STATISTICAL ANALYSIS PLAN FOR

"The Effect of Physician Ownership on Dialysis Outcomes"

PROTOCOL ID: HS-18-00530-1

SECONDARY ID: K08DK118213

Version #2

Last Modified: 3/11/2021

PI: Eugene Lin, MD, MS

Keck School of Medicine of the University of Southern California

ANALYTIC DATASETS

Facility level dataset

1) Merge the physician ownership dataset with DFC
    a. Primarily use 2017 data, but if a facility is missing data, salvage with 2018, 2016, 2015 years
2) Merge zip codes with census / ACS characteristics
3) Merge with the facility survey (from 2017)
4) Restrict to free-standing facilities where <= 15% of patients in 2017 were pediatric (< 18 years old)
5) Variables of interest
    a. Facility Characteristics
        i. Profit status
        ii. Chain vs. independent
        iii. Size of facility (Number of patients)
        iv. Staff : patient ratio
        v. ESRD Network
    b. Outcomes
        i. All quality measures from DFC dataset
6) Get distribution of patient characteristics for each facility
    a. Link each facility to a patient
        i. To assign a patient to a facility, first identify all dialysis patients with at least one month of dialysis in all of 2017 and meeting all of the following criteria
            1. Adults with valid 2728 form
            2. In the USA.
            3. Has Medicare FFS for that month of dialysis
        ii. If a patient has multiple facilities, assign that patient to the facility that holds the plurality share
        iii. If there are still ties in facilities, use the first observed facility that has the plurality
    b. For each patient, for the first observed month of dialysis, get:
        i. Age, Gender, Race, Ethnicity, Payer, Dual-Eligible (if Medicare)
        ii. Dialysis modality
        iii. (Do not exclude any patients for this analysis)
    c. Collapse back to the facility level, and get counts and % by patient characteristic (for continuous variables like age, we use age brackets, i.e., 18-44, 45-54, 55-64, 65-74, 75+)

Patient-month level dataset / inclusion criteria

1) This will be a patient-month dataset
2) Identify all patients with at least one month of dialysis in 2017 (Jan-Dec)
3) Inclusion criteria
    a. USA
    b. Adult in the current month
    c. Valid 2728 form (2005, 2015 form)

    d. Dialyzing in free-standing facility that has fewer than 15% pediatric
        i. Only keep if patient can be linked to a valid facility survey and can be linked to the ownership dataset
4) Must have Medicare FFS in the current month of dialysis
    a. Must have at least 12 months of Medicare FFS prior to the current month
    b. Create indicator variable for whether the patient-month has had 36+ months of Medicare FFS
5) Get patient comorbidities using the CCW algorithm
    a. Use diagnosis codes from claims in a 12 month lookback
    b. Sensitivity analysis: use diagnosis codes from claims in a 36 month lookback
6) Get the zipcode of residence for that month and merge with the ACS to get zipcode characteristics
7) Merge with facility file to get facility characteristics + DFC outcomes (we call these facility level outcomes)
8) Covariates of interest:
    a. Patient level
        i. Age
        ii. Sex
        iii. Race
        iv. Ethnicity
        v. Whether patient had a transplant previously
        vi. Whether the patient is incident (within the first 120 days of ESRD)
        vii. Number of years with ESRD
        viii. Comorbidities using CCW algorithm
    b. Facility characteristics
        i. Profit status
        ii. Chain vs. independent
        iii. Size of facility (Number of patients)
        iv. Staff : patient ratio
        v. ESRD Network
    c. Zipcode Characteristics
        i. Median income
        ii. High school graduation rate
        iii. % below poverty
    d. Monthly level fixed effects (an indicator for each month of the year)
9) Patient Level Outcomes
    a. Whether patient using home dialysis in the month
    b. Whether patient was receiving dialysis optimally (defined as either peritoneal dialysis, or hemodialysis through an AV fistula or graft)
        i. Whether patient was using a fistula
        ii. Whether patient was using a catheter for 3+ months, including the current month (i.e., look back two months to see whether the patient used a catheter in the other two months)
    c. Get crownweb outcomes (merge with the crownweb dataset by month/year)
        i. Create an indicator for whether crownweb data missing for that month
        ii. Hemoglobin < 10

iii. Hemoglobin > 12
iv. Calcium > 10.2
v. Phosphorus > 7

d. Get other outcomes
    i. Whether patient died in that month
    ii. Whether patient was hospitalized that month
    iii. Whether patient had a hospitalization that month followed by a 30-day readmission
        1. Patients are eligible for this only if the patient was discharged alive
    iv. Whether patient had a hospitalization that month followed by an unplanned 30-day readmission (https://qualitynet.cms.gov/inpatient/measures/readmission/resources#tab3 , "2018 All-Cause Hospital Wide Measure Updates and Specifications Report Hospital-Level 30-Day Risk-Standardized Readmission Measure – Version 7.0", pages 51 and 68)
    v. Whether patient had a blood transfusion that month
        1. Criteria 1
            a. From inpatient, outpatient, SNF claims
            b. Revenue code 0390 OR 0392 OR 0399 PLUS any of the following HCPCS/CPT
                i. P9010
                ii. P9011
                iii. P9016
                iv. P9021
                v. P9022
                vi. P9038
                vii. P9039
                viii. P9040
                ix. P9051
                x. P9054
                xi. P9056
                xii. P9057
                xiii. P9058
            c. Calculate number of units of blood by taking the sum of the units for (Revenue center code 0390 OR 0392 OR 0399) across all line items that correspond to that specific day.
                i. Sum across all days for a given month
        2. Criteria 2
            a. If Criteria 1 is missing, then determine whether a blood transfusion occurred for the entire claim (assign to the month of the through date of the claim)
            b. Look for value code 37
        3. Criteria 3

a. If Criteria 1 is missing, then determine whether a blood transfusion occurred for the entire claim (assign to the month of the through date of the claim)
b. Look for any of the following ICD procedure codes:
    i. ICD9
        1. 9903
        2. 9904
    ii. ICD10
        1. 30230H1, -P1, -N1
        2. 30233H1, -P1, -N1
        3. 30240H1, -P1, -N1
        4. 30243H1, -P1, -N1
        5. 30250H1, -P1, -N1
        6. 30253H1, -P1, -N1
        7. 30260H1, -P1, -N1
        8. 30263H1, -P1, -N1

vi. Whether a patient received an ESA and their dose (using CROWNWeb data)
    1. Consolidate to Darbepoetin, Erythropoietin alfa, Erythropoietin beta
    2. Make sure units are consistent (darbe in mcg, Epo alfa in units, Epo beta in units)
    3. Make sure the doses are normalized, to the same mean (i.e., we want to make sure we are looking at relative changes in doses within the same ESA)
        a. Take the mean doses of each of the ESA types, then divide all the doses by this mean. In other words, all means will be set at "1"
    4. For each month of dialysis create the following variables
        a. Whether ESA administered
        b. Type of ESA administered
        c. Dose
    5. For each dialysis chain, determine the ESA that is the majority of patient-months

ANALYSES

Facility Level Outcomes

1) Use facility level dataset (patient level covariates are collapsed to facility level using summary statistics for that facility
2) Dependent variable (left hand side)
    a. Facility level outcomes as specified above
3) Independent Variables:
    a. Whether facility is physician owned

   b. Regression 1: Control only for facility, geographic, and temporal level covariates
   c. Regression 2: Control for patient, facility, geographic, temporal level covariates.
4) Models: OLS for continuous outcomes, logit for binary outcomes, robust standard errors

Patient Level Regressions

1) Population: patient-month dataset
2) Dependent variable (left hand side):
   a. Patient level outcomes
3) Independent variables (right hand side)
   a. Whether patient is dialyzing in a facility that is physician owned (covariate of interest)
   b. Covariates above (patient, facility, zipcode, temporal characteristics)
4) Models (in order) for all outcomes except blood transfusion and ESA outcomes:
   a. Logit with patient level fixed effects, non-parametric bootstrap standard errors
   b. Logit with patient level fixed effects, robust standard errors
   c. OLS with patient level fixed effects, cluster robust standard errors at facility level
   d. OLS with patient level fixed effects, robust standard errors
5) Models (in order) for whether patient received a blood transfusion
   a. Outcomes
      i. Criteria 1 only
      ii. Criteria 1+2
      iii. Criteria 1+2+3
   b. Models
      i. Logit with patient level fixed effects, non-parametric bootstrap standard errors
      ii. Logit with patient level fixed effects, robust standard errors
      iii. OLS with patient level fixed effects, cluster robust standard errors at facility level
      iv. OLS with patient level fixed effects, robust standard errors
6) Models (in order) for # of units of blood transfused
   a. Outcomes
      i. Use the number of units of blood detected under Criteria 1. If a blood transfusion detected under Criteria 2 or 3, indicate that the patient received 1 unit of blood
      ii. Use the number of units of blood detected under Criteria 1. If a blood transfusion detected under Criteria 2 or 3, indicate that the patient received 1 unit of blood. Also include an indicator variable that the patient had a blood transfusion detected under criteria 2 or 3.
      iii. Use the number of units of blood detected under Criteria 1. If a blood transfusion detected under Criteria 2 or 3, indicate that the patient received 1 unit of blood. Also include an indicator variable that the patient had a blood transfusion detected under criteria 2 or 3 AND interact this indicator variable with the covariate for whether patient dialyzed in a facility that is physician owned.
   b. Models

        i. Poisson with patient level fixed effects, non-parametric bootstrap standard errors
        ii. Poisson with patient level fixed effects, robust standard errors
        iii. OLS with patient level fixed effects, cluster robust standard errors at facility level
        iv. OLS with patient level fixed effects, robust standard errors
7) Models (in order) for dose of ESA received
    a. Poisson with patient level fixed effects, non-parametric bootstrap standard errors
    b. Poisson with patient level fixed effects, robust standard errors
    a. OLS with patient level fixed effects (dose of ESA received), cluster robust standard errors at facility level
        i. Include covariate for chain
    b. OLS with patient level fixed effects (dose of ESA received), robust standard errors
        i. Include covariate for chain
    c. Split dataset into 3 separate datasets based on CHAIN of assigned dialysis facility. Specifically, assign each chain to the type of ESA that takes up the majority of all patient-months. Separately, run the following regression on each of the datasets
        i. Poisson with patient level fixed effects, non-parametric bootstrap standard errors
        ii. Poisson with patient level fixed effects, robust standard errors
        iii. OLS with patient level fixed effects (dose of ESA received), cluster robust standard errors at facility level
        iv. OLS with patient level fixed effects (dose of ESA received), robust standard errors

Subgroup analyses

1) Repeat the above regressions, but this time we have the following independent variables of interest (we also include all the other adjusters)
    a. Whether facility is physician owned (phys_own)
    b. Whether facility is LDO owned (ldo_own)
    c. Interaction term between these two (phys_own#ldo_own)
2) In addition to reporting the coefficients for these 3 terms, we also care about the following comparisons:
    a. FOR FACILITIES THAT ARE LDO OWNED: effect of physician ownership
        i. The parameter of interest is the sum of the following coefficients: phys_own + phys_own#ldo_own
    b. FOR FACILITIES THAT ARE NOT LDO OWNED: effect of physician ownership
        i. The parameter of interest is the following coefficient: phys_own
    c. Use the “lincom” command to estimate the sum and the standard errors of 2ai.