SUPPLEMENTAL MATERIAL

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	3	State specific objectives, including any prespecified hypotheses	3
Methods			
Study design	4	Present key elements of study design early in the paper	4
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	4
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	4 & 5
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	4 & 5
Bias	9	Describe any efforts to address potential sources of bias	4
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	4
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	6
		(*b*) Describe any methods used to examine subgroups and interactions	6
		(*c*) Explain how missing data were addressed	6
		(*d*) If applicable, explain how loss to follow-up was addressed	6
		(*e*) Describe any sensitivity analyses	7
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	8
		(b) Give reasons for non-participation at each stage	8
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8
		(b) Indicate number of participants with missing data for each variable of interest	20
		(c) Summarise follow-up time (eg, average and total amount)	8
Outcome data	15*	Report numbers of outcome events or summary measures over time	8&9
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted	9

		estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
		(*b*) Report category boundaries when continuous variables were categorized	8
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	9, 10 & 11
Discussion			
Key results	18	Summarise key results with reference to study objectives	11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	13 & 14
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	14
Generalisability	21	Discuss the generalisability (external validity) of the study results	14
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	1

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

Data S1.

```
#R Script for Exploratory analysis of POINT, Mac Grory et al. 2021.

#CONTENTS:
#0. General points and power calculations
#1. Input, inspection and merging of data
#2. Variables of interest
#3. Cleaning up variables of interest
#4. Recharacterizing variables
#5. Descriptive statistics
#6. Inferential statistics
#7. Table 1 ***TABLE 1***
#8. Kaplan-Meier curves ***FIGURE 1***
#9. Cox proportional hazards modelling
#--A. Unadjusted modelling
#--B. Adjusted modelling ***TABLE 2***
#--C. Interaction analyses
#1. Subgroup Analyses
#--1. Minor stroke only - KM/UA/A
#--2. TIA only - KM/UA/A
#--2. DAPT only - KM/UA/A ***TABLE 3***
#--2. SAPT only - KM/UA/A
#11. Sensitivity Analyses
#--11.1. Glucose as continuous variable ***FIGURE 2***
#--11.2. Propensity score-matched analysis
#12. Final sensivity analysis - infarct on imaging instead of adjudicated etiology

#0. General points

#Running lines 1-495 creates analysis dataset

#Packages:
library(powerSurvEpi)
library(haven)
library(dplyr)
library(doBy)
library(reshape)
library(table1)
library(survival)
library(survminer)
library(survMisc)
library(ggpubr)
library(ggplot2)
library(MatchIt)
library(ipw)
library(rms)
library(splines)
library(pROC)
library(coxphw)
library(Hmisc)
```

```
library('mgcv')
library(visreg)

#We did not "attach" data at any point to maintain clarity given the multiple datasets that were created in the course of the analysis.

#The POINT dataset is comprised of the following individual data files, stored as separate .SAS files:
#Form00 - Eligibility Form
#Form01 - Demographics
#Form02 - ABCD2 Score
#Form03 - Modified Rankin Scale
#Form04 - NIH Stroke Scale
#Form05 - Medical History
#Form06 - Prior Medications
#Form07 - Index TIA/Stroke Symptoms
#Form08 - Vital Signs
#Form10 - Randomization Form
#Form11 - Head CT/MRI Scan
#Form12 - Electrocardiogram
#Form13 - Carotid Imaging Results
#Form14 - Questionnaire for Verifying Stroke Free Status
#Form15 - Morisky Quesionnaire
#Form16 - Study Drug Compliance
#Form17 - End of Study
#Form18 - Concomitant Medications
#Form19 - SAE/Clinical Outcome Reporting Form
#Form20 - Final Diagnosis
#Form22 - Anciliary Biomarker study
#Pointoutcomes - All endpoints from both intention to treat and per protocol analysis

#Power Calculations
powerCT.default0(0.10,
        267,
        1.5,
        alpha = 0.05)

powerCT.default0(0.10,
        267,
        2,
        alpha = 0.05)

powerCT.default0(0.10,
        267,
        2.5,
        alpha = 0.05)

powerCT.default0(0.15,
        267,
        1.5,
```

```
        alpha = 0.05)

powerCT.default0(0.15,
        267,
        2,
        alpha = 0.05)

powerCT.default0(0.15,
        267,
        2.5,
        alpha = 0.05)

powerCT.default0(0.25,
        267,
        1.5,
        alpha = 0.05)

powerCT.default0(0.25,
        267,
        2,
        alpha = 0.05)

powerCT.default0(0.25,
        267,
        2.5,
        alpha = 0.05)


#1. Input, inspection and merging of data
#We required variables stored in "Form00", "Form 01", "Form02", "Form05", "Form20", and "pointoutcomes" for this study.
#The datasets of interest were loaded in to R as follows:
#OFFICE
form00 <- read_sas("C:/Users/bcm39/Desktop/POINT Sub-study/POINT DATA/POINT Datasets/form00.sas7bdat", NULL)
form01 <- read_sas("C:/Users/bcm39/Desktop/POINT Sub-study/POINT DATA/POINT Datasets/form01.sas7bdat", NULL)
form02 <- read_sas("C:/Users/bcm39/Desktop/POINT Sub-study/POINT DATA/POINT Datasets/form02.sas7bdat", NULL)
form05 <- read_sas("C:/Users/bcm39/Desktop/POINT Sub-study/POINT DATA/POINT Datasets/form05.sas7bdat", NULL)
form20 <- read_sas("C:/Users/bcm39/Desktop/POINT Sub-study/POINT DATA/POINT Datasets/form20.sas7bdat", NULL)
pointoutcomes <- read_sas("C:/Users/bcm39/Desktop/POINT Sub-study/POINT DATA/POINT
Datasets/pointoutcomes.sas7bdat", NULL)

#HOME
#form00 <- read_sas("POINT/POINT Datasets/form00.sas7bdat", NULL)
#form01 <- read_sas("POINT/POINT Datasets/form01.sas7bdat", NULL)
#form02 <- read_sas("POINT/POINT Datasets/form02.sas7bdat", NULL)
#form05 <- read_sas("POINT/POINT Datasets/form05.sas7bdat", NULL)
#form20 <- read_sas("POINT/POINT Datasets/form20.sas7bdat", NULL)
#pointoutcomes <- read_sas("POINT/POINT Datasets/pointoutcomes.sas7bdat", NULL)
```

```
#We visually inspected the data files before analysis
#View(form00)
#View(form01)
#View(form02)
#View(form05)
#View(form20)
#View(pointoutcomes)

#Then examined their structure
str(form00)
str(form01)
str(form02)
str(form05)
str(form20)
str(pointoutcomes)

#File merging
#Files were merged using "subject_id" as the linkage variable
merge1 <- left_join(form00, form01, by = "subject_id", copy = FALSE)
merge2 <- left_join(merge1, form02, by = "subject_id", copy = FALSE)
merge3 <- left_join(merge2, form05, by = "subject_id", copy = FALSE)
merge4 <- left_join(merge3, form20, by = "subject_id", copy = FALSE)
data <- left_join(merge4, pointoutcomes, by = "subject_id", copy = FALSE)

#For this project, there were 4,881 observation units or less in each data file and thus we did not need to rearrange from skinny to fat dataframes

#We manually inspected the new analysis data set
#View(data)


#2. Variables of interest
data$age #Age (Form 00)
data$F00Q28 #Serum glucose (Form 00)
data$F00Q48 #Glucose units (Form 00)
data$GENDER #Gender (Form01)
data$RACE #Race (Form 01)
data$ABCD2 #ABCD2 Score (Form 02)
data$F05Q01 #Congestive Heart Failure (Form 05)
data$F05Q02 #Atrial Fibrillation (Form 05)
data$F05Q03 #Ischemic Heart Disease (Form 05)
data$F05Q04 #Valvular Heart Disease (Form 05)
data$F05Q05 #Carotid stenosis/Endarterectomy/Stent/Angioplasty (Form 05)
data$F05Q06 #Hypertension (Form 05)
data$F05Q07 #Diabetes Mellitus (Form 05)
data$Smoke #Smoking status (Form 05)
data$F20Q01 #Final diagnosis of TIA (1) vs. Minor Stroke (2) (Form 20)
data$tx #Treatment assignment (from ITT analysis) - A: Placebo, B=Clopidogrel
data$itt_outcome_type4#Subsequent ischemic stroke (pointoutcomes)
```

```
data$itt_outcome_type4_days #Days from randomization to event (pointoutcomes)
data$F20Q04 #Infarct on imaging attributable to index event


#3. Cleaning up variables of interest
#Age (Form 00)
#Manual inspection:
data$age
summary(data$age)
sum(is.na(data$age)==FALSE)
sum(is.na(data$age)==TRUE)
#Age was initially stored as a factor.
table(data$age)
#There were also 73 patients with >89 listed as their age.
#For the purposes of this analysis we assigned the age 90 to them.
data$age[data$age==">89"] <- 90
table(data$age)
#We did this prior to conversion to a numeric variable to avoid introducing NAs
data$age <- as.numeric(data$age)
table(data$age)
summary(data$age)
hist(data$age)
sum(is.na(data$age)==FALSE)
sum(is.na(data$age)==TRUE)

#Serum glucose (Form 00)
data$F00Q28
str(data$F00Q28)
#Glucose is already stored as a numeric variable
sum(is.na(data$F00Q28)==FALSE)
sum(is.na(data$F00Q28)==TRUE)
#3 Subjects have missing information for this variable
#We will remove them from the analysis dataset
data[is.na(data$F00Q28),]
nrow(data)
data <- data[!(is.na(data$F00Q28)),]
nrow(data) #3 subjects have been excluded from this analysis
summary(data$F00Q28)
hist(data$F00Q28)
#Around 500 people have implausibly low glucose readings from inspecting the histogram
table(data$F00Q48)
#595 people have glucose stored in a different unit, explaining these apparently low readings
#From the data dictionary, serum glucose was stored in two units - mg/dl and mmol/L
#Before proceeding further with the analysis, we had to harmonize units
#Convert 2(mmol/L) to 1 (mg/dL)
#Conversion factor = 18.0182
data$F00Q28[data$F00Q48 == 2] <- (data$F00Q28[data$F00Q48 == 2]*18.0182)
hist(data$F00Q28)
str(data$F00Q28)
```

```
sort(data$F00Q28, decreasing=FALSE)
data$glucose <- data$F00Q28
#There remains one person with an implausible low glucose reading of 7.8
#However, on the basis of the information contained in the dataset, I cannot definitively state this is not real
#If this was a mmol/L measurement, it would still be below the threshold of 180mg/dl
#So for the main analyses this subject will still be classed as "not hypoglycemic"
#We then created a dummy variable (entitled "hyperglycemia")
#We dichotomized patients in to >= 180 ("1") or <180 ("0")
data$hyperglycemia <- ifelse(data$glucose >=180, "1", "0")
str(data$hyperglycemia)
plot(data$hyperglycemia,data$glucose)
sum(is.na(data$hyperglycemia))

#Gender (Form01)
data$GENDER
str(data$GENDER)
sum(is.na(data$GENDER)==TRUE)
data$GENDER <- as.factor(data$GENDER)
str(data$GENDER)
table(data$GENDER)
data$female <- data$GENDER
table(data$female)

#Ethnicity (Form 01)
#0 - Hispanic/Latino; 1 - Not hispanic or latino; 3 - Unknown
#We altered this to a dichotomous variable where subjects are classed as 1 (Hispanic/Latino) or 2(Not hispanic/latino)
data$ETHNIC
sum(is.na(data$ETHNIC)==FALSE)
sum(is.na(data$ETHNIC)==TRUE)
table(data$ETHNIC)
data$ETHNIC[data$ETHNIC==3] <- "1"
data$ETHNIC[data$ETHNIC==1] <- "2"
data$ETHNIC[data$ETHNIC==0] <- "1"
data$ETHNIC[data$ETHNIC==2] <- "0"
table(data$ETHNIC)
str(data$ETHNIC)
data$hispanic <- data$ETHNIC
data$hispanic <- as.factor(data$hispanic)
table(data$hispanic)
str(data$hispanic)

#Race (Form 01)
#0 - American Indian/Alaskan Native, 1 - Asian, 2 - Black/African American, 3 - Native Hawaiian, 4 - White, 5 - More than one race, 98 - Other, 99 - Unknown/not reported
str(data$RACE)
table(data$RACE)
sum(is.na(data$ETHNIC)==FALSE)
sum(is.na(data$ETHNIC)==TRUE)
#We altered this to a dichotomous variable where subjects are classed as "Black" or "Non-Black"
```

```
data$RACE[data$RACE==1] <- "0"
data$RACE[data$RACE==3] <- "0"
data$RACE[data$RACE==4] <- "0"
data$RACE[data$RACE==5] <- "0"
data$RACE[data$RACE==98] <- "0"
data$RACE[data$RACE==99] <- "0"
data$RACE[data$RACE==2] <- "1"
data$black <- data$RACE
str(data$black)
data$black <- as.factor(data$black)
str(data$black)
table(data$black)

#ABCD2 Score (Form 02)
data$ABCD2
str(data$ABCD2)
table(data$ABCD2)
sum(is.na(data$ABCD2[data$F20Q01==2]))
#1594 missing
sum(is.na(data$ABCD2[data$F20Q01==98]))
#132 missing
sum(is.na(data$ABCD2[data$F20Q01==1]))
#422 missing
#Very high volume of missing data in this variable, even among patients whose final adjudicated etiology was TIA so we chose to exclude it.

#Congestive Heart Failure (Form 05)
data$F05Q01
sum(is.na(data$F05Q01)==FALSE)
sum(is.na(data$F05Q01)==TRUE)
table(data$F05Q01)
#7 patients had "Unknown" CHF status
str(data$F05Q01)
data$F05Q01[data$F05Q01==2] <- "0"
table(data$F05Q01)
str(data$F05Q01)
data$F05Q01 <- as.factor(data$F05Q01)
str(data$F05Q01)
data$chf <- data$F05Q01
str(data$chf)

#Atrial Fibrillation (Form 05)
data$F05Q02
sum(is.na(data$F05Q02)==FALSE)
sum(is.na(data$F05Q02)==TRUE)
table(data$F05Q02)
#14 patients had "Unknown" AF status
str(data$F05Q02)
data$F05Q02[data$F05Q02==2] <- "0"
```

```
table(data$F05Q02)
str(data$F05Q02)
data$F05Q02 <- as.factor(data$F05Q02)
str(data$F05Q02)
data$af <- data$F05Q02
str(data$af)

#Ischemic Heart Disease (Form 05)
data$F05Q03
sum(is.na(data$F05Q03)==FALSE)
sum(is.na(data$F05Q03)==TRUE)
table(data$F05Q03)
#12 patients had "Unknown" CHF status
str(data$F05Q03)
data$F05Q03[data$F05Q03==2] <- "0"
table(data$F05Q03)
str(data$F05Q03)
data$F05Q03 <- as.factor(data$F05Q03)
str(data$F05Q03)
data$cad <- data$F05Q03
str(data$cad)

#Valvular Heart Disease (Form 05)
data$F05Q04
sum(is.na(data$F05Q04)==FALSE)
sum(is.na(data$F05Q04)==TRUE)
table(data$F05Q04)
#12 patients had "Unknown" Valvular heart disease status
str(data$F05Q04)
data$F05Q04[data$F05Q04==2] <- "0"
table(data$F05Q04)
str(data$F05Q04)
data$F05Q04 <- as.factor(data$F05Q04)
str(data$F05Q04)
data$valvedisease <- data$F05Q04
str(data$valvedisease)

#Carotid stenosis/Endarterectomy/Stent/Angioplasty (Form 05)
data$F05Q05
sum(is.na(data$F05Q05)==FALSE)
sum(is.na(data$F05Q05)==TRUE)
table(data$F05Q05)
#32 patients had "Unknown" carotid stenosis/endarterectomy/stent/angioplasty
str(data$F05Q05)
data$F05Q05[data$F05Q05==2] <- "0"
table(data$F05Q05)
str(data$F05Q05)
data$F05Q05 <- as.factor(data$F05Q05)
str(data$F05Q05)
```

```
data$carotiddisease <- data$F05Q05
str(data$carotiddisease)

#Hypertension (Form 05)
data$F05Q06
sum(is.na(data$F05Q06)==FALSE)
sum(is.na(data$F05Q06)==TRUE)
table(data$F05Q06)
#21 patients were "Unknown" hypertension status
str(data$F05Q06)
data$F05Q06[data$F05Q06==2] <- "0"
table(data$F05Q06)
str(data$F05Q06)
data$F05Q06 <- as.factor(data$F05Q06)
str(data$F05Q06)
data$htn <- data$F05Q06
str(data$htn)

#Diabetes Mellitus (Form 05)
data$F05Q07
sum(is.na(data$F05Q07)==FALSE)
sum(is.na(data$F05Q07)==TRUE)
table(data$F05Q07)
#9 patients were "Unknown" diabetes mellitus status
str(data$F05Q07)
data$F05Q07[data$F05Q07==2] <- "0"
table(data$F05Q07)
str(data$F05Q07)
data$F05Q07 <- as.factor(data$F05Q07)
str(data$F05Q07)
data$diabetes <- data$F05Q07
str(data$diabetes)

#Smoking status (Form 05)
data$smoke
str(data$smoke)
sum(is.na(data$smoke)==FALSE)
sum(is.na(data$smoke)==TRUE)
#4 patients had missing data on smoking
table(data$smoke)
#We considered active smoking as smoking (1) and past/never smoking as not smoking (0)
data$smoke[data$smoke==1] <- "0"
data$smoke[data$smoke==2] <- "1"
table(data$smoke)
str(data$smoke)
#The 4 patients had missing data on smoking were classified as "not smoking"
data$smoke[is.na(data$smoke)] <- "0"
data$smoke <- as.factor(data$smoke)
sum(is.na(data$smoke)==TRUE)
```

```
str(data$smoke)
data$smoking <- data$smoke
str(data$smoking)

#Final diagnosis of TIA (1) vs. Minor Stroke (2) (Form 20)
data$F20Q01
sum(is.na(data$F20Q01)==FALSE)
sum(is.na(data$F20Q01)==TRUE)
#3 patients had missing data
table(data$F20Q01)
str(data$F20Q01)
data$F20Q01[data$F20Q01==1] <- "0"
data$F20Q01[data$F20Q01==98] <- "0"
data$F20Q01[data$F20Q01==2] <- "1"
sum(is.na(data$F20Q01)==TRUE)
data$F20Q01[is.na(data$F20Q01)] <- 0
data$F20Q01 <- as.factor(data$F20Q01)
data$minorstroke <- data$F20Q01
table(data$minorstroke)
str(data$minorstroke)

#Treatment assignment (from ITT analysis) - A: Placebo, B=Clopidogrel
table(data$tx)
str(data$tx)
sum(is.na(data$tx)==FALSE)
sum(is.na(data$tx)==TRUE)
#No missing data on treatment assignment
data$tx[data$tx=="B"] <- "1"
data$tx[data$tx=="A"] <- "0"
table(data$tx)
str(data$tx)
data$tx <- as.factor(data$tx)
table(data$tx)
str(data$tx)
data$dapt <- data$tx
str(data$dapt)

#Subsequent ischemic stroke (pointoutcomes)
data$itt_outcome_type4
sum(is.na(data$itt_outcome_type4)==FALSE)
sum(is.na(data$itt_outcome_type4)==TRUE)
#No missing data on subsequent ischemic stroke
table(data$itt_outcome_type4)
str(data$itt_outcome_type4)
data$itt_outcome_type4 <- as.factor(data$itt_outcome_type4)
table(data$itt_outcome_type4)
str(data$itt_outcome_type4)
data$stroke <- data$itt_outcome_type4
str(data$stroke)
```

```
#Days from randomization to event (pointoutcomes)
data$itt_outcome_type4_days
str(data$itt_outcome_type4_days)
sum(is.na(data$itt_outcome_type4_days)==FALSE)
sum(is.na(data$itt_outcome_type4_days)==TRUE)
#No missing data on time to subsequent ischemic stroke
table(data$itt_outcome_type4_days)
data$itt_outcome_type4_days <- as.numeric(data$itt_outcome_type4_days)
str(data$itt_outcome_type4_days)
data$days <- data$itt_outcome_type4_days
str(data$days)


#4. Recharacterizing variables

data$age #Age (Form 00) --> NUMERIC
data$glucose #Serum glucose (Form 00) --> NUMERIC
data$hyperglycemia #Hyperglycemia (dummy variable) --> CHARACTER
data$female #Female sex (Form 00) --> FACTOR
data$black #Race (Form 01) --> FACTOR
data$hispanic #ETHNIC (Form 01) --> FACTOR
data$chf #Congestive Heart Failure (Form 05) --> FACTOR
data$af #Atrial Fibrillation (Form 05) --> FACTOR
data$cad #Ischemic Heart Disease (Form 05) --> FACTOR
data$valvedisease #Valvular Heart Disease (Form 05) --> FACTOR
data$carotiddisease #Carotid stenosis/Endarterectomy/Stent/Angioplasty (Form 05) --> FACTOR
data$htn #Hypertension (Form 05) --> FACTOR
data$diabetes #Diabetes Mellitus (Form 05) --> FACTOR
data$smoking #Smoking status (Form 05) --> FACTOR
data$minorstroke #Final diagnosis of TIA (1) vs. Minor Stroke (2) (Form 20) --> FACTOR
data$dapt #Treatment assignment (from ITT analysis) - A: Placebo, B=Clopidogrel --> FACTOR
data$stroke #Subsequent ischemic stroke (pointoutcomes) --> FACTOR
data$days #Days from randomization to event (pointoutcomes) --> NUMERIC

#data$age <- as.numeric(data$age)
#data$GENDER <- as.factor(data$GENDER)
#data$F05Q01 <- as.numeric(data$F05Q02)
#data$F05Q02 <- as.numeric(data$F05Q02)
#data$F05Q03 <- as.numeric(data$F05Q03)
#data$F05Q06 <- as.numeric(data$F05Q06)
#data$smoke <- as.numeric(data$smoke)
#data$itt_outcome_type4_days <- as.numeric(data$itt_outcome_type4_days)
#data$itt_outcome_type4 <- as.numeric(data$itt_outcome_type4)


#5. Descriptive statistics
#Age
summary(data$age [data$hyperglycemia == 1],)
```

```
summary(data$age [data$hyperglycemia == 0],)
mean(data$age, na.rm=TRUE)
sd(data$age, na.rm=TRUE)
#Sex
table(data$female)
table(data$female [data$hyperglycemia == 1])
table(data$female [data$hyperglycemia == 0])
#Race
table(data$black)
table(data$black [data$hyperglycemia == 1])
table(data$black [data$hyperglycemia == 0])
#Ethnicity
table(data$hispanic)
table(data$hispanic [data$hyperglycemia == 1])
table(data$hispanic [data$hyperglycemia == 0])
#Hypertension
table(data$htn)
table(data$htn [data$hyperglycemia == 1])
table(data$htn [data$hyperglycemia == 0])
#Diabetes mellitus
table(data$diabetes)
table(data$diabetes [data$hyperglycemia == 1])
table(data$diabetes [data$hyperglycemia == 0])
#Atrial fibrillation
table(data$af)
table(data$af [data$hyperglycemia == 1])
table(data$af [data$hyperglycemia == 0])
#CAD
table(data$cad)
table(data$cad [data$hyperglycemia == 1])
table(data$cad [data$hyperglycemia == 0])
#CHF
table(data$chf)
table(data$chf [data$hyperglycemia == 1])
table(data$chf [data$hyperglycemia == 0])
#Tobacco use
table(data$smoking)
table(data$smoking [data$hyperglycemia == 1])
table(data$smoking [data$hyperglycemia == 0])
#Index stroke
table(data$minorstroke)
table(data$minorstroke [data$hyperglycemia == 1])
table(data$minorstroke [data$hyperglycemia == 0])
#Treatment assignment
table(data$dapt)
table(data$dapt [data$hyperglycemia == 1])
table(data$dapt [data$hyperglycemia == 0])
```

```
#6. Creating table 1
table.data <- data

# Nice website with some advanced features https://benjaminrich.github.io/table1/vignettes/table1-examples.html
table.data$female <- factor(table.data$female, labels = c("Male","Female"))
table.data$black <- factor(table.data$black, labels = c("Non-Black","Black"))
table.data$hispanic <- factor(table.data$hispanic, labels = c("Non-Hispanic","Hispanic"))
table.data$hyperglycemia <- factor(table.data$hyperglycemia, labels = c("Normoglycemic","Hyperglycemic"))
label(table.data$age) <- "Age"
label(table.data$female) <- "Sex"
label(table.data$black) <- "Race"
label(table.data$hispanic) <- "Ethnicity"
label(table.data$htn) <- "Hypertension"
label(table.data$diabetes) <- "Diabetes Mellitus"
label(table.data$chf) <- "Congestive Heart Failure"
label(table.data$af) <- "Atrial Fibrillation"
label(table.data$cad) <- "Coronary Artery Disease"
label(table.data$valvedisease) <- "Valve Disease"
label(table.data$carotiddisease) <- "Carotid Disease"
label(table.data$smoking) <- "Smoking (active)"
label(table.data$minorstroke) <- "Minor Stroke"
label(table.data$dapt) <- "Dual Anti-platelet Therapy"
label(table.data$stroke) <- "Subsequent Stroke"
label(table.data$hyperglycemia) <- "Hyperglycemia"

# Creating table 1, comparing characteristics between the two groups
table1(~ table.data$age + table.data$female + table.data$black + table.data$hispanic + table.data$htn +
table.data$diabetes +
        table.data$chf + table.data$af + table.data$cad + table.data$valvedisease +
        table.data$carotiddisease + table.data$smoking + table.data$minorstroke +
        table.data$dapt + table.data$stroke| table.data$hyperglycemia, data=table.data,
        topclass="Rtable1-grid Rtable1-shade Rtable1-times")


#7. Inferential statistics
#A. Comparing continuous variables with a t-test
hist(data$age[data$hyperglycemia==0])
hist(data$age[data$hyperglycemia==1])
t.test(data$age~data$hyperglycemia, data=data, var.equal=TRUE, conf.level=0.95)

#7B. Comparing categorical variables with a Chi Squared test
x <- table(data$female, data$hyperglycemia)
chisq.test(x)
x <- table(data$black, data$hyperglycemia)
chisq.test(x)
x <- table(data$hispanic, data$hyperglycemia)
chisq.test(x)
x <- table(data$htn, data$hyperglycemia)
chisq.test(x)
```

```
x <- table(data$diabetes, data$hyperglycemia)
chisq.test(x)
x <- table(data$chf, data$hyperglycemia)
chisq.test(x)
x <- table(data$af, data$hyperglycemia)
chisq.test(x)
x <- table(data$cad, data$hyperglycemia)
chisq.test(x)
x <- table(data$valvedisease, data$hyperglycemia)
chisq.test(x)
x <- table(data$carotiddisease, data$hyperglycemia)
chisq.test(x)
x <- table(data$smoking, data$hyperglycemia)
chisq.test(x)
x <- table(data$minorstroke, data$hyperglycemia)
chisq.test(x)
x <- table(data$dapt, data$hyperglycemia)
chisq.test(x)
x <- table(data$stroke, data$hyperglycemia)
chisq.test(x)


#8. Kaplan-Meier curves for Main Analysis
time <- data$days
event <- data$stroke
event <- as.numeric(event)
#Changing property of hyperglycemia variable as a way of troubleshooting
data$hyperglycemia <- as.numeric(data$hyperglycemia)
group <- data$hyperglycemia
summary(time)
summary(event)
summary(group)
kmsurvival <- survfit(Surv(time,event) ~ 1, conf.type="none")
summary (kmsurvival)
#Getting estimates with 95% CIs for each group at 90 days
kmsurvivalestimate <- survfit(Surv(time,event) ~ 1)
summary (kmsurvivalestimate, times=90)
#Estimate
1-0.943
#Upper CI
1-0.936
#Lower CI
1-0.95
#Getting estimates with 95% CIs for hyperglycemia group
hyperglycemicgroup <- data[data$hyperglycemia ==1,]
time <- hyperglycemicgroup$days
event <- hyperglycemicgroup$stroke
event <- as.numeric(event)
kmsurvival <- survfit(Surv(time,event) ~ 1, conf.type="none")
```

```
summary (kmsurvival)
kmsurvivalestimate <- survfit(Surv(time,event) ~ 1)
summary (kmsurvivalestimate, times=90)
#Estimate
1-0.903
#Upper CI
1-0.878
#Lower CI
1-0.928
#Getting estimates with 95% CIs for normoglycemia group
normoglycemicgroup <- data[data$hyperglycemia ==0,]
time <- normoglycemicgroup$days
event <- normoglycemicgroup$stroke
event <- as.numeric(event)
kmsurvival <- survfit(Surv(time,event) ~ 1)
summary (kmsurvival)
plot(kmsurvival)
kmsurvivalestimate <- survfit(Surv(time,event) ~ 1)
summary (kmsurvivalestimate, times=90)
#Estimate
1-0.948
#Upper CI
1-0.942
#Lower CI
1-0.955
#kmsurvivalestimate <- survfit(Surv(time,event) ~ group)
#summary (kmsurvivalestimate, times=90)
#Curve for all patients in database
plot(kmsurvival)
#Curve stratified based on serum glucose (hyperglycemic or normoglycemic)
kmsurvival <- survfit(Surv(time,event) ~ group)
summary (kmsurvival)
plot(kmsurvival, fun="event", conf.type = "log")
####Repeat of graph with 4 lines 444444444444##################
kmsurvival <- survfit(Surv(time,event) ~ group)
summary (kmsurvival)
plot(kmsurvival, fun="event", conf.type = "log")
#Better way to incoporate graphics
#ggsurvplot(fit.km, data = ovarian2,
#          risk.table = TRUE,
#         surv.median.line = "hv")
#Comparing curves using the log Rank test
time <- data$days
event <- data$stroke
event <- as.numeric(event)
#Changing property of hyperglycemia variable as a way of troubleshooting
data$hyperglycemia <- as.numeric(data$hyperglycemia)
group <- data$hyperglycemia
survdiff(Surv(time,event) ~ group + data$diabetes, data=data)
```

```
#NB event has to be numeric and not a factor

#We then created an annotated and labelled figure with two components
#A. A large, labelled graph
plot(kmsurvival, fun="event", xlab="Days Since Randomization", ylab="Proportion of Patients With Subsequent Stroke",
lwd=1, ylim=c(0,1), col=c("red","blue"))
box (lwd=2)
axis(side=1, at = c(0,10,20,30,40,50,60,70,80,90))
axis(side=2, at = c(0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1))
legend("bottomright", c("Normoglycemia (<180mg/dl)", "Hyperglycemia (>180mg/dl)"), col=c("red", "blue"), lty=1)
#B. A small insert with less conspicious labelling and a smaller y access to magnify the area of interest
plot(kmsurvival, fun="event", col=c("red","blue"))
legend("bottomright", c("Normoglycemia (<180mg/dl)", "Hyperglycemia (>180mg/dl)"), col=c("red", "blue"), lty=1)
#We went on to do further manipulation of both graphs as follows:
#LARGE GRAPH
plot(kmsurvival, fun="event", col=c("red","red", "blue", "blue"), lwd=1, lty=c(1,5,1,5), ylim = c(0,1))
box(lwd=2)
axis(side=1, at = c(0,10,20,30,40,50,60,70,80,90))
axis(side=2, at = c(0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1))
legend("bottomright", c("Normoglycemia (<180mg/dl)", "Hyperglycemia (>180mg/dl)"), col=c("red", "blue"), lty=1)
#INSERT
plot(kmsurvival, fun="event", col=c("red", "blue"), lwd=2, ylim = c(0,0.14), axes = TRUE)
box(lwd=2)
axis(side=1, at = c(0,30,60,90))
axis(side=2, at = c(0,0.01,0.02,0.03,0.04,0.05,0.06,0.07,0.08,0.09,0.1,0.11,0.12,0.13))
slegend("bottomright", c("Normoglycemia (<180mg/dl)", "Hyperglycemia (>180mg/dl)"), col=c("red", "blue"), lty=1)


#9. Cox proportional hazards modeling
#A. Unadjusted Cox Model
time <- data$days
event <- data$stroke
group <- data$hyperglycemia
summary(time)
summary(event)
summary(group)
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ group, data=data, method="breslow")
summary(coxph)

#B. Adjusted Cox Model
#MODEl 1 - unadjusted
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia, method="breslow")
summary(coxph)

#MODEL 2 - Model 1 + Age, sex, race, ethnicity
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$age, method="breslow")
summary(coxph)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$age + data$female, method="breslow")
```

```
summary(coxph)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$age + data$female + data$black, method="breslow")
summary(coxph)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$age + data$female + data$black + data$hispanic,
method="breslow")
summary(coxph)

#MODEL 3 - Model 2 + treatment assignment and index event
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$age + data$female + data$black + data$hispanic +
data$dapt + data$minorstroke, method="breslow")
summary(coxph)

#MODEL 4 - Model 3 + vascular risk factors (excluding diabetes mellitus)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$age + data$female + data$black + data$hispanic +
data$dapt + data$minorstroke + data$htn + data$chf + data$af + data$cad + data$valvedisease + data$carotiddisease +
data$smoking, method="breslow")
summary(coxph)

#MODEL 5 - Model 4 + vascular risk factors (including diabetes mellitus)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$age + data$female + data$black + data$hispanic +
data$dapt + data$minorstroke + data$htn + data$chf + data$af + data$cad + data$valvedisease + data$carotiddisease +
data$smoking + data$diabetes, method="breslow")
summary(coxph)

#MODEL 5 with major hemorrhage as outcome
#Major hemorrhage is itt_outcome_type11 in point outcomes
#Time to major hemorrhage is itt_outcome_type11_days in point outcomes
time <- data$itt_outcome_type11_days
event <- data$itt_outcome_type11
group <- data$hyperglycemia
summary(time)
summary(event)
summary(group)
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$age + data$female + data$black + data$hispanic +
data$dapt + data$minorstroke + data$htn, method="breslow")
summary(coxph)

#MODEL 5 with composite as outcome
#Composite is itt_outcome_type1 in point outcomes
#Time to composite is itt_outcome_type1_days in point outcomes
time <- data$itt_outcome_type1_days
event <- data$itt_outcome_type1
group <- data$hyperglycemia
summary(time)
summary(event)
summary(group)
event <- as.numeric(event)
```

```
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$age + data$female + data$black + data$hispanic +
data$dapt + data$minorstroke + data$htn + data$chf + data$af + data$cad + data$valvedisease + data$carotiddisease +
data$smoking + data$diabetes, method="breslow")
summary(coxph)

#C. Interaction Analyses
hyperglycemicgroup <- data[data$hyperglycemia ==1,]
normoglycemicgroup <- data[data$hyperglycemia ==0,]
#1---> Hyperglycemia and DAPT
#OUTCOME 1 - subsequent stroke
table(hyperglycemicgroup$stroke, hyperglycemicgroup$dapt)
table(normoglycemicgroup$stroke, normoglycemicgroup$dapt)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia, data=data, method="breslow")
summary(coxph)
coxph <- coxph(Surv(time,event) ~ data$dapt, data=data, method="breslow")
summary(coxph)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$dapt, data=data, method="breslow")
summary(coxph)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$dapt + data$hyperglycemia*data$dapt, data=data,
method="breslow")
summary(coxph)

#Outcome 2 - major hemorrhage
#Major hemorrhage is itt_outcome_type11 in point outcomes
#Time to major hemorrhage is itt_outcome_type11_days in point outcomes
table(hyperglycemicgroup$itt_outcome_type11, hyperglycemicgroup$dapt)
table(normoglycemicgroup$itt_outcome_type11, normoglycemicgroup$dapt)
#In whole sample
time <- data$itt_outcome_type11_days
event <- data$itt_outcome_type11
group <- data$hyperglycemia
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia, data=data, method="breslow")
summary(coxph)
coxph <- coxph(Surv(time,event) ~ data$dapt, data=data, method="breslow")
summary(coxph)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$dapt, data=data, method="breslow")
summary(coxph)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$dapt + data$hyperglycemia*data$dapt, data=data,
method="breslow")
summary(coxph)
#In hyperglycemic group - Major hemorrhage/DAPT
time <- hyperglycemicgroup$itt_outcome_type11_days
event <- hyperglycemicgroup$itt_outcome_type11
group <- data$dapt
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ hyperglycemicgroup$dapt, data=hyperglycemicgroup, method="breslow")
summary(coxph)
#Won't work as 0 events in SAPT group
```

```
#In normoglycemic group - Major hemorrhage/DAPT
time <- normoglycemicgroup$itt_outcome_type11_days
event <- normoglycemicgroup$itt_outcome_type11
group <- data$dapt
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ normoglycemicgroup$dapt, data=normoglycemicgroup, method="breslow")
summary(coxph)

#Outcome 3 - subsequent ischemic stroke, myocardial infarction or vascular death
#Composite is itt_outcome_type1 in point outcomes
#Time to composite is itt_outcome_type1_days in point outcomes
table(hyperglycemicgroup$itt_outcome_type1, hyperglycemicgroup$dapt)
table(normoglycemicgroup$itt_outcome_type1, normoglycemicgroup$dapt)
#In whole sample
time <- data$itt_outcome_type1_days
event <- data$itt_outcome_type1
group <- data$hyperglycemia
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia, data=data, method="breslow")
summary(coxph)
coxph <- coxph(Surv(time,event) ~ group, data=data, method="breslow")
summary(coxph)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$dapt, data=data, method="breslow")
summary(coxph)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$dapt + data$hyperglycemia*data$dapt, data=data,
method="breslow")
summary(coxph)
#In hyperglycemic group
time <- hyperglycemicgroup$itt_outcome_type1_days
event <- hyperglycemicgroup$itt_outcome_type1
group <- hyperglycemicgroup$dapt
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ hyperglycemicgroup$dapt, data=hyperglycemicgroup, method="breslow")
summary(coxph)
#In normoglycemic group
time <- normoglycemicgroup$itt_outcome_type1_days
event <- normoglycemicgroup$itt_outcome_type1
group <- normoglycemicgroup$dapt
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ normoglycemicgroup$dapt, data=normoglycemicgroup, method="breslow")
summary(coxph)

#2---> Hyperglycemia and Stroke/TIA
hyperglycemicgroup <- data[data$hyperglycemia ==1,]
normoglycemicgroup <- data[data$hyperglycemia ==0,]
minorstroke <- data[data$minorstroke ==1,]
tiaother <- data[data$minorstroke ==0,]

#####################################################
```

```
#Hyperglycemia and Stroke/TIA interaction analysis
#OUTCOME 1 - subsequent stroke
table(minorstroke$stroke, minorstroke$hyperglycemia)
table(tiaother$stroke, tiaother$hyperglycemia)
#Whole sample
time <- data$days
event <- data$stroke
group <- data$hyperglycemia
summary(time)
summary(event)
summary(group)
event <- as.numeric(event)
#Unadjusted main analysis
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia, data=data, method="breslow")
summary(coxph)
#Model with hyperglycemia and stroke/tia
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$minorstroke, data=data, method="breslow")
summary(coxph)
#Otherwise unadjusted model with interaction term of hyperglycemia*minorstroke
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$minorstroke + data$hyperglycemia*data$minorstroke,
data=data, method="breslow")
summary(coxph)
#Adjusted model including interaction term of stroke/tia*hyperglycemia
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$minorstroke + data$hyperglycemia*data$minorstroke +
data$age + data$female + data$black + data$hispanic + data$dapt + data$htn + data$chf + data$af + data$cad +
data$valvedisease + data$carotiddisease + data$smoking + data$diabetes, method="breslow")
summary(coxph)
#Hyperglycemic group
time <- hyperglycemicgroup$days
event <- hyperglycemicgroup$stroke
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ hyperglycemicgroup$minorstroke + hyperglycemicgroup$age +
hyperglycemicgroup$female + hyperglycemicgroup$black + hyperglycemicgroup$hispanic + hyperglycemicgroup$dapt +
hyperglycemicgroup$htn + hyperglycemicgroup$chf + hyperglycemicgroup$af + hyperglycemicgroup$cad +
hyperglycemicgroup$valvedisease + hyperglycemicgroup$carotiddisease + hyperglycemicgroup$smoking +
hyperglycemicgroup$diabetes, method="breslow")
summary(coxph)
#Normoglycemic group
time <- normoglycemicgroup$days
event <- normoglycemicgroup$stroke
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ normoglycemicgroup$minorstroke + normoglycemicgroup$age +
normoglycemicgroup$female + normoglycemicgroup$black + normoglycemicgroup$hispanic +
normoglycemicgroup$dapt + normoglycemicgroup$htn + normoglycemicgroup$chf + normoglycemicgroup$af +
normoglycemicgroup$cad + normoglycemicgroup$valvedisease + normoglycemicgroup$carotiddisease +
normoglycemicgroup$smoking + normoglycemicgroup$diabetes, method="breslow")
summary(coxph)

#Hyperglycemia and Stroke/TIA interaction analysis
```

```
#OUTCOME 2 - major hemorrhage
#Major hemorrhage is itt_outcome_type11 in point outcomes
#Time to major hemorrhage is itt_outcome_type11_days in point outcomes
table(hyperglycemicgroup$itt_outcome_type11, hyperglycemicgroup$minorstroke)
table(normoglycemicgroup$itt_outcome_type11, normoglycemicgroup$minorstroke)
#In whole sample - Interation analysis not possible given 0 hemorrhages in hyperglycemia/minor stroke group
#In normoglycemic group - Major hemorrhage/minorstroke-TIA
time <- normoglycemicgroup$itt_outcome_type11_days
event <- normoglycemicgroup$itt_outcome_type11
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ normoglycemicgroup$minorstroke, method="breslow", data=normoglycemicgroup)
summary(coxph)
#Could only adjust for age, sex, race, ethnicity, treatment assignment and hypertension
coxph <- coxph(Surv(time,event) ~ normoglycemicgroup$minorstroke + normoglycemicgroup$age +
normoglycemicgroup$female + normoglycemicgroup$black + normoglycemicgroup$hispanic +
normoglycemicgroup$dapt + normoglycemicgroup$htn, method="breslow", data=normoglycemicgroup)
summary(coxph)

#Hyperglycemia and Stroke/TIA interaction analysis
#OUTCOME 3 - subsequent ischemic stroke, myocardial infarction or vascular death
#Composite is itt_outcome_type1 in point outcomes
#Time to composite is itt_outcome_type1_days in point outcomes
table(hyperglycemicgroup$itt_outcome_type1, hyperglycemicgroup$minorstroke)
table(normoglycemicgroup$itt_outcome_type1, normoglycemicgroup$minorstroke)
#In whole sample
time <- data$itt_outcome_type1_days
event <- data$itt_outcome_type1
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$minorstroke + data$hyperglycemia*data$minorstroke +
data$age + data$female + data$black + data$hispanic + data$dapt + data$htn + data$chf + data$af + data$cad +
data$valvedisease + data$carotiddisease + data$smoking + data$diabetes, method="breslow")
summary(coxph)
#In hyperglycemic group
time <- hyperglycemicgroup$itt_outcome_type1_days
event <- hyperglycemicgroup$itt_outcome_type1
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ hyperglycemicgroup$minorstroke + hyperglycemicgroup$age +
hyperglycemicgroup$female + hyperglycemicgroup$black + hyperglycemicgroup$hispanic + hyperglycemicgroup$dapt +
hyperglycemicgroup$htn + hyperglycemicgroup$chf + hyperglycemicgroup$af + hyperglycemicgroup$cad +
hyperglycemicgroup$valvedisease + hyperglycemicgroup$carotiddisease + hyperglycemicgroup$smoking +
hyperglycemicgroup$diabetes, method="breslow")
summary(coxph)
#In normoglycemic group
time <- normoglycemicgroup$itt_outcome_type1_days
event <- normoglycemicgroup$itt_outcome_type1
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ normoglycemicgroup$minorstroke + normoglycemicgroup$age +
normoglycemicgroup$female + normoglycemicgroup$black + normoglycemicgroup$hispanic +
normoglycemicgroup$dapt + normoglycemicgroup$htn + normoglycemicgroup$chf + normoglycemicgroup$af +
```

```
normoglycemicgroup$cad + normoglycemicgroup$valvedisease + normoglycemicgroup$carotiddisease +
normoglycemicgroup$smoking + normoglycemicgroup$diabetes, method="breslow")
summary(coxph)


#10. Subgroup analyses
#Variables used:
#---data$diabetes - Diabetes mellitus (1=Yes, 0=No)
#---data$F20Q01 - Adjudicated final etiology (2=minor stroke, 1=TIA)
#10-1 - TIA
#10-2 - minorstroke
#10-5 - Hypergylcemia only (for DAPT)
#10-6 - Normoglycemia only (for DAPT)

#SA 10-1 - Minor stroke only
#1. Create group
#2. Get number of events by creating table
#3. Do survival analysis
#4. Create two groups within (exposed/not exposed) to obtain KM estimates + 95% CI
#5. Plot K-M curves
#6. Add Log Rank test
#7. Make figure (NB include box(lwd=2))
#8. Cox model
#1. Create group
#Patients with minor stroke
minorstrokeonly <- data[data$minorstroke == 1,]
#2. Get number of events by creating table
table(minorstrokeonly$stroke, minorstrokeonly$hyperglycemia)
#3. Survival analysis
time <- minorstrokeonly$days
event <- minorstrokeonly$stroke
group <- minorstrokeonly$hyperglycemia
summary(time)
summary(event)
summary(group)
kmsurvival <- survfit(Surv(time,event) ~ 1)
summary (kmsurvival)
plot(kmsurvival)
plot(kmsurvival, fun="event")
kmsurvival <- survfit(Surv(time,event) ~ minorstrokeonly$hyperglycemia)
summary (kmsurvival)
plot(kmsurvival)
plot(kmsurvival, fun="event")
#4. Create two subgroups
#Subgroup 1 - Hyperglycemia in minorstrokeonly
minorstrokeonlyhyperglycemia <- minorstrokeonly[minorstrokeonly$hyperglycemia ==1,]
time <- minorstrokeonlyhyperglycemia$days
event <- minorstrokeonlyhyperglycemia$stroke
event <- as.numeric(event)
```

```
kmsurvival <- survfit(Surv(time,event) ~ 1)
summary (kmsurvival)
plot(kmsurvival)
kmsurvivalestimate <- survfit(Surv(time,event) ~ 1)
summary (kmsurvivalestimate, times=90)
#Estimate
1-0.885
#Upper CI
1-0.85
#Lower CI
1-0.921
#Subgroup 2 - Normoglycemia in minorstrokeonly
minorstrokeonlynormoglycemia <- minorstrokeonly[minorstrokeonly$hyperglycemia ==0,]
time <- minorstrokeonlynormoglycemia$days
event <- minorstrokeonlynormoglycemia$stroke
event <- as.numeric(event)
kmsurvival <- survfit(Surv(time,event) ~ 1)
summary (kmsurvival)
plot(kmsurvival)
kmsurvivalestimate <- survfit(Surv(time,event) ~ 1)
summary (kmsurvivalestimate, times=90)
#Estimate
1-0.923
#Upper CI
1-0.911
#Lower CI
1-0.935
#5. Create Kaplan-Meier curves
time <- minorstrokeonly$days
event <- minorstrokeonly$stroke
group <- minorstrokeonly$hyperglycemia
kmsurvival <- survfit(Surv(time,event) ~ group)
summary (kmsurvival)
#6. Add Log Rank test
event <- as.numeric(event)
survdiff(Surv(time,event) ~ group, data=minorstrokeonly)
#7. Make figure (NB include box(lwd=2))
#LARGE GRAPH
plot(kmsurvival, fun="event", col=c("red","blue"), ylim = c(0,1))
box(lwd=2)
axis(side=1, at = c(0,10,20,30,40,50,60,70,80,90), lwd=2)
axis(side=2, at = c(0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1), lwd=2)
#dlegend("bottomright", c("Normoglycemia (<180mg/dl)", "Hyperglycemia (>180mg/dl)"), col=c("red", "blue"), lty=1)
#---EXPORT TO POWERPOINT
#INSERT
plot(kmsurvival, fun="event", col=c("red","blue"), ylim = c(0,0.15), axes = FALSE, lty=1)
box(lwd=2)
axis(side=1, at = c(0, 30, 60, 90), lwd=2)
axis(side=2, at = c(0,0.01,0.02,0.03,0.04,0.05,0.06,0.07,0.08,0.09,0.1,0.11,0.12,0.13,0.14,0.15), lwd=2)
```

```
#legend("bottomright", c("SAPT", "DAPT"), col=c("red", "blue"), lty=1)
#---EXPORT TO POWERPOINT
#8. Cox model
#A. Unadjusted Cox model
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ minorstrokeonly$hyperglycemia, method="breslow")
summary(coxph)
#B. Adjusted Cox model (NB minorstroke not included as a covariate but DM included)
coxph <- coxph(Surv(time,event) ~ minorstrokeonly$hyperglycemia + minorstrokeonly$age + minorstrokeonly$female +
minorstrokeonly$black + minorstrokeonly$hispanic + minorstrokeonly$dapt + minorstrokeonly$htn +
minorstrokeonly$chf + minorstrokeonly$af + minorstrokeonly$cad + minorstrokeonly$valvedisease +
minorstrokeonly$carotiddisease + minorstrokeonly$smoking + minorstrokeonly$diabetes, method="breslow")
summary(coxph)

######################
#CLEAR ENVIRONMENT AND RESTART DATASET USING ONLY 2,330 PATIENTS WITH TIA THEN EXCLUDE 3 WITH MISSING
GLUCOSE DATA
#SA 10-2 - TIA only
# [X]
#1. Create group
#2. Get number of events by creating table
#3. Do survival analysis
#4. Create two groups within (exposed/not exposed) to obtain KM estimates + 95% CI
#5. Plot K-M curves
#6. Add Log Rank test
#7. Make figure (NB include box(lwd=2))
#8. Cox model
#1. Create group
#Have to re set up dataset and use data$F20Q01[data$F20Q01==1] for the 2,327 patients with TIA
#Patients with TIA
tiaonly <- data[data$F20Q01 == 1,]
tiaonly[is.na(tiaonly$F00Q28),]
nrow(tiaonly)
tiaonly <- tiaonly[!(is.na(tiaonly$F00Q28)),]
nrow(tiaonly) #3 subjects have been excluded from this analysis
#2. Get number of events by creating table
table(tiaonly$stroke, tiaonly$hyperglycemia)
#3. Survival analysis
time <- tiaonly$days
event <- tiaonly$stroke
group <- tiaonly$hyperglycemia
summary(time)
summary(event)
summary(group)
kmsurvival <- survfit(Surv(time,event) ~ 1)
summary (kmsurvival)
plot(kmsurvival)
plot(kmsurvival, fun="event")
kmsurvival <- survfit(Surv(time,event) ~ tiaonly$hyperglycemia)
```

```
summary (kmsurvival)
plot(kmsurvival)
plot(kmsurvival, fun="event")
#4. Create two subgroups
#Subgroup 1 - Hyperglycemia in tiaonly
tiaonlyhyperglycemia <- tiaonly[tiaonly$hyperglycemia ==1,]
time <- tiaonlyhyperglycemia$days
event <- tiaonlyhyperglycemia$stroke
event <- as.numeric(event)
kmsurvival <- survfit(Surv(time,event) ~ 1)
summary (kmsurvival)
plot(kmsurvival)
kmsurvivalestimate <- survfit(Surv(time,event) ~ 1)
summary (kmsurvivalestimate, times=90)
#Estimate
1-0.935
#Upper CI
1-0.902
#Lower CI
1-0.97
#Subgroup 2 - Normoglycemia in tiaonly
tiaonlynormoglycemia <- tiaonly[tiaonly$hyperglycemia ==0,]
time <- tiaonlynormoglycemia$days
event <- tiaonlynormoglycemia$stroke
event <- as.numeric(event)
kmsurvival <- survfit(Surv(time,event) ~ 1)
summary (kmsurvival)
plot(kmsurvival)
kmsurvivalestimate <- survfit(Surv(time,event) ~ 1)
summary (kmsurvivalestimate, times=90)
#Estimate
1-0.974
#Upper CI
1-0.967
#Lower CI
1-0.981
#5. Create Kaplan-Meier curves
time <- tiaonly$days
event <- tiaonly$stroke
group <- tiaonly$hyperglycemia
kmsurvival <- survfit(Surv(time,event) ~ group)
summary (kmsurvival)
#6. Add Log Rank test
event <- as.numeric(event)
survdiff(Surv(time,event) ~ group, data=tiaonly)
#7. Make figure (NB include box(lwd=2))
#LARGE GRAPH
plot(kmsurvival, fun="event", col=c("red","blue"), ylim = c(0,1))
box(lwd=2)
```

```
axis(side=1, at = c(0,10,20,30,40,50,60,70,80,90), lwd=2)
axis(side=2, at = c(0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1), lwd=2)
#dlegend("bottomright", c("Normoglycemia (<180mg/dl)", "Hyperglycemia (>180mg/dl)"), col=c("red", "blue"), lty=1)
#---EXPORT TO POWERPOINT
#INSERT
plot(kmsurvival, fun="event", col=c("red","blue"), ylim = c(0,0.15), axes = FALSE, lty=1)
box(lwd=2)
axis(side=1, at = c(0, 30, 60, 90), lwd=2)
axis(side=2, at = c(0,0.01,0.02,0.03,0.04,0.05,0.06,0.07,0.08,0.09,0.1,0.11,0.12,0.13,0.14,0.15), lwd=2)
#legend("bottomright", c("SAPT", "DAPT"), col=c("red", "blue"), lty=1)
#---EXPORT TO POWERPOINT
#8. Cox model
#A. Unadjusted Cox model
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ tiaonly$hyperglycemia, method="breslow")
summary(coxph)
#B. Adjusted Cox model (NB minorstroke not included as a covariate but DM included)
#WITH DIABETES
coxph <- coxph(Surv(time,event) ~ tiaonly$hyperglycemia + tiaonly$age + tiaonly$female + tiaonly$black +
tiaonly$hispanic + tiaonly$dapt + tiaonly$htn + tiaonly$chf + tiaonly$af + tiaonly$cad + tiaonly$valvedisease +
tiaonly$carotiddisease + tiaonly$smoking + tiaonly$diabetes, method="breslow")
summary(coxph)

############################################################
############################################################
### RE SET-UP DATASET FROM BEGINNING SKIPPING OVER PREVIOUS SECTION


#SA 10-3 - DAPT EFFECT in hyperglycemia
#1. Create group
#2. Get number of events by creating table
#3. Do survival analysis
#4. Create two groups within (exposed/not exposed) to obtain KM estimates + 95% CI
#5. Plot K-M curves
#6. Add Log Rank test
#7. Make figure (NB include box(lwd=2))
#8. Cox model

#For analysis of DAPT effects in those with/without hyperglycemia
#1. Create group
hyperglycemicgroup <- data[data$hyperglycemia ==1,]
#2. Get number of events by creating table (first term is on x axis, second term is on y axis)
table(hyperglycemicgroup$stroke, hyperglycemicgroup$dapt)
#3. Do survival analysis
time <- hyperglycemicgroup$days
event <- hyperglycemicgroup$stroke
group <- hyperglycemicgroup$dapt
summary(time)
summary(event)
```

```
summary(group)
kmsurvival <- survfit(Surv(time,event) ~ 1)
summary (kmsurvival)
plot(kmsurvival)
plot(kmsurvival, fun="event")
kmsurvival <- survfit(Surv(time,event) ~ group)
summary (kmsurvival)
plot(kmsurvival)
plot(kmsurvival, fun="event")
#4. Create two groups within (exposed/not exposed) to obtain KM estimates + 95% CI
#4A. DAPT + Hyperglycemia group
dapthyperglycemia <- hyperglycemicgroup[hyperglycemicgroup$dapt ==1,]
time <- dapthyperglycemia$days
event <- dapthyperglycemia$stroke
event <- as.numeric(event)
kmsurvival <- survfit(Surv(time,event) ~ 1)
summary (kmsurvival)
plot(kmsurvival)
kmsurvivalestimate <- survfit(Surv(time,event) ~ 1)
summary (kmsurvivalestimate, times=90)
#Estimate
1-0.895
#Upper CI
1-0.861
#Lower CI
1-0.932
#4B. SAPT + Hyperglycemia group
sapthyperglycemia <- hyperglycemicgroup[hyperglycemicgroup$dapt ==0,]
time <- sapthyperglycemia$days
event <- sapthyperglycemia$stroke
event <- as.numeric(event)
kmsurvival <- survfit(Surv(time,event) ~ 1)
summary (kmsurvival)
plot(kmsurvival)
kmsurvivalestimate <- survfit(Surv(time,event) ~ 1)
summary (kmsurvivalestimate, times=90)
#Estimate
1-0.911
#Upper CI
1-0.877
#Lower CI
1-0.946
#5. Plot Kaplan-Meier Curves:
time <- hyperglycemicgroup$days
event <- hyperglycemicgroup$stroke
group <- hyperglycemicgroup$dapt
kmsurvival <- survfit(Surv(time,event) ~ group)
summary (kmsurvival)
plot(kmsurvival, fun="event")
```

```
#LARGE GRAPH
plot(kmsurvival, fun="event", col=c("red","blue"), ylim = c(0,1))
box(lwd=2)
axis(side=1, at = c(0,10,20,30,40,50,60,70,80,90), lwd=2)
axis(side=2, at = c(0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1), lwd=2)
dlegend("bottomright", c("Normoglycemia (<180mg/dl)", "Hyperglycemia (>180mg/dl)"), col=c("red", "blue"), lty=1)
#---EXPORT TO POWERPOINT
#INSERT
plot(kmsurvival, fun="event", col=c("red","blue"), ylim = c(0,0.15), axes = FALSE, lty=1)
box(lwd=2)
axis(side=1, at = c(0, 30, 60, 90), lwd=2)
axis(side=2, at = c(0,0.01,0.02,0.03,0.04,0.05,0.06,0.07,0.08,0.09,0.1,0.11,0.12,0.13,0.14,0.15), lwd=2)
legend("bottomright", c("SAPT", "DAPT"), col=c("red", "blue"), lty=1)
#---EXPORT TO POWERPOINT
#6. Log Rank Test to compare curve and add as annotation to figure - NB event has to be numeric and not a factor
event <- as.numeric(event)
survdiff(Surv(time,event) ~ group, data=hyperglycemicgroup)
#7. Optimize in powerpoint
#8. Proportional Hazards Regression Modelling
#A. Unadjusted Cox model
time <- hyperglycemicgroup$days
event <- hyperglycemicgroup$stroke
group <- hyperglycemicgroup$dapt
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ group, method="breslow")
summary(coxph)
#B. Adjusted Cox model
#Maybe don't do for DAPT vs. SAPT comparisons
coxph <- coxph(Surv(time,event) ~ hyperglycemicgroup$dapt + hyperglycemicgroup$age + hyperglycemicgroup$female
+ hyperglycemicgroup$black + hyperglycemicgroup$minorstroke + hyperglycemicgroup$hispanic +
hyperglycemicgroup$htn + hyperglycemicgroup$chf + hyperglycemicgroup$af + hyperglycemicgroup$cad +
hyperglycemicgroup$valvedisease + hyperglycemicgroup$carotiddisease + hyperglycemicgroup$smoking +
hyperglycemicgroup$diabetes, method="breslow")
summary(coxph)

#SA 10-4 - DAPT EFFECT in normoglycemia
#1. Create group
#2. Get number of events by creating table
#3. Do survival analysis
#4. Create two groups within (exposed/not exposed) to obtain KM estimates + 95% CI
#5. Plot K-M curves
#6. Add Log Rank test
#7. Make figure (NB include box(lwd=2))
#8. Cox model
#1. Create group:
normoglycemicgroup <- data[data$hyperglycemia == 0,]
#2. Get number of events by creating table (first term is on x axis, second term is on y axis)
table(normoglycemicgroup$stroke, normoglycemicgroup$dapt)
#3. Do survival analysis
```

```
time <- normoglycemicgroup$days
event <- normoglycemicgroup$stroke
group <- normoglycemicgroup$dapt
summary(time)
summary(event)
summary(group)
kmsurvival <- survfit(Surv(time,event) ~ 1)
summary (kmsurvival)
plot(kmsurvival)
plot(kmsurvival, fun="event")
kmsurvival <- survfit(Surv(time,event) ~ group)
summary (kmsurvival)
plot(kmsurvival)
plot(kmsurvival, fun="event")
#4. Create two groups within (exposed/not exposed) to obtain KM estimates + 95% CI
#4A. DAPT + Normoglycemia group
daptnormoglycemia <- normoglycemicgroup[normoglycemicgroup$dapt ==1,]
time <- daptnormoglycemia$days
event <- daptnormoglycemia$stroke
event <- as.numeric(event)
kmsurvival <- survfit(Surv(time,event) ~ 1)
summary (kmsurvival)
plot(kmsurvival)
kmsurvivalestimate <- survfit(Surv(time,event) ~ 1)
summary (kmsurvivalestimate, times=90)
#Estimate
1-0.96
#Upper CI
1-0.951
#Lower CI
1-0.968
#4B. SAPT + Normoglycemia group
saptnormoglycemia <- normoglycemicgroup[normoglycemicgroup$dapt ==0,]
time <- saptnormoglycemia$days
event <- saptnormoglycemia$stroke
event <- as.numeric(event)
kmsurvival <- survfit(Surv(time,event) ~ 1)
summary (kmsurvival)
plot(kmsurvival)
kmsurvivalestimate <- survfit(Surv(time,event) ~ 1)
summary (kmsurvivalestimate, times=90)
#Estimate
1-0.937
#Upper CI
1-0.927
#Lower CI
1-0.948
#5. Plot K-M curves
time <- normoglycemicgroup$days
```

```
event <- normoglycemicgroup$stroke
group <- normoglycemicgroup$dapt
kmsurvival <- survfit(Surv(time,event) ~ group)
summary (kmsurvival)
plot(kmsurvival, fun="event")
#6. Add Log Rank test
event <- as.numeric(event)
survdiff(Surv(time,event) ~ group, data=normoglycemicgroup)
#7. Make figure (NB include box(lwd=2))
#LARGE GRAPH
plot(kmsurvival, fun="event", col=c("red","blue"), ylim = c(0,1))
box(lwd=2)
axis(side=1, at = c(0,10,20,30,40,50,60,70,80,90), lwd=2)
axis(side=2, at = c(0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1), lwd=2)
legend("bottomright", c("Normoglycemia (<180mg/dl)", "Hyperglycemia (>180mg/dl)"), col=c("red", "blue"), lty=1)
#---EXPORT TO POWERPOINT
#INSERT
plot(kmsurvival, fun="event", col=c("red","blue"), ylim = c(0,0.15), axes = FALSE, lty=1)
box(lwd=2)
axis(side=1, at = c(0, 30, 60, 90), lwd=2)
axis(side=2, at = c(0,0.01,0.02,0.03,0.04,0.05,0.06,0.07,0.08,0.09,0.1,0.11,0.12,0.13,0.14,0.15), lwd=2)
legend("bottomright", c("SAPT", "DAPT"), col=c("red", "blue"), lty=1)
#---EXPORT TO POWERPOINT
#8. Cox model
#B. Unadjusted Cox model
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ group, method="breslow")
summary(coxph)
#Adjusted Cox model
#Will not do it for this one as it is for DAPT/SAPT
#coxph <- coxph(Surv(time,event) ~ normoglycemicgroup$dapt + normoglycemicgroup$age +
normoglycemicgroup$female + normoglycemicgroup$black + normoglycemicgroup$hispanic +
normoglycemicgroup$minorstroke + normoglycemicgroup$htn + normoglycemicgroup$chf + normoglycemicgroup$af +
normoglycemicgroup$cad + normoglycemicgroup$valvedisease + normoglycemicgroup$carotiddisease +
normoglycemicgroup$smoking + normoglycemicgroup$diabetes, method="breslow")
#summary(coxph)

#11. Sensitivity analyses
#11.1 Glucose as continuous variable
#Base model assuming linear relationship between glucose and the hazard of subsequent stroke
#UNADJUSTED
time <- data$days
event <- data$stroke
event <- as.numeric(event)
survival <- Surv(time,event)
coxph <- coxph(survival ~ data$glucose, method="breslow")
summary(coxph)
#ADJUSTED
```

```
coxph <- coxph(survival ~ data$glucose + data$age + data$female + data$black + data$hispanic + data$dapt +
data$minorstroke + data$htn + data$chf + data$af + data$cad + data$valvedisease + data$carotiddisease +
data$smoking + data$diabetes, method="breslow")
summary(coxph)
rsq(coxph)
#Checking proportional hazards assumption
fit.coxph_zph <- cox.zph(coxph)
fit.coxph_zph
plot(fit.coxph_zph,var="data$glucose")
# Transform glucose as restrcited cubic spline  (5 knots between 0-1)
rcs_glucose <- rcs(data$glucose, quantile(data$glucose, c(0, .05, .275, .5, .725, .95, 1)))
rcscoxph <- coxph(survival ~ rcs_glucose + data$age + data$female + data$black + data$hispanic + data$dapt +
data$minorstroke + data$htn + data$chf + data$af + data$cad + data$valvedisease + data$carotiddisease +
data$smoking + data$diabetes, method="breslow")
summary(rcscoxph)
# likelihood ratio test for linearity
anova(coxph,rcscoxph, test="Chisq" )
#Figure 2 - plot of hazard of ischemic stroke vs serum blood glucose
glucose <- data$glucose
age <- data$age
sex <- data$female
black <- data$black
hispanic <- data$hispanic
dapt <- data$dapt
minorstroke <- data$minorstroke
htn <- data$htn
chf <- data$chf
af <- data$af
cad <- data$cad
valvedisease <- data$valvedisease
carotiddisease <- data$carotiddisease
smoking <- data$smoking
diabetes <- data$diabetes
dd <- datadist(glucose, age, sex, black, hispanic, dapt, minorstroke, htn, chf, af, cad, valvedisease, carotiddisease,
smoking, diabetes)
options(datadist="dd")
amod <- cph(survival ~ rcs(glucose,5) + age + sex + black + hispanic + dapt + minorstroke + htn + chf + af + cad +
valvedisease + carotiddisease + smoking + diabetes, x=TRUE, y=TRUE)
summary(amod5)
y <- Predict(amod,fun=exp, glucose)
theme_set(theme_bw())
ggplot(y, colfill="violetred3")+
  labs(x="Glucose (mg/dl)", y="Relative Hazard")+
  xlim(c(50,350))+
  theme(axis.title = element_text(size = 15, color = "black"), axis.text = element_text(size = 15))+
  theme(axis.line = element_line(size = 1))+
  geom_line(color = "firebrick", size=1.3)+
  geom_hline(yintercept = c(1), size=0.5, linetype="dashed")+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
```

```
      panel.background = element_blank(), axis.line = element_line(colour = "black"),
      panel.border = element_blank())+
  theme(axis.ticks.length=unit(0.25, "cm"))

#11.2. Propensity score matched analysis
data$hyperglycemia <- as.numeric(data$hyperglycemia)
psmodel <- glm(data$hyperglycemia ~ data$age + data$female + data$black + data$hispanic + data$dapt +
data$minorstroke + data$htn + data$chf + data$af + data$cad + data$valvedisease + data$carotiddisease +
data$smoking + data$diabetes, family=binomial, data=data)
summary(psmodel)
pscore <- psmodel$fitted.values
#Comparing characteristics before and after matching/PSM diagnostics
m.out <- matchit(data$hyperglycemia ~ data$age + data$female + data$black + data$hispanic + data$dapt +
data$minorstroke + data$htn + data$chf + data$af + data$cad + data$valvedisease + data$carotiddisease +
data$smoking + data$diabetes, family=binomial, data=data, caliper = 0.05, method = "nearest")
summary(m.out)
plot(m.out,type="hist")
plot(summary(m.out), xlim=c(0,2))

#Creating new object containing two matched groups
match1 <- match.data(m.out)

#Kaplan-Meier curves comparing propensity score-matched groups
time <- match1$days
event <- match1$stroke
group <- match1$hyperglycemia
summary(time)
summary(event)
summary(group)
kmsurvival <- survfit(Surv(time,event) ~ group)
summary(kmsurvival)
plot(kmsurvival)
#LARGE GRAPH
plot(kmsurvival, fun="event", col=c("red","blue"), ylim = c(0,1))
axis(side=1, at = c(0,10,20,30,40,50,60,70,80,90))
axis(side=2, at = c(0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1))
legend("bottomright", c("Normoglycemia (<180mg/dl)", "Hyperglycemia (>180mg/dl)"), col=c("red", "blue"), lty=1)
#INSERT
plot(kmsurvival, fun="event", col=c("red","blue"), ylim = c(0,0.15), axes = TRUE)
axis(side=1, at = c(0,10, 20, 30, 40, 50, 60, 70, 80, 90))
axis(side=2, at = c(0,0.01,0.02,0.03,0.04,0.05,0.06,0.07,0.08,0.09,0.1,0.11,0.12,0.13,0.14,0.15))
legend("bottomright", c("Normoglycemia (<180mg/dl)", "Hyperglycemia (>180mg/dl)"), col=c("red", "blue"), lty=1)
#Add in proportional hazards regression modelling
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ group, method="breslow")
summary(coxph)

#12.Final Sensitivity Analysis - Replacing final adjudicated etiology with infarct on imaging
#Infarct on imaging attributable to index event
```

```
data$F20Q04
str(data$F20Q04)
sum(is.na(data$F20Q04)==FALSE)
sum(is.na(data$F20Q04)==TRUE)
#5 subjects missing data on imaging attributable to index event
time <- data$days
event <- data$stroke
event <- as.numeric(event)
coxph <- coxph(Surv(time,event) ~ data$hyperglycemia + data$age + data$female + data$black + data$hispanic +
data$dapt + data$F20Q04 + data$htn + data$chf + data$af + data$cad + data$valvedisease + data$carotiddisease +
data$smoking + data$diabetes, method="breslow")
summary(coxph)
```

Table S1. Proportional hazards regression models performed separately in patients with and without hyperglycemia. The interaction term is derived from a model including all patients in the study sample including the term (hyperglycemia*final adjudicated etiology). Hazard ratios are for the association between minor stroke and the endpoint within the <180mg/dl and ≥180mg/dl strata.

Outcome	Minor stroke (n=2,304)	TIA/Other (n=2,574)	HR (95% CI)	P-value	P-value for interaction
Ischemic Stroke					
<180mg/dl	147/1,967	66/2,317	2.83 (2.11-3.80)[a]	<0.001	0.17
≥180mg/dl	37/337	17/257	1.84 (1.01-3.33)[a]	0.04	
Major Hemorrhage					
<180mg/dl	15/1,967	16/2,317	1.18 (0.58-2.39)[b]	0.65	-
≥180mg/dl	0/337	2/257	-	-	
Primary Endpoint[c]					
<180mg/dl	152/1,967	71/2,317	2.71 (2.04-3.60)[a]	<0.001	0.23
≥180mg/dl	40/337	18/257	1.90 (1.07-3.38)[a]	0.03	

a. Adjusted for age, sex, race, ethnicity, treatment assignment, hypertension, congestive cardiac failure, atrial fibrillation, coronary artery disease, valve disease, carotid disease, smoking and diabetes.
b. Adjusted for age, sex, race, ethnicity, treatment assignment and hypertension.
c. Subsequent ischemic stroke, myocardial infarction, ischemic vascular death.

Figure S1. Propensity-scores across hyperglycemic and normoglycemic subgroups before (LEFT PANELS) and after (RIGHT PANELS) the matching procedure.