Methods for Comparing Community Based, Real-world Chronic Disease Interventions using Electronic Record Data

Mary Masters, MPH, MS[1], Bob Harris, PhD[1], Lauren Greenberg, MPH[1], John Lynch, MPH[2], Karen Pasquale, MPH[2],
[1]Westat, Rockville, MD
[2]Connecticut Center for Primary Care, Inc., Farmington, CT:

November 17, 2014

Objectives: 1) develop a database in a real world primary care population and 2) apply these methods in a manner that resulted in an unbiased sample.

Study Questions: What steps are needed to develop an ambulatory care electronic health record database from which "real world" queries can improve the care of patients with multiple comorbidities (particularly hypertension) and guide primary care? How feasible is this using primary care and public health statistics?

Introduction: Statistical techniques applied to data from real world primary care patients offer a guide for treatment choices among patients with multiple chronic conditions. The assembly of a large sample of type II diabetic patient data from primary care electronic health records allowed analysis of a broad range of patients across combinations of demographics, co-morbidity, and concurrent medications.

Methods: Propensity score matching and adjustment for non-randomization of treatments were employed; the analysis considered large numbers of covariates, interventions, and outcomes. This database of type 2 diabetes patients included clinical data, physical and lab measures, concomitant medications, and *International Classification of Disease*, 9th revision coded medical problems. Methods of data cleaning included the review of frequencies to find extreme outlier values and combining similar categories to avoid small cell sizes.

Results: A cohort of 4,040 type 2 diabetes patients yielded statistically significant results when comparing the change in key indicator lab values across four oral treatment regimens. In covariate adjusted models, a diagnosis of hypertension made a significant contribution to predicting a change in HbA1c ($p=0.0016$) and the effective medication regimens differed among patients with and without hypertension.

Conclusions: Having a diagnosis of hypertension appeared to impact the change in HbA1c, which differed by treatment regimen. The more effective regimen for T2DM appeared to be MET-TZD for patients with hypertension and MET-SU for patients without hypertension.

Project Task Summary: This project spanned two years, the grant required that adequate progress was made in <u>year 1</u> before year 2 would be approved. Year 1 work involved building the database by: 1) First considering confidentiality issues, gaining institutional review approval which required using study specific IDS and blocking private health information, 2) next we reviewed the eHealth record data tables by topic to decide what would be the most relevant information, we limited the database to data merged from a) demographic characteristics; b) administrative/billing data, c) clinical problems and diagnoses, d) interventions, medications, and processes, e) lab findings, and f) office physical measures. Where relevant start and stop dates were available in the system, 3) Next we built a pretest file that was data for 25 adult patients and extracted from these data tables. The start date was the beginning of the care center's launching of the eHealth system. Patient data was examined for duplicates and for inconsistencies, 4) we described the coding schemes and the variables and created a data dictionary which ultimately included the categories of responses/results for each variable or the acceptable ranges. This later became our edit rules dictionary, outlier data was reviewed. We were approved for year 2 and we proceeded with more work with the database as shown in Figure 2.

Statistical Methods

Propensity Score modeling: Since patients were not randomly assigned to treatment groups, there could be significant differences in patients comprising the four treatment groups, leading to selection bias when comparing outcomes among groups. The use of propensity score modeling was proposed by Rosenbaum and Rubin (1983) as a means for addressing bias in observational studies due to differences in baseline characteristics between treatment groups. In brief, the propensity score is the conditional probability that a subject will receive one treatment versus another, derived from a logistic regression model with treatment as the dependent variable and various covariates as predictors. The propensity score can be included as a continuous-scaled covariate in modeling assessing the relationship of treatment with the outcome measure of interest to address possible imbalances between comparison groups. The goal is to make the comparison groups as similar as possible with respect to the measured covariates, based on the assumption that subjects with similar propensity scores have similar values for the covariates. The PS approach offers the added benefit of enabling adjustment for potential bias due to a large number of covariates with a small number of propensity score variables. In this investigation of four treatment groups, the model would yield three PS scores for each subject; the three PS scores would represent each subject's predicted probability of receiving one of the comparison treatments versus the referent treatment (Non-Met) based on their covariate pattern.

Covariates were screened for possible inclusion in the PS model using simple bivariate analyses based on covariate values obtained at baseline. This included Fisher's exact test for categorical covariates and the nonparametric Wilcoxon test for continuous-scaled covariates. The relationship between treatment group and over 200 covariates were examined, including demographic variables, laboratory measures, office measures, and the presence of

comorbid conditions at baseline. Variables associated with treatment regimen at an alpha level of <0.1 were considered as candidates for inclusion in the PS modeling. Given the very large sample size and consequent ample power to detect differences, an alpha level of 0.1 was considered a reasonable target for selecting covariates for this pilot investigation. Given the large number of variables explored in the PS modeling, the screening process also considered the distributional characteristics of the covariates in an attempt to avoid mathematical difficulties and excessive losses to the study population; for categorical measures adequacy of cell sizes for each treatment group was defined as at least 20 observations per cell in the cross-tabulated data and variables with >10% missing values were typically dropped from consideration.

As is true of multivariable modeling in general, propensity score analysis typically involves an iterative process that begins by considering all candidate covariates in the model and assessing how well the model behaves, such as whether or not it fails to converge to a solution. The logistic regression modeling employed stepwise selection using SAS Proc LOGISTIC to generate the final model. Ideally, the distribution of the propensity scores generated from the model should overlap among the treatment groups if the propensity scores are to satisfy their intended goal of adjusting for potential bias due to non-randomization of the treatment prescribed. Box-and-whisker plots were used to facilitate assessment of the overlap in each of the four propensity scores according to the actual treatment received.

Modeling of change in hemoglobin A1c: Multivariable linear regression modeling (SAS Proc GLMSELECT and Proc GLM) was used to examine the relationship between change in A1c (from regimen start until regimen end) and treatment group, adjusted for propensity scores and other covariates. Two "base" models were defined: 1) a model including treatment group and propensity scores, and 2) a model that adjusted for propensity scores, age and gender; the latter model was used when assessing the influence of other covariates and interactions.

Rather than screen every potential covariate for inclusion in the A1c modeling, we focused on those variables that had been considered as candidates for the propensity score model, and thus known to have a significant relationship with treatment group, as well as adequate cell sizes and relatively few missing values. These variables were then screened for further consideration by assessing the significance of their relationship with change in A1c. The results of these screening tests were used to identify variables for further consideration in modeling the change in A1c, from which a sampling of covariates was selected for final evaluation.

Models were fit to the data to assess main effects of the covariates and interactions of the covariates with treatment regimen. A significant main effect of gender would indicate that change in A1c differed significantly between males and females. A significant interaction of gender with treatment would indicate that the effect of gender on change in A1c differed according to the treatment prescribed or equivalently, that the effect of the prescribed treatment on change in A1c differed according to gender. It is possible to generate estimated mean changes in A1c for main effects and interactions of treatment with categorical covariates. For interactions of treatment with continuous-scaled covariates plots were generated to illustrate the relationships, as the effect of the treatment would vary over levels of the continuous covariate.

As a final step in the analysis, change in A1c was modeled with the significant covariates and significant interaction terms identified previously included as candidates for stepwise selection in the regression model. Analyses used SAS statistical software, 9.2 or later (SAS Institute Inc. Cary, NC).
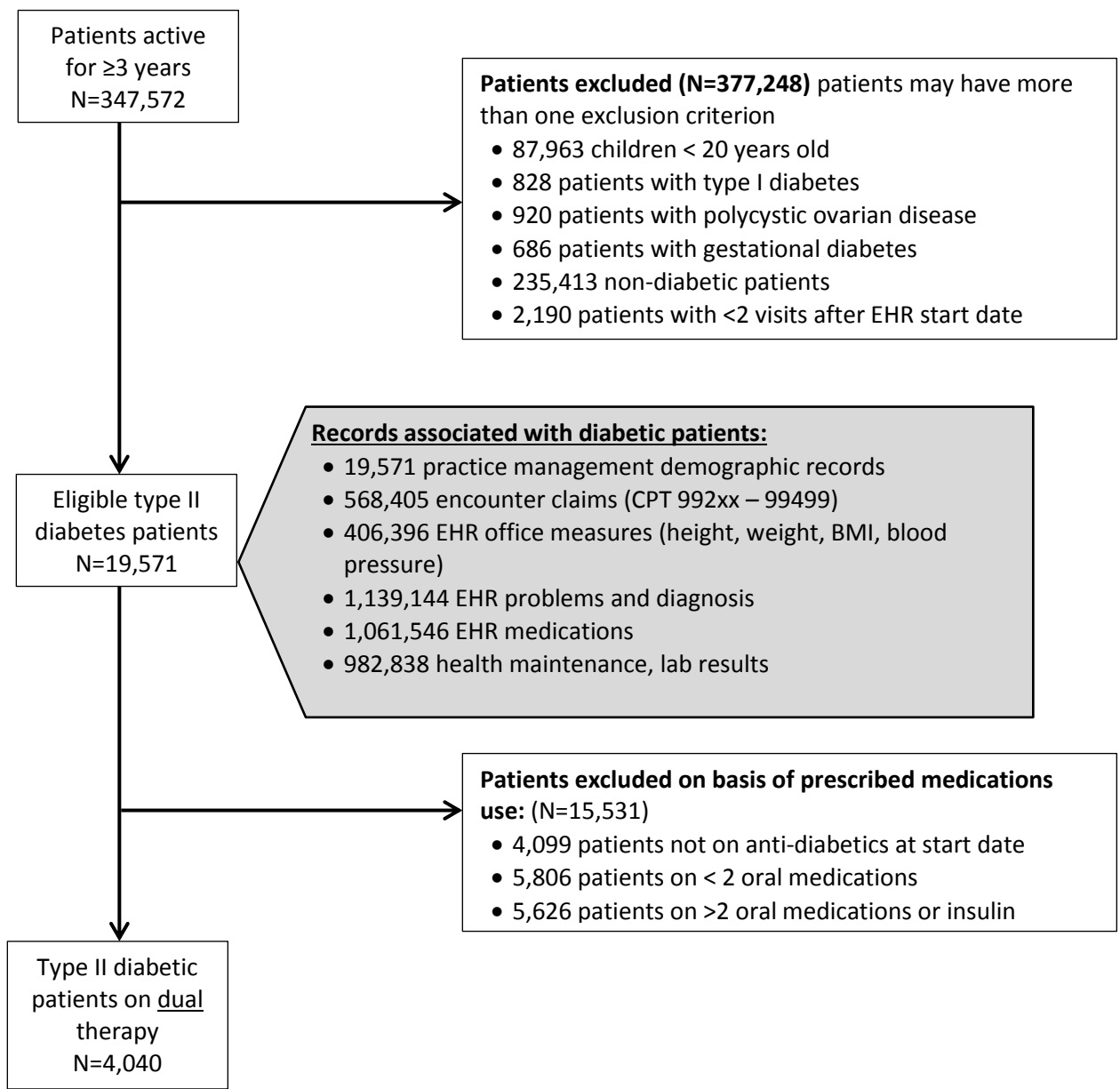
Figure 1:   Study Population Derivation

```
┌─────────────────┐
│ Patients active │
│ for ≥3 years    │        ┌──────────────────────────────────────────────────────────┐
│ N=347,572       │        │ Patients excluded (N=377,248) patients may have more       │
└─────────────────┘        │ than one exclusion criterion                               │
         │                 │   • 87,963 children < 20 years old                         │
         │──────────────►  │   • 828 patients with type I diabetes                      │
         │                 │   • 920 patients with polycystic ovarian disease           │
         │                 │   • 686 patients with gestational diabetes                 │
         │                 │   • 235,413 non-diabetic patients                          │
         │                 │   • 2,190 patients with <2 visits after EHR start date     │
         │                 └──────────────────────────────────────────────────────────┘
         ▼
┌─────────────────┐        ┌──────────────────────────────────────────────────────────┐
│ Eligible type II│        │ Records associated with diabetic patients:                 │
│ diabetes patients│◄─────│   • 19,571 practice management demographic records         │
│ N=19,571        │        │   • 568,405 encounter claims (CPT 992xx – 99499)           │
└─────────────────┘        │   • 406,396 EHR office measures (height, weight, BMI, blood│
         │                 │     pressure)                                              │
         │                 │   • 1,139,144 EHR problems and diagnosis                   │
         │                 │   • 1,061,546 EHR medications                              │
         │                 │   • 982,838 health maintenance, lab results                │
         │                 └──────────────────────────────────────────────────────────┘
         │
         │                 ┌──────────────────────────────────────────────────────────┐
         │                 │ Patients excluded on basis of prescribed medications       │
         │──────────────►  │ use: (N=15,531)                                            │
         │                 │   • 4,099 patients not on anti-diabetics at start date     │
         │                 │   • 5,806 patients on < 2 oral medications                 │
         │                 │   • 5,626 patients on >2 oral medications or insulin        │
         │                 └──────────────────────────────────────────────────────────┘
         ▼
┌─────────────────┐
│ Type II diabetic│
│ patients on dual│
│ therapy         │
│ N=4,040         │
└─────────────────┘
```

Figure 2. Selected Characteristics and Categories for Propensity Score Modeling

Characteristic	Total	Antidiabetic Regimen			
		Met-SU	Met-TZD	Met-DPP4	Non-Met
Age group:*+					
20-49	708 (17.5)	273 (14.4)	172 (19.5)	234 (26.9)	29 (7.5)
50-64	1552 (38.4)	688 (36.2)	371 (42.1)	391 (44.9)	102 (26.4
65-79	1358 (33.6)	703 (37.0)	277 (31.4)	206 (23.7)	172 (44.4)
≥80	422 (10.4)	237 (12.5)	61 (6.9)	40 (4.6	84 (21.7)
Gender: *					
Female	1792 (44.4)	817 (43.0)	376 (42.7)	409 (47.0)	190 (49.1)
Male	2248 (55.6)	1084 (57.0)	505 (57.3)	462 (53.0)	197 (50.9)
Race: *					
White	3075 (76.1)	1464 (77.0)	626 (71.1)	668 (76.7)	317 (81.9)
Not White	965 (23.9)	437 (23.0)	255 (28.9)	203 (23.3)	70 (18.1)
'Baseline' hemoglobin A1c percent:*					
Mean / median	7.4 / 7.1	7.5 / 7.2	7.0 / 6.7	7.8 / 7.3	7.2 / 7
Diastolic blood pressure: *					
Mean / median	75.3 / 76	75.4 / 76	74.5 / 74	77.2 / 78	72.3 / 72
Primary insurance: *					
Government	2043 (50.6)	1075 (56.5)	388 (44.0)	302 (34.7)	278 (71.8)
Private	1997 (49.4)	826 (43.5)	493 (56.0)	569 (65.3)	109 (28.2)
Days in study: *					
Mean / median	491.7 / 503	474.9 / 490	513.8 / 526	494.8 / 507	517.0 / 526

* Included in the propensity score model.

**Met-SU** = Metformin plus Sulfonylureas   **Met-DPP4** = Metformin plus Dipeptidyl Pepitidase-4
Met-TZD = Metformin plus Thiazolamides   **Non-Met** = Sulfonylureas plus either DPP4 or TZD

Figure 3. Predicted mean change (±95% confidence interval) in A1c by treatment regimen and hypertension diagnosis from adjusted model

(Significant interaction of hypertension diagnosis with treatment regimen: p=0.007)
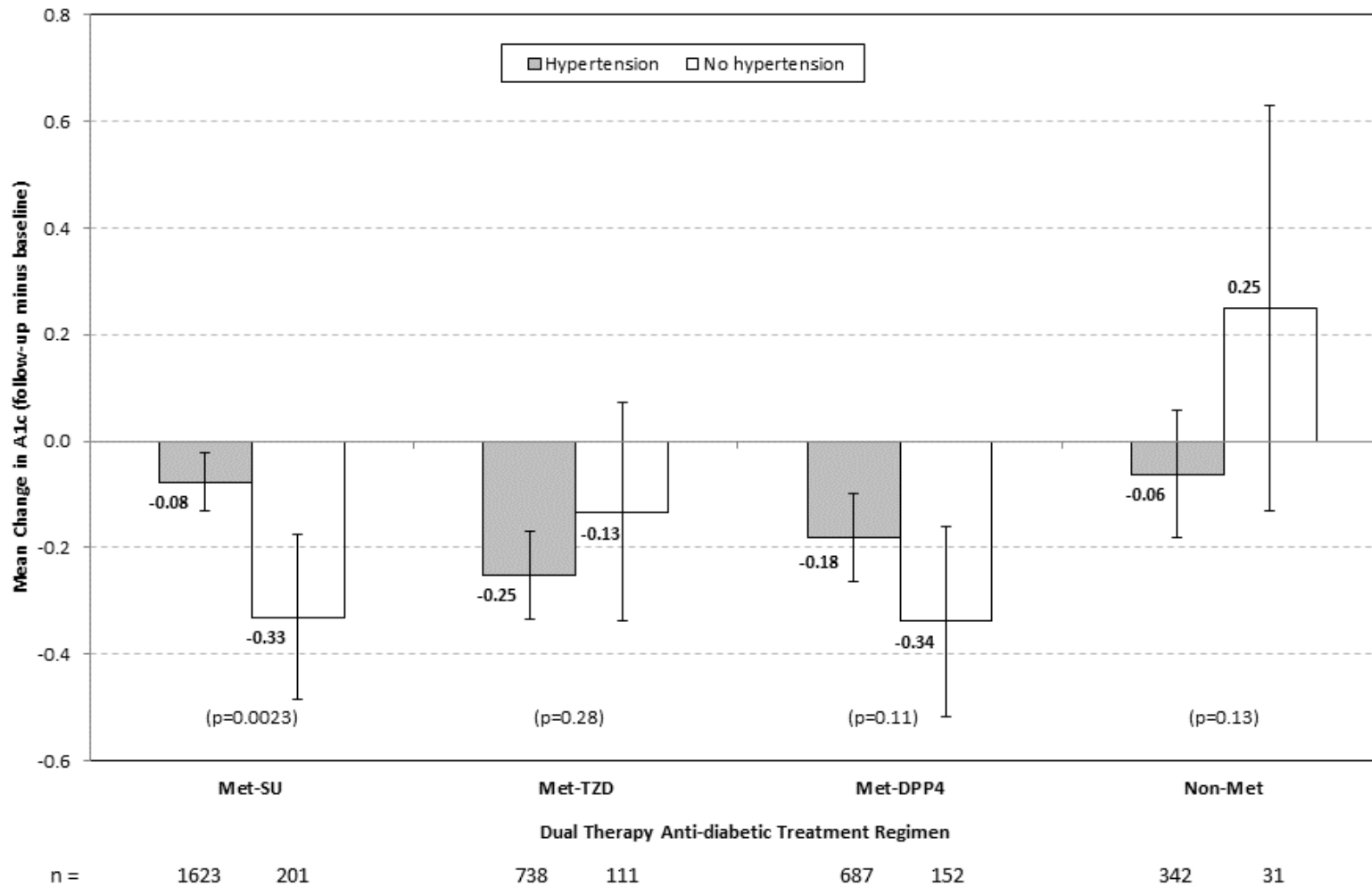
Predicted Change in A1c: Significant Covariate-Treatment Interaction (Baseline Hgb A1c)
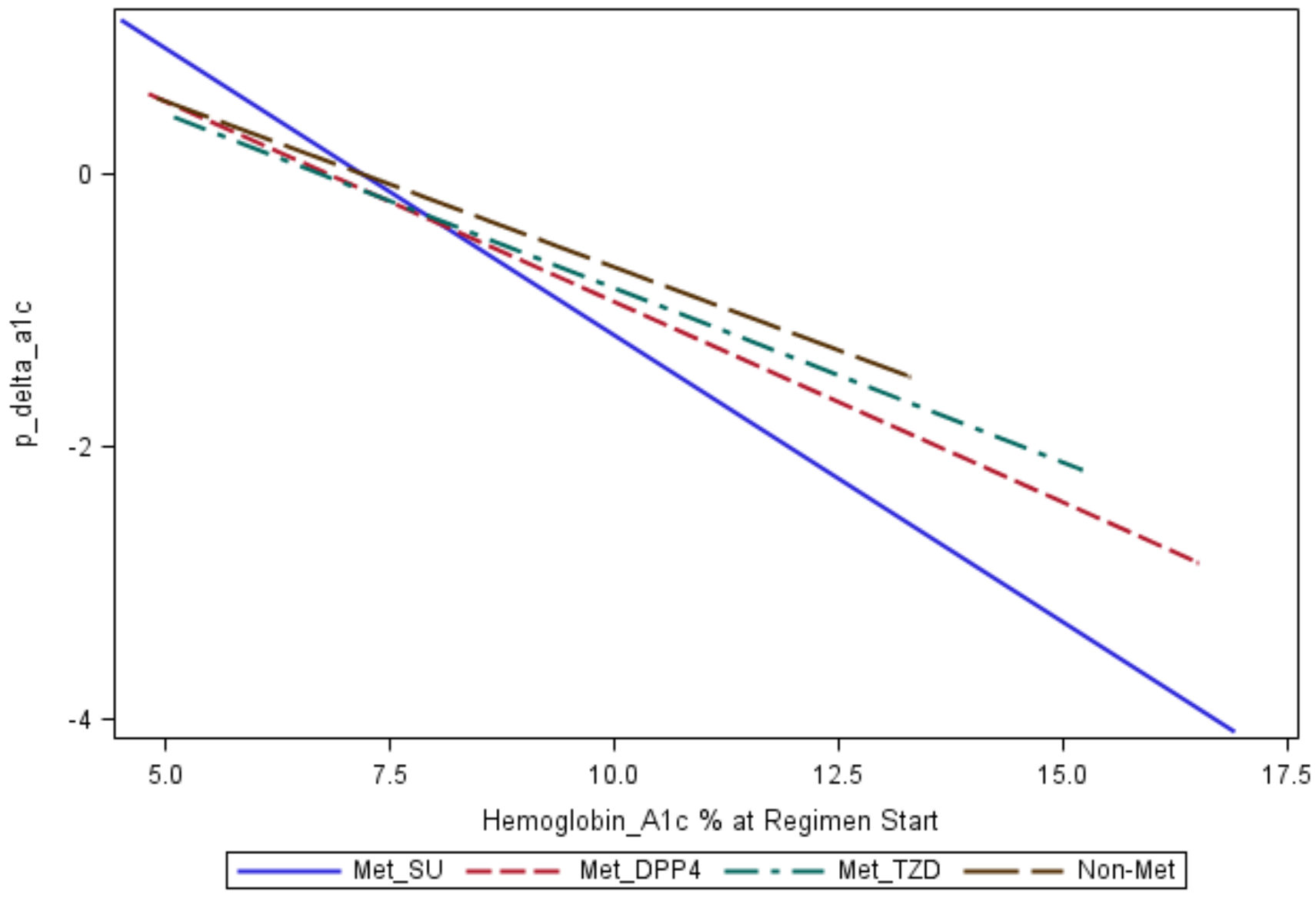
SUMMARY:

In statistical modeling, a significant interaction of 'x1' with 'x2' in predicting the outcome, 'y', indicates that the effect of 'x1' on 'y' varies over levels of 'x2' or equivalently that the effect of 'x2' on 'y' varies over levels of 'x1'. For example, the effect of a medication on a particular disease may differ depending on whether or not a certain comorbidity is present (interaction of medication with a categorical-scaled covariate) or depending on the patient's serum cholesterol level (interaction of medication with a continuous-scaled covariate).

The above figures illustrate the significant interactions that were detected among the targeted covariates that were explored in the BICEP analyses. The first two figures illustrate interactions of treatment regimen with categorical-scaled covariates, while the last three figures illustrate interactions with continuous-scaled covariates. Note that while the model-based predicted change in A1c (follow-up measure minus regimen start measure) is plotted along the y-axis in all plots, the appearance of the plots differs according to the scale of measure of each covariate.

Interactions of treatment regimen with categorical-scaled covariates

For interactions of categorical-scaled covariates with treatment regimen, the figures show the mean change in A1c (± 95% confidence interval [CI]) for each treatment regimen for those with a diagnosis of hypertension (or anemia/blood disorders) compared to those without such a diagnosis. When the interaction contributes significantly to the model, we expect there will be differences in the mean change in A1c within one or more of the treatment groups when the change in A1c is compared between those with and without the diagnosis. Equivalently, we would expect there to be differences among those with (or without) a diagnosis of hypertension when the mean change in A1c is compared across treatment regimens.

Hypertension:
- The mean change in A1c differed between those with and without hypertension who were receiving Met-SU (p<0.0001); a borderline difference was observed between hypertensives and non-hypertensives among those receiving Met_DPP4 (p=0.06)
- Among patients with hypertension, the change in A1c differed between those receiving Met-SU and those receiving Met-TZD or Met DPP4
- Among patients without a diagnosis of hypertension, change in A1c differed between those receiving Met-SU and Non-Met regimens

Interactions of treatment regimen with continuous-scaled covariates

For interactions of continuous-scaled covariates with treatment regimen, the figures display the change in A1c according to covariate level, with separate lines plotted for each treatment regimen. If the interaction does not contribute to the model, we expect these lines to essentially be parallel, with the main effect of treatment corresponding to the distance between the lines and the main effect of the covariate represented by the slope of the lines. Conversely, if the interaction does contribute significantly to the model, we expect the lines

to be non-parallel; this can take on the appearance of the lines crossing one another at some point or simply diverging from one another in a 'v-like' shape.

Baseline Hemoglobin A1c: (Significance of interaction in model: $p<0.0001$)
- The change in A1c appears fairly similar for the Met-DPP4, Met-TZD and Non-Met regimens (similar slopes), while the Met-SU regimen appears to be associated with larger declines in A1c among those with higher baseline A1c levels and larger increases in A1c among those with lower baseline A1c levels
- Language referring to the "association" of treatment with change in A1c is used, since we cannot necessarily imply a causal relationship between the treatment regimen and change in A1c

KEY REFERENCES

1. Pace WD, Cifuentes M, Robert J, et al., An Electronic Practice-Based Network for Observational Comparative Effectiveness Research, *Ann Intern Med.* 2009;151:338-340.

2. Tinetti ME, and Studenski, SA. Comparative Effectiveness Research and Patients with Multiple Chronic Conditions, *N Engl J Med* 2011, 364:2478-2481.

3. Desai JR, Pingsheng W, Nichols, GA, et al. Diabetes and Asthma Case Identification, Validation, and Representativeness When Using Electronic Health Data to Construct Registries for Comparative Effectiveness and Epidemiologic Research, *Med Care* 2012, 50 (7) S30-S35.

4. D'Agostino RB, and D'Agostino RB. Estimating treatment effects using observational data, *JAMA* 2007; 297(3): 314-316.

5. Nichols GA, Desai J, Elston Lafata J, et al. Construction of a Multisite DataLink Using Electronic Health Records for the Identification, Surveillance, Prevention, and Management of Diabetes Mellitus: The SUPREME-DM Project. *Prev Chronic Dis* 2012; 9:110311.

6. Malik, S., Billimek, J., Greenfield, S., et al., Patient Complexity and Risk Factor Control Among Multimorbid Patients With Type 2 Diabetes Results From the R2D2C2 Study, *Med Care, 51*: 180-185, 2013.

Contact information: marymasters@westat.com