

# Statistical Analysis Plan

**PI: Matt Swedlund**

**Date: 2017-04-04**

## Introduction

Hypertension Best Practice Alerts (BPA's) were active in the UW Health medical records system from approximately November 2013 through November 2015. For patients visiting the healthcare system during this period, if an elevated blood pressure (BP) was recorded, then a "Smart Set" link would pop-up on their record (the "Smart Set" would "fire"), which could be clicked to go to a recommended set of actions.

During the BPA active period, three separate implementation of the BPA were used:

BPA #	BPA ID	BPA name	first fire date	last fire date
1	1002351	UWOP B ACO ELEVATED BLOOD PRESSURE READING	2013-11-18	2014-03-21
2	10023456	UWOP B ELEVATED BLOOD PRESSURE PRIMARY CARE	2014-05-07	2014-12-23
3	10123456	ZZUWOP B ACO ELEVATED BP OR GOAL NOT MET - RETIRED	2014-12-23	2015-11-24

Differences exist between the three BPA's above in both the logic that would trigger a "fire", and to whom the "Smart Set" link would be displayed. As such, BPA 3 will be the BPA period of interest in these analyses as it was the most recent, more revised, and longest active BPA compared to the other two.

For the remainder of this document the "BPA" is always in reference to BPA 3 above. The logic for when this BPA "fires" is:

- patient is  $\geq$  18 years
- AND** either one of the following occurs during an appointment or office visit with the family medicine, internal medicine, geriatrics, or gerontology departments:
- the patient did not have a BP goal on their problem list AND their BP was  $\geq$  140 mmHg systolic or 90 mmHg diastolic
  - the patient did have a BP goal on their problem list AND their BP was  $\geq$  their systolic or diastolic goal

If the logic was satisfied, then the "Smart Set" link would open in the general BPA section of the health record, when the cart is opened.

## General Research Objectives

The primary object of these analyses is to help assess the relation between the BPA active period and patients' return for followup.

The secondary object of these analysis is to help assess the relation between the BPA active period and patients' BP control.

## Study Population and Sampling Strategy

The encounter records to sample from in these analyses will be comprised of records where:

- The subject was  $\geq 18$  years of age at the time of the encounter
- The encounter occurred on or between the dates in the following date pairs:
  - 2013-05-01 and 2013-11-17 (control period: before any BPA's)
  - 2014-12-24 and 2015-11-23 (treatment period: during BPA)
  - 2015-11-25 and 2016-05-15 (control period: after any BPA's)
- The encounter type was for an appointment, office visit, or nurse check-in with the family medicine, internal medicine, geriatrics, or gerontology departments

All such above records will be referred to as the "master encounter set."

### Primary analysis sampling

From the master encounter set, the records associated with nurse check-in's will be excluded. The remaining records to then include for the sampling will be those that:

- If the record occurred in the BPA period, that the EHR indicate that the BPA "fired"
- If the record did not occur in the BPA period, that the patient had a BP reading **AND** either one of the following occurred:
  - the patient did not have a BP goal on their problem list **AND** their BP was  $\geq 140$  mmHg systolic or 90 mmHg diastolic
  - the patient did have a BP goal on their problem list **AND** their BP was  $\geq$  their systolic or diastolic goal

After the above inclusion / exclusion criteria has been applied, the remaining records will be referred to as the "primary encounter set."

From the primary encounter set, for each unique subject in the set, one encounter will be randomly selected, uniformly, for inclusion in the primary analysis.

### Secondary analysis sampling

From the primary encounter set, the following additional inclusion criteria will be applied for each encounter:

- within the master encounter set, another encounter occurred between and including the dates 2 weeks to 3 months (14 to 90 days) after the encounter of interest in the primary set
- at least one of these master encounters in that period has a BP reading

After the above additional inclusion criteria has been applied, the remaining records will be referred to as the "secondary encounter set".

From the secondary encounter set, for each unique subject in the set, one encounter will be randomly selected, uniformly, for inclusion in the secondary analysis.

# **Outcome Variables**

## **Primary analysis outcome**

### **Return for BP follow-up indicator (binary)**

- For each encounter sampled, did the master encounter set contain another encounter for that subject which:
  - occurred between and including the dates 2 weeks to 3 months (14 to 90 days) after the sampled encounter
  - had a BP reading

## **Secondary analysis outcome**

### **BP systolic and diastolic pressures (continuous)**

- For each encounter sampled, for that subject, the BP reading from the encounter in the master encounter set that occurred closest to the sampled encounter, for the period between and including the date 2 weeks after to 3 months (14 to 90 days) after the sampled encounter. If multiple BP reading occur in this encounter, the last BP reading of the visit will be used.

### **BP control indicator (binary)**

- For the BP reading used above:
  - For subjects negative for CKD (see below):
    - \* systolic < 140 AND diastolic < 90
  - For subjects positive for CKD (see below):
    - \* systolic < 130 AND diastolic < 80

# **Covariates**

For each of the encounters sampled above, for both the primary and secondary samplings:

## **Treatment of the sampled encounter (categorical)**

- Did the encounter of interest occur “before”, “during” or “after” the BPA period

### **Glomerular Filtration Rate (GFR) (continuous)**

- The closest GFR value within 365 days of the sampled encounter, if available. When ties are present, use the value before the sampled encounter.

### **Chronic kidney disease (CKD) indicator (binary)**

- Did the subject have a GFR < 60

### **Hypertension diagnosis indicator (binary)**

- Was the subject on the UW Health “HTN” registry

#### **Diabetes indicator (binary)**

- Was the subject on the UW Health “Diabetes” registry

#### **Type II diabetes indicator (binary)**

- If the subject had diabetes, as indicated above, were they type II diabetic

#### **Body Mass Index (BMI) (continuous)**

- BMI measurement

#### **Weight groups (categorical)**

- Discrete groups based upon BMI:
  - $\text{BMI} < 30$
  - $\text{BMI} \in [30, 40)$
  - $\text{BMI} \geq 40$

#### **Age (continuous)**

- Age in years

#### **Sex (binary)**

- Male or female

#### **Race (categorical)**

- Race specified in the EHR

#### **Ethnicity (categorical)**

- Ethnicity specified in the EHR

#### **Insurance type (categorical)**

- Type of insurance: Medicaid, Medicare, private, or none

#### **Year of the sampled encounter (integer)**

- Year of the sampled encounter

#### **Month of the sampled encounter (categorical)**

- Month of the sampled encounter

## Statistical Methods

### Baseline Comparisons (Table 1)

All subject characteristics listed in the “Outcome Variables” and “Covariates” sections above will be reported (except encounter date) for both the data sampled for the primary analysis and secondary analysis, separately. The primary sample will be stratified by the follow-up outcome indicator; the secondary sample will be stratified by the BP control outcome indicator.

Numeric items will be summarized with the mean and standard deviation as well as median and inter-quartile range (IQR), with Mann-Whitney-Wilcoxon tests used for assessing differences in central tendency between the strata. Categorical items will be summarized by frequency and percent, with Fisher’s exact tests being used for assessing frequency differences between the strata.

### Primary Analysis

Logistic regression will be utilized to examine the association between the BPA active / control periods and follow-up visits where BP was measured. Four models will be fit. The outcome in all models will be the indicator for the “return for BP follow-up”.

1. An unadjusted analysis will only include the binary indicator of if the sampled encounter was in the BPA period.
2. An adjusted analysis will also include the covariates age, sex, race, ethnicity, CKD indicator, weight group category, HTN indicator, HTN medications indicator, Diabetes indicator, and insurance category.
3. Secular / seasonal variants of both unadjusted and adjusted models will also be fit; these models will also include a linear component for secular trend (year of sampled encounter) as well as sine and cosine seasonal components (month of sampled encounter) for both 12 month and 6 month periods (with all periods having integer multiples of  $2\pi$  occurring in the month of January). For each of these secular / seasonal variants, Bayesian Information Criterion (BIC) will be used to assess if any of these 5 additional terms should be removed from the final such model fitted.. Thus a total of 4 models will be fit to aid in the primary analysis.

### Secondary Analysis

The same general modeling setup used in the primary analysis will be used to aid in the secondary analysis. The outcome for the logistic regression models will instead be the BP control indicator for the sampled encounter. In addition to the 4 models from the crossing of unadjusted / adjusted and exclusion / inclusion of the seasonal-secular trends, 4 additional models will also be fit using these same structures, but with the addition of the estimated conditional probability of return-for-followup. This additional covariate will be determined by using the model fit from the primary analysis from the model with the same associated structure without this covariate, and estimating conditional probabilities for return-for-followup on the subjects in the secondary analysis. Thus a total of 8 models will be fit to aid in the secondary analysis.

### Additional Tables

Tables of the regression estimates, confidence intervals, and p-values will be provided separately for the primary and secondary analyses.