MLR is kind of going south here and we’re making money over here,” seems to be consistent with our modeling results. Again, now we’re talking magnitude. Now, the other thing I would say is we have looked at this enough to understand what we believe is going on when we see differentials and partial year enrollment.
and different conditions either seeming to be consistent throughout the year in costs, so it
doesn’t matter if you’re a partial year, your per month cost seems to be about the same.
Tending to be chronic conditions with a regular treatment pattern may be the same prescription
drug that you’re taking all year long. So, things are lining up with what we would expect based
on the types of conditions. So, clinically and kind of what we think is going to go on, in a partial
year model, we’re seeing the kind of effects we would expect to see. Does that mean it is exactly
correct for this population, no, but again you want to be directionally correct and start
narrowing down on kind of getting to a, you know, this order of magnitude to getting a more
precise order of magnitude. Everything is about increasing precision through time, so lack of a
perfect data set doesn’t mean we can’t get better, and then we look at advances through time
that will make us even more accurate. So, this is a step-wise way of getting to a better answer.

And we were informed that our phone participants had trouble hearing the answer for Ashley
Smithson’s question, so, we’re going to repeat it if you could. That question was: Could you model the
partial year enrollment factors to be a multiplier instead of an additive factor?

So, in the presentation that CCIIO gave today, they talked about three methodologies for
modeling partial enrollment. You can think of kind of two extremes. One is just incorporating a
dummy variable into the model based on enrollment duration. The other is to calibrate a
separate models for a partial and full enrollment and other things equal that will get not only
the partial enrollment, the duration groups right, but by HCC and HHS, although as Krutika said
there are issues with that, too, in terms of stability of the statistical modeling. There are
intermediate types of methodologies that one can use. You can incorporate interactions into a
single model where you would look at the HCCs that are most important for partial enrollees
and then you can interact those HCCs with the partial enrollment binary indicators and then, in
our work with CCIIO and with Medicare, there’s also the option of, like the caller is saying that
you can identify an enrollment duration adjuster and basically use it outside the models of
multiplicative adjustment. So, those are all, you know, possible adjustments.

Thank you very much for that response. Our next question is from Kendall Johnson of Blue Cross/Blue
Shield. That question is: With respect to the enrollment duration for partial year adjustment, did you
consider other groupings of months that are more heavily weighted toward the short duration
enrollments that are typical of SEP enrollees such as one-to-two months, three-to-six and seven-to-12
months?

So, we group them generally by thirds of the years, but we haven’t tried weighting earlier more
heavily if that’s what the commenter is mentioning and that is something we could consider in
our future analysis.

Again, remembering that, as mentioned, I think, you know, with this data set, particularly there
are very small sample sizes on the partial years to begin with, so the more we start slicing down
to fewer and fewer months, the smaller those sample sizes get and the more error you get in
the prediction.
Thank you for the response. Our next question is from Derek Skoog of Crestwater Healthkeepers. The question is: What were the select few drugs or diagnosis that had high predictive power?

I think the Hepatitis C drugs in particular added a lot of the incremental increase in our square that we saw, so that’s one that stands out the most. So, of course, there’re quite expensive drugs.

Thank you for that response. Our next question is from Barbara Holmes of Community First Health Plan. That question is: We cater to the silver level Medicare plan so we tend to attract them with our marketing and we noticed that many of our members lost their subsidies temporarily which eventually meant the loss of health insurance because they couldn’t afford the higher premium, so we have a lot more partial enrollments currently in the silver subsidized plan. So, the question is: How is this accounted for in the 2015 model and what is the plan to account for it going forward?

So, I think the whole point of this presentation is that, you know, we started out with a model that was built on full year enrollment. We have been examining partial year. We have not yet completed using the modeling of partial year and determining modeling approach for partial year. We announced the methodology for 2015 that does not have partial year. So, for 2015 it remains the same. We will work through the modeling as quickly as we can and try to get something implemented here. I think the struggle here is finding the right way to incorporate this.

Thank you. Our next question is from Kevin Delaney of Kaiser Permanente Ohio. The question is: One of the principles of the diagnosis phase HCC model is to encourage more complete and accurate documentation and coding of conditions within the proposal to incorporate prescription data to impute conditions. Couldn’t this violate that principle and release plans of the responsibility to complete an accurate code condition?

Well, so, I think first of all it depends on what kind of model you implement. So, when you look at these hybrid models, we’re using both conditions and prescription drugs. There’re only certain conditions that we’re even thinking about prescription drugs being an imputation for. So, I don’t honestly think that this is going to be a substitute. In a hybrid model, in some of these cases, it’s going to do better to have the condition and the drug and in some cases some that we’re looking at, on the severity only, you need the condition to even get the drug. So, it depends on what you implement with a prescription drug model as to whether or not it somewhat relieves the need to get conditions. But, in general, you’re going to need to be able to document health conditions because we’re not looking at that large a set of drugs. So, this all starts with diagnosis-based Risk Adjustment with a slight addition of drugs and I don’t think health plans sit there and get that granular on documenting conditions. If you’re trying to document conditions, you’re trying to document people’s health status and get a full roster of diagnosis codes and you’re not going to say, “Well, ignore diabetes because I probably could get that with insulin..” If you know you got diabetes, you’re going to submit it. So, I honestly don’t
Risk Adjustment
HHS-Operated Risk Adjustment Methodology Meeting
Webinar - Remote Access
March 31, 2016, 9:00 a.m. – 4:30 p.m. ET

see with the limited use of prescription drugs and the hybrid approach to prescription drugs that there’s much that would discourage documenting conditions.

Thank you for that response. Our next question is from Sharon Easterlin of Recovery Analytics, LLC. And the question is: Ambulatory documentation and coding is undergoing increased focus as physicians have performed code assignments in the clinical setting. Guidelines are not utilized impacting patients not seen in the hospital and providers are changing that. How is that being considered?

Could you repeat the question?

Was this a question that was submitted here on a card or did it come in through the chat?

This actually came in through WebEx.

Okay, we may need more clarification on the question.

Okay. So, we’ll move on to our next question at this time, then, and that comes from Gabriel McGlamery of Florida Blue. The question is: If CMS uses pharmacy data, shouldn’t they generate a prescription to HCC crosswalk retrospectively?

I’m not sure what the retrospective component of it is. What we are proposing is different ways of creating HCCs with prescription drugs. We are not proposing that we would use this in a model that already exists. So, if the question is would you change how you generate HCCs for 2015 in a model that was calibrated without prescription drugs, the answer is no, because when you add prescription drugs to the model, the variables change. So, we will have a different model when we add drugs and we will have a calibration that then includes the drugs and we will pay consistent with our calibration and that’s consistent with everything we do in Risk Adjustment throughout all of our programs.

Thank you very much for the response. Our next question here is from Tom Hoffman of Prime Therapeutics. Please provide a specific example of gaming the system risk associated with adding prescription drug claims to the Risk Adjustment methodology. Are there any drug costs that are so low-cost while also impacting the risk score enough that an issuer might be motivated to increase utilization?

I think we gave a presentation, Kelly gave the example of diabetes and treatments that, you know, may be on the margin between diet and exercise and drug treatments which are not all, at least they think, expensive and so it, you know, financial considerations could make a difference there at the margin and most of these cases there will be some marginal patients where there may be a borderline decision between whether the drug is used or not, it’s up to the patient, so for some patients it may be very clearly indicated to have the drug and some very clearly indicated not to get the drug treatment, but there are often some patients where there may be a gray area and those are the types of situations that we would be concerned about.
Thank you for that response. Our next question is from Catherine Liang from Partners Health Care. And, that question is: Could you elaborate on the high-risk enrollee pooling idea and potential sources of funding for the pool?

There’s a number of ways that this could be funded. I mean conceptually what you do is you are figuring out anybody’s cost in this high-risk pool and then allocating those so, you know, conceptually you could know the total cost. You could add up nationally the total cost and pools on a national basis. If it’s on a state basis, you’d add up all the costs in the pool and then you would allocate a share based on either a percent of premium, percent of enrollment, et cetera, et cetera in the overall market that you’re doing this in and then you basically can see, you’ll know what each issuer has paid already, what kind of costs they’ve incurred in this and you will know what their theoretical costs should be and it’s just a simple subtraction problem at that point in time. So you either have suffered a lot of high cost claims and your percent of enrollments are very small, or you suffered no high cost claims, you had a large percent of enrollment, you’d end up paying in, then, that’s conceptually at least how it works.

Thank you for that response. Our next question comes from Jonathan McCloud of Blue Cross/Blue Shield of Michigan. The question is: Can you clarify the difference in infant male to infant female?

So, in the infant model the model assumes that you’re females, and so for male there’s an additive coefficient ticket added.

Yeah. just to elaborate, the males, both the males and the females, are included in calibrating the model, it’s just as critique as there’s males are controlling for the other factors in the model are more expensive, so it’s an additional adjustment for them, but it’s not, it’s not, you know, only calibrated on females, it’s both genders together.

Thank you for your responses. The next question is from Kevin Delaney, Kaiser Permanente.

The question is: How much of the increase in predictive power of the prescription models is due to the addition of Hepatitis C drug?

Yeah we haven’t, we haven’t run models with just adding one drug pair at a time, but just kind of anecdotally from, from what we have seen, you do get sort of the, the largest impact from the Hep C drug and HHC pairing. I think it’s still on the order of about half a percentage point, but I would have to check that to be sure, but it’s, it’s kind of in, in that magnitude.

Thank you for that response. The next question comes from Myja Peterson of New Mexico Health Connection. The question is: How will, or will the model be changed to address potential avoidance by health plan of healthy non-HCC enrollees?

Well, the model currently is predicting costs for all enrollees. So in fact, most of what we’ve seen in terms of people criticizing the model is that it under-predicts the cost of the no-HCCs, but, so, I’m not clear what the question is intending to answer, but the model is built on based
demographic factors. Of course, most of the population is a no HCC population in this population. When you deal with Medicare, those would have a lot of time of experience with this. We were dealing with 55, 60 percent morbidity rates and, and we don’t have anything like that in this model. So, the model is primarily predicting the cost of the healthy people, and predicting the cost of the much fewer people that have diseases on top of that.

Thank you for that response. The next question is from Gaurish Chandrashekhar of Harvard Pilgrim Health Care. The question is: As part of duration factor wait, was there consideration of plan year versus calendar year deductible plan, and if not, can this be considered?

Well, so it’s an interesting concept. The simple answer is no, we have not done that yet. I think the more complicated answer is to do that would require using EDGE data, and to use EDGE data we even have to change EDGE data because we would have to, now I’m assuming this was a small group question, because in the individual market, plan year should be what it is. I’m not sure if it really matters anyway, because a person’s benefit starts when they enroll. If they have a short year, I’m not sure if it matters that they enroll starting in July. And you know, the plan benefit year, the renewal cycle would be January, because they’re in the individual market, so they get cycled over, they’d be on a short year, but assuming they stayed in they would be in through June. I’m not even sure that it really matters, but to do that we’d have to collect some other variable that not only totals what the person’s enrollment date was, but what the plan year was, and I’m still not convinced that, that is any different. I think everybody has their own plan year based on their dates of enrollment to begin with.

Thank you. The next question is from Aaron Solomon of Blue Cross, Blue Shield of California. This question is on high risk pooling: How will the large claims be validated? Carriers may renegotiate or incentivize large claims from providers in exchange for reductions elsewhere. The gaming of this risk pool likely outweighs any predicted gains additionally, while carriers can avoid sick members, it is difficult to see how they could avoid members with catastrophic claims. Does this problem currently exist?

Well, so I think that question points out both sides of this problem. One is that, there’s one argument that the catastrophic costs are unavoidable, so this is not issuer behavior, this is actual, when you get one of these high cost cases you’re stuck with a high cost case and you’re stuck with paying this cost. And with the no annual limits, no lifetime limits, these costs can be exceptionally high compared to what carriers have experienced in the past. On the flip side, some gaming could occur. I’m not sure to what degree you could do this, it would be something to certainly look at to see if there was any evidence of what was happening. I think part of the disincentive for that is having some kind of coinsurance rate at which you calculate this, so you don’t pay 100% of costs above a certain level. I think that would potentially incentivize some really bad behavior. I think there has to be skin in the game, or at any level this is a risk-based program. So, I think as long as you retain sufficient risk, it would be hard to negotiate a contract with a specific provider that could get you there on your high cost claim. It would be a very
complicated negotiation and you’d be having to reprice things retrospectively, again, a lot of these are unpredictable cases, they can happen to anybody. Certainly, maybe somebody that had a large enough risk pool could figure out an elaborate game to plan is something we should certainly be concerned about. But you’d have to be pretty elaborate and I think you’d have to be pretty large to pull that off.

Thank you for your response. Our next question is from Lisa DiSalvo of Altegra Health. The question is: Could you elaborate on the high-risk enrollee pooling idea and potential sources of funding for the pool?

So, I think, I think we answered that already, but thank you.

So, she did have a follow-up then on: Can you elaborate on why you’re considering a separate funding pool for Rx HCCs rather than incorporating into the PLRS?

Oh, I don’t think that’s what we’re proposing, so this would be part of the risk model. All we’re doing is breaking them out and showing the specifics that go on with these HCCs, but these would be inherently, and that’s why we talk about hybrid models. I mean, very clearly in the hybrid model, it’s all a part of one model, but we’ll develop one PLRS. And this gets back to Gabriel’s question about, you know, retroactively applying this. This is why you can’t retroactively apply this, it’s because you’re, you’re now coming up with a new model, that’s comprehensive in scope, including both diagnoses codes and prescription drugs. It’s a new way of looking at the model.

Thank you for the response. The next question is from Eric Hedrick. And the question is: With respect to slide 57 of the presentation, did you hold the HCC factors constant when calculating the factors associated with the enrollment duration?

So, they weren’t held constant, but enrollment duration factors were added to the model. Oh, excuse me. So, the HCC factors would not have been held constant in this case. We added dummy variables essentially on top of the HCC variables.

But this is, this is with each one of these things we proposed, these are complicated modeling efforts. But every time we’re proposing a change to the model, this is why you can’t just say, “Oh, well just add that factor into your existing model,” that’s not, what we really do is we relook at the data. So, when you’re attributing costs to something, those costs were already there in the model, so every one of these people’s costs has been accounted for without partial year, without prescription drugs, what have you. So, they exist somewhere. Once you attribute these costs to some variable that did not previously exist, some other variable gave up some cost component for it to get there, because the model in the end is going to account for the total cost of the population. So, every time we propose one of these changes, we’re not saying we’ll hold that model constant. No, it’s remodeled with these new variables in it and that will by definition change the value of the underlying HCCs that existed in the model without these new variants.
Transcript
Risk Adjustment
HHS-Operated Risk Adjustment Methodology Meeting
Webinar - Remote Access
March 31, 2016, 9:00 a.m. – 4:30 p.m. ET

Thank you for that response. Our next question is from Nathan Graff of WellMark Blue Cross Blue Shield and the question is for the conquest of a nationwide pool for high risk enrollees to give issuers a better idea of the implication, how many enrollees in 2014 had more than $1 million in claims or how much PMPM claim amounts nationally represent all claim amounts greater than 1 million?

...slightly and it’s around a percent or two. It’s not more than that in MarketScan.

Yeah and again, we don’t have a specific report I think. We do have EDGE server data. We could clear the EDGE server data and we would have to write some very specific reporting once we come down on parameters and we’d probably would try to get some intelligence on the EDGE server. I’m looking at Sri now and planning just more work for Sri, but one of our programmers; that’s another guess we have, is we have our folks from, a couple folks from Accenture here as well, who program our system and later on today you’ll get to meet Sri, who has been heroic in his efforts to get the EDGE server running, but we would essentially have to write a query to the EDGE server that would have a very specific parameter on it and look for those costs. It is a knowable fact, it is not a known fact right now because we don’t; again we only get summarized reports and this was not a relevant questions to us. The only thing we know is how many people went above our $250,000 reinsurance cap. That’s the only reporting we have right now, but we could modify our reports such that we could gain this data.

Thank you for that response. And the next question here is from Seth Quiggle of Health Partners and the question is with regard to the pooling concept, there are significant variations in provider fee levels across the country, sometimes even within a city. This could inappropriately impact the pooling transfers. So, are there any plans to address this?

Well I think that question is probably the argument for a national pool. If we were to do it at state level, situations like that, would more adversely impact the other Risk Adjustment issuers in that state, market or risk pool.

Thank you for your response. The next question is from Michael Norton of Health Connector and the question is how would the high-risk enrollee pool interact with reinsurance purchased by the issuer?

Well, I think this would be an issuer; this would have to be an issuer consideration. So, I think first we’d have to propose the parameters for such a pool. There have been suggestions when talking about this, the parameters be set that it’s above the outer limits of commercial reinsurance, which is one way to do that. Another way is it might fall somewhere within the current one which would maybe make people think, rethink how much they’re reinsuring for. So, there are a number of different ways this could happen depending on where the parameters are set, but once the parameters are known, I would think it would influence a carrier’s commercial reinsurance purchasing decisions.
Thank you for the response and the next question comes from Mary Boatright of Advantage Healthcare Management. The question is: With regard to the high risk pooling. Have you seen other diagnoses or conditions that might be impacted besides Hepatitis C? Specifically, what about members with EFRD?

I think there’s, so Hepatitis C was mentioned as one of the, kind of leading model improvement areas for the prescription drug model in terms of severity indicator that in the hybrid model, I don’t think it was singled out as, like a single high cost driver of high risk pools and we were not proposing the presence of certain conditions and not others for the high risk pool. We were treating that as there can be any of a number of conditions that just have some kind of unpredictable launch point for a long hospital stay and exceptionally high costs and we’re not trying to associate those things. This again, at a high enough level, these are just the unpredictable, uncontrollable cases.

Yeah and often times, these enrollees have multiple conditions, not a single condition.

Thank you for the response. The next question is from Dean Ratzlaff from Optima Health and the question is: Has the assumption that RXCs impute an HCC been empirically tested? So, for example, the White Paper gives count where an RXC imputes an HCC. Have the medical records for such members been reviewed to substantiate or negate the imputed HCC? There’s concern that the relationship between a prescription and a particular medical condition may not be strong enough to warrant an imputation-type model.

Well so, the data that we use is MarketScan. It’s de-identified. It is a commercial data set. We would have no idea and should have no idea who the members are that have these conditions or these prescriptions. So, the simple answer is no, there is no way we’re looking at a medical record to verify these conditions and I think this is also part of the cost-sharing notes we have on any of these. So, I’ll let Greg talk about that, but I think we’ve cautioned that, you know that might not always be true; that you’re use of a specific drug is predicting a specific condition.

That’s correct, that we haven’t attempted to validate with medical records. One thing we have looked at is people that are in two consecutive years or so MarketScan; when say they have the drug without associated diagnosis that we’re attempting to impute. We’ve looked at those people the following year and a number of those people do have the diagnosis written down in the following year. I forget the exact percentage and it varies from drug to drug, but that’s what we tried to look at that association a little bit. And we’ve looked at the association of drugs and diagnoses in the data we have, but we don’t have a gold standard of a medical record or some kind of external definitive diagnosis or whether someone does or doesn’t have a diagnosis. So, no, we don’t have an old standard comparison there.

Thank you for the response. The next question is again from Seth Quiggle of Health Partners and this is a follow-up for you on pooling. So, will there be any opportunity to reprice the high cost claimants on a standardized fee schedule to eliminate the impact of high fee levels at certain providers?
Is this another one from the web or is this somebody that’s here?

It’s from the web.

All right, I think we need some more clarification on that one. If this is another kind of version of the gaming question, I think with all of these approaches, with any change we make in the model, there are reasons that these things are not traditionally included in models is the gaming side of things is definitely worried about, but I’m not clear if that, if they’re talking about overall repricing to push things, to push costs towards the high risk people and away from the lower risk people or whether this is repricing to eliminate high risk people. It wasn’t clear, but if it’s a variation on that theme, it’s the same answer as before. I think you have to definitely have coinsurance rates. There has to be skin in the game. We have to figure out what that needs to be to protect against this.

Thank you for your response and at this time we’ve covered all the questions that we’ve received for this session. We did receive some questions on recalibration, which we will handle in the next session upcoming.