

# Descriptive Analysis of Cohort

## Contents

Basic Statistics	1
How long until people switch?	2
How old are patients when they start taking an antihypertensive?	3
How many antihypertensives are these people on?	4
When do people get their blood pressure taken?	5
Blood Pressure Visist Counts by Switchers vs. Non-Switchers	9
Diagnoses Codes Prior to First Prescription of a Single Antihypertensive	13
How do new diagnoses vary after the first prescription?	17
Number of Drug Orders	20
Blood Pressure Values	21

## Basic Statistics

```
cohort = qfetch(db, "select * from cohort")
cohort$switched = as.logical(cohort$switched)
derp_refresh = TRUE
```

```
dim(cohort)
```

```
## [1] 4349 6
```

```
summary(cohort)
```

```
##   patient_id      t0      tf      switched
## Min.   : 894   Min.   : 5660   Min.   : 7578   Mode :logical
## 1st Qu.: 516678 1st Qu.:17868   1st Qu.:18534  FALSE:3524
## Median :1007675 Median :21819   Median :22905  TRUE :825
## Mean   :1011136 Mean   :21523   Mean   :22406  NA's :0
## 3rd Qu.:1502858 3rd Qu.:25418   3rd Qu.:26354
## Max.   :2010926 Max.   :32461   Max.   :32872
##                                     NA's   :3524
## grp          tdx
```

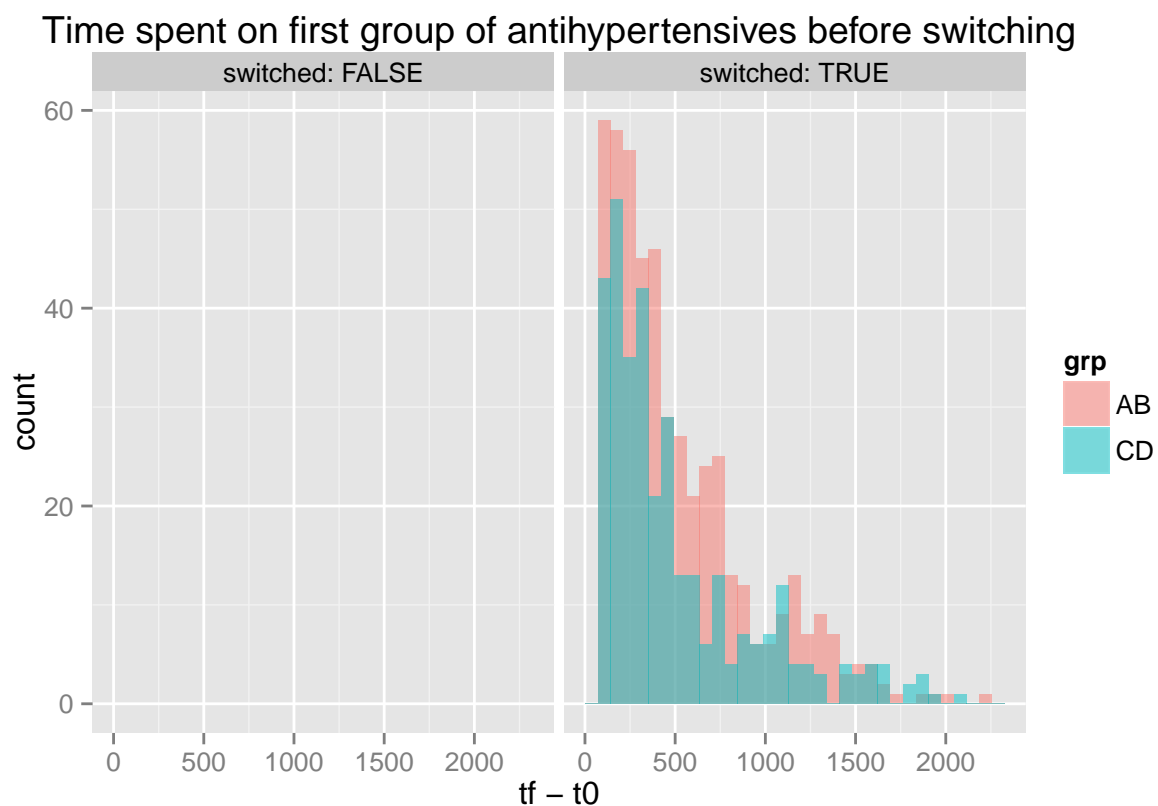
```
## AB:2983   Min.    : 5660
## CD:1366   1st Qu.:17834
##           Median  :21769
##           Mean    :21476
##           3rd Qu.:25361
##           Max.    :32434
##
```

```
table(cohort$switched, cohort$grp)
```

```
##
##           AB   CD
## FALSE 2493 1031
##  TRUE   490  335
```

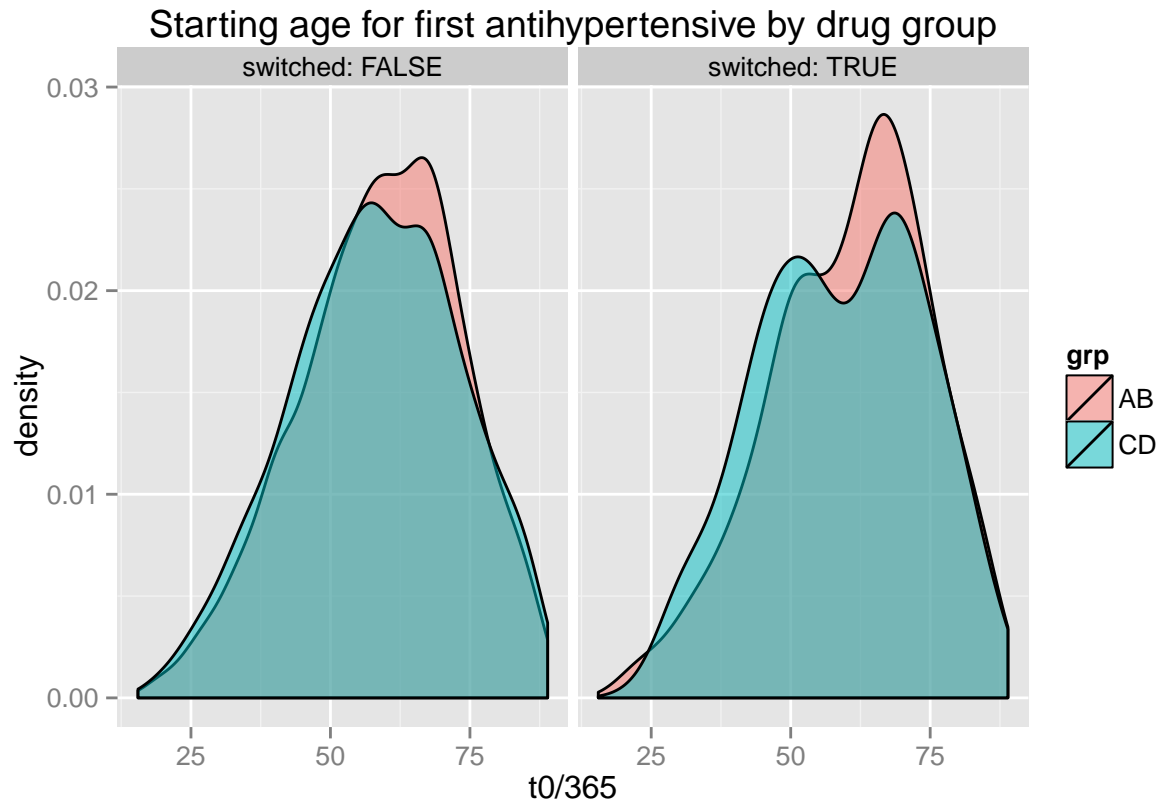
## How long until people switch?

```
ggplot(cohort, aes(x=tf-t0, fill=grp)) +
  geom_histogram(alpha=.5, position="identity") +
  facet_grid(. ~ switched, labeller='label_both') +
  ggtitle("Time spent on first group of antihypertensives before switching")
```



## How old are patients when they start taking an antihypertensive?

```
ggplot(cohort, aes(x=t0/365, fill=grp)) +  
  geom_density(alpha=.5, position="identity") +  
  facet_grid(. ~ switched, labeller='label_both') +  
  ggtitle("Starting age for first antihypertensive by drug group")
```



```
t.test(cohort$t0[cohort$switched], cohort$t0[!cohort$switched])
```

```
##  
## Welch Two Sample t-test  
##  
## data: cohort$t0[cohort$switched] and cohort$t0[!cohort$switched]  
## t = 2.2177, df = 1246.2, p-value = 0.02676  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 52.24936 853.73312  
## sample estimates:  
## mean of x mean of y  
## 21890.18 21437.19
```

A non-significant difference in age- switchers are about 9 mo older when first prescribed.

## How many antihypertensives are these people on?

```
kable(qfetch(db, "
SELECT COUNT(patient_id) AS n_patients, n AS n_initial_drugs FROM (
SELECT COUNT(DISTINCT o0.ingr_set_id) AS n, patient_id FROM (
SELECT o.*
from cohort c
JOIN ht_rx o
  ON c.patient_id = o.patient_id
  AND c.t0 = o.age_at_start_time_in_days
) o0
GROUP BY o0.patient_id) cnt
GROUP BY cnt.n
"), caption="Number of different antihypertensives at first antihypertensive perscription")
```

Table 1: Number of different antihypertensives at first antihypertensive perscription

n_patients	n_initial_drugs
4305	1
266	2
5	3

```
kable(qfetch(db, "
SELECT n_drugs, COUNT(DISTINCT patient_id) AS n_patients
FROM
  (SELECT COUNT(DISTINCT ingr_set_id) AS n_drugs, patient_id
  from ht_rx
  GROUP BY patient_id
  ) AS inter
GROUP BY n_drugs
ORDER BY n_drugs
"), caption="Number of different antihypertensives taken over lifetime")
```

Table 2: Number of different antihypertensives taken over lifetime

n_drugs	n_patients
1	2475
2	1325
3	488
4	180
5	68
6	21
7	14
8	3
9	2

```

kable(qfetch(db, "
SELECT n_drugs, COUNT(DISTINCT patient_id) AS n_patients
FROM
  (SELECT COUNT(DISTINCT o.ingr_set_id) AS n_drugs, o.patient_id
   from ht_rx o
   join cohort c
   WHERE o.age_at_start_time_in_days < c.tf
   GROUP BY o.patient_id
  ) AS inter
GROUP BY n_drugs
ORDER BY n_drugs
"), caption="Number of different antihypertensives taken over first group window")

```

Table 3: Number of different antihypertensives taken over first group window

n_drugs	n_patients
1	2477
2	1327
3	487
4	179
5	67
6	20
7	14
8	3
9	2

## When do people get their blood pressure taken?

Takeaway: Most measurements happen right before the first drug and about 180 days after the first drug, consistent with the behavior we would expect from physicians diagnosing high blood pressure and prescribing an antihypertensive.

```

bp = merge(qfetch(db, "select * from bp"), cohort, by='patient_id')
bp$patient_id = factor(bp$patient_id)
bp = group_by(bp, patient_id, grp, switched, t0)

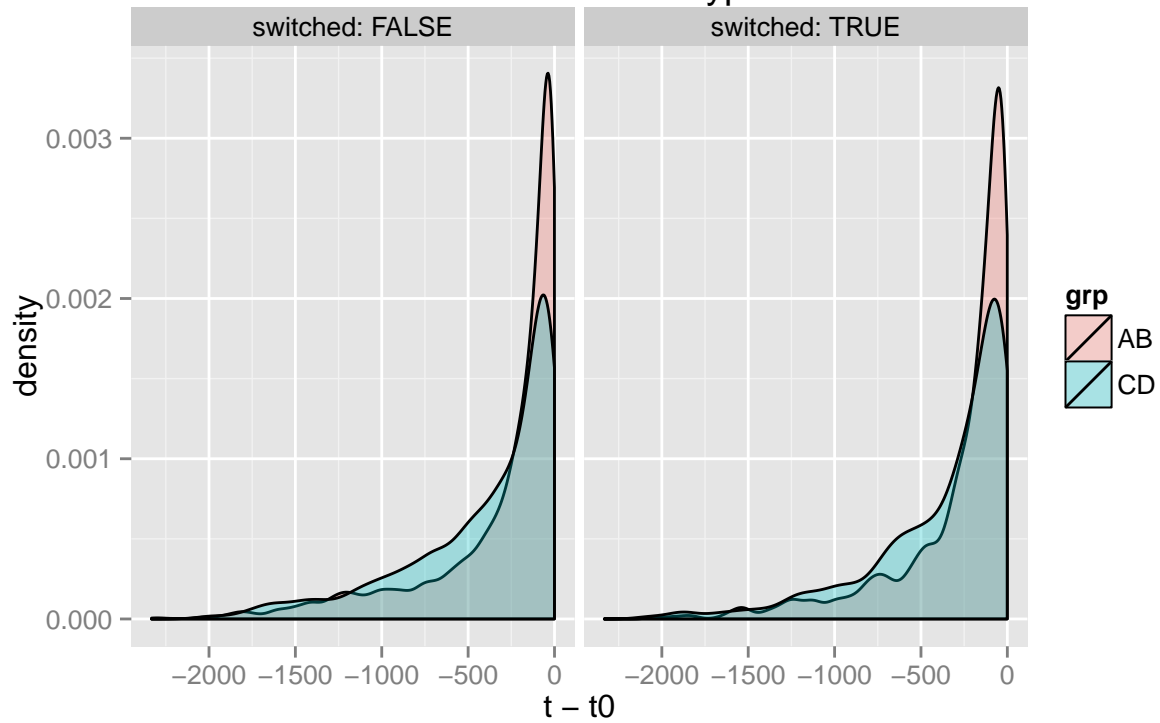
```

```

ggplot(bp[bp$t<bp$t0,], aes(x=t-t0, fill=grp)) +
  geom_density(position="identity", alpha=0.3) +
  facet_grid(. ~ switched, labeller='label_both') +
  ggtitle("Times of all previous blood pressure measurements\nrelative to time of first antihypertensive")

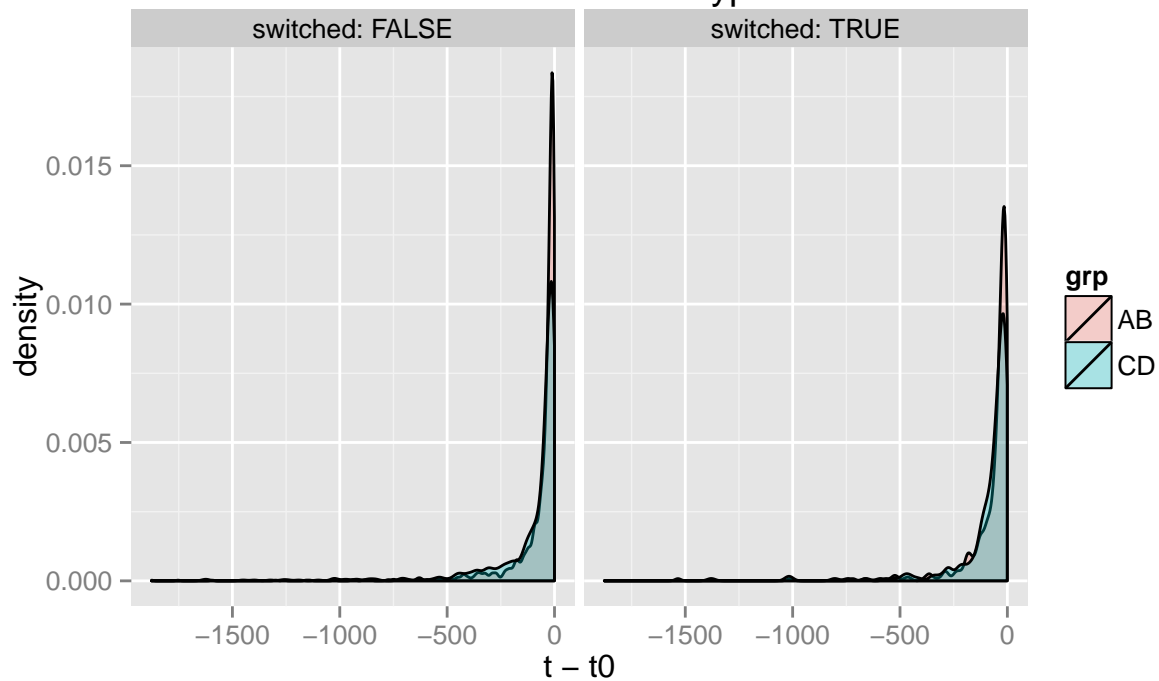
```

## Times of all previous blood pressure measurements relative to time of first antihypertensive



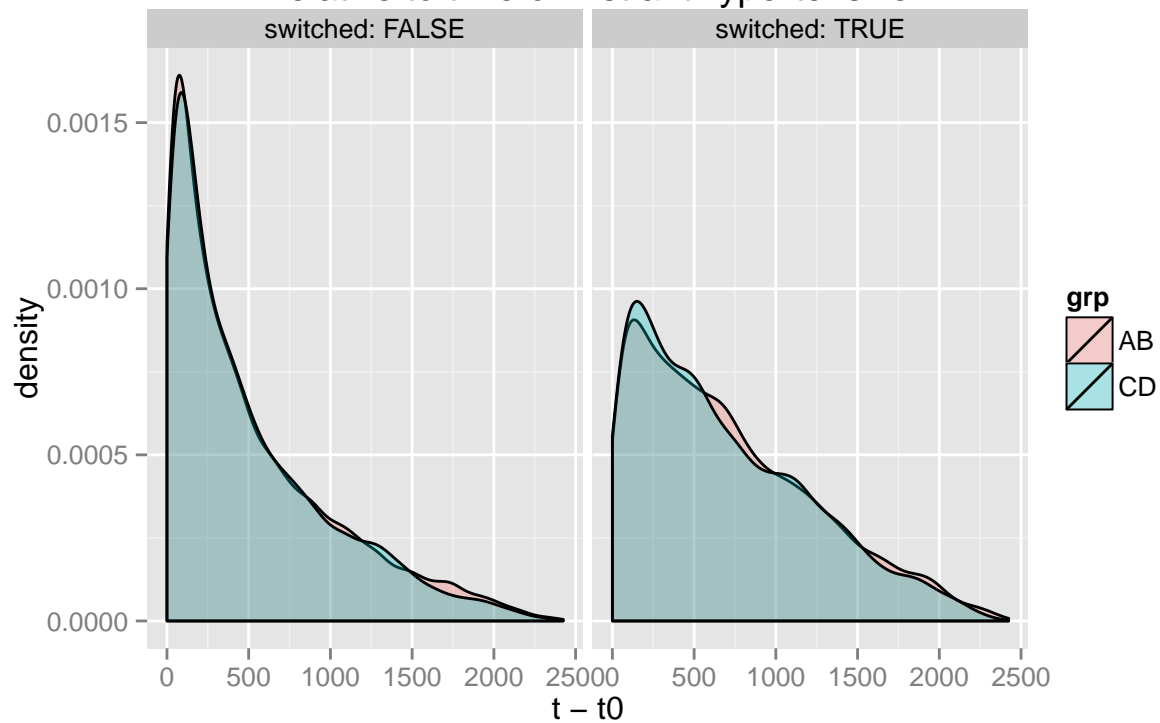
```
bp_latest_prior = summarize(bp[bp$t < bp$t0,], t=max(t))
ggplot(bp_latest_prior, aes(x=t-t0, fill=grp)) +
  geom_density(position="identity", alpha=0.3) +
  facet_grid(. ~ switched, labeller='label_both') +
  ggtitle("Times of latest blood pressure measurement prior to the\nfirst perscription of a single anti")
```

Times of latest blood pressure measurement prior to the first perscription of a single antihypertensive for each patient, relative to time of first antihypertensive



```
ggplot(bp[bp$t>bp$t0,], aes(x=t-t0, fill=grp)) +  
  geom_density(position="identity", alpha=0.3) +  
  facet_grid(. ~ switched, labeller='label_both') +  
  ggtitle("Times of all subsequent blood pressure measurements\n relative to time of first antihyperten")
```

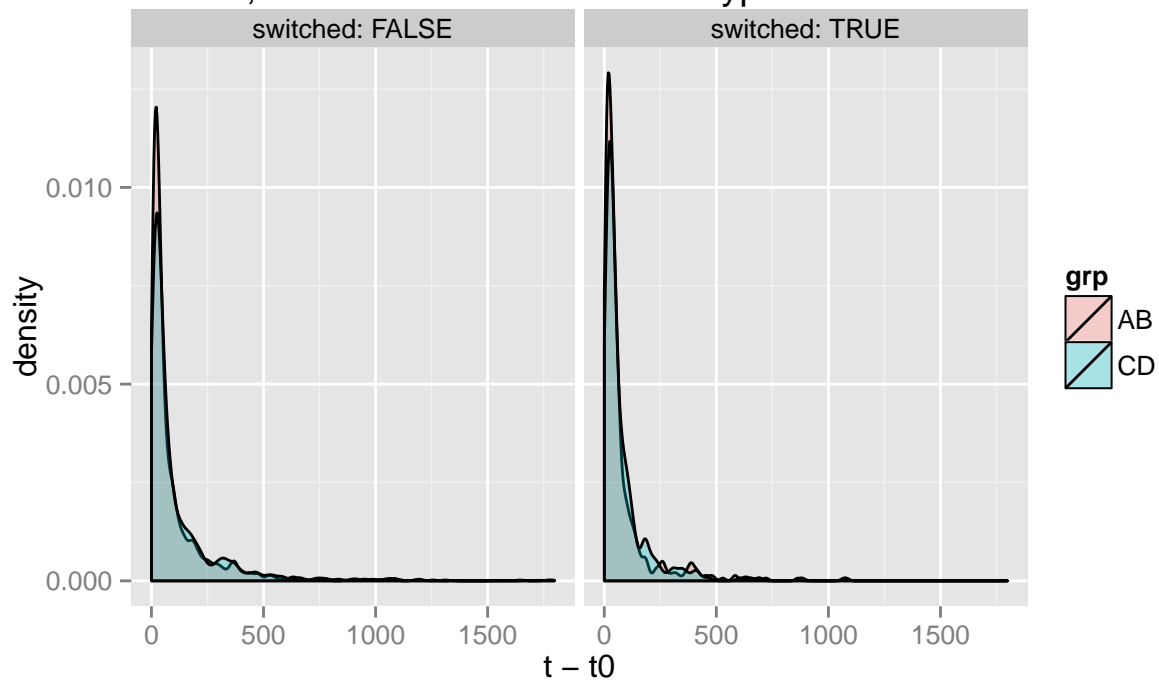
## Times of all subsequent blood pressure measurements relative to time of first antihypertensive



```
bp_earliest_subsequent = summarize(bp[bp$t>bp$t0,], t=min(t))
ggplot(bp_earliest_subsequent, aes(x=t-t0, fill=grp)) +
  geom_density(position="identity", alpha=0.3) +
  facet_grid(. ~ switched, labeller='label_both') +
  ggtitle("Times of earliest blood pressure measurement subsequent\n to the first perscription of a sing")
```

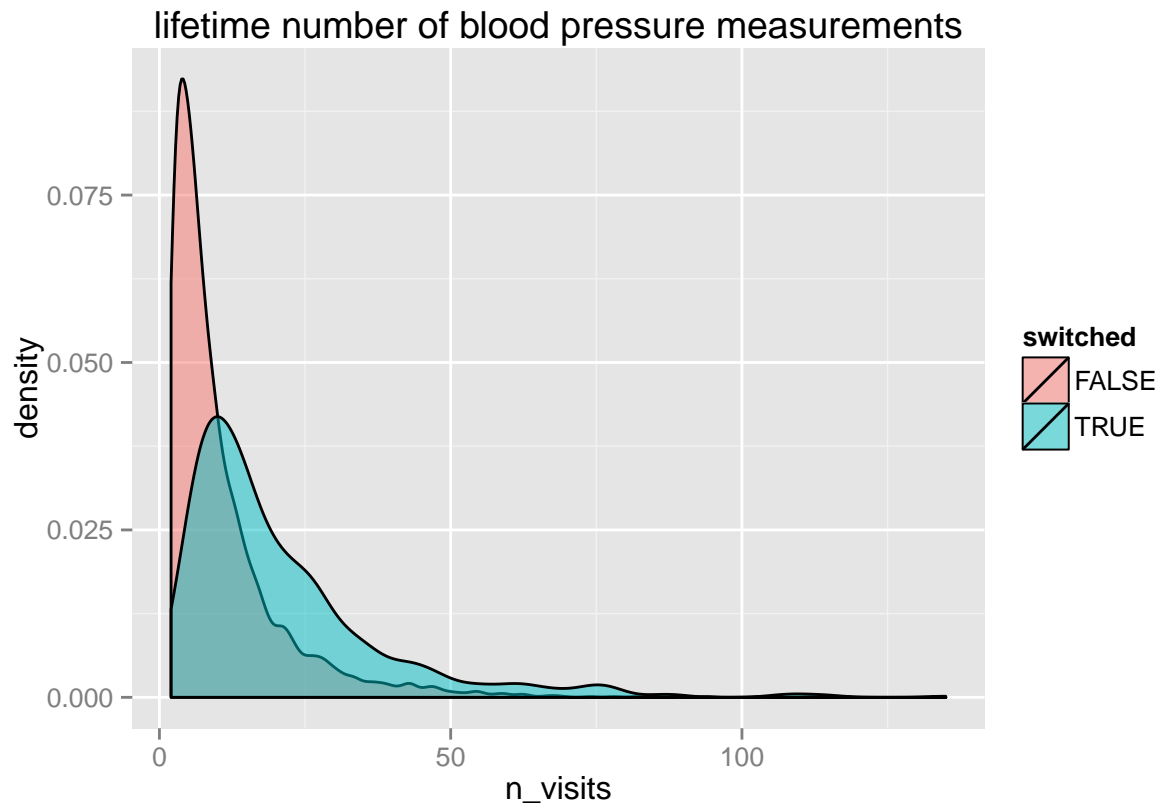


Times of earliest blood pressure measurement subsequent to the first perscription of a single antihypertensive for each patient , relative to time of first antihypertensive



## Blood Pressure Visist Counts by Switchers vs. Non-Switchers

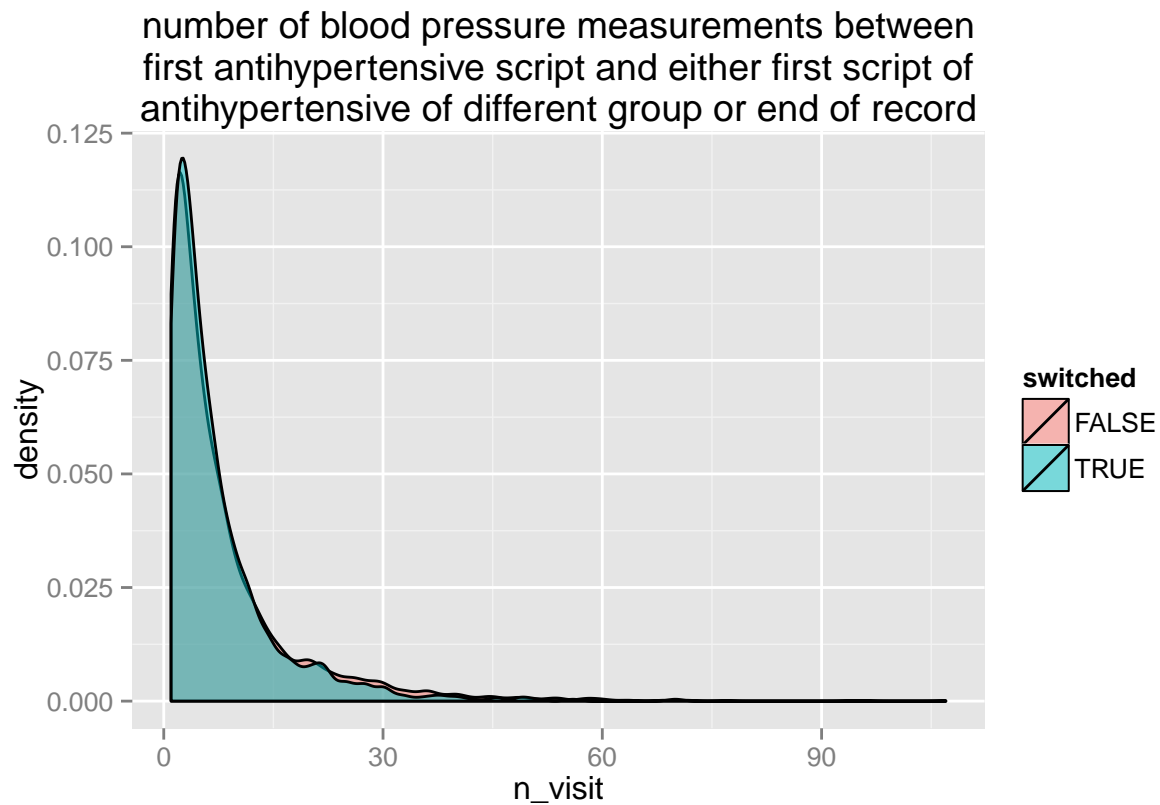
```
n_bp_visits = summarize(bp, n_visits=n())
ggplot(n_bp_visits, aes(x=n_visits, fill=switched)) +
  geom_density(alpha=0.5, position='identity') +
  ggtitle("lifetime number of blood pressure measurements")
```



Turns out that people who switch have more visits during their lifetimes where they get their blood pressure taken...

```

window_bp_visits = summarize(filter(bp, t>t0&(is.na(tf)|(t<tf))), n_visit=n())
ggplot(window_bp_visits, aes(x=n_visit, fill=switched)) +
  geom_density(alpha=0.5, position='identity') +
  ggtitle("number of blood pressure measurements between\nfirst antihypertensive script and either first
  
```



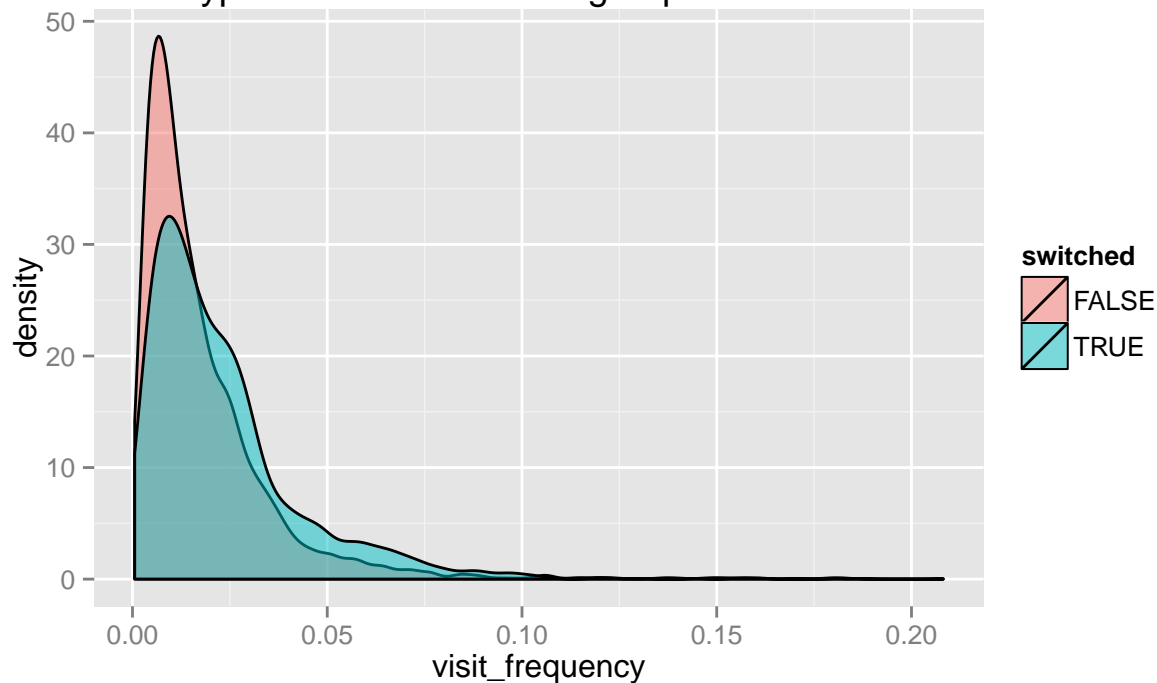
Within their windows, the number of visits are about the same, but this is deceptive because the switchers have much, much shorter windows and so they have a much greater density of visits right after being prescribed their first antihypertensive.

```

window_bp_visits = summarize(filter(bp, t>t0&(is.na(tf)|t<tf)), visit_frequency=n()/(max(t)-first(t0))
ggplot(window_bp_visits, aes(x=visit_frequency, fill=switched)) +
  geom_density(alpha=0.5, position='identity') +
  ggtitle("time density of blood pressure measurements between\nfirst antihypertensive script and either")

```

time density of blood pressure measurements between first antihypertensive script and either first script of antihypertensive of different group or end of record

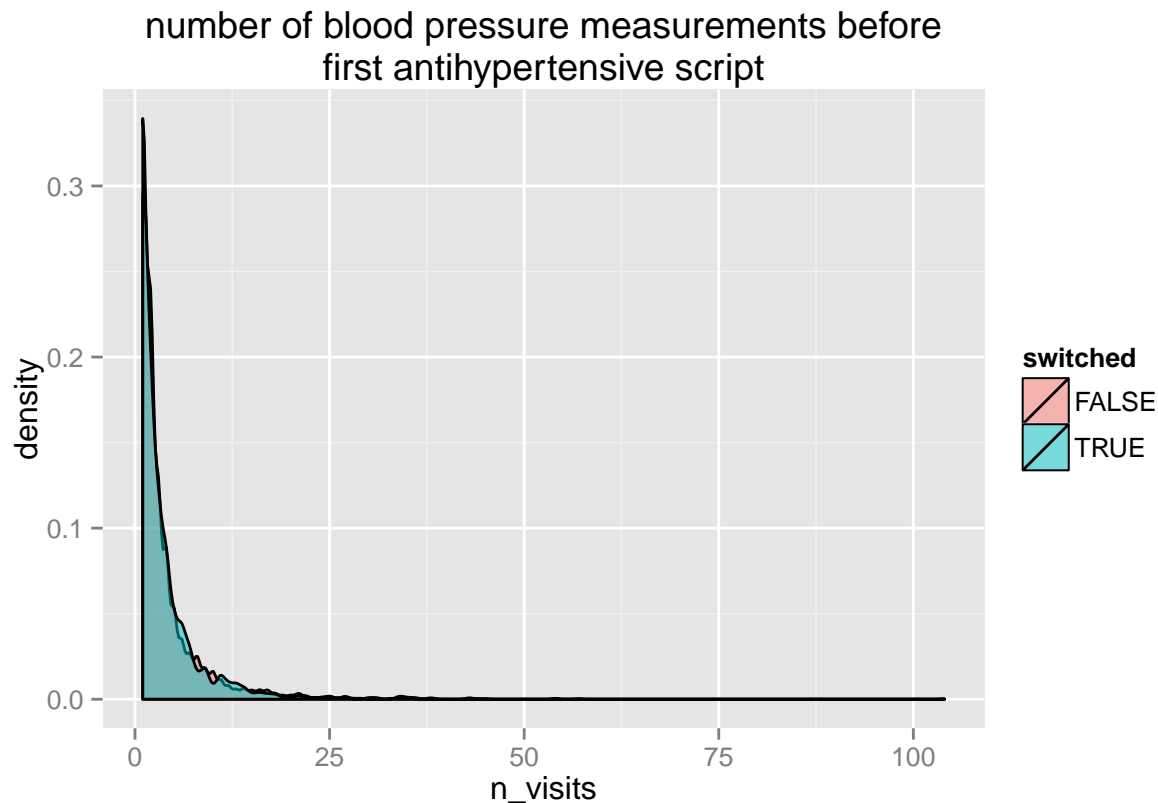


```
wilcox.test(window_bp_visits$visit_frequency[window_bp_visits$switched],
            window_bp_visits$visit_frequency[!window_bp_visits$switched])
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: window_bp_visits$visit_frequency[window_bp_visits$switched] and window_bp_visits$visit_frequen
## W = 2052500, p-value < 2.2e-16
## alternative hypothesis: true location shift is not equal to 0
```

Here we see that the frequency of visits from the time of first antihypertension perscription to the time of last visit (for non-switchers) or to the time of switching (for switchers) is different between switchers and non-switchers. It could be the case that more of the switchers are better adherents and go in more often and are thus more likeley to be switched.

```
pre_bp_visits = summarize(filter(bp, t<t0), n_visits=n())
ggplot(pre_bp_visits, aes(x=n_visits, fill=switched)) +
  geom_density(alpha=0.5, position='identity') +
  ggtitle("number of blood pressure measurements before\nfirst antihypertensive script")
```



```
wilcox.test(pre_bp_visits$n_visits[pre_bp_visits$switched],
            pre_bp_visits$n_visits[!pre_bp_visits$switched])
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: pre_bp_visits$n_visits[pre_bp_visits$switched] and pre_bp_visits$n_visits[!pre_bp_visits$switched]
## W = 1718600, p-value = 0.8136
## alternative hypothesis: true location shift is not equal to 0
```

It's not possible to tell the two groups apart on the basis of visit count prior to their first antihypertensive script.

## Diagnoses Codes Prior to First Prescription of a Single Antihypertensive

```
dbSendQuery(db, "
  create table if not exists _dx_pre
  (index(patient_id), index(code), index(grp), index(switched))
  select distinct cohort.patient_id, cohort.grp, cohort.switched,
    codes.code
  from codes
  join cohort
  on codes.patient_id = cohort.patient_id
```

```

where codes.t <= cohort.t0 + 1
AND codes.ctype != 'CPT'
")

```

```
## <MySQLResult:-326921224,0,0>
```

```

count_analyse = function(df, N1, N2) {
  df$odds_ratio = (df[,1]/(N1-df[,1]))/(df[,2]/(N2-df[,2]))
  df$log_odds = log(df$odds_ratio)
  se = sqrt( 1/df[,1] + 1/(N1-df[,1]) + 1/df[,2] + 1/(N2-df[,2]) )
  df$p = p.adjust(2*pnorm(abs(df$log_odds)/se, lower.tail=F), method="BH")
  df$str = sapply(df$code, icd9Explain, doCondense=FALSE, brief=T)

  result=list()

  df = df[order(df[,1]+df[,2], decreasing=T),]
  row.names(df) = NULL
  result$n = df[1:20, ]

  df = df[order(df$p),]
  row.names(df) = NULL
  result$or = df[1:20, ]

  return(result)
}

```

```

dx_grp = qfetch(db, "
select dx_ab.n as n_ab, dx_cd.n as n_cd, dx_ab.code from (
  SELECT COUNT(patient_id) n, code
  FROM _dx_pre
  WHERE grp = 'AB'
  GROUP BY code) dx_ab
join (
  SELECT COUNT(patient_id) n, code
  FROM _dx_pre
  WHERE grp = 'CD'
  GROUP BY code) dx_cd
on dx_ab.code = dx_cd.code
WHERE dx_ab.n > 10
AND dx_cd.n > 10
")
N_ab = sum(cohort$grp=='AB')
N_cd = sum(cohort$grp=='CD')

result = count_analyse(dx_grp, N_ab, N_cd)
kable(result$n, caption="Diagnoses before first antihypertensive prescription, ordered by total patients affected")

```

Table 4: Diagnoses before first antihypertensive prescription, ordered by total patients affected

n_ab	n_cd	code	odds_ratio	log_odds	p	str
2907	1344	401.9	0.6261161	-0.4682195	0.1601070	Hypertension NOS

n_ab	n_cd	code	odds_ratio	log_odds	p	str
1187	391	272.4	1.6480571	0.4995971	0.0000000	Hyperlipidemia NEC/NOS
728	517	V70.0	0.5301539	-0.6345879	0.0000000	Routine medical exam
717	121	250.00	3.2556914	1.1804047	0.0000000	DMII wo cmp nt st uncntr
377	160	530.81	1.0904211	0.0865640	0.5302859	Esophageal reflux
346	175	786.50	0.8929758	-0.1131958	0.4004481	Chest pain NOS
373	134	V72.84	1.3139361	0.2730273	0.0577104	Preop exam unspcf
386	59	414.00	3.2925997	1.1916774	0.0000000	Cor ath unsp vsl ntv/gft
289	136	311	0.9702116	-0.0302411	0.8605387	Depressive disorder, not elsewhere classified
281	121	278.00	1.0700522	0.0677074	0.6786204	Obesity NOS
255	145	719.46	0.7871246	-0.2393687	0.1096186	Joint pain-l/leg
253	146	724.2	0.7743991	-0.2556679	0.0853415	Lumbago
257	129	244.9	0.9040392	-0.1008826	0.5184800	Hypothyroidism NOS
202	177	V76.12	0.4879316	-0.7175800	0.0000000	Screen mammogram NEC
237	137	780.79	0.7742463	-0.2558653	0.0940710	Malaise and fatigue NEC
221	153	786.2	0.6343632	-0.4551336	0.0007547	Cough
196	140	729.5	0.6158593	-0.4847367	0.0006302	Pain in limb
204	120	V76.51	0.7622166	-0.2715245	0.0951880	Screen malig neop-colon
198	123	477.9	0.7184656	-0.3306374	0.0351411	Allergic rhinitis NOS
205	109	300.00	0.8510016	-0.1613412	0.3283476	Anxiety state NOS

`kable(result$or, caption="Diagnoses before first antihypertensive prescription, ordered by adjusted p-value")`

Table 5: Diagnoses before first antihypertensive prescription, ordered by adjusted p-value

n_ab	n_cd	code	odds_ratio	log_odds	p	str
717	121	250.00	3.2556914	1.1804047	0.0000000	DMII wo cmp nt st uncntr
728	517	V70.0	0.5301539	-0.6345879	0.0000000	Routine medical exam
386	59	414.00	3.2925997	1.1916774	0.0000000	Cor ath unsp vsl ntv/gft
1187	391	272.4	1.6480571	0.4995971	0.0000000	Hyperlipidemia NEC/NOS
202	177	V76.12	0.4879316	-0.7175800	0.0000000	Screen mammogram NEC
140	125	692.9	0.4888920	-0.7156136	0.0000016	Dermatitis NOS
146	17	250.02	4.0837256	1.4070097	0.0000034	DMII wo cmp uncntrld
74	77	466.0	0.4258437	-0.8536830	0.0000162	Acute bronchitis
82	81	611.72	0.4484192	-0.8020267	0.0000285	Lump or mass in breast
113	100	733.90	0.4984599	-0.6962321	0.0000391	Bone & cartilage dis NOS
158	127	780.52	0.5456400	-0.6057958	0.0000394	Insomnia NOS
143	118	796.2	0.5325376	-0.6301018	0.0000394	Elev bl pres w/o hypertn
111	13	424.1	4.0224716	1.3918965	0.0000846	Aortic valve disorder
36	46	E849.9	0.3505407	-1.0482784	0.0001048	Accident in place NOS
96	85	V72.31	0.5011349	-0.6908799	0.0001862	Routine gyn examination
158	31	278.01	2.4085641	0.8790308	0.0002960	Morbid obesity
84	76	V76.2	0.4918211	-0.7096402	0.0003150	Screen mal neop-cervix
159	121	V04.81	0.5793172	-0.5459051	0.0003419	Vaccin for influenza
133	24	428.0	2.6094444	0.9591373	0.0004608	CHF NOS
164	122	784.0	0.5932112	-0.5222048	0.0006156	Headache

This reflects current guidelines. Ask experts.

```

dx_switch = qfetch(db, "
select dx_switch.n as n_switch, dx_stay.n as n_stay, dx_switch.code from (
  SELECT COUNT(patient_id) n, code
  FROM _dx_pre
  WHERE switched = TRUE
  GROUP BY code) dx_switch
join (
  SELECT COUNT(patient_id) n, code
  FROM _dx_pre
  WHERE switched = FALSE
  GROUP BY code) dx_stay
on dx_switch.code = dx_stay.code
WHERE dx_switch.n > 10
AND dx_stay.n > 10
")
N_switch = sum(cohort$switched)
N_stay = sum(!cohort$switched)

result=count_analyse(dx_switch, N_switch, N_stay)
kable(result$n, caption="Diagnoses before first antihypertensive prescription, ordered by total patients affected")

```

Table 6: Diagnoses before first antihypertensive prescription, ordered by total patients affected

n_switch	n_stay	code	odds_ratio	log_odds	p	str
810	3441	401.9	1.3025283	0.2643072	0.7450550	Hypertension NOS
330	1248	272.4	1.2158120	0.1954121	0.1925754	Hyperlipidemia NEC/NOS
249	996	V70.0	1.0972222	0.0927817	0.6801440	Routine medical exam
164	674	250.00	1.0491253	0.0479568	0.9049207	DMII wo cmp nt st uncnt
104	433	530.81	1.0296964	0.0292640	0.9433726	Esophageal reflux
119	402	786.50	1.3090285	0.2692853	0.1925754	Chest pain NOS
74	433	V72.84	0.7034009	-0.3518283	0.1699581	Preop exam unspcf
83	362	414.00	0.9770741	-0.0231928	0.9613448	Cor ath unsp vsl ntv/gft
81	344	311	1.0064235	0.0064029	0.9865350	Depressive disorder, not elsewhere classified
90	312	278.00	1.2605965	0.2315851	0.3377425	Obesity NOS
91	309	719.46	1.2899350	0.2545918	0.2864824	Joint pain-l/leg
73	326	724.2	0.9522827	-0.0488934	0.9144868	Lumbago
69	317	244.9	0.9233514	-0.0797454	0.8866058	Hypothyroidism NOS
89	290	V76.12	1.3485101	0.2990004	0.1925754	Screen mammogram NEC
86	288	780.79	1.3075853	0.2681822	0.2661826	Malaise and fatigue NEC
94	280	786.2	1.4898183	0.3986541	0.1265790	Cough
74	262	729.5	1.2268019	0.2044107	0.4727918	Pain in limb
72	252	V76.51	1.2415101	0.2163285	0.4543865	Screen malig neop-colon
62	259	477.9	1.0243552	0.0240633	0.9637103	Allergic rhinitis NOS
47	267	300.00	0.7369275	-0.3052658	0.3283020	Anxiety state NOS

```

kable(result$or, caption="Diagnoses before first antihypertensive prescription, ordered by adjusted p-value")

```



Table 7: Diagnoses before first antihypertensive prescription, ordered by adjusted p-value

n_switch	n_stay	code	odds_ratio	log_odds	p	str
15	14	593.2	4.6428571	1.5353299	0.0102126	Cyst of kidney, acquired
94	280	786.2	1.4898183	0.3986541	0.1265790	Cough
74	211	780.52	1.5471441	0.4364107	0.1265790	Insomnia NOS
11	13	369.9	3.6496881	1.2946417	0.1265790	Visual loss NOS
21	41	729.81	2.2188751	0.7970004	0.1692590	Swelling of limb
74	433	V72.84	0.7034009	-0.3518283	0.1699581	Preop exam unspcf
25	184	719.45	0.5672554	-0.5669456	0.1699581	Joint pain-pelvis
49	137	401.1	1.5610938	0.4453867	0.1699581	Benign hypertension
45	118	250.02	1.6652542	0.5099778	0.1699581	DMII wo cmp uncntrld
43	114	211.3	1.6447929	0.4976145	0.1699581	Benign neoplasm lg bowel
31	76	274.9	1.7713111	0.5717200	0.1699581	Gout NOS
23	51	682.9	1.9529363	0.6693340	0.1699581	Cellulitis NOS
19	39	784.2	2.1064771	0.7450169	0.1699581	Swelling in head & neck
13	22	537.89	2.5484774	0.9354961	0.1699581	Gastroduodenal dis NEC
330	1248	272.4	1.2158120	0.1954121	0.1925754	Hyperlipidemia NEC/NOS
119	402	786.50	1.3090285	0.2692853	0.1925754	Chest pain NOS
89	290	V76.12	1.3485101	0.2990004	0.1925754	Screen mammogram NEC
73	226	789.00	1.4166000	0.3482596	0.1925754	Abdmnal pain unspcf site
66	199	692.9	1.4529168	0.3735731	0.1925754	Dermatitis NOS
42	115	789.07	1.5900716	0.4637791	0.1925754	Abdmnal pain generalized

There's very little information predicting switching behavior from previous diagnoses.

## How do new diagnoses vary after the first prescription?

```
dbSendQuery(db, "
create table if not exists _dx_post
(index(patient_id), index(code), index(grp), index(switched))
select distinct cohort.patient_id, cohort.grp, cohort.switched,
codes.code
from codes
join cohort
on codes.patient_id = cohort.patient_id
where codes.t > cohort.t0 + 1
and codes.t < cohort.t0 + 365
AND codes.ctype != 'CPT'
")
```

```
## <MySQLResult:1,0,0>
```

```
dbSendQuery(db, "
create table if not exists _dx_post_new
(index(patient_id), index(code), index(grp), index(switched))
select post.*
from _dx_post post
```

```

left join _dx_pre pre
on post.patient_id = pre.patient_id
and post.code = pre.code
where pre.code is NULL
")

```

```
## <MySQLResult:713479368,0,1>
```

```

refer= 1
dx_grp = qfetch(db, "
select dx_ab.n as n_ab, dx_cd.n as n_cd, dx_ab.code from (
  SELECT COUNT(patient_id) n, code
  FROM _dx_post_new
  WHERE grp = 'AB'
  GROUP BY code) dx_ab
join (
  SELECT COUNT(patient_id) n, code
  FROM _dx_post_new
  WHERE grp = 'CD'
  GROUP BY code) dx_cd
on dx_ab.code = dx_cd.code
WHERE dx_ab.n > 10
AND dx_cd.n > 10
")
N_ab = sum(cohort$grp=='AB')
N_cd = sum(cohort$grp=='CD')

result = count_analyse(dx_grp, N_ab, N_cd)
kable(result$or, caption="Diagnoses right after first antihypertensive prescription, ordered by adjusted p-value")

```

Table 8: Diagnoses right after first antihypertensive prescription, ordered by adjusted p-value

n_ab	n_cd	code	odds_ratio	log_odds	p	str
16	22	276.8	0.3294420	-1.1103548	0.0409197	Hypopotassemia
75	16	V45.89	2.1761090	0.7775384	0.0892939	Post-proc states NEC
21	23	692.9	0.4139829	-0.8819306	0.0892939	Dermatitis NOS
61	12	285.9	2.3555213	0.8567621	0.0923049	Anemia NOS
20	20	790.21	0.4542693	-0.7890650	0.1385362	Impaired fasting glucose
95	30	V70.0	1.4649123	0.3817954	0.5020927	Routine medical exam
40	29	790.29	0.6266184	-0.4674175	0.5020927	Abnormal glucose NEC
16	14	790.4	0.5207762	-0.6524350	0.5020927	Elev transaminase/l dh
143	50	272.4	1.3252676	0.2816144	0.5476201	Hyperlipidemia NEC/NOS
53	16	268.9	1.5262372	0.4228054	0.6222867	Vitamin D deficiency NOS
46	13	564.00	1.6300778	0.4886277	0.6222867	Constipation NOS
18	14	782.1	0.5862684	-0.5339776	0.6222867	Nonspecif skin erupt NEC
56	18	786.50	1.4327905	0.3596239	0.6832515	Chest pain NOS
25	17	790.6	0.6706638	-0.3994873	0.6832515	Abn blood chemistry NEC
16	12	216.9	0.6084710	-0.4968061	0.6832515	Benign neoplasm skin NOS
16	12	466.0	0.6084710	-0.4968061	0.6832515	Acute bronchitis
48	16	V72.84	1.3798978	0.3220094	0.7250610	Preop exam unspcf
30	19	784.0	0.7202310	-0.3281833	0.7250610	Headache

n_ab	n_cd	code	odds_ratio	log_odds	p	str
35	11	786.05	1.4624707	0.3801273	0.7250610	Shortness of breath
19	13	789.07	0.6671598	-0.4047257	0.7250610	Abdmnal pain generalized

These are many ADEs that are associated with each drug group. Check with expert.

```
refer=1
dx_switch = qfetch(db, "
select dx_switch.n as n_switch, dx_stay.n as n_stay, dx_switch.code from (
  SELECT COUNT(patient_id) n, code
  FROM _dx_post_new
  WHERE switched = TRUE
  GROUP BY code) dx_switch
join (
  SELECT COUNT(patient_id) n, code
  FROM _dx_post_new
  WHERE switched = FALSE
  GROUP BY code) dx_stay
on dx_switch.code = dx_stay.code
WHERE dx_switch.n > 10
AND dx_stay.n > 10
")
N_switch = sum(cohort$switched)
N_stay = sum(!cohort$switched)

result=count_analyse(dx_switch, N_switch, N_stay)
kable(result$or, caption="Diagnoses right after first antihypertensive prescription, ordered by adjusted p-value")
```

Table 9: Diagnoses right after first antihypertensive prescription, ordered by adjusted p-value

n_switch	n_stay	code	odds_ratio	log_odds	p	str
18	20	276.8	3.907807	1.3629763	0.0012755	Hypopotassemia
18	32	366.16	2.434015	0.8895421	0.0377516	Senile nuclear cataract
11	14	593.9	3.388031	1.2202489	0.0377516	Renal & ureteral dis NOS
22	46	786.2	2.071471	0.7282590	0.0518829	Cough
18	35	719.46	2.223473	0.7990705	0.0518829	Joint pain-l/leg
23	52	530.81	1.914828	0.6496280	0.0542900	Esophageal reflux
21	46	780.79	1.974854	0.6804945	0.0542900	Malaise and fatigue NEC
12	20	789.07	2.585978	0.9501037	0.0542900	Abdmnal pain generalized
17	38	724.2	1.930107	0.6575754	0.0950537	Lumbago
12	23	366.9	2.246751	0.8094852	0.0950537	Cataract NOS
11	20	702.19	2.367568	0.8618631	0.0950537	Other sborheic keratosis
19	45	V72.84	1.822470	0.6001929	0.1020897	Preop exam unspcf
17	42	564.00	1.744283	0.5563438	0.1675355	Constipation NOS
16	39	786.09	1.767297	0.5694515	0.1675355	Respiratory abnorm NEC
11	25	787.02	1.891351	0.6372916	0.2047374	Nausea alone
11	25	787.91	1.891351	0.6372916	0.2047374	Diarrhea
12	29	719.41	1.778852	0.5759684	0.2302034	Joint pain-shlder
14	36	311	1.672558	0.5143541	0.2393742	Depressive disorder, not elsewhere classified
19	54	285.9	1.514796	0.4152811	0.2530914	Anemia NOS
11	27	250.02	1.750250	0.5597588	0.2530914	DMII wo cmp uncntrld

These are ADEs that people got and caused them to switch!

## Number of Drug Orders

```
os = qfetch(db, "  
SELECT COUNT(*) AS n, order_status FROM (  
SELECT o.*  
FROM cohort c  
JOIN ht_rx o  
  ON c.patient_id = o.patient_id  
  AND c.t0 = o.age_at_start_time_in_days  
  ) o0  
GROUP BY o0.order_status  
ORDER BY n DESC  
")  
kable(os, caption="Order Statuses of First Antihypertensive Orders")
```

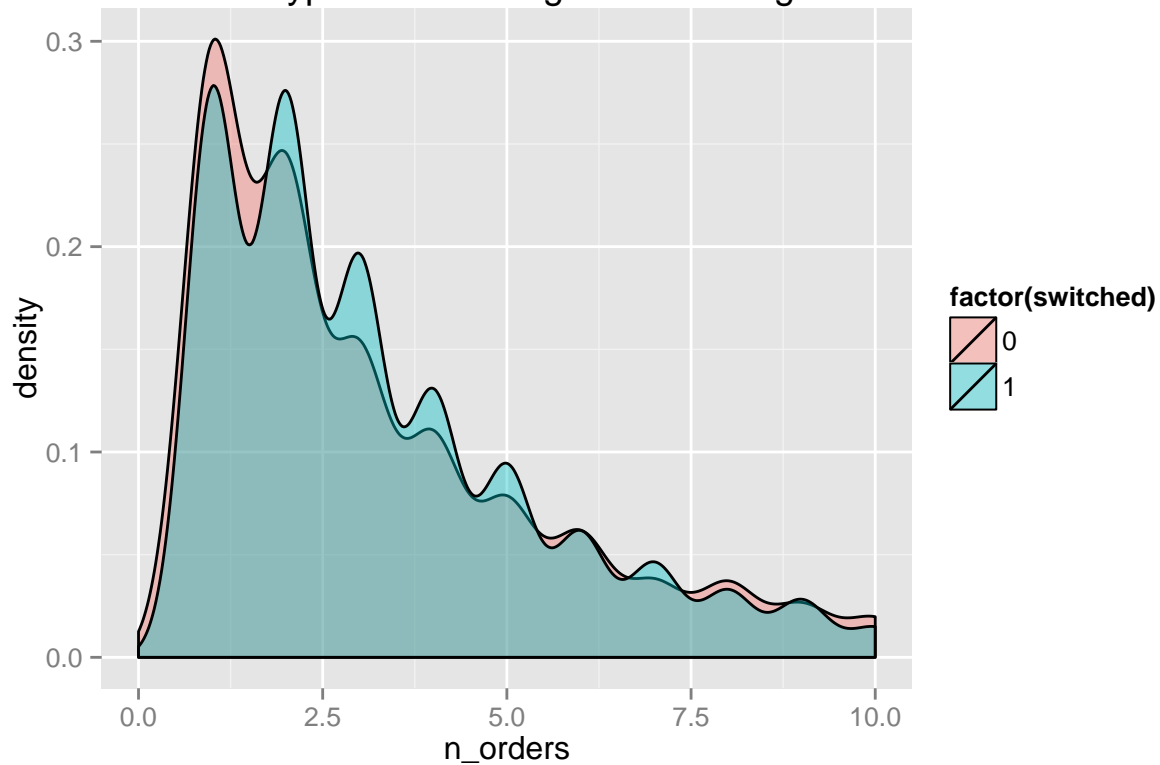
Table 10: Order Statuses of First Antihypertensive Orders

n	order_status
3703	Sent
1106	Discontinued
31	Suspend
29	Completed

Need to figure out what's going on with these "discontinued" orders, since I don't know if they can possibly be refills, considering they're the first ever order of an antihypertensive. Is it possible that they're a "formulary change" from a totally different drug, or an add-on to a prior drug?

```
window_orders = qfetch(db, "  
SELECT COUNT(*) AS n_orders, o.patient_id, c.grp, c.switched  
  FROM cohort c  
  JOIN ht_rx o  
    ON c.patient_id = o.patient_id  
  WHERE o.age_at_start_time_in_days < c.tf  
  OR c.tf is NULL  
  GROUP BY o.patient_id  
")  
ggplot(window_orders, aes(x=n_orders, fill=factor(swapped))) +  
  geom_density(alpha=0.4, binwidth=1, position='identity') +  
  # facet_grid(. ~ grp, labeller='label_both') +  
  xlim(0,10) +  
  ggtitle("Number of antihypertensive drug orders during the window")
```

## Number of antihypertensive drug orders during the window

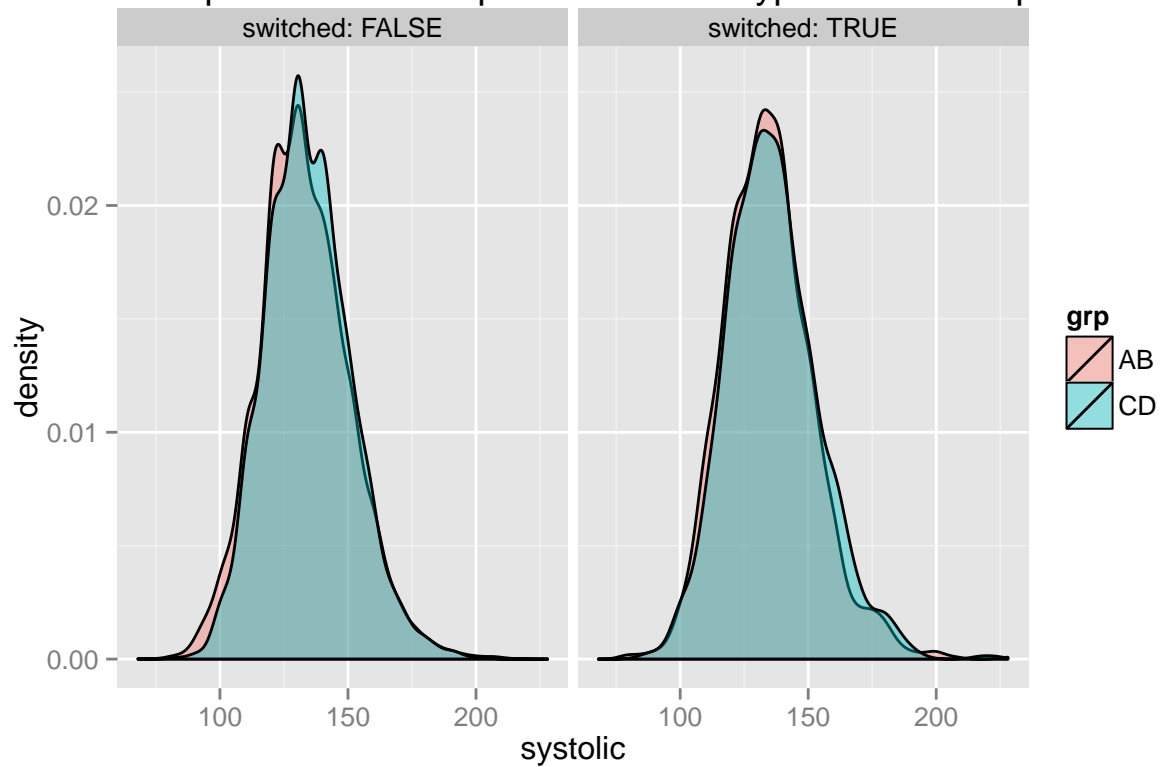


Switchers are getting more orders of their drugs before they switch over than non-switchers are getting for the rest of their lives!

## Blood Pressure Values

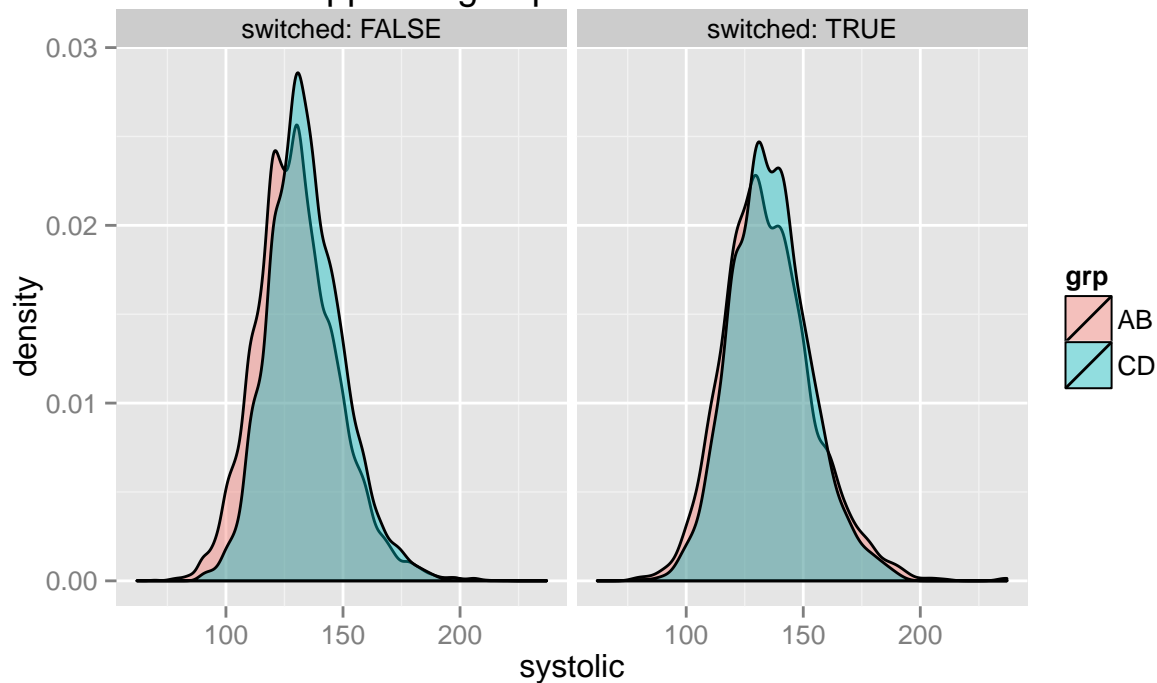
```
ggplot(bp[bp$t-bp$t0<0,], aes(x=systolic, fill=grp)) +  
  geom_density(alpha=0.4, position='identity') +  
  facet_grid(. ~ switched, labeller='label_both') +  
  ggtitle("Blood pressure values prior to first antihypertensive script")
```

## Blood pressure values prior to first antihypertensive script



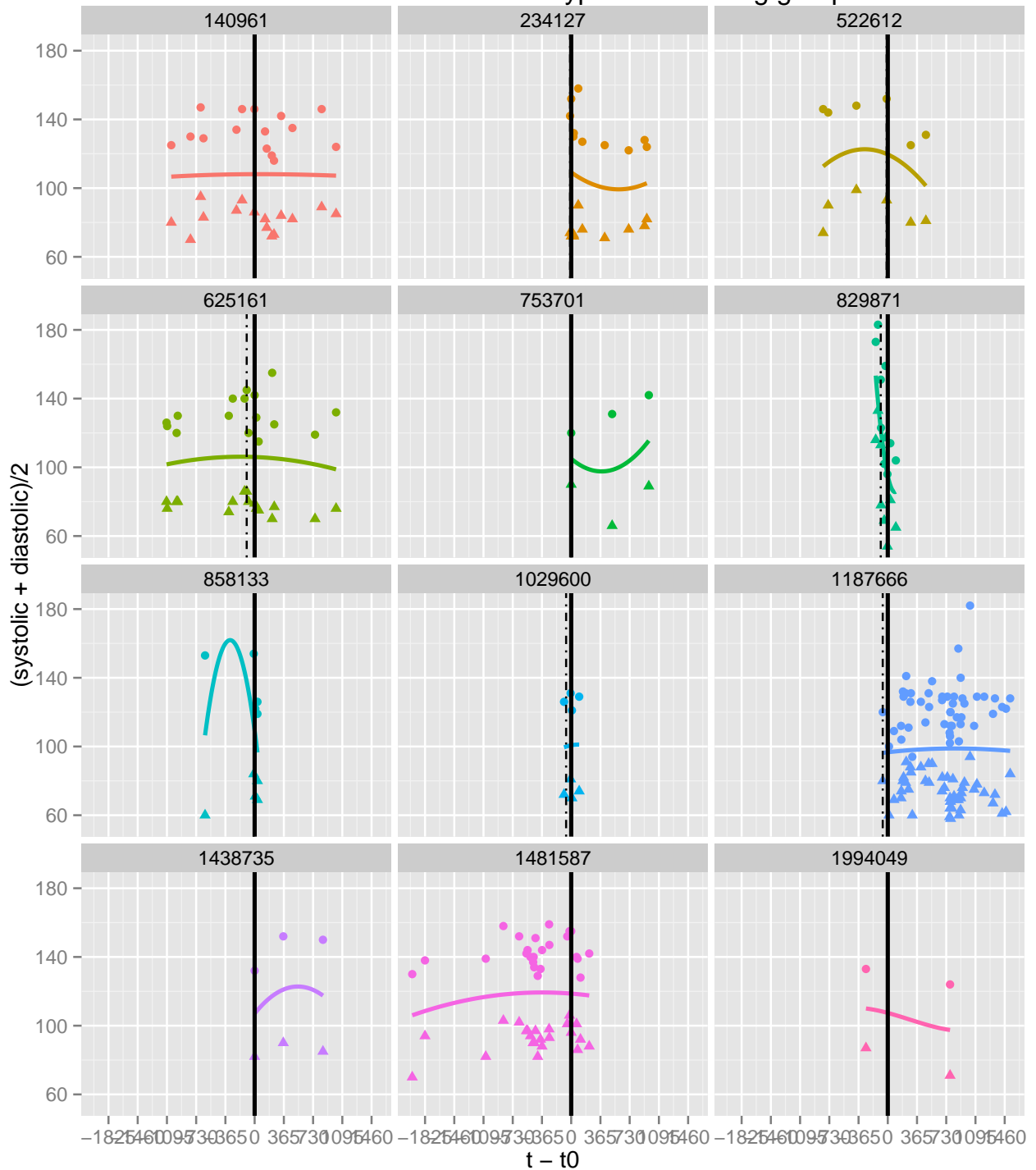
```
ggplot(bp[(bp$t-bp$t0>30)&(is.na(bp$tf)|bp$t<bp$tf)], aes(x=systolic, fill=grp)) +  
  geom_density(alpha=0.4, position='identity') +  
  facet_grid(. ~ switched, labeller='label_both') +  
  ggtitle("Blood pressure values subsequent to first antihypertensive\nscript + 30 days, but before pre
```

Blood pressure values subsequent to first antihypertensive script + 30 days, but before prescription of antihypertensive of the opposite group or before end of record



```
# non-switchers
samp_patients = sample(unique(bp$patient_id[!bp$switched]), 12, replace=F)
ggplot(bp[bp$patient_id %in% samp_patients,],
  aes(x=t-t0, group=patient_id, color=patient_id)) +
  facet_wrap(~patient_id, nrow=4, ncol=3) +
  stat_smooth(aes(y=(systolic+diastolic)/2),
    alpha=1, se=F, size=1, method='loess', span=10) +
  geom_point(aes(y=systolic), size=2, shape=16) +
  geom_point(aes(y=diastolic), size=2, shape=17) +
  geom_vline(xintercept=0, size=1, color='black') +
  geom_vline(aes(xintercept=tdx-t0), linetype=4) +
  theme(legend.position="none") +
  scale_y_continuous(minor_breaks = seq(20, 200, 20),
    breaks = seq(20, 200, 40)) +
  scale_x_continuous(breaks = seq(-20*365, 20*365, 365)) +
  ggtitle('Blood pressure trajectories for patients who did not switch or add antihypertensive drug gr
```

### Blood pressure trajectories for patients who did not switch or add antihypertensive drug groups



```
# switchers
samp_patients = sample(unique(bp$patient_id[bp$switched]), 12, replace=F)
ggplot(bp[bp$patient_id %in% samp_patients,],
  aes(x=t-t0, group=patient_id, color=patient_id)) +
  facet_wrap(~patient_id, nrow=4, ncol=3) +
  stat_smooth(aes(y=(systolic+diastolic)/2),
```

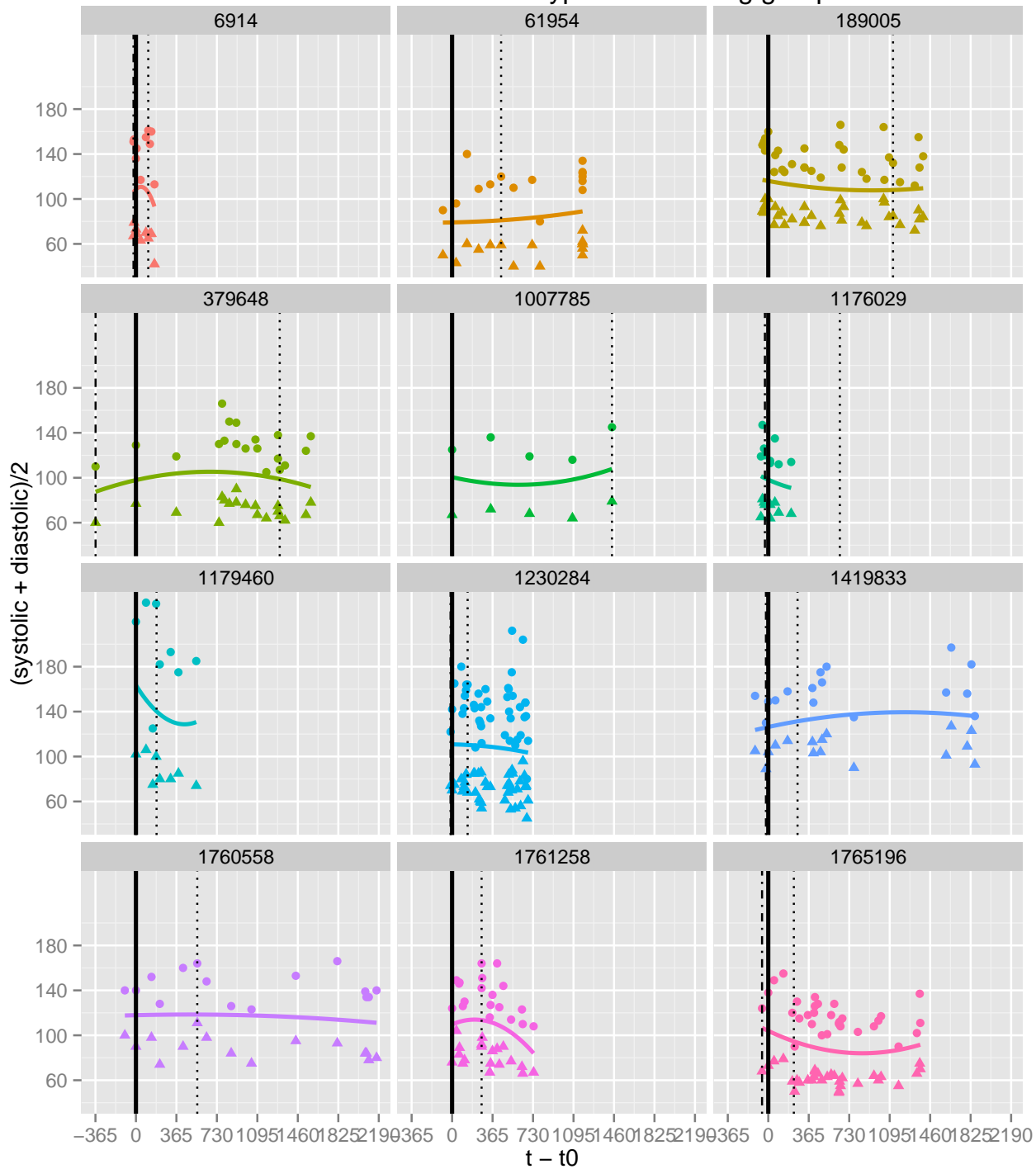


```

    alpha=1, se=F, size=1, method='loess', span=10) +
geom_point(aes(y=systolic), size=2, shape=16) +
geom_point(aes(y=diastolic), size=2, shape=17) +
geom_vline(xintercept=0, size=1, color='black') +
geom_vline(aes(xintercept=tf-t0), linetype=3) +
geom_vline(aes(xintercept=tdx-t0), linetype=4) +
theme(legend.position="none") +
scale_y_continuous(minor_breaks = seq(20 , 200, 20),
                   breaks = seq(20, 200, 40)) +
scale_x_continuous(breaks = seq(-20*365, 20*365, 365)) +
ggtitle('Blood pressure trajectories for patients who\nswitched or added antihypertensive drug groups

```

### Blood pressure trajectories for patients who switched or added antihypertensive drug groups

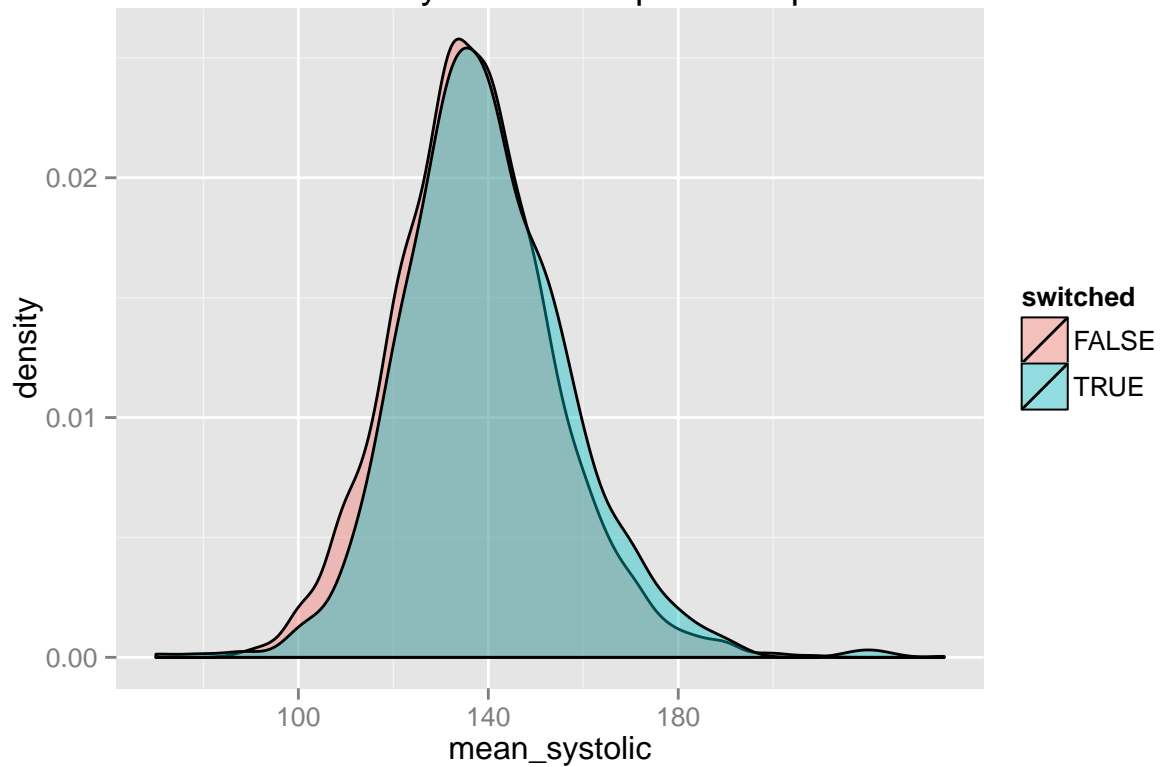


```

mean_pre = summarize(filter(bp, t<t0+1), mean_systolic = mean(systolic))
ggplot(mean_pre, aes(x=mean_systolic, fill=switched)) +
  geom_density(alpha=0.4, position='identity') +
  ggtitle("Patients' maximum systolic blood pressure prior to first rx") +
  scale_x_continuous(minor_breaks = seq(20, 200, 20),
    breaks = seq(20, 200, 40))

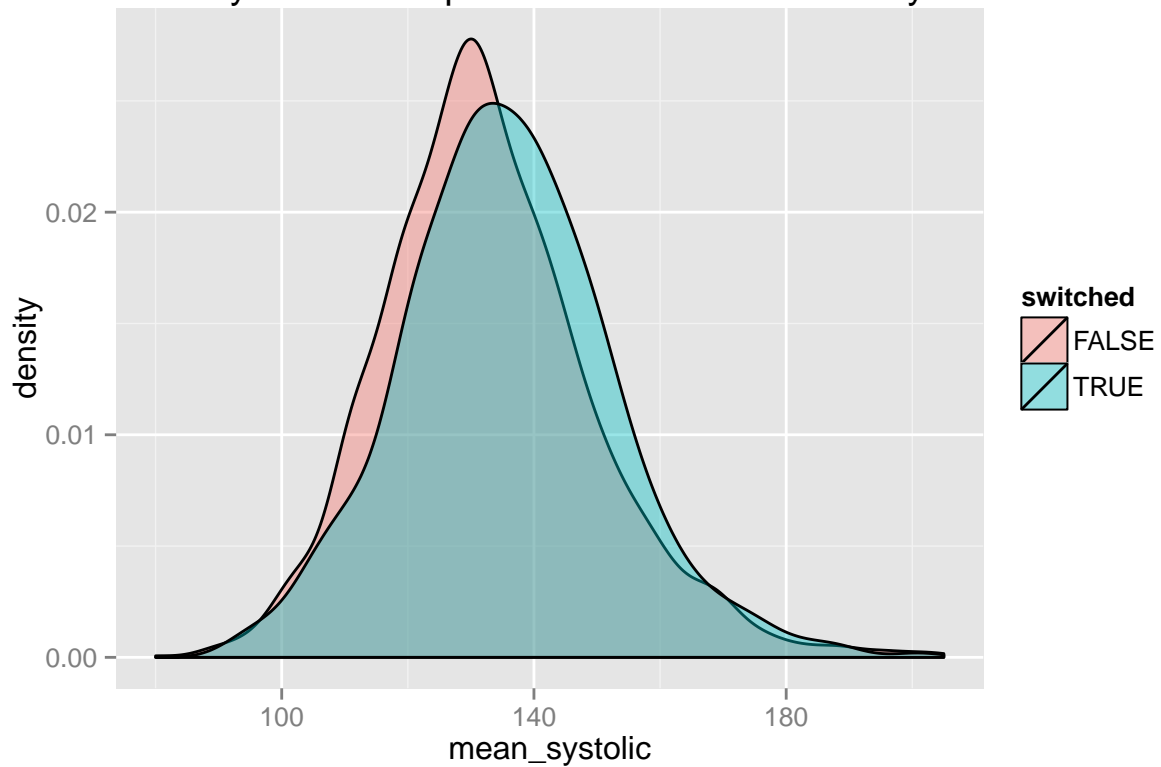
```

Patients' maximum systolic blood pressure prior to first rx



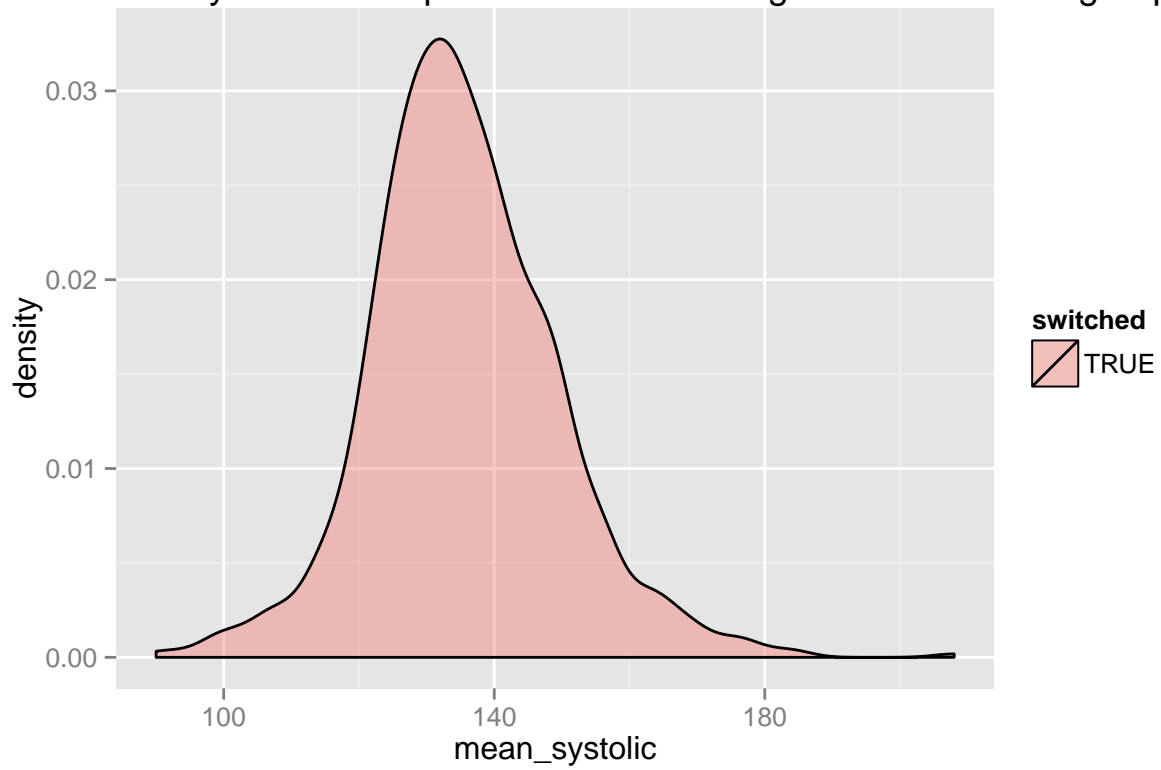
```
mean_window = summarize(filter(bp, t0+30<=t, t<=t0+90), mean_systolic = mean(systolic))
ggplot(mean_window, aes(x=mean_systolic, fill=switched)) +
  geom_density(alpha=0.4, position='identity') +
  ggtitle("Patients' mean systolic blood pressure between 30-90 days after first rx") +
  scale_x_continuous(minor_breaks = seq(20 , 200, 20),
                    breaks = seq(20, 200, 40))
```

patients' mean systolic blood pressure between 30–90 days after first rx



```
mean_post_switch = summarize(filter(bp, tf<t), mean_systolic = mean(systolic))
ggplot(mean_post_switch, aes(x=mean_systolic, fill=switched)) +
  geom_density(alpha=0.4, position='identity') +
  ggtitle("Patients' mean systolic blood pressure after starting an rx of different group") +
  scale_x_continuous(minor_breaks = seq(20 , 200, 20),
                    breaks = seq(20, 200, 40))
```

patients' mean systolic blood pressure after starting an rx of different group



What we see here is that before first treatment, the distribution of mean SBP is the same between patients who go on to switch or add drugs and those who don't. After treatment begins, we see that the group of patients who *don't* add or switch drugs in the future have better controlled their blood pressure, and that patients who *do* add or switch antihypertensives have better controlled their BP after the new prescription.