Final Installment

Asthma Treatment Pattern Longitudinal Analysis

Chong Kim

supervised by

Dr. Matthew Strand

November 29, 2016

The Department of Clinical Pharmacy (DOCP) has provided funding for this educational project but has not conducted the research or written this report. While the authors have worked on the best information available to them, neither DOBB/DOCP nor the authors shall in any event be liable for any loss, damage or injury howsoever suffered directly or indirectly in relation to the report or the research on which it is based.

Reference herein to trade names and proprietary products without stating that they are protected does not imply that they may be regarded as unprotected and thus free for general use. No endorsement of named products is intended nor is it any criticism implied of other alternative, but unnamed, products.

Contents

List of Figures

List of Tables

A Changes and Additions

A.1 Data Management

Change: Previously the time since diagnoses was a continuous measure that had counts of days. The new time variable, that allows for estimable standard errors, is measured in years not days. This allowed for the model to have stable standard errors that were estimable.

Inspection: Check whether or not the data sets created for sensitivity analysis (n=300, 1000, and 5000) had the appropriate number of unique patients within the data.

A.2 Modeling

Addition 1: Sensitivity analyses using 300, 1k and 5k results in different significance values for demographic variables.

Addition 2: Sensitivity analyses using multiple sample data sets of size 1k results in age significance robust to sampling.

Addition 3: 5k sample indicates significance in age, region, and cci along with time. Association between treatment steps and these variables seem to be strongly associated based on p-value (<0.0001).

Addition 4: Added clustering component (random intercept for subjects) in the hypotheses section.

Final Data use: Using the 5k data set for hypothesis 1 and hypothesis 2a. The objective function was not obtained using the 5k data for hypothesis 2b so using 1k data for hypothesis 2b.

B Hypotheses

Hypothesis 1: We hypothesize that the likelihood of asthma treatment step will be different for different patient characteristics.

Hypothesis 1a: *The likelihood of asthma treatment step (ordinal variable) will be different for different patient characteristics.*

Hypothesis 2: We hypothesize that the likelihood of asthma exacerbation event will be different for different treatment patterns.

Hypothesis 2a: *The likelihood of asthma exacerbation event (yes/no) will be different for different treatment patterns.*

Hypothesis 2b: *The likelihood of asthma exacerbation event (nominal variable) will be different for different treatment patterns.*

C Statistical Methods

C.1 Hypothesis 1

Various treatment step patterns will be associated with patient characteristics as described in the covariates section using multinomial regression methods that account for the longitudinal (correlated) data structure of the treatment step patterns. Specifically, the regression model will be built using a generalized linear mixed model with a cumulative logit link and with multinomial distribution. There will be minimal model selection as we are going to include all of the covariates mentioned in the covariates section.

C.1.1 Model 1: Model with All Covariates

Model with all covariates(baseline age, gender, chronic comorbidity index, time, baseline cost, and treatment step). Results should indicate the relationship between the baseline characteristics of subjects and the outcome, asthma treatment step. Repeated measure $\mathbf{R_i}$ side random effects are not supported for multinomial distribution.

$$logit(P(Y_i < j)) = \theta_j - \beta_1 baseage_i - \beta_2 cci_i - \beta_3 time_{il} - \beta_4 basecost_i - gender_k - b_{i0} \quad (1)$$

$$i = 1, .., n; j = 1, ..., J-1 \ where \ J = 6 \ steps; k = 1, 2; l = 1, ..., r_i;$$

$$\theta_j = Threshold \ variable; \ b \sim N(0, \sigma_b^2)$$

C.1.2 Model 2: Polynomial Trend

Predictor variables(baseline age, gender, chronic comorbidity index, time, baseline cost). Results should indicate the relationship between the baseline characteristics of subjects and the outcome, asthma treatment step.

$$logit(P(Y_i < j)) = \theta_j - \beta_1 baseage_i - \beta_2 cci_i - \beta_3 time_{il} - \beta_4 basecost_i - gender_k - b_{i0} \quad (2)$$
$$i = 1, .., n; j = 1, ..., J-1 \ where \ J = 6 \ steps; k = 1, 2; l = 1, ..., r_i;$$
$$\theta_j = Threshold \ variable; \ b \sim N(0, \sigma_b^2)$$

C.2 Hypothesis 2a

Asthma exacerbation events will be associated primarily with the asthma treatment steps along with additional covariates of interest using multinomial regression methods that account for the longitudinal (correlated) data structure of the asthma exacerbation event (yes/no). Specifically, the regression model will be built using a generalized linear mixed model with a logit link and with binomial distribution.

C.2.1 Model 1:Simple Model with all covariates

Model with all covariates(baseline age, gender, chronic comorbidity index, time, baseline cost, and treatment step). Results should indicate the relationship between the baseline characteristics of subjects and the outcome, asthma exacerbation events. Repeated measure $\mathbf{R_i}$ side random effects are not supported for multinomial distribution.

$$\log(\frac{p}{1-p}) = \beta_0 + \beta_1 baseage_i + \beta_2 cci_i + \beta_3 basecost_i + \beta_4 time_{ij} + gender_k \quad (3)$$
$$\beta_5 TS_{ij} + \beta_6 TS_{ij} * time_{ij} + b_{i0} i = 1, .., n; k = 1, 2; j = 1, ..., r_i;$$
$$b \sim N(0, \sigma_b^2)$$

C.2.2 Model 2: Modeling polynomial trend for time

Model 2 + Polynomial variables for time. Results should indicate the relationship between the baseline characteristics of subjects and the outcome, asthma treatment step.

$$\log(\frac{p}{1-p}) = \beta_0 + \beta_1 baseage_i + \beta_2 cci_i + \beta_3 basecost_i + \beta_4 time_{il} + \beta_5 time_{il}^2 + \quad (4)$$
$$gender_k + \beta_5 TS_{ij} + \beta_6 TS_{ij} * time_{ij} + b_{i0}$$
$$i = 1, .., n; k = 1, 2; j = 1, ..., r_i;\ b \sim N(0, \sigma_b^2)$$

C.3 Hypothesis 2b

Asthma exacerbation events will be associated primarily with the asthma treatment steps along with additional covariates of interest using multinomial regression methods that account for the longitudinal (correlated) data structure of the asthma exacerbation event (yes/no). Specifically, the regression model will be built using a generalized linear mixed model with a generalized logit link and with multinomial distribution.

C.3.1 Model 1: Model with all covariates

Predictor variables(baseline age, gender, chronic comorbidity index, time, baseline cost, and treatment step). Results should indicate the relationship between the baseline characteristics of subjects and the outcome, asthma treatment step.

$$\log(\frac{p_{ijc}}{p_{ij1}}) = \beta_0 + \beta_1 baseage_i + \beta_2 cci_i + \beta_3 basecost_i + \beta_4 time_{ij} + gender_k \quad (5)$$
$$\beta_5 TS_{ij} + \beta_6 TS_{ij} * time_{ij} + b_{i0}\ for\ c = 2, 3, .., 5\ i = `, .., n; k = 1, 2; j = 1, ..., r_i;\ b \sim N(0, \sigma_b^2)$$

C.3.2 Model 2: Modeling polynomial trend for time

Model 1 + Polynomial variables for time. Results should indicate the relationship between the baseline characteristics of subjects and the outcome, asthma treatment step.

$$\log(\frac{p_{ijc}}{p_{ij1}}) = \beta_0 + \beta_1 baseage_i + \beta_2 cci_i + \beta_3 basecost_i + \beta_4 time_{il} + \beta_5 time_{il}^2 + \tag{6}$$
$$gender_k + \beta_5 TS_{ij} + \beta_6 TS_{ij} * time_{ij} + b_{i0}$$
$$for\ c = 2, 3, .., 5;\ i = `, .., n; k = 1, 2; l = 1, ..., r_i; b \sim N(0, \sigma_b^2)$$

D New Results from Changes

D.1 Hypothesis 1

D.1.1 Final Output for Hypothesis 1

Type III Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
year	1	124E3	145.87	<.0001
year*year	1	124E3	258.99	<.0001
age	1	124E3	35.40	<.0001
gender	2	124E3	1.52	0.2197
region	3	124E3	7.76	<.0001
CCI	1	124E3	16.80	<.0001
Insurance	5	124E3	1.15	0.3302

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error
Intercept	pat_id	0.4407	0.01600

Solutions for Fixed Effects

Effect	Trt_Step	region	gender	Insurance	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Intercept	5				-5.1055	0.7528	4988	-6.78	<.0001	0.05	-6.5814	-3.6296
Intercept	4				-2.7194	0.7525	4988	-3.61	0.0003	0.05	-4.1946	-1.2441
Intercept	3				-1.8802	0.7525	4988	-2.50	0.0125	0.05	-3.3554	-0.4049
Intercept	2				-1.1881	0.7525	4988	-1.58	0.1144	0.05	-2.6634	0.2871
Intercept	1				-0.4434	0.7525	4988	-0.59	0.5557	0.05	-1.9187	1.0318
year					0.1299	0.01076	124E3	12.08	<.0001	0.05	0.1088	0.1510
year*year					-0.03290	0.002045	124E3	-16.09	<.0001	0.05	-0.03691	-0.02890
age					0.004176	0.000702	124E3	5.95	<.0001	0.05	0.002800	0.005552
gender			1		0.05459	0.7522	124E3	0.07	0.9421	0.05	-1.4196	1.5288
gender			2		0.01030	0.7519	124E3	0.01	0.9891	0.05	-1.4634	1.4840
gender			0		0	.	.	.	.	.	.	.
region		1			0.04728	0.03328	124E3	1.42	0.1554	0.05	-0.01794	0.1125
region		2			0.1471	0.03168	124E3	4.64	<.0001	0.05	0.08503	0.2092
region		3			0.09719	0.04321	124E3	2.25	0.0245	0.05	0.01250	0.1819
region		4			0	.	.	.	.	.	.	.
CCI					0.08480	0.02069	124E3	4.10	<.0001	0.05	0.04425	0.1254
Insurance				2	0.09773	0.1807	124E3	0.54	0.5886	0.05	-0.2564	0.4519
Insurance				3	-0.05397	0.05893	124E3	-0.92	0.3597	0.05	-0.1695	0.06153
Insurance				4	0.04027	0.1571	124E3	0.26	0.7976	0.05	-0.2676	0.3481
Insurance				5	0.07852	0.03789	124E3	2.07	0.0382	0.05	0.004267	0.1528
Insurance				6	0.03086	0.1769	124E3	0.17	0.8615	0.05	-0.3159	0.3776
Insurance				1	0	.	.	.	.	.	.	.

D.2 Hypothesis 2a

D.2.1 Final Output for Hypothesis 2a

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
year	1	124E3	203.90	<.0001
year*year	1	124E3	67.45	<.0001
year*year*year	1	124E3	49.53	<.0001
Trt_Step	5	124E3	164.07	<.0001
year*Trt_Step	5	124E3	13.26	<.0001
age	1	124E3	4.67	0.0307
gender	2	124E3	18.64	<.0001
region	3	124E3	10.80	<.0001
CCI	1	124E3	7.47	0.0063
Insurance	5	124E3	1.52	0.1804

Covariance Parameter Estimates			
Cov Parm	Subject	Estimate	Standard Error
Intercept	pat_id	0.6861	0.02759

Solutions for Fixed Effects												
Effect	region	Trt_Step	gender	Insurance	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Intercept					-2.2507	1.3134	4988	-1.71	0.0867	0.05	-4.8255	0.3242
year					-0.5039	0.03976	124E3	-12.67	<.0001	0.05	-0.5818	-0.4260
year*year					0.1455	0.01772	124E3	8.21	<.0001	0.05	0.1108	0.1802
year*year*year					-0.01452	0.002064	124E3	-7.04	<.0001	0.05	-0.01857	-0.01048
Trt_Step		1			-0.5011	0.04302	124E3	-11.65	<.0001	0.05	-0.5855	-0.4168
Trt_Step		2			-1.1734	0.06440	124E3	-18.22	<.0001	0.05	-1.2996	-1.0471
Trt_Step		3			-0.8962	0.06151	124E3	-14.57	<.0001	0.05	-1.0168	-0.7756
Trt_Step		4			-0.9159	0.06251	124E3	-14.65	<.0001	0.05	-1.0384	-0.7934
Trt_Step		5			0.7312	0.09054	124E3	8.08	<.0001	0.05	0.5536	0.9087
Trt_Step		0			0	.	.	.	.	.	.	.
year*Trt_Step		1			-0.1476	0.02208	124E3	-6.68	<.0001	0.05	-0.1908	-0.1043
year*Trt_Step		2			-0.1822	0.03695	124E3	-4.93	<.0001	0.05	-0.2546	-0.1098
year*Trt_Step		3			-0.07320	0.02990	124E3	-2.45	0.0144	0.05	-0.1318	-0.01460
year*Trt_Step		4			-0.02069	0.02759	124E3	-0.75	0.4534	0.05	-0.07477	0.03340
year*Trt_Step		5			-0.03407	0.03666	124E3	-0.93	0.3527	0.05	-0.1059	0.03779
year*Trt_Step		0			0	.	.	.	.	.	.	.
age					0.002119	0.000981	124E3	2.16	0.0307	0.05	0.000197	0.004042
gender			1		0.6799	1.3130	124E3	0.52	0.6045	0.05	-1.8934	3.2533
gender			2		0.8962	1.3127	124E3	0.68	0.4948	0.05	-1.6766	3.4690
gender			0		0	.	.	.	.	.	.	.
region	1				-0.2196	0.04609	124E3	-4.76	<.0001	0.05	-0.3099	-0.1293
region	2				-0.1748	0.04362	124E3	-4.01	<.0001	0.05	-0.2602	-0.08926
region	3				-0.2723	0.06076	124E3	-4.48	<.0001	0.05	-0.3914	-0.1532
region	4				0	.	.	.	.	.	.	.
CCI					0.07688	0.02812	124E3	2.73	0.0063	0.05	0.02176	0.1320
Insurance				2	-0.2150	0.2665	124E3	-0.81	0.4198	0.05	-0.7374	0.3073
Insurance				3	0.07467	0.08320	124E3	0.90	0.3695	0.05	-0.08841	0.2377
Insurance				4	-0.2697	0.2258	124E3	-1.19	0.2324	0.05	-0.7124	0.1729
Insurance				5	0.1116	0.05222	124E3	2.14	0.0327	0.05	0.009210	0.2139
Insurance				6	0.08206	0.2532	124E3	0.32	0.7459	0.05	-0.4143	0.5784
Insurance				1	0	.	.	.	.	.	.	.

D.3 Hypothesis 2b

D.3.1 Final Output for Hypothesis 2b

Covariance Parameter Estimates				
Cov Parm	Subject	Group	Estimate	Standard Error
Intercept	pat_id	Event 1	3.3033	0.5279
Intercept	pat_id	Event 2	0	.
Intercept	pat_id	Event 3	1.0201	0.1430
Intercept	pat_id	Event 4	1.0423	0.1224

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
year	4	21246	14.39	<.0001
year*year	4	21246	6.53	<.0001
year*year*year	4	21246	4.53	0.0012
Trt_Step	20	21246	13.49	<.0001
year*Trt_Step	20	21246	1.55	0.0554
age	4	21246	3.71	0.0051
gender	4	21246	4.47	0.0013
CCI	4	21246	6.08	<.0001

E Initial Descriptives

	Baseline Treatment Steps							
	0	1	2	3	4	5	p	test
n	781	112	39	38	26	4		
age (mean (sd))	31.42 (17.95)	28.00 (18.32)	22.26 (16.32)	29.89 (17.87)	33.19 (18.94)	42.00 (19.13)	0.013	
pat_region (%)							0.206	
E	219 (28.0)	25 (22.3)	6 (15.4)	10 (26.3)	9 (34.6)	2 (50.0)		
MW	254 (32.5)	40 (35.7)	13 (33.3)	15 (39.5)	9 (34.6)	1 (25.0)		
S	229 (29.3)	28 (25.0)	15 (38.5)	5 (13.2)	5 (19.2)	1 (25.0)		
W	79 (10.1)	19 (17.0)	5 (12.8)	8 (21.1)	3 (11.5)	0 (0.0)		
Insurance (%)							0.191	
1	665 (85.1)	90 (80.4)	34 (87.2)	29 (76.3)	24 (92.3)	3 (75.0)		
2	3 (0.4)	1 (0.9)	0 (0.0)	1 (2.6)	1 (3.8)	0 (0.0)		
3	20 (2.6)	8 (7.1)	2 (5.1)	1 (2.6)	0 (0.0)	1 (25.0)		
4	3 (0.4)	1 (0.9)	0 (0.0)	1 (2.6)	0 (0.0)	0 (0.0)		
5	86 (11.0)	11 (9.8)	3 (7.7)	6 (15.8)	1 (3.8)	0 (0.0)		
6	4 (0.5)	1 (0.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)		
baseline_cost (mean (sd))	2361.08 (7009.85)	1506.44 (2521.98)	1142.77 (1332.55)	886.63 (1378.18)	1734.34 (3036.80)	1940.89 (504.53)	0.455	
CCI (mean (sd))	0.30 (0.62)	0.18 (0.71)	0.18 (0.56)	0.05 (0.23)	0.19 (0.49)	0.25 (0.50)	0.074	

F Graphical Visualization

Figure 1: Correlation Matrix of Continuous Variables

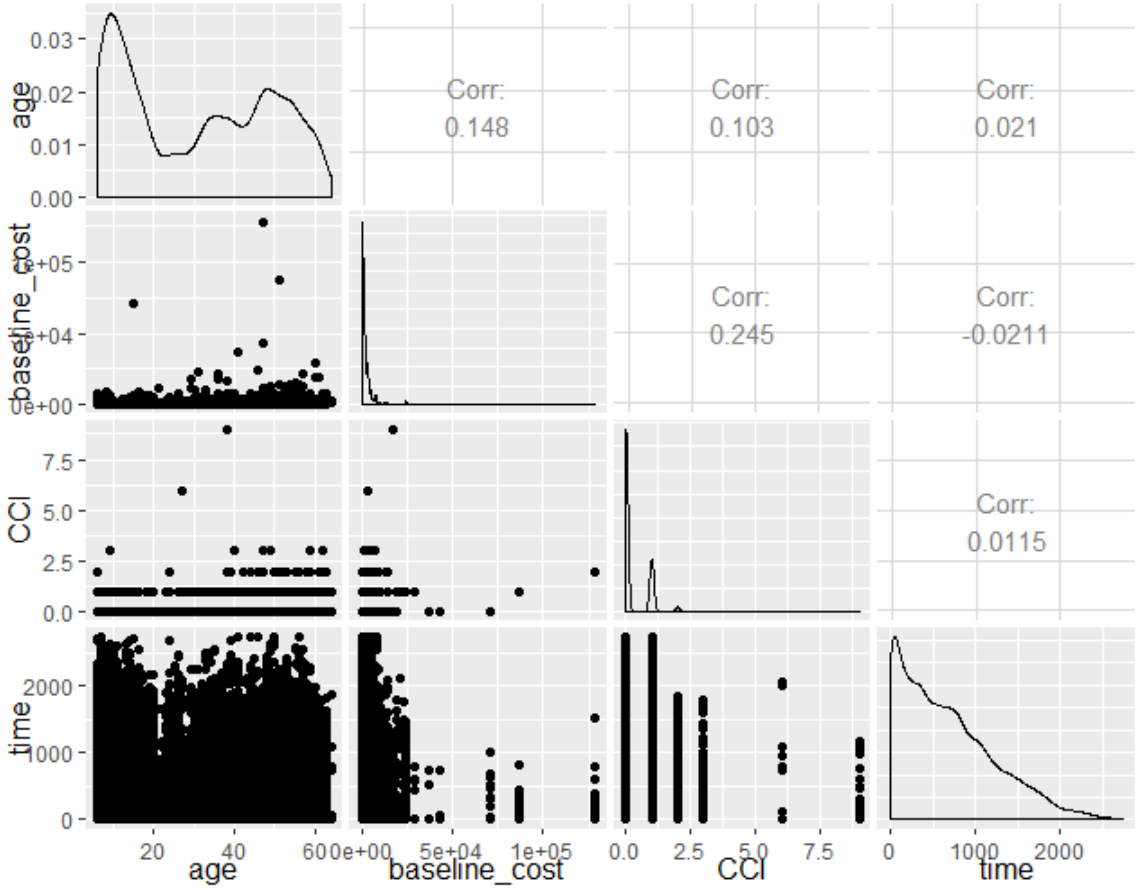


Figure 2: Histogram of Continuous Variables by Event as Binary

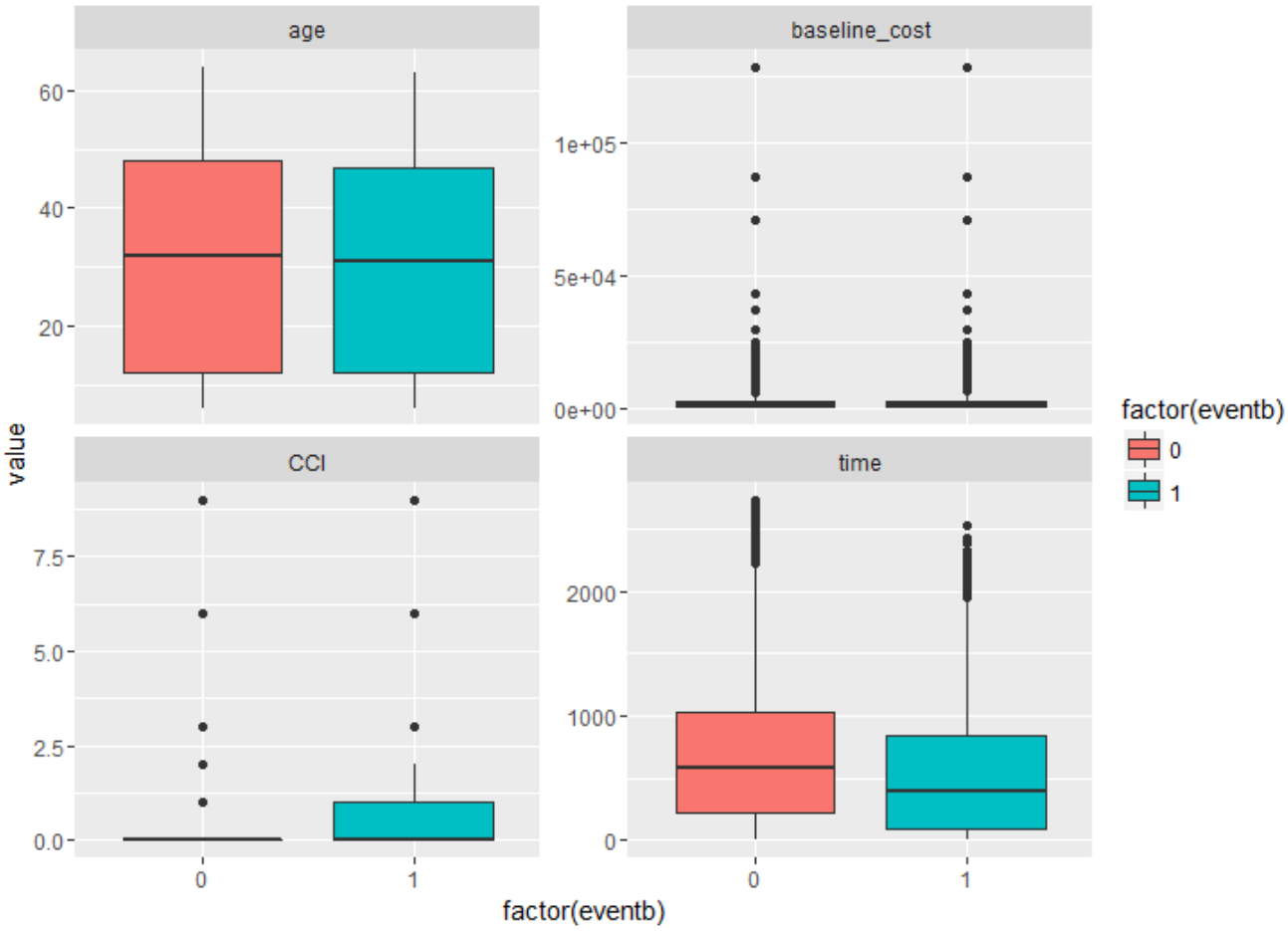


Figure 3: Bubble Diagram of Treatment Step by Event as Binary

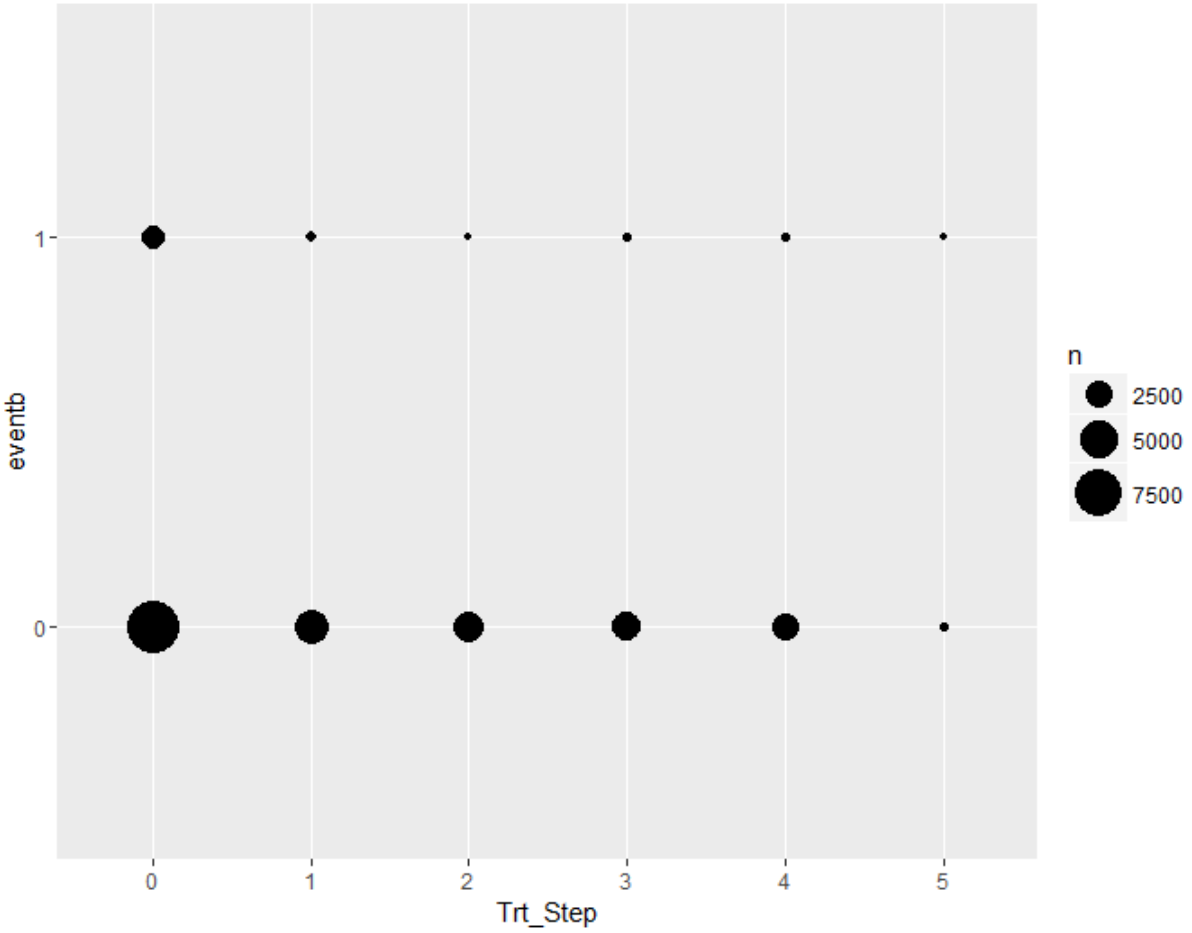
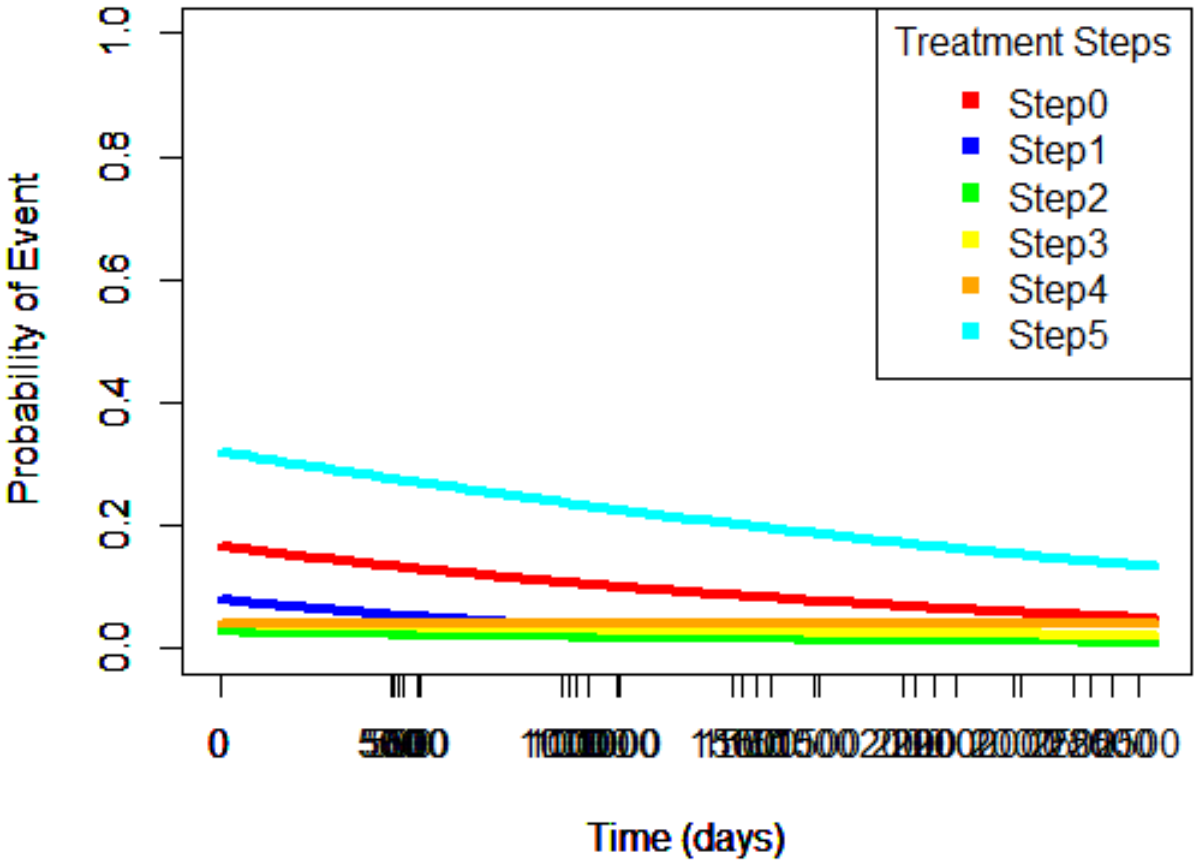


Figure 4: Probability of Event by Time



G Appendix

G.1 Description of Data

There are currently 3 data sets.

- Patient characteristic file
- Exposure file
- Outcome file

Study period: January 2006 Ðecember 2013

Population: All patients with asthma within most recent 10% random sample database housed by CePOR

Patient entry criteria:

1. evidence of asthma at least two recorded diagnosis for asthma (icd 9:493.xx) within 1 year.
2. having at least 24 months continuous eligibility after index date (follow-up period) and at least 6 months continuous eligibility before index date (pre period) environment.
3. Age 6 64 at index.

Patient exclusion criteria:

1. diagnosed with cystic fibrosis (ICD 9; 277.0x) or chronic obstructive pulmonary diseases (491.xx, 492.xx, 494.xx and 496.xx) or respiratory tract cancer (160.xx 164.xx or 231.xx) or bronchopulmonary dysplasia (770.7x) or respiratory distress syndrome (769.xx).
2. one of the following diagnoses: Addison disease (255.4x), glomerulonephritis (580.xx to 582.xx), multiple sclerosis (340.xx), polymyositis/dermatomyositis (710.3x, 710.4x), rheumatoid arthritis (714.xx), scleroderma (710.1x), Sjogren disease (710.2x), systemic lupus erythematosus (710.0x), uveitis (360.11, 363.20 364.3x), vitiligo (709.01), Wegener granulomatosis (446.4x), Primary systemic vasculitis (447.6x), Crohn?s disease (555.0x to 555.2x,

555.9x), Ulcerative colitis (556.0x to 556.6x, 556.8x, or 556.9x), Chronic eosinophilic pneumonia (518.3x), Idiopathic pulmonary fibrosis (516.3x or 515.xx), minimal change disease (581.3x), autoimmune hepatitis (571.42), Myasthenia Gravis (358.0x), Muscular dystrophy (359.0x, 359.1x, or 359.21), Still?s disease (714.2x), Churg Strauss syndrome (446.4x), Polymyalgia rheumatica (725.xx)

Index date: The first date of asthma diagnosis with 6 months eligibility before the index date (NOTE: asthma diagnosis can occur prior to the index date, however, these diagnoses would not have 6 months eligibility prior to the dates.)

G.2 Variables

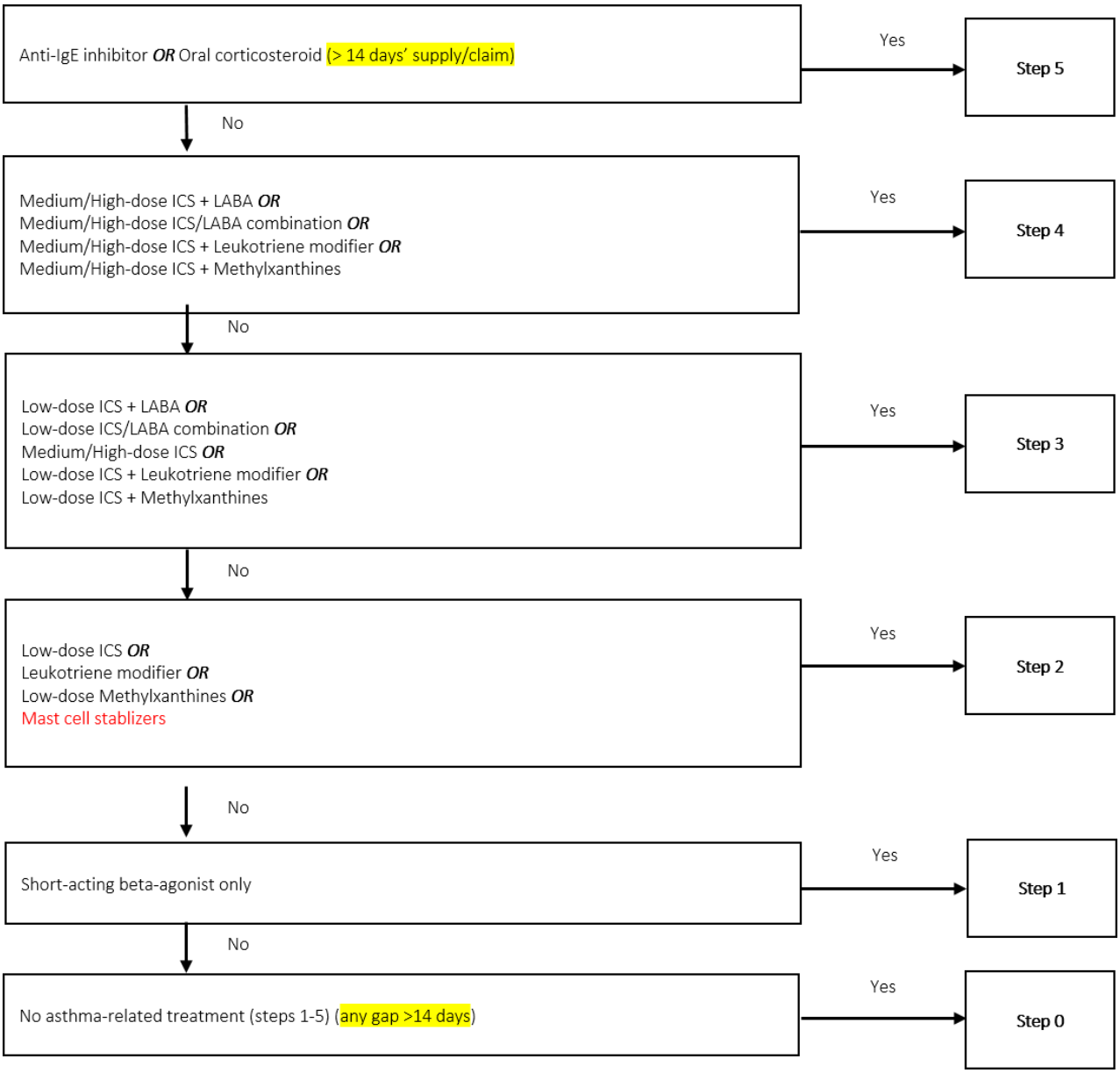
Outcomes: The primary outcome of interest is a nominal variable that includes:

1. No event.
2. Asthma-related hospitalization.
3. Asthma-related emergency department visit.
4. Asthma-related outpatient exacerbation.
5. Outpatient visit for lower respiratory tract infection treated with antibiotics.

(See codes and definition in excel) For patients who have at least one event, we would like to have data on type of event in numeric format and number of days after the index date in numeric format. For those patients with at least one event, only the event category that are observed need to be included i.e. patients who have hospitalization and outpatient exacerbation would only have event 1 and 3 included in the outcome file. For patient who do NOT have any events during the follow-up time, we would like to include them and have event = 0 and date = missing.

Exposure Variable: The primary exposure variable is treatment steps which is an ordinal variable characterized as:

Figure 5


Anti-IgE inhibitor OR Oral corticosteroid (> 14 days' supply/claim)
Yes
Step 5
No
Medium/High-dose ICS + LABA OR
Medium/High-dose ICS/LABA combination OR
Medium/High-dose ICS + Leukotriene modifier OR
Medium/High-dose ICS + Methylxanthines
Yes
Step 4
No
Low-dose ICS + LABA OR
Low-dose ICS/LABA combination OR
Medium/High-dose ICS OR
Low-dose ICS + Leukotriene modifier OR
Low-dose ICS + Methylxanthines
Yes
Step 3
No
Low-dose ICS OR
Leukotriene modifier OR
Low-dose Methylxanthines OR
Mast cell stablizers
Yes
Step 2
No
Short-acting beta-agonist only
Yes
Step 1
No
No asthma-related treatment (steps 1-5) (any gap >14 days)
Yes
Step 0