inovalon™
healthcare empowered

THE IMPACT OF DUAL ELIGIBLE POPULATIONS ON CMS FIVE-STAR QUALITY MEASURES: CONTROLLING FOR PLAN (PBP) CHARACTERISTICS

October 20th 2014

Submitted by Inovalon on behalf of the Duals Data Collaboration Advisory Panel:

- **Cigna-HealthSpring**—Dr. Paige Huber Reichert, Senior Medical Director of Quality,
- **Wellcare**—Elizabeth Goodman, Vice President of Public Policy and Government Affairs
- **Healthfirst**—Joyce Chan, Assistant Vice President of Clinical Performance
- **Gateway Health**—Dr. Austin Ifedirah, Vice President, Medicare & Strategic Planning
- **BCBS Minnesota & Blue Plus**—Dr. Dave Pautz, Senior Medical Director, Government Programs
- **Health Care Services Corporation (HCSC)**—Susan Sommers, Divisional Vice President, Medicaid Programs
- **SNP Alliance**—Richard Bringewatt, President, National Health Policy Group and President
- **Medicaid Health Plans of America (MHPA)**—Jeffrey M Myers, President and CEO

Table of Contents

Introduction and Objectives

A 2013 Inovalon study titled "The Impact of Dual Eligible Populations on CMS Five-Star Quality Measures and Member Outcomes in Medicare Advantage Health Plans" presented new quantitative evidence, based on in-depth member-level analysis of 1.6 million Medicare Advantage (MA) plan beneficiaries, that dual eligible members performed worse on 9 of 10 Star quality measures examined.

The 2013 study provided new information demonstrating that beyond the influences of demographics and severity of disease, there remain fundamental differences in outcomes in dual eligible beneficiaries that are unexplained. The objective of this follow-up investigation is to better understand the factors underlying poor performance of dual eligible members served by MA plans enumerated in the earlier study. This study was conducted by the Inovalon research division in collaboration with multiple industry partners, including Cigna-HealthSpring, Wellcare, Healthfirst, Gateway Health, Blue Cross Blue Shield Minnesota and Blue Plus, Health Care Services Corporation, the Special Needs Plan (SNP) Alliance, and Medicaid Health Plans of America (MHPA).

In 2014, the National Quality Forum released draft recommendations that for the first time pointed to the need for risk adjustment of quality measures to account for the impact of SES factors on outcomes in order to make correct and fair inferences about quality and improve outcomes in this vulnerable population. The report noted that a lack of available data on SES factors has to date limited the ability to scientifically test the validity and feasibility of these factors as potential risk adjustors to the quality measures. This study utilizes new data sources to enable testing the impact of various SES factors on a set of eighteen quality measures.

On September 9, 2014, CMS issued a Request for Information on, "*Data on Differences in Medicare Advantage (MA) and Part D Star Rating Quality Measurements for Dual-Eligible versus Non-Dual-Eligible Enrollees.*" One specific request was, "Analysis of the difference in measurement scores between dual and non-dual … enrollees in the same contract and/or plan for all contracts under a parent organization for the Star Ratings measures. Analyses would be more helpful if all enrollees from all contracts under a parent organization are included in the analysis."

The research presented in this report is part of a larger study "*2014 Dual Eligible Collaboration, An Investigation of Medicare Advantage Dual Eligible Member Level Performance on CMS Five-Star Quality Measures*." This large scale study has three major components, including retrospective contract level analyses, member level analyses, and in-depth multivariate analyses. The study protocol has been reviewed by Chesapeake IRB and determined not to require IRB oversight as per Department of Health and Human Services regulations 45 CFR 46.

This report presents one subset of the study, the multivariate analyses. The purpose of this phase of the study is to address CMS' request using a large retrospective database of MA members to examine variation in Star Measure performance level between duals and non-duals after controlling for variation due to Plan Benefit Package (PBP) within contract. Although previous studies have examined this question at the contract level, this is the first large-scale study to look at Star Measure performance at the individual member level and the first to control for the effects of individual plans within contracts.

The specific objectives of these analyses were to:

- Examine Star rating performance at the individual member level (i.e., the dependent variable is member level outcome on measure).
- Investigate systematic differences in Star performance in dual and non-dual status members ("within" effect)
  - After controlling for PBP; and
  - After controlling for both PBP and percentage of dual members in the PBP ("between" effect and "contextual" effect, modeled using two different but statistically equivalent approaches).

- Investigate generalizability and consistency of findings by employing multiple sampling techniques and modeling methods.

Study Data and Methods

Data Sources

The study utilized member-level MA data extracted from Inovalon's Medical Outcomes Research for Effectiveness and Economics Registry (MORE^2 Registry®). MORE² is a large nationally representative and statistically de-identified administrative claims database. The database includes longitudinal patient-level data for more than 110 million individuals from a broad range of data sources across all payer types (commercial, Medicare, and Medicaid), geographic regions (capturing virtually all U.S. counties), health care settings (inpatient and outpatient services), and provider specialties. A base population of 2,207,940 MA members in 81 separate MA contracts with 364 individual plans (PBPs) in 2013 was extracted for this study. The data include age, gender, race/ethnicity and comprehensive information on diseases/diagnoses, chronic conditions, and medical and pharmacy utilization.

CMS Monthly Membership Reports (MMR) were utilized to identify members' original reason for entitlement, the amount of low income drug subsidy received, institutional status and Medicaid dual eligible status.

Outcome Measures

Dependent variables in the study were 8 Star Measures and 10 Star Display measures, selected because they can be calculated at the member level using administrative claims data. The measures studied are shown in Table 1.

All Cause Readmissions (PCR) is the only measure among those evaluated that has been risk-adjusted. We analyzed this measure 3 different ways to provide maximum insight into performance of dual versus non-dual plan members:

- *Unadjusted*: The dependent variable is an indicator of whether the member was readmitted (1) or not readmitted (0). If dual status of the individual member (within effect) is significant, it means duals have higher rates of observed readmissions compared to non-duals.
- *Adjusted*: Member's predicted probability to be re-hospitalized (risk of readmission or "expected" rate) is included as a covariate in the model—if dual status of the individual member (within effect) is significant, it shows duals have a higher risk of readmission after controlling for age, gender and HCC risk factors which are captured in the risk adjusted or "expected" rate.
- *Expected*: Member's expected readmission rate is treated as the dependent variable in the model—if dual status of the individual member (within effect) is significant, it shows duals have a higher expected risk of readmission than non-duals.

Measure Acronym	Measure Name	Weighting/Risk Adjustment	Higher Score is:
ART	Rheumatoid Arthritis Management–C19	None	Better
BPD	Diabetes Treatment–D11	Fraction of year enrolled	Better
HRM	High Risk Medication—D10	Fraction of year enrolled	Worse
MA-C	Medication Adherence for Cholesterol (Statins)—D14	Fraction of year enrolled	Better
MA-D	Medication Adherence for Diabetes Medications—D12	Fraction of year enrolled	Better
MA-H	Medication Adherence for Hypertension (RAS Antagonists)—D13	Fraction of year enrolled	Better
OMW	Osteoporosis Management in Women who had a Fracture—C13	None	Better
PCR	Plan All-Cause Readmissions—C22	Risk-adjusted at discharge by age, gender, comorbid conditions	Worse
AAP	Access to Primary Care Doctor Visits–DMC12	None	Better
AMM	Antidepressant Medication Management (6 months) –DMC03	None	Better
BCS	Breast Cancer Screening–DMC28	None	Better
DDI	Drug-Drug Interactions–DMD06	Fraction of year enrolled	Worse
IET-E	Engagement of Alcohol or other Drug Treatment–DMC19	None	Better
IET-I	Initiation of Alcohol or other Drug Treatment–DMC18	None	Better
PBH	Continuous Beta-Blocker Treatment–DMC04	None	Better
PCE-B	Pharmacotherapy Management of COPD Exacerbation-Bronchodilator–DMC17	None	Better
PCE-S	Pharmacotherapy Management of COPD Exacerbation-Systemic Corticosteroid–DMC16	None	Better
SPR	Testing to Confirm Chronic Obstructive Pulmonary Disease–DMC07	None	Better

Table 1: Star Measures Evaluated

Explanatory (Independent) Variables

Each member was categorized as either dual or non-dual. In exploratory analyses not reported here, we investigated a three-group categorization: always dual, never dual and partial-year dual. The partial-year group was very small (3.6% of the population in 2013) and the partial-year group typically yielded results similar to the always dual group. Therefore, the decision was made to use the "ever dual" method for all subsequent analyses. A member was categorized as dual if s/he was on Medicaid for at least one month during the enrollment period. Member dual status is the individual or "within" predictor. For each PBP, the total number of dual eligible members was divided by the plan total membership, resulting in PBP % dual. This is the between or "contextual" predictor.

It should be noted that the number of members included in the analysis of any given measure can change, depending on whether or not that member was included in the denominator for that measure based on the measure specifications. The percent dual variable is treated as a plan characteristic, calculated from the entire membership, and does not change from one measure to another.

Sample selection

Because the goal of this study phase was to examine the effect of dual status *within* PBP, CMS suggested in early conversations about the study that the analysis should include PBPs with a minimum percentage of members in both groups (duals and non-duals). Subsequently, the CMS RFI indicated that analyses including all members/PBPs may also be of interest. Four subsets of PBPs with different relative proportions of dual vs. non-dual members were examined:

1. At least 30% of each group
2. At least 20% of each group
3. At least 10% of each group
4. All PBPs regardless of % dual

The counts of PBPs and members by PBP % dual membership is shown in Table 2. It shows that analyzing the sample of PBPs with at least 30% of both duals and non-duals resulted in eliminating over 90% of the data. This is because most plans tend to be clustered at the ends; they have either relatively few duals or a large portion of duals. The 30% group was deemed not representative of PBPs generally because of its small size and was dropped from the analyses.

	PBPs		Members	
% dual	N	%	N	%
0	24	6.6	832	<0.1
1 – 9	140	38.5	932,719	42.2
10 – 19	74	20.3	492,485	22.3
20 – 29	34	9.3	171,153	7.8
30 – 69	29	8.0	197,544	9.0
70 – 79	1	0.3	107	<0.1
80 – 89	1	0.3	621	<0.1
90 – 99	51	14.0	411,616	18.6
100	10	2.8	863	<0.1
All	**364**	**100.0**	**2,207,940**	**100.0**

Table 2: Number of PBPs and Members by PBP Percent Dual

Green: At least 30% of each—very small sample of PBPs/Members
Gray: At least 20% of each
Blue: At least 10% of each

NOTE: Adding in the Blue and Grey highlighted PBPs tends to add non-duals disproportionately

Statistical Analysis

This study employs a set of statistical approaches that seek to estimate the relative impact of both individual member and group characteristics on outcome. Members of a group are more similar to other members of the same group and less similar to members of a different group, and statistical methods must take these within-group inter-correlations into account (Feaster et al 2011). Known by different names depending on discipline (e.g., multilevel models, random effects or mixed models, hierarchical linear models), these statistical strategies share the goal of modeling data when individuals are nested within higher order categories (Snijders & Bosker, 1999). They are commonly used in disciplines such as healthcare outcomes research (where patients are nested within hospitals, doctors or plans), educational evaluation (where students are nested within schools and school districts) and organizational psychology (employees are nested within employers).

In the present study, the observations are the individual members of health plans (PBPs) (i.e., nested within PBPs). Members and health plans are both characterized in terms of dual status: members are categorized as dual eligible or not and PBPs are categorized by the percent of their membership that is dual eligible.

This analysis tested three different types of mixed models. All are parameterized to account for the non-independence of data from members of the same plan by including a random effects term for PBP.

1. Model 1 examines the (fixed) effect on outcomes of the member's dual status alone ('*within*' *effect*). The question addressed by this model is: "*Do duals and non-duals differ on any outcome measures after controlling for PBP?"*

2. Model 2 adds an effect for percent of plan members who are dual. This is a general mixed model that is appropriate for a variety of analytical situations where variables are measured at both the individual and group membership levels. The question addressed by this model is: *"After controlling for the plan's % dual ('contextual/between' effect) and PBP random effect, Is there a still a statistically significant difference between duals and non-duals ('within' effect)?"*

3. Model 3 is a more specific form of Model 2 which is appropriate when both the individual and the group are measured on the same variable. It is formulated such that the within-plan (individual) and between-plan (group or contextual) effects can be independently estimated using a technique known as group mean centering (Feaster et al, 2011). In this model, the individual member dual status is transformed into a deviation from the mean of the group of which it is a member in order to assure the two variables are independent and remove any multicollinearity due to the fact that plan percent dual is a roll up of dual status of the individual members of the plan. This model thus provides separate coefficient estimates for the between and contextual effects.

In the description of the models, the following notation is used:

y_{ij} : Response (dependent) variable member *j* from PBP *i*, typically 1 if the member is included in that measure's numerator and 0 otherwise

D_{ij} : Dual status of member *j* from PBP *i*. It is equal to 1 if the member is dual eligible and 0 otherwise

D_Bar_i : Plan % dual membership for PBP *i*: category identifier (model 2) or continuous proportion of dual eligible members (model 3)

β_0 : Overall intercept

β_W : Within-plan, or individual, effect of dual status ("within")

β_B : Between-plan effect of % dual membership ("between")

u_i : PBP-specific random effect

e_{ij} : Individual error term

Model 1, which tests only for the main effect of members' dual status on outcome, controlling for the PBP random effect, is defined as:

$$y_{ij} = \beta_0 + \beta_W D_{ij} + u_i + e_{ij}$$

Model 2, which includes member dual status and adds PBP % dual, tests for the main effects of members' duals status controlling for PBP % dual and the PBP random effect. Using PBP data that are grouped into categories based on % dual, the model is defined as:

$$y_{ij} = \beta_0 + \beta_W D_{ij} + (\beta_1 D\_Bar_{i1} + \cdots + \beta_{k-1} D\_Bar_{ik-1}) + u_i + e_{ij}$$

In Model 2, PBP % dual is specified as dummy variables, defined as D_Bar_{ik}= 1 for PBP i with % dual enrollment from kth group and D_Bar_{ik}= 0 otherwise, to capture a possible non-linear relationship between PBP % dual and the outcome. Five groups were used: < 10%, 10 – <20%, 20 – <30%, 30 – < 90% and 90 – 100%.

This model can also be estimated treating % dual as a continuous variable. However, including % dual defined as a continuous variable results in precisely the same coefficient estimates as shown in the contextual effect coefficient estimates in Model 3, thus those results were not shown here.

Model 3, which transforms the member's dual status to a deviation from the group plan percent dual (group mean centering) in order to independently estimate the within (individual) and between effects, is defined as:

$$y_{ij} = \beta_0 + \beta_W (D_{ij} - D_{Bar_i}) + \beta_B D\_Bar_i + u_i + e_{ij}$$

The contextual effect $(\beta_B - \beta_W)$ can be tested by testing the equality of the within and between coefficients:

H_0: $\beta_W = \beta_B$
H_1: $\beta_W \neq \beta_B$

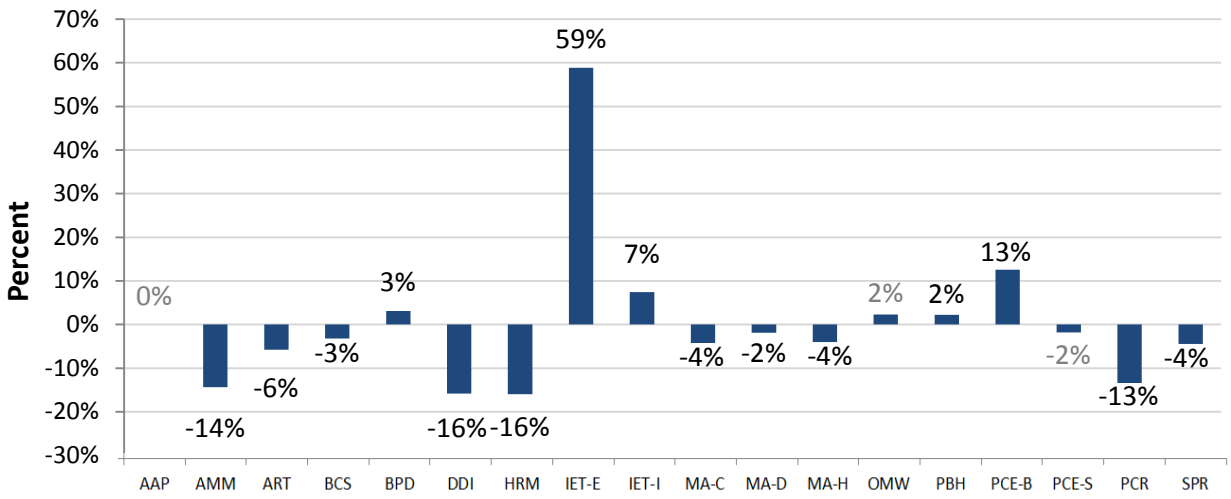
Failing to reject the null hypothesis implies that there is no contextual effect of PBP.

Results

A summary of the 18 Star measures studied is shown in Figure 1. The bars represent the overall difference in rate between duals and non-duals, expressed as a percentage of the non-dual rate. A column below 0% indicates *duals perform worse* than non-duals; above 0% indicates *duals perform better*. (Note that the signs of the inverse measures—DDI, HRM and PCR—where a higher rate indicates worse performance were reversed so that a percent difference below 0 always indicates worse performance by duals.)

Duals perform significantly worse on 10 of 18 measures (56%) and on 6 of the 8 current Star measures (75%). Duals perform significantly better on 5 of 18 measures (28%), all related to medication treatment, including two related to alcohol/drug/substance abuse. Duals perform similarly on Access to Primary Care Visits (AAP) and two other measures (OMW & PCE-S) (17%).

Figure 1: 2013 Star Rating Comparison
Percent Difference in Average Star Ratings Duals vs. Non-Duals



NOTE: **Gray highlighted %** are not statistically significantly different (AAP, OMW, PCE-S)

Model 1*:* Detailed results of the analysis for all measures and all samples are shown in Tab1. Model 1 in Excel document titled "INV Dual Study PBP Analysis Tables 10 20 2014 v 1 0 0".

Duals perform significantly worse on 6 measures in all 3 samples and on 7 measures using the two most inclusive samples: "all members" and "at least 10% dual and non-dual". These are: Breast Cancer Screening (BCS), Drug-Drug Interactions (DDI), High Risk Medications (HRM), Medication Adherence for Hypertension (MA-H), Osteoporosis Management in Women with Fracture (OMW), and All Cause Readmissions (PCR), generally for all three specifications. Use of Spirometry Testing COPD (SPR) is significant using the two largest sample methods.

Duals perform significantly better on only 2 measures in all 3 samples and on 3 measures using the two most inclusive samples: "all members" and "at least 10% dual and non-dual". These are: Adult's Access to Preventive/Ambulatory Care Services (AAP) and Pharmacotherapy Management of COPD Exacerbation-Bronchodilator (PCE-B). Persistence of Beta-Blocker Treatment after Heart Attack (PBH) is significant using the two largest sample methods.

Model 2*:* Detailed results of the analysis for all measures and all samples are shown in Tab2. Model 2 in Excel document titled "INV Dual Study PBP Analysis Tables 10 20 2014 v 1 0 0".

Results for the effect of dual status are identical to the simple model. In other words, even after controlling for the plan percent dual membership, the effect of individual members' dual status persists for the same measures and in the same direction. For some measures, % dual is also significantly associated with outcome.

Model 3*:* Detailed results of the analysis for all measures and all samples are shown in Tab3. Model 3 in Excel document titled "INV Dual Study PBP Analysis Tables 10 20 2014 v 1 0 0".

This model independently estimates the individual dual status difference from plan % dual at the member level, and thus evaluates the true contextual effect of plan % dual. Results for Model 3 are identical to Models 1 & 2.

Comparison of Member Dual Status ("Within" effect) results*:* With minor exceptions, the 3 models (see Tab 4. Compare 3 models) and 3 sampling methods (see Tab 5. Compare 3 samples) are highly consistent.

Results from Plan % dual (Between/Contextual) Effect*:* Models 2 and 3 included a term to analyze the relationship between measures and Plan % dual. As described in the Methods section, this variable was categorized into 5 groups in Model 2 and treated as a continuous variable in Model 3. Model 2 tests for an overall difference between % dual groups allowing for a nonlinear relationship and Model 3 tests specifically for a linear effect.

The results for this effect for Model 2 are shown in Tab 2 in the set of columns labeled "Plan % Dual Contextual/Between Effect." A significant F indicates some difference in measure rate between groups. The results for Model 3 are shown in Tab 3 in the set of columns labeled "Contextual Effect (Between – Within)." A significant t indicates a linear trend between plan % dual and measures rate, and the sign of the estimate indicates its direction. A positive slope (estimate) suggests that plans with more duals tend to perform better on the measure, and a negative slope suggests plans with higher % dual membership tend to perform worse.

The following measures were significantly related to plan % dual for at least one sample in both Models 2 and 3:

- AAP Adults' Access to Preventive/Ambulatory Health Services
- AMM Antidepressant Medication Management-Effective Continuation Phase Treatment
- BPD Diabetes Treatment
- DDI Drug-Drug Interactions
- HRM High Risk Medication
- MA-C Medication Adherence for Cholesterol (Statins)
- MA-D Medication Adherence for Diabetes Medications
- MA-H Medication Adherence for Hypertension (RAS antagonists)

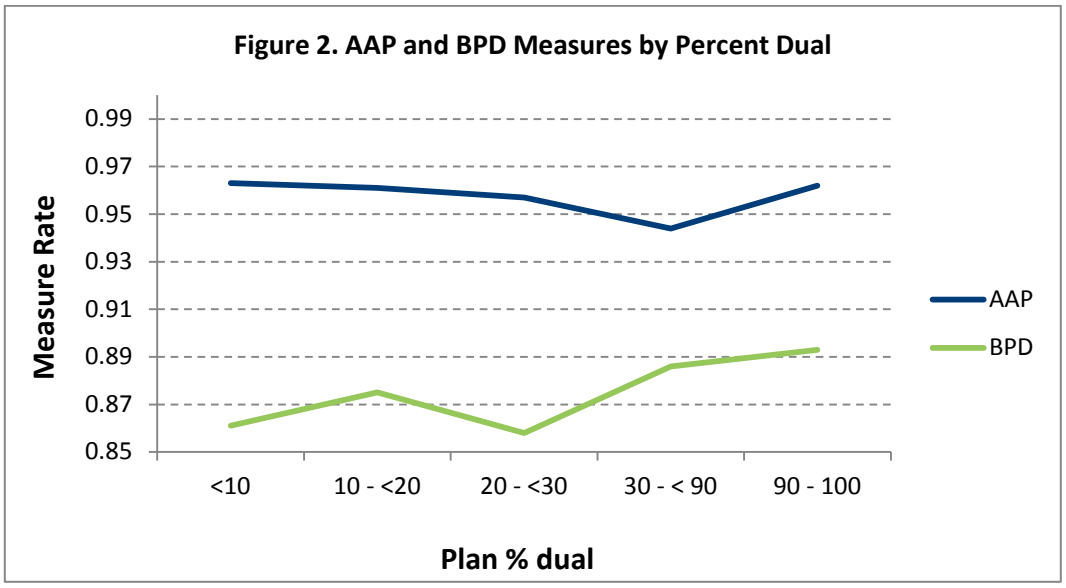
No other measures were significant for both models.

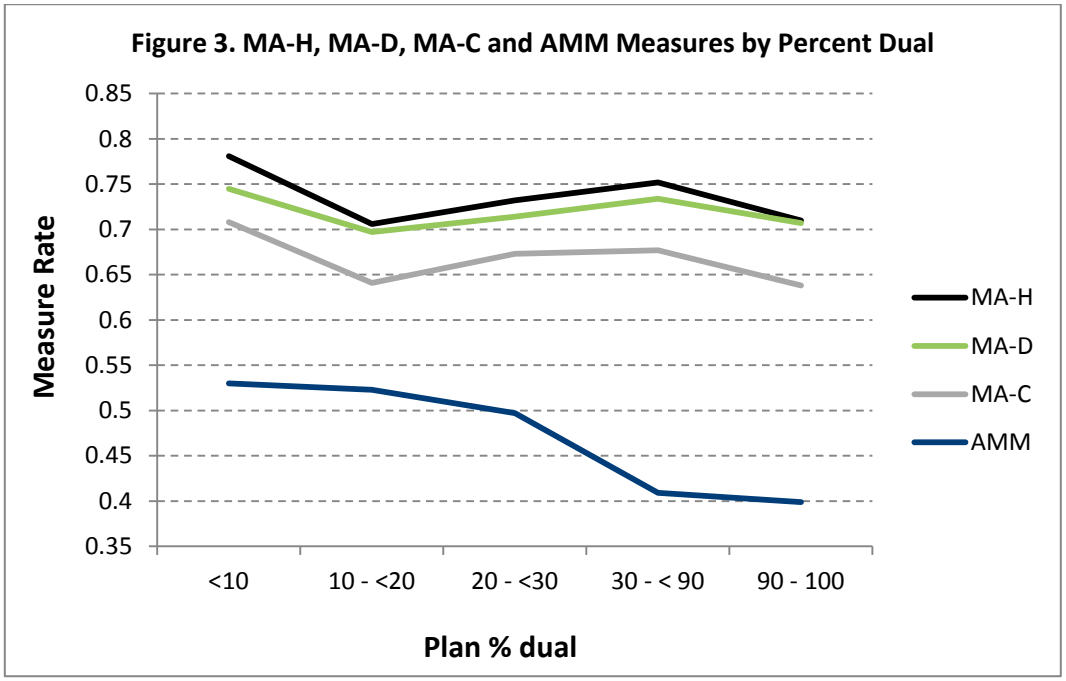
For four of the measures with a significant contextual effect, both plan % dual and member dual status were significant: AAP, DDI, HRM and MA-H. The parameter estimates are independent of each other, however, so both the plan level and the member level variables significantly and independently impact the outcome. In other words, for these 4 measures, the dual status of the member is significant, even after controlling for the (significant but independent) effect of plan % dual.

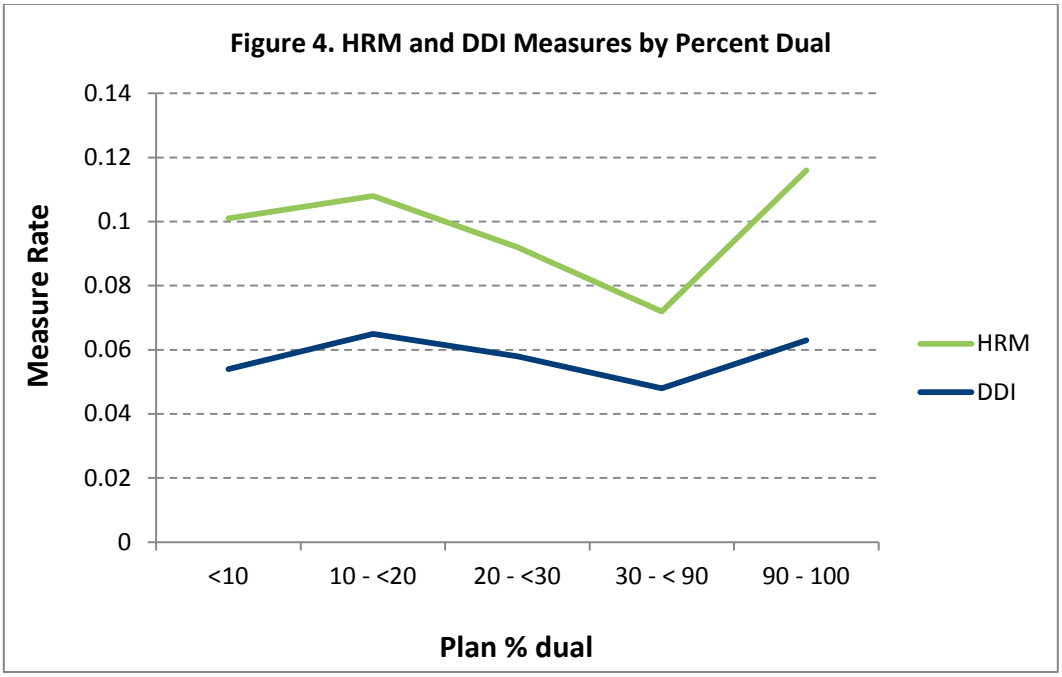
For four of the measures with a significant within effect on dual status of the member—BCS, OMW, PCE-B and PCR—only the within effect was a significant predictor (after controlling for the random effect of PBP). The contextual effect made no significant contribution to the models, indicating plan percent dual does not impact the difference in outcomes between dual and non-dual members. This is particularly important for PCR, the only risk-adjusted measure. Even after applying the risk adjustment to the measure, member dual status remains significantly related to the likelihood of readmission, in contrast to plan % dual, which is not significant.

Graphs of the average measure rate for each % dual group are shown below. AAP and BPD are graphed together in Figure 2 because they are on approximately the same scale (Y-axis range). AMM and the four MA measures are shown together in Figure 3. For all these measures, a higher score is better. DDI and HRM are graphed together in Figure 4 because these are inverse measures, where a higher score is worse.

Visual inspection of these graphs indicates that the relationship between Plan % dual, after controlling for member dual status and the random effect of PBP, may not be a simple linear relationship. In these data, it may best be characterized as quadratic or even cubic. The primary focus of this study has been on the within effect, however, so a more rigorous investigation of the nature of this relationship has not been completed at this point.







Conclusions

Across three sampling methods and three statistical models, nine measures show a statistically significant difference for member dual status after controlling for PBP. Duals perform significantly worse for 7 of the 9 measures. This is particularly important for PCR, where even the risk adjustment for age, gender and chronic conditions does not fully correct for the impact of dual status—a dual member is at higher risk for readmission compared to a non-dual member with same demographic characteristics and same chronic conditions. The contextual effect of plan percent dual is not significant for PCR, indicating a higher proportion of duals in a plan does not impact outcomes in dual members differentially.

There are some measures for which plan % dual membership has a significant effect on Star measure outcomes. It is important to note that even in these cases, it is not an either/or effect—there is a contextual effect related to plan percent dual membership and a separate, independent dual status (within) effect. Further study is underway to determine whether this effect is linear (i.e., do plans serving a high proportion of dual members tend toward lower scores) or non-linear.

Future Analyses

The next phase of this study will incorporate additional covariates in the models as suggested by the CMS RFI: "*If submitters are interested in more in-depth analyses, CMS would suggest using a multivariate model (e.g., logistic regression) to explore the relationship between dual/non-dual status and scores on the Star Ratings measures. These models allow for additional control variables (e.g., contract, comorbidities and health status) to explore these relationships.*"

We plan to evaluate demographic factors (e.g., age, gender, race/ethnicity), socioeconomic and sociodemographic (SES) factors (e.g., income, education, household size, low income drug subsidy), clinical characteristics (i.e., chronic conditions and diagnoses), community resource availability (e.g., primary care and mental health professional shortage area, number of physicians per 10,000 people, rural and isolated rural areas versus urban and large metropolitan areas), and contract characteristics (e.g., plan type, age of contract).

Inovalon Project Team

Christie Teigland, Ph.D., Director Statistical Research (Principal Investigator)
Dr. Paige Killian, Vice President, Clinical Analytics
Jon Bumbaugh, Senior Director, Research, Development & Analytics
Shirley Doyle, Senior Director, Care Management Outcomes Reporting
Zulkarnain Pulungan, Ph.D., Senior Health Economics & Outcomes Researcher
Karl Kilgore, Ph.D., Senior Health Economics & Outcomes Researcher
Charles Maris, Ph.D., Senior Manager, Research Initiatives
Ping Chen, MD, MS, Health Services Researcher
Jeanette Hunsberger, Statistician
Barton Jones, MS, Health Data Scientist
Matthew McClellan, Health Data Scientist
Sandhya Mehta, Ph.D., Quantitative Healthcare Analyst
Alexis Parente, Ph.D., Health Services Researcher

References

1. Feaster D, Brincks A, Robbins M, Szapocznik J. Multilevel models to identify contextual effects on individual group member outcomes: a family example. Fam Process. 2011; 50:167-183.
2. Snijders T, Bosker R. Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. Sage Publishers: 1999.