**Supplemental Material**

# A. SUPPLEMENTARY FIGURES

## Figure A.1. Simplistic Representation of the Problem and Solution

|  |  |  |
| --- | --- | --- |
| **Figure A.1. Step-by-step description of the Problem and Solution** | | |
| **Step** | **Design** | **Simplistic illustration of dataset** |
| **A.1.1.** | Target Trial | Graphical user interface, application, Teams  Description automatically generated |
| **A.1.2.** | Naïve **aligned** wide dataset, crude | Graphical user interface, application, Teams  Description automatically generated |
| **A.1.3.** | **Misaligned** wide dataset, crude | Graphical user interface, application, Teams  Description automatically generated |
| **A.1.4.** | Cloned, aligned, crude | Initiate Treatment within 7 Days Clone |
| Graphical user interface, application, timeline, Teams  Description automatically generated |
| Initiate Treatment within 7 Days Clone |
| Graphical user interface, application, timeline, Teams  Description automatically generated |

**Legend:** This list of figures illustrates the problem of immortal time-bias.

**Target trial** – Such a trial does not exist in real life. We illustrated four patients: patients 2 and 4 are randomized to receiving benzodiazepine within 7 days, and we will know that information about randomization at admission. Patients 1 and 3 are randomized to not receive benzodiazepines within 7 days. They are then followed from the admission until they either die or until the end of the study. Here, they are censored because the study ended at 30 days.

**Step 1 Naïve aligned, crude.** Depicts crude naïve analysis **including the immortal time**, therefore **introducing immortal time bias**. Patient 4, the follow-up starts at admission, but the person is only exposed starting on day six. In this approach, some of the early deaths are misattributed to the unexposed group (i.e., patient 4 must survive up until day six to be able to receive the exposure and be classified as exposed).

**Step 2 Misaligned, crude.** Depicts crude or naïve analysis **excluding the immortal time to avoid introducing that bias.** Patient 4, this approach ignores the survival of this person for up to day six and then starts the follow-up at the exposure initiation day, while the unexposed start follow-up continues to be at admission. In this approach, the exposure may be a selective group of patients that are evolving with complications and are about to have the highest mortality rate in the subsequent days.

**Step 3 Cloned, aligned, crude.** First, we create pseudo-observations or ‘clones’ for each patient and then assign each of those clones to one of the two treatment strategies at the time of the hospital admission. Next, we proceed to artificially censor those who deviate from the assignment strategy. This is to ensure that clones follow their assigned strategy after the time zero. For instance, patient 2 in the “Initiate Treatment within 7 Days Clone dataset” was assigned to start treatment and started within the period that the patient was supposed to start. However, patient 3 was supposed to start treatment within seven days but deviated from the assigned strategy and needs to be censored. Additional illustrative examples can be found in previously published peer-reviewed publications.6-10

## Figure A.2. Standardized survival curves stratified by age

|  |
| --- |
| **A.2.1.**  A picture containing diagram  Description automatically generated |
|
| **A.2.2.**  **Chart, funnel chart  Description automatically generated** |

**Legend:** Blue: Benzodiazepine initiated within seven days post-AIS admission. Red: Benzodiazepine non-initiators. Shaded areas: 95% confidence intervals constructed using bootstrap with 500 replications. A.2.1.: patients 65-74 years of age. A.2.2.: patients ≥ 74 years of age.

## Figure A.3. Standardized survival curves across categories of stroke severity

|  |  |
| --- | --- |
| **A.3.1.**  Diagram  Description automatically generated with low confidence | **A.3.3.**  **Chart  Description automatically generated** |
| **A.3.2.**  **Chart  Description automatically generated** |

**Legend:** Blue: Benzodiazepine initiated within seven days post-AIS admission. Red: Benzodiazepine non-initiators. Shaded areas: 95% confidence intervals constructed using bootstrap with 500 replications. A.3.1.: minor stroke. A.3.2.: moderate stroke. A.3.3.: moderate to severe stroke.

# B. SUPPLEMENTARY TABLES

## Table B.1. ICD-9 and ICD-10 for Key Covariates

|  |  |  |
| --- | --- | --- |
| AIS-Associated Symptoms | ICD-9 | ICD-10 |
| Sleep disturbance | 780.5 | G47.9 |
| Anxiety, dissociative and somatoform disorders | 300.xx, 300.2, 308.xx | 300.2, F40, F41 |

**Legend:** AIS, acute ischemic stroke

## Table B.2. Model Parameters for Estimating Benzodiazepine Initiation Weights

|  |  |  |  |
| --- | --- | --- | --- |
| Dataset | Term | Estimate | Standard Error |
| Model Parameters | Intercept | -4.68 | 0.45 |
| Age | -0.01 | 0.01 |
| Race: white | 0.33 | 0.14 |
| NIHSS | 0.01 | 0.01 |
| Prescription Count | 0.11 | 0.00 |
| CMO Status | 2.02 | 0.14 |

**Table B.3. Time-varying additional characteristics of patients, by benzodiazepine initiator versus non-benzodiazepine initiator (with start of follow-up aligned at admission)**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Benzodiazepine initiator (N = 389) | Benzodiazepine non-initiator  (N = 2195) | SMD |
| In-hospital Measures of Stroke Severity and Complications (recorded during first day of admission) a (%) | | | |
| Seizure-Like Event (SLE) | 54 (13.9) | 368 (16.8) | 0.080 |
| Delirium | 28 (7.2) | 128 (5.8) | 0.055 |
| EEG | 6 (1.5) | 41 (1.9) | 0.025 |
| Observed Large Vessel Occlusion | 77 (30.3) | 336 (28.9) | 0.031 |
| Prescription Count, Mean (SD) | 15.09 (12.80) | 7.00 (10.57) | 0.689 |
| IV injection of tissue plasminogen activator (tPA) | 34 (8.7) | 138 (6.3) | 0.093 |
| Endovascular thrombectomy (EVT) Performed | 13 (3.3) | 49 (2.2) | 0.067 |
| Computed tomography (CT/CAT) Scan | 240 (61.7) | 1330 (60.6) | 0.023 |
| Magnetic resonance imaging (MRI) of the brain | 207 (53.2) | 1118 (50.9) | 0.046 |
| Comfort Measures Only Determination During Admission, aka CMO (%) | | | |
| CMO % |  |  | 0.544 |
| Determined CMO during day 0 or 1 | 28 (7.2) | 31 (1.4) |  |
| Determined CMO during 2 or after | 72 (18.5) | 111 (5.1) |  |
| Not CMO during Stroke admission | 289 (74.3) | 2053 (93.5) |  |

**Legend:** This table describes patient characteristics among benzodiazepine initiators and versus non-initiators, after standardization by age, race, NIHSS and prescription count on the day of admission. SD: standard deviation; SMD: standardized mean difference; EEG: electroencephalogram; ED: emergency department; FRI: fall-related injuries; SLE: seizure-like event.

a In-hospital measures of stroke severity and complications are time-varying covariates – the values will be updated daily for everyone for the main analysis. For simplicity, we presented in the table above just the values obtained during the first day of admission. Overall SMD threshold for considering substantial different among two groups: 0.2. We noted NIHSS (consequently CMO status (more often benzodiazepines are given to CMO patients), and in-hospital prescription count (greater count among benzodiazepine initiators).

## Table B.4. Characteristics of Patients, restricted to Mild Stroke Severity

|  |  |  |  |
| --- | --- | --- | --- |
|  | Benzodiazepine initiator  (N = 188) | Benzodiazepine non-initiator  (N = 1,223) | SMD |
| Socio-Demographic Characteristics (recorded at admission) | | | |
| Age, mean (SD) | 75.93 (7.58) | 76.90 (8.08) | 0.125 |
| Female (%) | 103 (54.8) | 523 (42.8) | 0.242 |
| Non-White (%) | 19 (10.2) | 171 (14.5) | 0.130 |
| Ethnicity Hispanic or Latino (%) | 1 (0.5) | 11 (0.9) | 0.046 |
| Primary Insurance Medicare or other government (vs private) (%) | 151 (80.3) | 989 (80.9) | 0.016 |
| Baseline Medication Use (recorded during the 90 days before admission) | | | |
| Prescription Count, Mean (SD) | 8.41 (31.90) | 5.05 (18.01) | 0.130 |
| Categories of Medication use (%) |  |  | 0.176 |
| No prescription recorded | 121 (64.4) | 860 (70.3) |  |
| 1-4 drugs | 24 (12.8) | 168 (13.7) |  |
| 5-9 drugs | 14 (7.4) | 62 (5.1) |  |
| >9 drugs | 29 (15.4) | 133 (10.9) |  |
| Baseline Clinical Characteristics (recorded during 12 months before admission) | | | |
| Charlson Comorbidity Score, mean (SD) | 1.43 (1.88) | 1.20 (1.69) | 0.127 |
| Pertinent Comorbid Conditions (%) |  |  |  |
| Sleep disturbance, insomnia | 8 (4.3) | 36 (2.9) | 0.070 |
| Anxiety, dissociative, somatoform disorders | 15 (8.0) | 68 (5.6) | 0.096 |
| Baseline Health-Resource Utilization (recorded during 12 months before admission), % | | | |
| Hospitalization | 48 (25.5) | 306 (25.0) | 0.012 |
| Emergency Room | 20 (10.6) | 145 (11.9) | 0.039 |
| Fall-Related Injury | 13 (6.9) | 128 (10.5) | 0.126 |
| Seizure-Like Event | 8 (4.3) | 82 (6.7) | 0.108 |
| EEG | 4 (2.1) | 14 (1.1) | 0.078 |
| In-hospital Measures of Stroke Severity and Complications (recorded during first day of admission) a (%) | | | |
| Seizure-Like Event | 20 (10.6) | 200 (16.4) | 0.168 |
| Delirium | 9 (4.8) | 44 (3.6) | 0.059 |
| EEG | 2 (1.1) | 25 (2.0) | 0.079 |
| Observed Large Vessel Occlusion | 13 (10.6) | 83 (12.9) | 0.072 |
| Prescription Count, Mean (SD) | 12.79 (10.04) | 6.20 (8.61) | 0.705 |
| IV injection of tissue plasminogen activator (tPA) | 3 (1.6) | 22 (1.8) | 0.016 |
| Endovascular thrombectomy (EVT) Performed b | 2 (1.1) | 9 (0.7) | 0.035 |
| Computed tomography (CT/CAT) Scan | 107 (56.9) | 672 (54.9) | 0.040 |
| Magnetic resonance imaging (MRI) of the brain | 106 (56.4) | 631 (51.6) | 0.096 |
| Comfort Measures Only Determination During Admission, aka CMO (%) | | | |
| CMO % |  |  | 0.239 |
| Determined CMO during day 0 or 1 | 1 (0.5) | 0 (0.0) |  |
| Determined CMO during 2 or after | 10 (5.3) | 18 (1.5) |  |
| Not CMO during Stroke admission | 177 (94.1) | 1205 (98.5) |  |

**Legend:** SD: standard deviation; SMD: standardized mean difference; EEG: electroencephalogram; CMO: Comfort Measures Only. Footnotes: Sample restricted to Mild Stroke Severity (0-4)**.** a In-hospital measures of stroke severity and complications are time-varying covariates – the values will be updated daily for everyone for the main analysis. **b** No prescription recorded: the prescription information was a) missing from the MGB structured health system data warehouse, b) the patient was not taking any prescription drug, c) the patient was taking prescription drugs given elsewhere (e.g., over the counter, prescribed and recorded in other healthcare system), d) other unknown reason. For simplicity, we presented in Table B.3 just the values obtained during the first day of admission. Notes: Overall SMD threshold for considering substantial different among two groups: 0.2. We noted imbalances for in-hospital prescription count (greater count among benzodiazepines initiators), frequency of brain MRI and seizure-like events (more often among benzodiazepines initiators). On average, younger patients received benzodiazepines more often. Benzodiazepine indication among minor stroke survivors appears to be often used for keeping people still in imaging studies (agitation, anxiety).

## Table B.5. Characteristics of Patients, restricted to Moderate Stroke Severity

|  |  |  |  |
| --- | --- | --- | --- |
|  | Benzodiazepine initiator  (N = 131) | Benzodiazepine non-initiator  (N = 768) | SMD |
| Socio-Demographic Characteristics (recorded at admission) | | | |
| Age, mean (SD) | 78.78 (8.26) | 79.10 (8.64) | 0.038 |
| Female (%) | 78 (59.5) | 412 (53.6) | 0.119 |
| Non-White (%) | 9 (7.3) | 129 (18.2) | 0.334 |
| Ethnicity Hispanic or Latino (%) | 0 (0.0) | 15 (2.1) | 0.208 |
| Primary Insurance Medicare or other government (vs private) (%) | 107 (81.7) | 633 (82.5) | 0.022 |
| Baseline Medication Use (recorded during the 90 days before admission) | | | |
| Prescription Count, Mean (SD) | 8.97 (29.73) | 3.91 (18.47) | 0.204 |
| Categories of Medication use (%) |  |  | 0.276 |
| No prescription recorded | 96 (73.3) | 630 (82.0) |  |
| 1-4 drugs | 10 (7.6) | 65 (8.5) |  |
| 5-9 drugs | 7 (5.3) | 21 (2.7) |  |
| >9 drugs | 18 (13.7) | 52 (6.8) |  |
| Baseline Clinical Characteristics (recorded during 12 months before admission) | | | |
| Charlson Comorbidity Score, mean (SD) | 1.09 (1.82) | 1.01 (1.66) | 0.045 |
| Pertinent Comorbid Conditions (%) |  |  |  |
| Sleep disturbance, insomnia | 4 (3.1) | 16 (2.1) | 0.061 |
| Anxiety, dissociative, somatoform disorders | 8 (6.1) | 28 (3.6) | 0.114 |
| Baseline Health-Resource Utilization (recorded during 12 months before admission), % | | | |
| Hospitalization | 28 (21.4) | 167 (21.7) | 0.009 |
| Emergency Room | 19 (14.5) | 72 (9.4) | 0.159 |
| Fall-Related Injury | 21 (16.0) | 86 (11.2) | 0.141 |
| Seizure-Like Event | 5 (3.8) | 45 (5.9) | 0.095 |
| EEG | 0 (0.0) | 6 (0.8) | 0.125 |
| EEG (Long-term) | 0 (0.0) | 1 (0.1) | 0.051 |
| In-hospital Measures of Stroke Severity and Complications (recorded during first day of admission) a (%) | | | |
| Seizure-Like Event | 22 (16.8) | 146 (19.0) | 0.058 |
| Delirium | 17 (13.0) | 68 (8.9) | 0.133 |
| EEG | 2 (1.5) | 12 (1.6) | 0.003 |
| Observed Large Vessel Occlusion | 35 (40.7) | 170 (42.7) | 0.041 |
| Prescription Count, Mean (SD) | 16.19 (13.52) | 7.51 (11.67) | 0.687 |
| IV injection of tissue plasminogen activator (tPA) | 23 (17.6) | 88 (11.5) | 0.174 |
| Endovascular thrombectomy (EVT) Performed b | 7 (5.3) | 25 (3.3) | 0.103 |
| Computed tomography (CT/CAT) Scan | 86 (65.6) | 523 (68.1) | 0.052 |
| Magnetic resonance imaging (MRI) of the brain | 66 (50.4) | 400 (52.1) | 0.034 |
| Comfort Measures Only Determination During Admission, aka CMO (%) | | | |
| CMO % |  |  | 0.519 |
| Determined CMO during day 0 or 1 | 9 (6.9) | 12 (1.6) |  |
| Determined CMO during 2 or after | 30 (22.9) | 64 (8.3) |  |
| Not CMO during Stroke admission | 92 (70.2) | 692 (90.1) |  |

**Legend:** SD: standard deviation; SMD: standardized mean difference; EEG: electroencephalogram; CMO: Comfort Measures Only. Footnotes: Sample restricted to Moderate Stroke Severity (5-15)**.** a In-hospital measures of stroke severity and complications are time-varying covariates – the values will be updated daily for everyone for the main analysis. **b** No prescription recorded: the prescription information was a) missing from the MGB structured health system data warehouse, b) the patient was not taking any prescription drug, c) the patient was taking prescription drugs given elsewhere (e.g., over the counter, prescribed and recorded in other healthcare system), d) other unknown reason. For simplicity, we presented in Table B.4 just the values obtained during the first day of admission. Notes: Overall SMD threshold for considering substantial different among two groups: 0.2. We noted imbalances for in-hospital prescription count (greater count among benzodiazepines initiators), frequency of brain MRI and seizure-like events (more often among benzodiazepines initiators). On average, younger patients received benzodiazepines more often. Benzodiazepine indication among minor stroke survivors appears to be often used for keeping people still in imaging studies (agitation, anxiety).

## Table B.6. Characteristics of Patients, restricted to Moderate-to-Severe Stroke Severity

|  |  |  |  |
| --- | --- | --- | --- |
|  | Benzodiazepine initiator  (N = 70) | Benzodiazepine non-initiator  (N = 204) | SMD |
| Socio-Demographic Characteristics (recorded at admission) | | | |
| Age, mean (SD) | 81.87 (8.78) | 78.81 (8.89) | 0.347 |
| Female (%) | 42 (60.0) | 103 (50.5) | 0.192 |
| Non-White (%) | 6 (9.7) | 25 (13.6) | 0.122 |
| Ethnicity Hispanic or Latino (%) | 1 (1.6) | 7 (3.8) | 0.135 |
| Primary Insurance Medicare or other government (vs private) (%) | 56 (80.0) | 161 (78.9) | 0.027 |
| Baseline Medication Use (recorded during the 90 days before admission) | | | |
| Prescription Count, Mean (SD) | 4.30 (15.08) | 2.87 (12.95) | 0.102 |
| Categories of Medication use (%) |  |  | 0.303 |
| No prescription recorded | 58 (82.9) | 171 (83.8) |  |
| 1-4 drugs | 4 (5.7) | 14 (6.9) |  |
| 5-9 drugs | 0 (0.0) | 6 (2.9) |  |
| >9 drugs | 8 (11.4) | 13 (6.4) |  |
| Baseline Clinical Characteristics (recorded during 12 months before admission) | | | |
| Charlson Comorbidity Score, mean (SD) | 0.70 (1.53) | 0.74 (1.43) | 0.024 |
| Pertinent Comorbid Conditions (%) |  |  |  |
| Sleep disturbance, insomnia | 2 (2.9) | 3 (1.5) | 0.095 |
| Anxiety, dissociative, somatoform disorders | 1 (1.4) | 5 (2.5) | 0.074 |
| Baseline Health-Resource Utilization (recorded during 12 months before admission), % | | | |
| Hospitalization | 9 (12.9) | 32 (15.7) | 0.081 |
| Emergency Room | 5 (7.1) | 10 (4.9) | 0.094 |
| Fall-Related Injury | 8 (11.4) | 13 (6.4) | 0.178 |
| Seizure-Like Event | 4 (5.7) | 6 (2.9) | 0.137 |
| EEG | 1 (1.4) | 2 (1.0) | 0.041 |
| EEG (Long-term) | 70 (100.0) | 204 (100.0) | <0.001 |
| In-hospital Measures of Stroke Severity and Complications (recorded during first day of admission) a (%) | | | |
| Seizure-Like Event | 12 (17.1) | 22 (10.8) | 0.184 |
| Delirium | 2 (2.9) | 16 (7.8) | 0.223 |
| EEG | 2 (2.9) | 4 (2.0) | 0.058 |
| Observed Large Vessel Occlusion | 29 (64.4) | 83 (69.2) | 0.100 |
| Prescription Count, Mean (SD) | 19.20 (16.40) | 9.90 (15.24) | 0.588 |
| IV injection of tissue plasminogen activator (tPA) | 8 (11.4) | 28 (13.7) | 0.069 |
| Endovascular thrombectomy (EVT) Performed b | 4 (5.7) | 15 (7.4) | 0.066 |
| Computed tomography (CT/CAT) Scan | 47 (67.1) | 135 (66.2) | 0.021 |
| Magnetic resonance imaging (MRI) of the brain | 35 (50.0) | 87 (42.6) | 0.148 |
| Comfort Measures Only Determination During Admission, aka CMO (%) | | | |
| CMO % |  |  | 1.094 |
| Determined CMO during day 0 or 1 | 18 (25.7) | 19 (9.3) |  |
| Determined CMO during 2 or after | 32 (45.7) | 29 (14.2) |  |
| Not CMO during Stroke admission | 20 (28.6) | 156 (76.5) |  |

**Legend:** SD: standard deviation; SMD: standardized mean difference; EEG: electroencephalogram; CMO: Comfort Measures Only. Footnotes: Sample restricted to Moderate to Severe Stroke Severity (15-19)**.** a In-hospital measures of stroke severity and complications are time-varying covariates – the values will be updated daily for everyone for the main analysis. **b** No prescription recorded: the prescription information was a) missing from the MGB structured health system data warehouse, b) the patient was not taking any prescription drug, c) the patient was taking prescription drugs given elsewhere (e.g., over the counter, prescribed and recorded in other healthcare system), d) other unknown reason. For simplicity, we presented in Table B.5 just the values obtained during the first day of admission. Notes: Overall SMD threshold for considering substantial different among two groups: 0.2. We noted imbalances for in-hospital prescription count (greater count among benzodiazepines initiators), frequency of brain MRI and seizure-like events (more often among benzodiazepines initiators). On average, younger patients received benzodiazepines more often. Benzodiazepine indication among minor stroke survivors appears to be often used for keeping people still in imaging studies (agitation, anxiety).

## Table B.7. Proportion of Benzodiazepine by Type in Benzodiazepine-Initiators

|  |  |
| --- | --- |
| Benzodiazepine | Percentage a |
| Lorazepam | 89.16 |
| Clonazepam | 6.09 |
| Alprazolam | 5.19 |
| Diazepam | 2.26 |

**Legend:** The proportion of benzodiazepines used, within the window of 7 after stroke admission. Among 389 benzodiazepine initiators. Benzodiazepines used with a frequency of less than 2% were omitted.

## Table B.8. Pattern of Anticonvulsant Use Post-Benzodiazepine Initiation

|  |  |
| --- | --- |
| Benzodiazepine | Frequency |
| Carbamazepine | 2 (3.2) |
| Phenobarbital | 3 (4.8) |
| Phenytoin | 3 (4.8) |
| Valproic Acid | 2 (3.2) |
| Gabapentin | 27 (43.5) |
| Lamotrigine | 2 (3.2) |
| Levetiracetam | 26 (41.9) |

**Legend:** The pattern of the first non-Benzodiazepine anticonvulsant dispensation on or after the day of benzodiazepine dispensation, within the window of 7 after stroke admission. Among 389 benzodiazepine initiators, 28 (14.81%) received a second anticonvulsant on or after the day of benzodiazepine initiation.

## Table B.9. Main Results, Stratified by Age Groups

|  |  |  |  |
| --- | --- | --- | --- |
| Standardized (Addressing Selection + Confounding), 65-74 years | | | |
| Do not initiate | 0.017 | 4.630 | 0.040 |
| Initiate | 0.013 | 0.003 | 0.025 |
| Risk Difference | -0.004 | -0.031 | 0.011 |
| Hazard Ratio | 1.349 | 0.576 | >50 |
| Standardized (Addressing Selection + Confounding), Restricted to ≥ 74 years | | | |
| Do not initiate | 0.024 | 3.058 | 0.051 |
| Initiate | 0.037 | 8.350 | 0.090 |
| Risk Difference | 0.013 | -0.023 | 0.077 |
| Hazard Ratio | 2.048 | 0.336 | >50 |

## Table B.10. Main Results, Stratified by Stroke Severity

|  |  |  |  |
| --- | --- | --- | --- |
| Standardized (Addressing Selection + Confounding), Restricted to Minor Stroke | | | |
| Do not initiate | 0.003 | 4.311 | 0.008 |
| Initiate | 0.004 | 6.763 | 0.010 |
| Risk Difference | 8.161 | -0.005 | 0.009 |
| Hazard Ratio | 1.122 | 0.150 | >50 |
| Standardized (Addressing Selection + Confounding), Restricted to Moderate Stroke | | | |
| Do not initiate | 0.030 | 0.008 | 0.054 |
| Initiate | 0.028 | 0.014 | 0.044 |
| Risk Difference | -0.002 | -0.027 | 0.021 |
| Hazard Ratio | 1.400 | 0.684 | 3.755 |
| Standardized (Addressing Selection + Confounding), Restricted to Severe Stroke | | | |
| Do not initiate | 0.135 | 0.037 | 0.236 |
| Initiate | 0.115 | 0.060 | 0.297 |
| Risk Difference | -0.021 | -0.122 | 0.1998 |
| Hazard Ratio | 1.433 | 0.563 | 6.889 |

**Table B.11. The RECORD Statement**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Item No. | | STROBE items.51 | Location where items are reported | RECORD items | Location in manuscript where items are reported |
| Title and abstract | | | | | |
|  | 1 | (a) Indicate the study’s design with a commonly used term in the title or the abstract  (b) Provide in the abstract an informative and balanced summary of what was done and what was found | (a): Abstract (page 4)  (b): Abstract (page 4) | RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included.  RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract.  RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract. | 1.1: Abstract  1.2: Abstract  1.3 Abstract |
| Introduction | | | | | |
| Background rationale | 2 | Explain the scientific background and rationale for the investigation being reported | Introduction |  |  |
| Objectives | 3 | State specific objectives, including any prespecified hypotheses | Introduction |  |  |
| Methods | | | | | |
| Study Design | 4 | Present key elements of study design early in the paper | Methods |  |  |
| Setting | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection | Methods |  |  |
| Participants | 6 | *(a) Cohort study* - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  *Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  *Cross-sectional study* - Give the eligibility criteria, and the sources and methods of selection of participants  *(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed  *Case-control study* - For matched studies, give matching criteria and the number of controls per case | (a): Methods  (b) NA | RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided.  RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided.  RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of patients with linked data at each stage. | 6.1: Methods  6.2: NA  6.3: NA |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable. | Methods | RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided. | Methods |
| 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. | Methods |  |  |
| Bias | 9 | Describe any efforts to address potential sources of bias | Methods | Bias |  |
| Study size | 10 | Explain how the study size was arrived at | Methods |  |  |
| Quantitative variables | 11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why | Methods |  |  |
| Statistical methods | 12 | (a) Describe all statistical methods, including those used to control for confounding  (b) Describe any methods used to examine subgroups and interactions  (c) Explain how missing data were addressed  (d) *Cohort study* - If applicable, explain how loss to follow-up was addressed  *Case-control study* - If applicable, explain how matching of cases and controls was addressed  *Cross-sectional study* - If applicable, describe analytical methods taking account of sampling strategy  (e) Describe any sensitivity analyses | (a): Methods  (b): Methods  (c): Methods  (d): Methods (e): Methods |  |  |
| Data access and cleaning methods |  |  |  | RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population.  RECORD 12.2: Authors should provide information on the data cleaning methods used in the study. | 12.1: Methods  12.2: Methods |
| Linkage |  |  |  | RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided. | NA |
| Results |  |  |  |  |  |
| Participants | 13 | (a) Report the numbers of patients at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed)  (b) Give reasons for non-participation at each stage.  (c) Consider use of a flow diagram | (a): Results; Figure 1; Table 2.  (b): Figure 1  (c): Figure 1 | RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram. | 13.1 Results; Methods |
| Descriptive data | 14 | (a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders  (b) Indicate the number of participants with missing data for each variable of interest  (c) *Cohort study* - summarise follow-up time (*e.g.*, average, and total amount) | (a): Results  (b): Results  (c): Results |  |  |
| Outcome data | 15 | *Cohort study* - Report numbers of outcome events or summary measures over time  *Case-control study* - Report numbers in each exposure category, or summary measures of exposure  *Cross-sectional study* - Report numbers of outcome events or summary measures | Results |  |  |
| Main results | 16 | (a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included  (b) Report category boundaries when continuous variables were categorized  (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful period | (a): Results  (b): Results:  (c): Results |  |  |
| Other analyses | 17 | Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses | Results |  |  |
| Discussion | | |  |  |  |
| Key results | 18 | Summarise key results with reference to study objectives | Discussion |  |  |
| Limitations | 19 | Discuss limitations of the study, considering sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias | Discussion | RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported. | Discussion |
| Interpretation | 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence | Discussion |  |  |
| Generalisability | 21 | Discuss the generalisability (external validity) of the study results | Discussion |  |  |
| 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 | Acknowledgments |  |  |
| Accessibility of protocol, raw data, and programming code |  |  |  | RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code. | Supplementary tables, figures, and text. |

\*Checklist is protected under Creative Commons Attribution ([CC BY](http://creativecommons.org/licenses/by/4.0/)) license.

# C. SUPPLEMENTARY TEXT

## C.1. Operational Definitions for Measures of Interest

### C.1.1. Socio-demographic factors

We used data from the Enterprise Data Warehouse (EDW) to obtain several demographic factors such as age, sex, race, ethnicity, language, and addresses. We obtained a series of clinical factors and then derived some validated summary measures, such as the Charlson Comorbidity Index (CCI,) which can predict a patient’s mortality for short and long term by categorizing a range of comorbidities (i.e., a total of 22 conditions such as heart disease). We derived the CCI from the Get with The Guidelines (GWTG) and EDW datasets (baseline outpatient and in-hospital data). The CCI was based on the International Classification of Diseases (ICD) diagnosis codes.52

### C.1.2. Stroke Severity

Stroke severity is a strong predictor of anticonvulsant initiation, seizures, and mortality.53 Factors of stroke severity include cortical infarction and stroke extension,54 neuroimaging traits (e.g., infarct volume and location, diffusion-perfusion mismatch, poor collateral blood flow, development of cerebral edema in non-lacunar ischemic stroke), and ischemic stroke mechanism. We used the validated National Institutes of Health stroke severity score (NIHSS), which is a summary measure of stroke severity and may be associated with benzodiazepine initiation and mortality. We obtained NIHSS from the GWTG dataset (in-hospital data). The NIHSS score is defined as the sum of 15 individually evaluated elements, and ranges from 0 to 42. Stroke severity scores can be used as a continuous measure or categorized as no stroke symptoms (0), minor stroke (NIHSS 1-4), moderate stroke (NIHSS 5-15), moderate to severe stroke (NIHSS 16-20).55,56

### C.1.3. Prescription Patterns

Polypharmacy is a major risk factor for adverse drug reactions, and has been associated with benzodiazepine use and mortality.57 We examined data from 90 days prior to stroke admission to estimate several measures of prescription patterns prior to AIS.57 We also divided the prescription count into 4 categories; no prescription recorded, 1 to 4 drugs, 5 to 9 drugs and more than 9 drugs. No benzodiazepines were included in any prescription count category. If a patient did not have any prescription recorded, the prescription information was either a) missing from the MGB structured health system data warehouse, b) the patient was not taking any prescription drug, c) the patient was taking prescription drugs given elsewhere (e.g., over the counter, prescribed and recorded in other healthcare system), d) other unknown reason.

### C.1.4. Healthcare Utilization

We examined several measures of healthcare utilization, including visit frequency and location and institutionalization (e.g., frequency of Emergency Department admissions). We obtained discharge status, length of stay, dementia, restlessness, claustrophobia, delirium, tremor, deep vein thrombosis (DVT), long-term EEG monitoring, aspiration pneumonia, intubation, tracheostomy, and percutaneous endoscopic gastrostomy (PEG), for each stroke admission.

### C.1.5. ED visits

An ED visit is a marker of health resource utilization, time-varying severity (which could represent drug adverse effects, disease complications, decompensated comorbid conditions, etc.). We obtained a baseline ED visit variable with a count of ED visits during the 6 months prior to the stroke admission date.

### C.1.6. Technical Section for Addressing Immortal-Person Time

In this study, the trial is about "start treatment within the first seven days after admission" in the same pattern that we would have seen people start treatment in real-life, with everyone starting on day seven if they haven't already done so. In this approach, we first clone the population. Therefore, there is no table 1 to illustrate differences across the two groups, they are identical (one clone is assigned treatment and the other clone is not). Then, we apply censoring weights as they violate one of the protocols. At that time, we use the baseline and time-dependent covariates that affect the change in strategy. In this design, results may be sensitive to when during the seven days people start treatment. We provide more details in the next paragraphs:

***C.1.7. Treatment Strategies***

We defined the following treatment strategies: a) initiate benzodiazepines within seven days of admission, or b) do not treat with benzodiazepines during these seven days. In the target trial emulation, we obtained information on benzodiazepine use from inpatient and outpatient pharmacy claims data. Unlike in a randomized trial, we could not know which treatment strategy the patient had been assigned to until the day of the prescription (for those exposed) or seven days post-AIS (for those unexposed). Therefore, for patients who died within seven days without initiation, we could not know if they would have received treatment had they not died. Thus, for the seven days post-AIS, follow-up days until treatment initiation or death count towards both treatment strategies. To carry out such counting, we duplicated the dataset, creating “clones” of each patient so that each clone would contribute to both treatment strategies until their strategy is known. The follow-up of a clone is censored when its treatment strategy is violated, i.e., clones assigned to no-initiation were censored if they initiated treatment within those seven days and clones assigned to initiation were censored if they did not initiate by day seven. At most one clone remains in the dataset after the first seven days of follow-up. Lastly, the generated pseudo-population of clones for each treatment strategy is weighted by inverse probability weights to correct for the fact that the same patient does not adhere to both treatment strategies.33-37 To mimic randomization, these weights also account for the non-random treatment initiation.21,38-40

This "cloning-censoring-weighting” approach has been used in previous studies33 and avoids a common methodological problem in observational data in the presence of staggered treatment initiation. Solving this problem requires aligning start of follow-up and exposure assignment. There are two possible approaches that correspond to two different target trials. First, our proposed target trial, where both treatment assignment and follow-up start at baseline (i.e., hospital admission), the first time at which treatment can be initiated. In a randomized trial the assigned treatment strategy would be known at that time, even if no treatment was initiated that day; in observational studies, assigning patient “clones” to each treatment strategy allows the time-zero alignment, as explained above.33-37 Second, we could have emulated a trial where patients are randomized each day throughout the first week post-AIS as they become eligible (e.g., a new indication). There would be seven time-zeros when follow-up would start for those assigned to initiating and not initiating on that day in the target trial and, as well as in the observational emulation. Both approaches help avoid selection and immortal time bias by ensuring that the start of follow-up and treatment assignment are aligned, as they would be in a randomized trial.

Alternative traditional approaches to deal with grace periods for exposure initiation have included the following: First, if benzodiazepine initiators are simply compared to non-initiators and the day of AIS admission is considered time zero, the start of follow-up would not be aligned with exposure initiation unless treatment initiation occurs exclusively at baseline. This is generally not the case, so to be treated, patients have already survived several days. The treated group would therefore have no deaths during the first days of follow-up, a bias that is referred to as an“immortal time bias”.21,22,41 More generally, this bias arises in naïve analyses which use post-baseline information to define exposure strategies.22 Second, an analysis that instead started follow-up for both the treated and untreated groups after the seven-day treatment initiation window would be missing deaths in both groups that occurred during that window. If mortality differed between groups, they would no longer be comparable, even in a randomized trial (with randomization at admission). Excluding the first week of follow-up would miss potential acute effects of benzodiazepines and would deplete the sample of the most susceptible patients. Third, starting follow-up of exposed patients the day of treatment initiation and of unexposed ones the day of admission would also be biased in the presence of mortality trends during the first days post-AIS, since those initiating treatment later would have a different baseline risk.

***C.1.7.1. Cloning and Censoring:*** In the “Initiate Treatment within seven days” dataset, we create a copy of the original dataset but keep data points on clones that started treatment within the grace period and patient clones that were censored at the end of the grace period because they did not start treatment within grace period (censor unless it is during grace period). In the “Do not Initiate Treatment within seven days” dataset, we create a copy of the original dataset but keep data points on clones that never started treatment, and clones that started treatment, before they started (i.e., they are being censored for starting, censor if start treatment any time during grace period). Then, we create a cloned dataset that consists of the two datasets combined (i.e., cloned, and censored, and now ready to proceed with weighting).

***C.1.7.2. Weighting:*** In original data, we fit a weight model among people yet to start treatment (model for treatment initiation). Then, in the cloned dataset, we apply weights [Pr (uncensored at time t | uncensored at time t – 1]. In the treatment arm: the weight contribution is 1 during grace period because Pr (uncensored | grace period) = 1 even if the patient does not start treatment.

Patients who have started treatment within the grace period (i.e., protocol compliant) are therefore uncensored at the end of the grace period (e.g., as illustrated in Supplementary Figure S1-C, individual 2), but they need to receive an upweight to account for those who deviated from protocol (i.e., those who did not start treatment but were supposed to start, based on their assigned strategy – as illustrated in Supplementary Figure S1-C, individual 3). After the grace period (so any other days), the patients cannot be censored because they already started treatment, so the weight is 1.

In the no treatment arm, patients can get censored during grace period for starting treatment (e.g., as illustrated in Supplementary Figure S1-D, individuals 2 and 4), then they receive a weight [Pr (no treatment)]. These weights are updated daily [Pr (uncensored at time t | uncensored at time t - 1, history) x Pr (uncensored at time t - 1 | uncensored at time t - 2, history) x etc.]. These inverse-probability weights allow for adjustment for the fact that the same patient does not adhere to both treatment strategies and therefore must be censored from one of them.34,36-39

***C.1.7.3. Weight creation and Model specifications:*** First, we defined the model for treatment initiation among patients yet to start treatment and we predict Pr(untreated at time t | untreated at time t - 1): *Numerator: Logit (A/1-A) = B0 + B1\*(Age) + B2\*(Race)*. *Denominator: Logit (A/1-A) = B0 + B1\*(Age) + B2\*(Race) + B3\*(Rx\_count) + B5\*(CMO\_status).* Next, we estimate the weights = 1 / Pr (uncensored at time t | baseline & time-varying baseline variables) and the stabilized weights = (numerator product of treatment weights)/(denominator product of treatment weights). Finally, we define the outcome models (logistic regression), that use stabilized weights in the cloned data and predicts death hazard (int\_surv = 1 – haz) within each day: *Logit (Death/1-Death) = B0 + B1\*(Date\_post\_adm ) + B2\*I(Date\_post\_adm\*Date\_post\_adm) + B3\*(A\*Date\_post\_adm).* Then, we obtain average risk and average survival over each treatment group for the 30 days (i.e., pooled logistic regression, surv = cumprod(int\_surv) and risk = 1 – surv). In an additional step with arguable assumptions, we approximate the hazard ratio for the first 30 days with outcome model with constant treatment effect, that also uses stabilized weights in the cloned data: Logit (D/1-D) = B0 + B1\*(Date\_post\_adm ) + B2\*I(Date\_post\_adm\*Date\_post\_adm) + B3\*(A).

### C.2. Statistical Code

---

title: "Stroke analysis – Please email [lidia.moura@mgh.harvard.edu](mailto:lidia.moura@mgh.harvard.edu) for questions."

---

```{r}

library(tidyverse)

library(survival)

library(missForest)

condition <- "Age > 0"

new\_dir <- paste0("results-", Sys.Date(), "-(", condition, ")")

dir.create(new\_dir)

wide\_data\_file <- "wide\_data\_file.csv"

long\_data\_file <- "long\_data\_file.csv"

last\_day <- 30

grace\_day <- 7

n\_boot <- 500

```

```{r}

wide\_data <- read\_csv(wide\_data\_file)

long\_data <- read\_csv(long\_data\_file)

long\_data\_longer <- long\_data %>%

right\_join(expand(long\_data, MRN, Date\_post\_adm = 0:last\_day), by = c("MRN", "Date\_post\_adm")) %>%

arrange(MRN, Date\_post\_adm) %>%

group\_by(MRN) %>%

fill(jc\_admitdate, .direction = "down") %>%

mutate(Date = jc\_admitdate + Date\_post\_adm)

dat <- left\_join(long\_data\_longer, wide\_data, by = "MRN") %>%

filter(Date <= Death\_surtime) %>%

filter(!! rlang::parse\_expr(condition)) %>%

group\_by(MRN) %>%

arrange(Date) %>%

# initA = 1 if the day they started medication

mutate(initA = as.numeric(!is.na(Medication\_Date) & Date == Medication\_Date),

# datA = # of days post baseline they started medication

dayA = ifelse(Date == Medication\_Date, Date\_post\_adm, NA),

# A = 1 if they have already initiated medication (even if not currently taking it)

A = as.numeric(cumsum(initA) > 0),

# Alag = 1 if they had already initiated medication by yesterday

Alag = lag(A, default = 0),

# event = 1 on the day of death

event = as.numeric(Death\_Status == 1 & Date == Date\_Of\_Death),

# ltfu = 1 if last day of follow-up is < 30 but not due to death

ltfu = as.numeric(row\_number() == n() &

Date\_post\_adm != last\_day &

Death\_Status == 0),

rx\_count = PrescriptionCount\_SingleDay,

CMO\_status = CMO\_time\_varying,

Race\_b = Race\_clean) %>%

# for time-varying variables, use last value carried forward

fill(rx\_count, CMO\_status, .direction = "down") %>%

# day at which medication started can be the same in every row

fill(dayA, .direction = "downup") %>%

ungroup()

# data used to standardize survival curves

baseline\_vars <- long\_data %>%

right\_join(expand(long\_data, MRN, Date\_post\_adm = 0:last\_day), by = c("MRN", "Date\_post\_adm")) %>%

left\_join(wide\_data, by = "MRN") %>%

mutate(

rx\_count = PrescriptionCount\_SingleDay,

CMO\_status = CMO\_time\_varying,

Race\_b = Race\_clean

) %>%

filter(!! rlang::parse\_expr(condition))

```

# Analysis function

```{r}

# this is the function used for the analysis

# requires a long, cloned dataset (orig\_dat) and the baseline variables

# for standarization (baseline\_vars)

analysis\_function <- function(orig\_dat, baseline\_vars, boot = FALSE, trunc\_q = 0.975, ...) {

# sample MRNs if part of the bootstrap

if (boot) {

MRNs <- unique(orig\_dat$MRN)

boot\_MRNs <- tibble(old\_MRN = sample(MRNs, replace = TRUE),

MRN = 1:length(MRNs))

dat <- right\_join(boot\_MRNs, orig\_dat, by = c("old\_MRN" = "MRN"))

} else { # or else just use the original data

dat <- orig\_dat

}

# these are the people that started treatment within the grace period

# and those that were censored at the end of the grace period because they didn't

start\_txt <- dat %>%

# censor unless it's during grace period

# or if someone started treatment w/in grace period

filter(Date\_post\_adm < grace\_day | dayA < grace\_day) %>%

# indicator for randomization arm: start treatment within grace period

mutate(txt = 1)

# these are the people that never started treatment

# or those that did, before they started (ie they are being censored for starting)

never\_txt <- dat %>%

# censor if start treatment any time during follow-up

# since once A is 1, is always 1

filter(A == 0) %>%

mutate(txt = 0)

# combine the data

clones <- bind\_rows(start\_txt, never\_txt)

# in original data, fit models among people yet to start treatment (Alag = 0)

# ie model for treatment initiation

num\_mod <- glm(A ~ Age + Race\_b + NIHSS ,

family = binomial(), data = dat,

subset = Alag == 0)

denom\_mod <- glm(A ~ Age + Race\_b + NIHSS + rx\_count + CMO\_status,

family = binomial(), data = dat,

subset = Alag == 0)

weighted\_dat <- clones %>%

# only use complete cases

filter(!if\_any(c(Age, Race\_b, SVI\_imputed, NIHSS, rx\_count), is.na)) %>%

# pr(uncensored at time t | uncensored at time t - 1)

mutate(pnum = predict(num\_mod, newdata = ., type = "response"),

pdenom = predict(denom\_mod, newdata = ., type = "response"),

pnum\_cens = predict(num\_cens, newdata = ., type = "response"),

pdenom\_cens = predict(denom\_cens, newdata = ., type = "response"),

numCont = case\_when(

# in the txt arm, weight contribution is 1 during grace period

# b/c pr(uncensored | grace period) = 1 even if you don't start

# (in the don't start txt arm, can get censored during grace period for starting txt)

txt == 1 & Date\_post\_adm < grace\_day ~ 1,

# at the end of the grace period, these people are UNCENSORED

# so must upweight them to account for those who didn't start

# this should ONLY include people who have A = 1

txt == 1 & Date\_post\_adm == grace\_day ~ pnum,

# after the grace period (so any other days),

# can't be censored because already started

txt == 1 & Date\_post\_adm > grace\_day ~ 1,

# in no txt arm, always will be p(no txt)

# because can always be censored for starting txt

txt == 0 ~ 1 - pnum

),

# same logic in the denominator

denomCont = case\_when(

txt == 1 & Date\_post\_adm < grace\_day ~ 1,

txt == 1 & Date\_post\_adm == grace\_day ~ pdenom,

txt == 1 & Date\_post\_adm > grace\_day ~ 1,

txt == 0 ~ 1 - pdenom

),

# censoring weights

# always probability of not being censored (last day doesn't count)

numCont\_cens = ifelse(Date\_post\_adm == last\_day, 1, 1 - pnum\_cens),

denomCont\_cens = ifelse(Date\_post\_adm == last\_day, 1, 1 - pdenom\_cens)

) %>%

group\_by(MRN,txt) %>%

# pr(uncensored at time t | uncensored at time t - 1, history) x

# pr(uncensored at time t - 1 | uncensored at time t - 2, history) x etc.

# this only matters for the no txt arm, who keep not taking txt

mutate(num\_prod = cumprod(numCont),

denom\_prod = cumprod(denomCont),

num\_cens\_prod = cumprod(numCont\_cens),

denom\_cens\_prod = cumprod(denomCont\_cens),

# wt = 1 / prob (uncensored at time t | bl & tv vars)

stabw = (num\_prod) / (denom\_prod))

# truncate weights - for full adjustment

tau <- quantile(weighted\_dat$stabw, trunc\_q, na.rm = TRUE)

weighted\_dat$stabw[weighted\_dat$stabw > tau] <- tau

outcome\_mod <- glm(event ~ Date\_post\_adm + I(Date\_post\_adm\*Date\_post\_adm) +

txt\*Date\_post\_adm,

data = weighted\_dat, weights = stabw, family = quasibinomial())

predictions <- bind\_rows(baseline\_vars,

baseline\_vars, .id = "txt") %>%

# turn txt from 1, 2 to 0, 1

mutate(txt = as.numeric(txt) - 1) %>%

# get predicted hazard

mutate(haz = predict(outcome\_mod, newdata = ., type = "response"),

# and survival within a certain day

int\_surv = 1 - haz) %>%

group\_by(MRN, txt) %>%

mutate(surv = cumprod(int\_surv),

risk = 1 - surv) %>%

# only group by treatment to average risk over the treatment group

ungroup() %>%

group\_by(txt, Date\_post\_adm) %>%

# remove some missing values because missing covariates

summarise(average\_risk = mean(risk, na.rm = TRUE),

average\_survival = mean(surv, na.rm = TRUE),

.groups = "drop")

# don't keep the data unless it's the main analysis

if (boot) clones <- NULL

# approximate HR with outcome model with constant treatment effect

HR\_mod <- glm(event ~ Date\_post\_adm + I(Date\_post\_adm\*Date\_post\_adm) +

txt,

data = weighted\_dat, weights = stabw, family = quasibinomial())

list(predictions = predictions, clones = clones,

tau = tau,

denom\_mod = broom::tidy(denom\_mod),

HR\_mod = broom::tidy(HR\_mod))

}

```

# Main analysis

```{r}

main\_res <- analysis\_function(dat, baseline\_vars, boot = FALSE, trunc\_q = 0.975)

# survival estimates

predicted\_surv <- main\_res$predictions

# the cloned data for checking

clones <- main\_res$clones

# the quantile of weights at which they were truncated

tau <- main\_res$tau

# the model for the weights

denom\_mod <- main\_res$denom\_mod

# the model for the outcome

HR\_mod <- main\_res$HR\_mod

```

```{r}

count(clones, txt) %>% kable()

```

```{r}

count(clones, txt, MRN) %>% count(txt) %>% kable()

```

```{r}

person\_days <- clones %>%

group\_by(MRN) %>%

# in either arm did they ever die, get lost to follow-up, or start med

mutate(ever\_event = max(event), ever\_ltfu = max(ltfu), ever\_A = max(A)) %>%

ungroup() %>%

count(MRN, txt, ever\_event, ever\_ltfu, ever\_A) %>%

pivot\_wider(names\_from = txt, values\_from = n,

names\_prefix = "txt\_", values\_fill = 0)

person\_days %>% head %>% kable()

```

```{r}

median(person\_days$txt\_0); median(person\_days$txt\_1)

```

```{r}

# run the analysis function n\_boot times

boot\_res <- map(1:n\_boot, analysis\_function,

orig\_dat = dat, baseline\_vars = baseline\_vars,

boot = TRUE, trunc\_q = 0.975

)

```

```{r}

# calculate confidence intervals

boot\_res\_t <- transpose(boot\_res)

boot\_predicted\_surv <- bind\_rows(boot\_res\_t$predictions, .id = "boot")

boot\_tau <- flatten\_dbl(boot\_res\_t$tau)

boot\_denom\_mod <- bind\_rows(boot\_res\_t$denom\_mod, .id = "boot")

boot\_HR\_mod <- bind\_rows(boot\_res\_t$HR\_mod, .id = "boot")

# for the survival curves

surv\_CIs <- boot\_predicted\_surv %>%

group\_by(txt, Date\_post\_adm) %>%

summarise(lci\_risk = quantile(average\_risk, .025),

uci\_risk = quantile(average\_risk, .975),

lci\_survival = quantile(average\_survival, .025),

uci\_survival = quantile(average\_survival, .975),

.groups = "drop")

# for 30-day risk and risk differences

dif\_CIs <- boot\_predicted\_surv %>%

filter(Date\_post\_adm == last\_day) %>%

select(-average\_survival) %>%

pivot\_wider(names\_from = txt,

values\_from = average\_risk, names\_prefix = "risk\_") %>%

mutate(risk\_dif = risk\_1 - risk\_0) %>%

summarise(lci\_risk\_dif = quantile(risk\_dif, .025),

uci\_risk\_dif = quantile(risk\_dif, .975),

lci\_risk\_1 = quantile(risk\_1, .025),

lci\_risk\_0 = quantile(risk\_0, .025),

uci\_risk\_1 = quantile(risk\_1, .975),

uci\_risk\_0 = quantile(risk\_0, .975))

all\_surv <- left\_join(predicted\_surv, surv\_CIs, by = c("txt", "Date\_post\_adm")) %>%

rename\_with(str\_remove, starts\_with("average"), "average\_")

all\_difs <- predicted\_surv %>%

filter(Date\_post\_adm == last\_day) %>%

select(-average\_survival, -Date\_post\_adm) %>%

pivot\_wider(names\_from = txt,

values\_from = average\_risk, names\_prefix = "risk\_") %>%

mutate(risk\_dif = risk\_1 - risk\_0) %>%

bind\_cols(dif\_CIs)

# for the outcome model ~ hazard ratios

all\_HRs <- boot\_HR\_mod %>%

filter(term == "txt") %>%

summarise(lci\_est = quantile(estimate, .025),

uci\_est = quantile(estimate, .975),

.groups = "drop") %>%

bind\_cols(filter(HR\_mod, term == "txt")) %>%

transmute(HR = exp(estimate),

lci\_HR = exp(lci\_est),

uci\_HR = exp(uci\_est))

```

# D. SUPPLEMENTARY REFERENCES

1. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med 2015;12:e1001885.PMC4595218

2. Hall RE, Porter J, Quan H, Reeves MJ. Developing an adapted Charlson comorbidity index for ischemic stroke outcome studies. BMC Health Serv Res 2019;19:930.PMC6892203

3. Zelano J. Poststroke epilepsy: update and future directions. Ther Adv Neurol Disord 2016;9:424-35.PMC4994782

4. Bryndziar T, Sedova P, Kramer NM, et al. Seizures Following Ischemic Stroke: Frequency of Occurrence and Impact on Outcome in a Long-Term Population-Based Study. J Stroke Cerebrovasc Dis 2016;25:150-6.PMC4957641

5. Hung LC, Sung SF, Hsieh CY, et al. Validation of a novel claims-based stroke severity index in patients with intracerebral hemorrhage. J Epidemiol 2017;27:24-9.PMC5328736

6. Kogan E, Twyman K, Heap J, Milentijevic D, Lin JH, Alberts M. Assessing stroke severity using electronic health record data: a machine learning approach. BMC Med Inform Decis Mak 2020;20:8.PMC6950922

7. Moore KL, Patel K, Boscardin WJ, Steinman MA, Ritchie C, Schwartz JB. Medication burden attributable to chronic co-morbid conditions in the very old and vulnerable. PLoS One 2018;13:e0196109.PMC5912775