No effect of remoteness on clinical outcomes following myocardial infarction: an analysis of 43,927 myocardial infarctions in Victoria, Australia

Protocol
June 9, 2023
https://github.com/cardiopharmnerd/outcomeremote

Adam Livori
PhD Student
adam.livori1@monash.edu

Center for Medicine Use and Safety, Faculty of Pharmacy and Pharmaceutical Sciences, Monash University, Melbourne, Australia

Contents

1 Preface

This is the protocol for the paper No effect of remoteness on clinical outcomes following myocardial infarction: an analysis of 43,927 myocardial infarctions in Victoria, Australia.
This protocol details the data analysis methods that were undertaken from a linked dataset provided by the Victorian Department of Health as the source of Victorian Admitted Episodes Dataset data for this study, and the Centre for Victorian Data Linkage (Victorian Department of Health) for the provision of data linkage to NDI and MBS. This study was approved by the Human Research and Ethics Committees from the Australian Institute for Health and Welfare (AIHW) (EO2018/4/468) and Monash University (14339).
The original cohort formation and data cleaning steps are noted in a separate protocol from a previous study To generate this document, the Stata package texdoc was used, which is available from: `http://repec.sowi.unibe.ch/stata/texdoc/` (accessed 14 November 2022). The final Stata do file and this pdf are available here . The do file was originally coded and then exported from the Secure Unified Research Environment (SURE). Therefore, when reproducing the code, use the do file rather than copying from the LaTex document.

2 Abbreviations

- A: Accessible (ARIA cateogry with values from 1.85 to 3.50)
- ABS: Australian Bureau of Statistics
- AF: Atrial fibrillation
- AIHW: Australian Institute of Health and Welfare
- ARIA: Accessability/remoteness index of Australia
- CABG: Coronary Artery Bypass Graft
- DM: Diabetes melitus
- GLM: Generalised linear model
- HA: Highly accessible (ARIA cateogry with values from 0.00 to 1.84)
- HF: Heart failure
- HT: Hypertension
- ICD: International classification of disease (ICD-10 Australian Modified codes were present in this dataset)
- IRSD: Index of relative socio-economic disadvantage
- IS: Ischaemic stroke
- MA: Moderately acceessible (ARIA cateogry with values from 3.51 to 5.80)
- MACE: Major adverse cardiovascular event

- MBS: Medicare benefits scheme
- MI: Myocardial infarction
- NDI: National Death Index
- NSTEMI: Non-ST elevation myocardial infarction
- PCI: Perceutaneous Coronary Intervention
- SLA: Statistical local area
- STEMI: ST elevation myocardial infarction
- VAED: Victorian admitted episode dataset
- WHO: World Health Organisation

3 Introduction

MI is the leading cause of death and clinical burden worldwide [1]. The outcomes following MI can differ when comparing people living in metropolitan and non-metrolpolitan areas. Remoteness was found to be a driver for increase death in with a 2009-2012 cohort of people from Victoria [2]. This study analysed a cohort of 43,927 MI admissions in Victoria, Australia between 2012-2018 and included multiple adjustments for socioeconomic status, age, sex, cardiovascular comorbidity, and revascularisation strategy.

The following protocol lists the steps that were taken to analyse this cohort using a linked dataset from the VAED, MBS and NDI. Outcomes of interest were 1-year 4-point MACE (defined as admission for MI, stroke; heart failure; or CV death), all-cause mortality, and each individual components of MACE.

4 Identifying outcomes in cohort

4.1 Recurrent MI

We will use an analysis dataset created from the protocol of a previous study to analyse MI admissions within Victoria, Australia. We used ICD codes present in the dataset to define each outcome

```
set rmsg on
use MI_cohort_ndi, clear
br
keep ppn admdate sepdate
bysort ppn (admdate) : gen nonfatal_MI = admdate[_n+1]
format nonfatal_MI %td
keep if nonfatal_MI !=.
gen nonfatal_MI_tag = 1
save nonfatal_MI, replace
```

4.2 Stroke

```
use MI_cohort_raw, clear
br
keep ppn admdate sepdate tdiag1
gen stroke_tag = 0
replace stroke_tag = 1 if inrange(tdiag1, "I60","I698")
ta stroke_tag
gen stroke = admdate if stroke_tag==1
format stroke %td
drop tdiag1
keep if stroke_tag == 1
hist stroke
save stroke, replace
```

```
quietly {
forval i = 1/16 {
use MI_cohort_ndi, clear
bysort ppn (admdate) : keep if _n == `i´
merge 1:m ppn using stroke
keep if _merge==3
keep ppn admdate sepdate stroke stroke_tag
drop if stroke <= sepdate
save stroke_merge`i´, replace
}
}
clear
forval i = 1/16 {
append using stroke_merge`i´
}
br
count if stroke < sepdate
bysort ppn admdate (stroke) : keep if _n==1
save stroke_clean, replace
```

4.3 HF hospitalisation

```
use MI_cohort_raw, clear
br
keep ppn admdate sepdate tdiag1
gen HF_adm_tag = 0
replace HF_adm_tag = 1 if inrange(tdiag1, "I50","I509")
ta HF_adm_tag
gen HF_adm = admdate if HF_adm_tag ==1
```

```
format HF_adm %td
drop tdiag1
keep if HF_adm_tag == 1
hist HF_adm
save HF_adm, replace

quietly {
forval i = 1/16 {
use MI_cohort_ndi, clear
bysort ppn (admdate) : keep if _n == `i´
merge 1:m ppn using HF_adm
keep if _merge==3
keep ppn admdate sepdate HF_adm_tag HF_adm
drop if HF_adm <= sepdate
save HF_adm_merge`i´, replace
}
}
clear
forval i = 1/16 {
append using HF_adm_merge`i´
}
br
count if HF_adm < sepdate
bysort ppn admdate (HF_adm) : keep if _n==1
save HF_adm_clean, replace
```

4.4 Cardiovascular death

This outcome was taken using date of death from the merged NDI dataset and the primary underlying cause of death being of a cardiovascular ICD codes I10 to I99.

```
use MI_NDI_cause, clear
br
keep ppn admdate sepdate underlying_cause_of_death othercauses deathdate
ta underlying_cause_of_death
gen CV_death_tag = 0
replace CV_death_tag = 1 if inrange(underlying_cause_of_death,"I10","I99")
gen CV_death = deathdate if CV_death_tag == 1
format CV_death %td
keep ppn admdate sepdate CV_death CV_death_tag
keep if CV_death_tag == 1
save CV_death, replace
```

4.5 Merge outcome data

The outcomes data was merged into an analysis dataset with MI admission, remoteness, comorbidity, revascularisation method, and socioeconomic status.

```
use MI_cohort_ndi, clear
br
keep if inrange(sepdate,td(1/7/2012),td(30/6/2017))
gen death_time = deathdate-sepdate
gen hosp_mort = 0
replace hosp_mort = 1 if death_time < 1
replace hosp_mort = 2 if death_time >= 1 & death_time <=90
ta hosp_mort
drop death_time
drop sepmode
save MI_ADS0, replace

use MI_sla_match_VIC_noblanks, clear
keep ppn admdate ARIA Mean
br
rename Mean mean_ARIA
save MI_ADS1, replace
```

```
use MI_cohort_SES, clear
keep ppn admdate IRSD_score quint postcode
save MI_ADS2, replace

use MI_comorbid, clear
bysort ppn admdate : keep if _n==1
save MI_ADS3, replace

use MI_proc_clean, clear
drop sepdate
save MI_ADS4, replace

use CV_death, clear
drop CV_death_tag
save MI_ADS5, replace

use nonfatal_MI, clear
drop nonfatal_MI_tag
save MI_ADS6, replace

use stroke_clean, clear
drop stroke_tag
save MI_ADS7, replace

use HF_adm_clean, clear
drop HF_adm_tag
save MI_ADS8, replace

use MI_ADS0, clear
br
forval i = 1/8 {
merge 1:1 ppn admdate using MI_ADS`i´
drop if _merge == 2
drop _merge
}
br
replace agegroup = substr(agegroup,1,2)
destring agegroup, replace force
ta agegroup
gen dead90 = 0
replace dead90 = 1 if deathdate <= sepdate + 90
gen dead365 = 0
replace dead365 = 1 if deathdate <= sepdate + 365
*Remove MIs with no ARIA
count if mean_ARIA == .
drop if mean_ARIA == .

save MI_ADS_ALL, replace
```

4.6 Overall MACE

Following the creation of inidivudal dates for events within MACE, we then created a field for MACE that returned the earliest date of an event within the MACE composite outcome.

```
use MI_ADS_ALL, clear
br
drop if hosp_mort == 1
ta ARIA
rename ARIA ARIA1
label define ARIA1 1 "HA" 2 "A" 3 "MA"
encode ARIA1, gen(ARIA) label(ARIA1)
ta ARIA
ta ARIA, nolabel
ta ARIA dead90
ta ARIA dead365
gen id = _n
gen MACEdate = min(CV_death,nonfatal_MI,stroke,HF_adm,deathdate,sep+365)
format MACEdate %td
gen MACE = 0
replace MACE = 1 if MACEdate == CV_death | MACEdate == nonfatal_MI ///
 | MACEdate == stroke | MACEdate == HF_adm
foreach i in CV_death nonfatal_MI stroke HF_adm {
```

```
gen `i´count = 0
replace `i´count = 1 if `i´ !=. & `i´ <= sepdate+365
}
save MI_ADS_ALL_MACE, replace
```

5 Analysis

5.1 Results

5.1.1 Cohort demographics

The table provides information on cohort demographics by remoteness category.

```
use MI_ADS_ALL_MACE, clear
*create age groups of 10
forval i = 3/8 {
replace agegroup = `i´0 if agegroup == `i´5
}
ta agegroup
br
gen tahelp = 1

ta tahelp ARIA, matcell(B1)
ta STEMI ARIA if STEMI == 1, matcell(B23)
ta STEMI ARIA if STEMI == 0, matcell(B24)
ta sex ARIA, matcell(B2)
ta agegroup ARIA, matcell(B3)
ta HT ARIA if HT==1, matcell(B5)
ta AF ARIA if AF==1, matcell (B6)
ta DM ARIA if DM==1, matcell(B7)
ta HF ARIA if HF==1, matcell (B8)
ta PCI_tag ARIA if PCI_tag==1, matcell (B10)
ta CABG_tag ARIA if CABG_tag==1, matcell (B11)
ta dead365 ARIA if dead365 == 1, matcell (B13)
ta MACE ARIA if MACE ==1, matcell (B14)
ta CV_deathcount ARIA if CV_deathcount == 1, matcell (B15)
ta nonfatal_MIcount ARIA if nonfatal_MIcount == 1, matcell (B16)
ta strokecount ARIA if strokecount == 1, matcell (B17)
ta HF_admcount ARIA if HF_admcount ==1, matcell (B18)

matrix analpoparia = (B1\B23\B24\B2\B3\B5\B6\B7\B8\B10\B11\B13\B14\B15\B16\B17\B18)
matrix list analpoparia
clear
svmat analpoparia
br

rename (analpoparia1 analpoparia2 analpoparia3) (HA A MA)
order HA A MA
egen total = max(HA+A+MA)
gen totalproportion = string((100 * (HA+A+MA) / total), "%3.0f")+"%"
replace total = HA+A+MA

foreach i in HA A MA {
egen `i´total = max(`i´)
gen `i´proportion1 = (100 * `i´ / `i´total)
replace `i´proportion1 = `i´total / total * 100 if `i´proportion1 == 100
gen `i´proportion = string(`i´proportion1, "%3.0f")+"%"
drop `i´proportion1
}

gen id = _n
gen demoname = ""
replace demoname = "STEMI"  if _n == 2
replace demoname = "NSTEMI" if _n == 3
replace demoname = "Male" if _n == 4
replace demoname = "Female" if _n== 5
replace demoname = "Aged 30-39" if _n== 6
```

```
replace demoname = "Aged 40-49" if _n== 7
replace demoname = "Aged 50-59" if _n== 8
replace demoname = "Aged 60-69" if _n== 9
replace demoname = "Aged 70-79" if _n==10
replace demoname = "Aged 80+" if _n== 11
replace demoname = "Hypertension" if _n == 12
replace demoname = "Atrial fibrillation" if _n== 13
replace demoname = "Diabetes mellitus" if _n== 14
replace demoname = "Heart failure" if _n== 15
replace demoname = "PCI within 90 days of MI" if _n== 16
replace demoname = "CABG within 90 days of MI" if _n== 17
replace demoname = "All-cause mortality" if _n == 18
replace demoname = "Major adverse cardiovascular event" if _n == 19
replace demoname = "Cardiovascular death" if _n == 20
replace demoname = "Myocardial infarction" if _n == 21
replace demoname = "Stroke" if _n == 22
replace demoname = "Heart failure admission" if _n == 23

drop HAtotal Atotal MAtotal id
order demoname total totalproportion HA HAproportion A Aproportion MA MAproportion
foreach i in total HA A MA {
gen `i´s = string(`i´)
drop `i´
rename `i´s `i´
}
foreach i in total HA A MA {
gen `i´s = `i´ + " " + "(" + `i´proportion + ")"
drop `i´ `i´proportion
rename `i´s `i´
}
save survdemo, replace

*Create IRSD table data
{
use MI_ADS_ALL_MACE, clear
drop if hosp_mort != 0
su(IRSD_score), detail
matrix B = (r(p50)\r(p25)\r(p75))
su(IRSD_score) if ARIA == 1, detail
matrix B = (B,(r(p50)\r(p25)\r(p75)))
su(IRSD_score) if ARIA == 2, detail
matrix B = (B,(r(p50)\r(p25)\r(p75)))
su(IRSD_score) if ARIA == 3, detail
matrix B = (B,(r(p50)\r(p25)\r(p75)))
mat list B
clear
svmat B
foreach i in B1 B2 B3 B4 {
gen `i´s = string(`i´)
drop `i´
gen `i´ = `i´ + " (" + `i´[_n+1] + "-" + `i´[_n+2] + ")"
drop `i´s
}
drop if _n != 1
rename (B1 B2 B3 B4 ) (total HA A MA)
gen demoname = "Median IRSD score (IQR)"
order demoname
save irsdsurv, replace
}
*append IRSD table data
{
use survdemo, clear
br
append using irsdsurv
gen id = _n
replace id = 17.5 if id == 24
sort id
drop id

gen category = ""
```

```
order category demoname total HA A MA
replace category = "Myocardial infarction (MI) diagnosis" if _n == 2
replace category = "Baseline characteristics" if _n == 4
replace category = "Co-morbidities" if _n == 12
replace category = "Revascularisation strategy" if _n == 16
replace category = "Socio-economic status" if _n == 18
replace category = "Clinical outcome" if _n == 19
save analpopariastemi_outcomes, replace
}
```

5.2 Parametric survival analysis

A generalized linear model (GLM) with a log-link function and Poisson distribution was constructed for each outcome separately, using log of person-time as the offset and stratified by STEMI or NSTEMI. These models included spline effects of ARIA, age, time since MI, a binary effect of sex, cardiovascular comorbidities, and revascularization strategy, and IRSD as a continuous variable. Two models were constructed; one with predictions from the models made at mean values of the co-variates with respect to stratified NSTEMI and STEMI sub-cohorts; and one with means of co-variates from the total analysed cohort.
This analysis was broken up into several steps:

1. Create prediction set for NSTEMI and STEMI stratified cohorts
2. Create prediction set for total cohort
3. Create fail dates across different ouctomes
4. Perform survival analysis and predicted IRR (means for stratified cohorts)
5. Perform survival analysis and predicted IRR (means for total cohort)
6. Plot predicted IRR and confidence intervals with respect to remoteness

5.2.1 Create prediction set for NSTEMI and STEMI stratified cohorts

Using the analysis dataset, we created two separate datasets for NSTEMI and STEMI that lists the range of possible ARIA values along with the means values of the following co-variates:

NSTEMI

- Mean age 71.46
- Sex 1.36 [male = 1 female = 2]
- IRSD score 997.83
- Presence of hypertension 0.65 (binary)
- Presence of AF 0.16 (binary)
- Presence of heart failure 0.18 (binary
- Presence of diabetes 0.32 (binary)

- Pesence of ischaemic stroke 0.01 (binary)
- Received PCI within 90 days of admission 0.32 (binary)
- Received CABG within 90 days of admission 0.11 (binary)

STEMI

- Mean age 64.68
- Sex 1.26 [male = 1 female = 2]
- IRSD score 999.73
- Presence of hypertension 0.54 (binary)
- Presence of AF 0.12 (binary)
- Presence of heart failure 0.15 (binary
- Presence of diabetes 0.23 (binary)
- Pesence of ischaemic stroke 0.01 (binary)
- Received PCI within 90 days of admission 0.73 (binary)
- Received CABG within 90 days of admission 0.08 (binary)

```
use MI_ADS_ALL_MACE, clear
gen agespline = agegroup + 2.5 if agegroup != 85
replace agespline = agegroup + 5 if agegroup == 85
save MI_ADS_ALL_MACE, clear
forval ii = 0/1 {
use MI_ADS_ALL_MACE, clear
foreach i in agespline sex IRSD_score HT AF HF DM IS CABG_tag PCI_tag {
su(`i´) if STEMI == `ii´
local m`i´ = r(mean)
}
clear
set obs 49
gen mean_ARIA = (_n-1)/10
br
foreach i in agespline sex IRSD_score HT AF HF DM IS CABG_tag PCI_tag {
gen `i´ = `m`i´´
}
gen time =  5
mkspline times=time, cubic knots(0 2 5 10)
mkspline ARIAS= mean_ARIA, cubic knots(0 0.05 0.5 2)
mkspline ages = agespline, cubic knots(32.5 62.5 72.5 77.5)
gen py = 1
save outcomeset_`ii´, replace
}
}
```

5.2.2 Create prediction set for total cohort

We then created a similar dataset, but using the means of the total analysed cohort from the following co-variates:

- Mean age 69.74
- Sex 1.33 [male = 1 female = 2]
- IRSD score 998.31
- Presence of hypertension 0.62 (binary)
- Presence of AF 0.14 (binary)
- Presence of heart failure 0.17 (binary
- Presence of diabetes 0.30 (binary)
- Pesence of ischaemic stroke 0.01 (binary)
- Received PCI within 90 days of admission 0.43 (binary)
- Received CABG within 90 days of admission 0.10 (binary)

```
use MI_ADS_ALL_MACE, clear
foreach i in agespline sex IRSD_score HT AF HF DM IS CABG_tag PCI_tag {
su(`i´)
local m`i´ = r(mean)
}
clear
set obs 49
gen mean_ARIA = (_n-1)/10
br
foreach i in agespline sex IRSD_score HT AF HF DM IS CABG_tag PCI_tag {
gen `i´ = `m`i´´
}
gen time =  5
mkspline times=time, cubic knots(0 2 5 10)
mkspline ARIAS= mean_ARIA, cubic knots(0 0.05 0.5 2)
mkspline ages = agespline, cubic knots(32.5 62.5 72.5 77.5)
gen py = 1
save outcomeset, replace
```

5.2.3 Create fail dates across different ouctomes

Using the outcome information, we created faildate date variables and fail tags for each outcome to be used in building survival models

```
use MI_ADS_ALL_MACE, clear
gen faildate_365_death = min(deathdate,sepdate+365)
gen fail_365_death = dead365
gen faildate_365_mace = min(CV_death,nonfatal_MI,stroke,HF_adm,deathdate,sep+365)
gen fail_365_mace = 0
replace fail_365_mace = 1 if faildate_365_mace == CV_death | faildate_365_mace == nonfatal_MI ///
 | faildate_365_mace == stroke | faildate_365_mace == HF_adm
gen faildate_365_cvd = min(CV_death, deathdate, sep+365)
gen fail_365_cvd = 0
replace fail_365_cvd = 1 if faildate_365_cvd == CV_death
```

```
gen faildate_365_mi = min(nonfatal_MI, deathdate, sep+365)
gen fail_365_mi = 0
replace fail_365_mi = 1 if faildate_365_mi == nonfatal_MI
gen faildate_365_stroke = min(stroke, deathdate, sep+365)
gen fail_365_stroke = 0
replace fail_365_stroke = 1 if faildate_365_stroke == stroke
gen faildate_365_hf = min(HF_adm, deathdate, sep+365)
gen fail_365_hf = 0
replace fail_365_hf = 1 if faildate_365_hf == HF_adm
save MI_ADS_ALL_MACE_FAIL, replace
```

5.2.4 Perform survival analysis and predicted IRR (means for stratified cohorts)

We ran analysis across each outcome, firslty creating spline for remoteness, time to event, and age, as well as creating an offset for person-years.
We initially built a model with an interaction of ARIA and time to event, however when reviewing the Akaike Information Criterion the model was a better fit without the interaction. (see deactivated code noted by * before the line)
We constructed a glm using poisson distrbution and a loglink function, noting the variance m̄ean, demonstrating equipdisperion and meeting the assumption for Poisson distrubution.
Finally we used the model for each outcome predict incident rate ratios as the means of covariates previously specified.

```
foreach i in death mace cvd mi stroke hf {
forval ii = 0/1 {
use MI_ADS_ALL_MACE_FAIL, clear
stset faildate_365_`i´, fail(fail_365_`i´) entry(sepdate) origin(sepdate) scale(30.417) id(id)
*create splines for remoteness
mkspline ARIAS = mean_ARIA, cubic knots(0 0.05 0.5 2)
*split times to allow for interaction variable of time to be created
stsplit time, at(0(1)12)
mkspline times = time, cubic knots(0 1 3 6)
*create splines for ages
mkspline ages = agegroup, cubic knots(32.5 62.5 72.5 77.5)
*offset creation
gen py = (_t - _t0)/12

*build model with time as an interaction
*poisson _d c.ARIAS*##c.times* c.ages* i.sex IRSD_score HT AF DM CA CPD IS CABG_tag PCI_tag, exposur
> e(py) irr
*estat ic
*This model was not as efficient when time and ARIA interact, therefore proportional hazards over ti
> me can be assumed.
*Run analysis without the use of a interaction term for time and ARIA.
poisson _d c.ARIAS* c.times* c.ages* i.sex c.IRSD_score i.HF i.HT i.AF i.DM i.IS i.CABG_tag i.PCI_ta
> g if STEMI == `ii´, exposure(py) irr
estat ic

*Calculate prediction data for figure for NSTEMI and STEMI means
use outcomeset_`ii´, clear
*create prediction variable from model
predict A, ir
*create SE to build confidence intervals
predict B, stdp
replace A = A * 1000
gen ll = (exp(ln(A)-1.96*B))
gen ul = (exp(ln(A)+1.96*B))
gen STEMI = `ii´
keep mean_ARIA A ul ll STEMI

save `i´_`ii´, replace
}
}
```

5.2.5 Perform survival analysis and predicted IRR (means for total cohort)

The same process as above was carried out, but using the prediction dataset of the means of co-variates for the total analysed cohort; noting that we still stratified analysis by diagnosis.

```
foreach i in death mace cvd mi stroke hf {
forval ii = 0/1 {
use MI_ADS_ALL_MACE_FAIL, clear
stset faildate_365_`i´, fail(fail_365_`i´) entry(sepdate) origin(sepdate) scale(30.417) id(id)
*create splines for remoteness
mkspline ARIAS = mean_ARIA, cubic knots(0 0.05 0.5 2)
*split times to allow for interaction variable of time to be created
stsplit time, at(0(1)12)
mkspline times = time, cubic knots(0 1 3 6)
*create splines for ages
mkspline ages = agegroup, cubic knots(32.5 62.5 72.5 77.5)
*offset creation
gen py = (_t - _t0)/12

*build model with time as an interaction
*poisson _d c.ARIAS*##c.times* c.ages* i.sex IRSD_score HT AF DM CA CPD IS CABG_tag PCI_tag, exposur
> e(py) irr
*estat ic
*This model was not as efficient when time and ARIA interact, therefore proportional hazards over ti
> me can be assumed.
*Run analysis without the use of a interaction term for time and ARIA.
poisson _d c.ARIAS* c.times* c.ages* i.sex c.IRSD_score i.HF i.HT i.AF i.DM i.IS i.CABG_tag i.PCI_ta
> g if STEMI == `ii´, exposure(py) irr
estat ic

*Calculate prediction data for figure for NSTEMI and STEMI means
use outcomeset, clear
*create prediction variable from model
predict A, ir
*create SE to build confidence intervals
predict B, stdp
replace A = A * 1000
gen ll = (exp(ln(A)-1.96*B))
gen ul = (exp(ln(A)+1.96*B))
gen STEMI = `ii´
keep mean_ARIA A ul ll STEMI

save `i´_`ii´_total, replace
}
```

5.2.6 Plot predicted IRR and confidence intervals with respect to remoteness

Using the predcition datasets created for both models above across each of the six outcomes, we plotted predicited incident rates with 95% confidence intervals with respect to ARIA.

NSTEMI and STEMI models

```
*NSTEMI
foreach i in death mace cvd mi stroke hf {
use `i´_0, clear

twoway ///
(rarea ul ll mean_ARIA if STEMI == 0, col(navy%30) fintensity(inten80) lwidth(none)) ///
(line A mean_ARIA if STEMI == 0, col(navy)) ///
, graphregion(color(white)) ///
xtitle("mean ARIA score", size(large)) ytitle("Predicted incidence rate*", size(large)) ///
legend(order(2 "NSTEMI") position(6) ring(0) row(1) col(2) region(lcolor(white) color(none))) ///
yscale(log range(2 80)) ylabel(2 "2" 5 "5" 10 "10" 20 "20" 50 "50" 100 "100", angle(0) format(%9.0f)
> ) xscale(range(0 5)) ///
title("`i´", placement(west) color(black) size(large))
graph save "Graph" fig_log`i´_0, replace
}

*STEMI
```

```
foreach i in death mace cvd mi stroke hf {
use `i´_1, clear
twoway ///
(rarea ul ll mean_ARIA if STEMI == 1, col(dkorange%30) fintensity(inten80) lwidth(none)) ///
(line A mean_ARIA if STEMI == 1 , col(dkorange)) ///
, graphregion(color(white)) ///
xtitle("mean ARIA score", size(large)) ytitle("Predicted incidence rate*", size(large)) ///
legend(order(2 "STEMI") position(6) ring(0) row(1) col(2) region(lcolor(white) color(none))) ///
yscale(log range(2 80)) ylabel(2 "2" 5 "5" 10 "10" 20 "20" 50 "50" 100 "100", angle(0) format(%9.0f)
> ) xscale(range(0 5)) ///
title("`i´", placement(west) color(black) size(large))
graph save "Graph" fig_log`i´_1, replace
}
*Combine all graphs
{
graph combine ///
fig_logdeath_1.gph ///
fig_logdeath_0.gph ///
fig_logmace_1.gph ///
fig_logmace_0.gph ///
fig_logcvd_1.gph ///
fig_logcvd_0.gph ///
fig_logmi_1.gph ///
fig_logmi_0.gph ///
fig_logstroke_1.gph ///
fig_logstroke_0.gph ///
fig_loghf_1.gph ///
fig_loghf_0.gph ///
, altshrink cols(2) graphregion(color(white)) xsize(4.5) ysize(5)
graph export "G:\Adam\Project 2  - location and MI outcomes\Results\Fig_logoutcomes_STEMI_NSTEMI.pdf
> ", as(pdf) name("Graph") replace
}
}
```

Total cohort model

```
foreach i in death mace cvd mi stroke hf {
use `i´_0_total, clear
append using `i´_1_total
twoway ///
(rarea ul ll mean_ARIA if STEMI == 1, col(dkorange%30) fintensity(inten80) lwidth(none)) ///
(line A mean_ARIA if STEMI == 1 , col(dkorange)) ///
(rarea ul ll mean_ARIA if STEMI == 0, col(navy%30) fintensity(inten80) lwidth(none)) ///
(line A mean_ARIA if STEMI == 0, col(navy)) ///
, graphregion(color(white)) ///
xtitle("mean ARIA score") ytitle("Predicted incidence rate per 1000 person-years") ///
legend(order(4 "NSTEMI" 2 "STEMI") position(6) ring(0) row(1) col(2) region(lcolor(white) color(none
> ))) ///
yscale(log range(2 80)) ylabel(2 "2" 5 "5" 10 "10" 20 "20" 50 "50" 100 "100", angle(0) format(%9.0f)
> ) xscale(range(0 5)) ///
title("`i´", placement(west) color(black) size(medium))
graph save "Graph" fig_log`i´, replace
}
*Combine all graphs
{
graph combine ///
fig_logdeath.gph ///
fig_logmace.gph ///
fig_logcvd.gph ///
fig_logmi.gph ///
fig_logstroke.gph ///
fig_loghf.gph ///
, altshrink cols(2) graphregion(color(white)) xsize(4.5)
graph export "G:\Adam\Project 2  - location and MI outcomes\Results\Fig_logoutcomes_total.pdf", as(p
> df) name("Graph") replace
}
```

5.3 Sensitivity analysis

To account for possible intra-sample clustering due to people have repeat admissions for MI throughout the dataset (8151 admissions), a sensitivty analysis was conducted, replicating the parametric survival analysis but only using the first instance of MI in the cohort.

```
use MI_ADS_ALL_MACE, clear
br
keep if priorMI == 0
bysort ppn : ppncheck = _n
ta ppncheck
drop ppncheck
save MI_ADS_FIRSTMIONLY_MACE, replace
```

5.3.1 Create prediction set for NSTEMI and STEMI stratified cohorts

Using the analysis dataset, we created two separate datasets for NSTEMI and STEMI that lists the range of possible ARIA values along with the means values of the following co-variates:

NSTEMI

- Mean age 70.81
- Sex 1.36 [male = 1 female = 2]
- IRSD score 999.59
- Presence of hypertension 0.63 (binary)
- Presence of AF 0.15 (binary)
- Presence of heart failure 0.17 (binary
- Presence of diabetes 0.29 (binary)
- Pesence of ischaemic stroke 0.01 (binary)
- Received PCI within 90 days of admission 0.34 (binary)
- Received CABG within 90 days of admission 0.12 (binary)

STEMI

- Mean age 64.53
- Sex 1.26 [male = 1 female = 2]
- IRSD score 999.95
- Presence of hypertension 0.53 (binary)
- Presence of AF 0.12 (binary)
- Presence of heart failure 0.15 (binary
- Presence of diabetes 0.22 (binary)

- Pesence of ischaemic stroke 0.01 (binary)
- Received PCI within 90 days of admission 0.75 (binary)
- Received CABG within 90 days of admission 0.08 (binary)

```
use MIS_ADS_FIRSTMIONLY_MACE, clear
gen agespline = agegroup + 2.5 if agegroup != 85
replace agespline = agegroup + 5 if agegroup == 85
save MI_ADS_FIRSTMIONLY_MACE, clear
forval ii = 0/1 {
use MI_ADS_ALL_MACE, clear
foreach i in agespline sex IRSD_score HT AF HF DM IS CABG_tag PCI_tag {
su(`i´) if STEMI == `ii´
local m`i´ = r(mean)
}
clear
set obs 49
gen mean_ARIA = (_n-1)/10
br
foreach i in agespline sex IRSD_score HT AF HF DM IS CABG_tag PCI_tag {
gen `i´ = `m`i´´
}
gen time =  5
mkspline times=time, cubic knots(0 2 5 10)
mkspline ARIAS= mean_ARIA, cubic knots(0 0.05 0.5 2)
mkspline ages = agespline, cubic knots(32.5 62.5 72.5 77.5)
gen py = 1
save FIRSTMIoutcomeset_`ii´, replace
}
```

5.3.2 Create prediction set for total cohort

We then created a similar dataset, but using the means of the total analysed cohort from the following co-variates:

- Mean age 69.03
- Sex 1.33 [male = 1 female = 2]
- IRSD score 999.69
- Presence of hypertension 0.60 (binary)
- Presence of AF 0.14 (binary)
- Presence of heart failure 0.16 (binary
- Presence of diabetes 0.27 (binary)
- Pesence of ischaemic stroke 0.01 (binary)
- Received PCI within 90 days of admission 0.46 (binary)
- Received CABG within 90 days of admission 0.11 (binary)

```
use MI_ADS_ALL_MACE, clear
foreach i in agespline sex IRSD_score HT AF HF DM IS CABG_tag PCI_tag {
su(`i´)
local m`i´ = r(mean)
}
clear
set obs 49
gen mean_ARIA = (_n-1)/10
br
foreach i in agespline sex IRSD_score HT AF HF DM IS CABG_tag PCI_tag {
gen `i´ = `m`i´´
}
gen time =  5
mkspline times=time, cubic knots(0 2 5 10)
mkspline ARIAS= mean_ARIA, cubic knots(0 0.05 0.5 2)
mkspline ages = agespline, cubic knots(32.5 62.5 72.5 77.5)
gen py = 1
save FIRSTMIoutcomeset, replace
```

5.3.3 Create fail dates across different ouctomes

Using the outcome information, we created faildate date variables and fail tags for each outcome to be used in building survival models

```
use MI_ADS_ALL_MACE, clear
gen faildate_365_death = min(deathdate,sepdate+365)
gen fail_365_death = dead365
gen faildate_365_mace = min(CV_death,nonfatal_MI,stroke,HF_adm,deathdate,sep+365)
gen fail_365_mace = 0
replace fail_365_mace = 1 if faildate_365_mace == CV_death | faildate_365_mace == nonfatal_MI ///
 | faildate_365_mace == stroke | faildate_365_mace == HF_adm
gen faildate_365_cvd = min(CV_death, deathdate, sep+365)
gen fail_365_cvd = 0
replace fail_365_cvd = 1 if faildate_365_cvd == CV_death
gen faildate_365_mi = min(nonfatal_MI, deathdate, sep+365)
gen fail_365_mi = 0
replace fail_365_mi = 1 if faildate_365_mi == nonfatal_MI
gen faildate_365_stroke = min(stroke, deathdate, sep+365)
gen fail_365_stroke = 0
replace fail_365_stroke = 1 if faildate_365_stroke == stroke
gen faildate_365_hf = min(HF_adm, deathdate, sep+365)
gen fail_365_hf = 0
replace fail_365_hf = 1 if faildate_365_hf == HF_adm
save MI_ADS_FIRSTMIONLY_MACE_FAIL, replace
```

5.3.4 Perform survival analysis and predicted IRR (means for stratified cohorts)

We ran analysis across each outcome, firslty creating spline for remoteness, time to event, and age, as well as creating an offset for person-years.
We initially built a model with an interaction of ARIA and time to event, however when reviewing the Akaike Information Criterion the model was a better fit without the interaction. (see deactivated code noted by * before the line)
We constructed a glm using poisson distrbution and a loglink function, noting the variance m̄ean, demonstrating equipdisperion and meeting the assumption for Poisson distrubution.
Finally we used the model for each outcome predict incident rate ratios as the means of covariates previously specified.

```
foreach i in death mace cvd mi stroke hf {
forval ii = 0/1 {
use MI_ADS_ALL_MACE_FAIL, clear
stset faildate_365_`i´, fail(fail_365_`i´) entry(sepdate) origin(sepdate) scale(30.417) id(id)
*create splines for remoteness
```

```
mkspline ARIAS = mean_ARIA, cubic knots(0 0.05 0.5 2)
*split times to allow for interaction variable of time to be created
stsplit time, at(0(1)12)
mkspline times = time, cubic knots(0 1 3 6)
*create splines for ages
mkspline ages = agegroup, cubic knots(32.5 62.5 72.5 77.5)
*offset creation
gen py = (_t - _t0)/12

*build model with time as an interaction
*poisson _d c.ARIAS*##c.times* c.ages* i.sex IRSD_score HT AF DM CA CPD IS CABG_tag PCI_tag, exposur
> e(py) irr
*estat ic
*This model was not as efficient when time and ARIA interact, therefore proportional hazards over ti
> me can be assumed.
*Run analysis without the use of a interaction term for time and ARIA.
poisson _d c.ARIAS* c.times* c.ages* i.sex c.IRSD_score i.HF i.HT i.AF i.DM i.IS i.CABG_tag i.PCI_ta
> g if STEMI == `ii´, exposure(py) irr
estat ic

*Calculate prediction data for figure for NSTEMI and STEMI means
use FIRSTMIoutcomeset_`ii´, clear
*create prediction variable from model
predict A, ir
*create SE to build confidence intervals
predict B, stdp
replace A = A * 1000
gen ll = (exp(ln(A)-1.96*B))
gen ul = (exp(ln(A)+1.96*B))
gen STEMI = `ii´
keep mean_ARIA A ul ll STEMI

save FIRSTMI`i´_`ii´, replace
}
}
```

5.3.5 Perform survival analysis and predicted IRR (means for total cohort)

The same process as above was carried out, but using the prediction dataset of the means of co-variates for the total analysed cohort; noting that we still stratified analysis by diagnosis.

```
foreach i in death mace cvd mi stroke hf {
forval ii = 0/1 {
use MI_ADS_FIRSTMIONLY_MACE_FAIL, clear
stset faildate_365_`i´, fail(fail_365_`i´) entry(sepdate) origin(sepdate) scale(30.417) id(id)
*create splines for remoteness
mkspline ARIAS = mean_ARIA, cubic knots(0 0.05 0.5 2)
*split times to allow for interaction variable of time to be created
stsplit time, at(0(1)12)
mkspline times = time, cubic knots(0 1 3 6)
*create splines for ages
mkspline ages = agegroup, cubic knots(32.5 62.5 72.5 77.5)
*offset creation
gen py = (_t - _t0)/12

*build model with time as an interaction
*poisson _d c.ARIAS*##c.times* c.ages* i.sex IRSD_score HT AF DM CA CPD IS CABG_tag PCI_tag, exposur
> e(py) irr
*estat ic
*This model was not as efficient when time and ARIA interact, therefore proportional hazards over ti
> me can be assumed.
*Run analysis without the use of a interaction term for time and ARIA.
poisson _d c.ARIAS* c.times* c.ages* i.sex c.IRSD_score i.HF i.HT i.AF i.DM i.IS i.CABG_tag i.PCI_ta
> g if STEMI == `ii´, exposure(py) irr
estat ic

*Calculate prediction data for figure for NSTEMI and STEMI means
use FIRSTMIoutcomeset, clear
*create prediction variable from model
```

```
predict A, ir
*create SE to build confidence intervals
predict B, stdp
replace A = A * 1000
gen ll = (exp(ln(A)-1.96*B))
gen ul = (exp(ln(A)+1.96*B))
gen STEMI = `ii´
keep mean_ARIA A ul ll STEMI
save FIRSTMI`i´_`ii´_total, replace
}
```

5.3.6 Plot predicted IRR and confidence intervals with respect to remoteness

Using the predcition datasets created for both models above across each of the six outcomes, we plotted predicited incident rates with 95% confidence intervals with respect to ARIA.

NSTEMI and STEMI models

```
*NSTEMI
foreach i in death mace cvd mi stroke hf {
use FIRSTMI`i´_0, clear
twoway ///
(rarea ul ll mean_ARIA if STEMI == 0, col(navy%30) fintensity(inten80) lwidth(none)) ///
(line A mean_ARIA if STEMI == 0, col(navy)) ///
, graphregion(color(white)) ///
xtitle("mean ARIA score", size(large)) ytitle("Predicted incidence rate*", size(large)) ///
legend(order(2 "NSTEMI") position(6) ring(0) row(1) col(2) region(lcolor(white) color(none))) ///
yscale(log range(2 80)) ylabel(2 "2" 5 "5" 10 "10" 20 "20" 50 "50" 100 "100", angle(0) format(%9.0f)
> ) xscale(range(0 5)) ///
title("`i´", placement(west) color(black) size(large))
graph save "Graph" fig_log`i´_0, replace
}
*STEMI
foreach i in death mace cvd mi stroke hf {
use FIRSTMI`i´_1, clear
twoway ///
(rarea ul ll mean_ARIA if STEMI == 1, col(dkorange%30) fintensity(inten80) lwidth(none)) ///
(line A mean_ARIA if STEMI == 1 , col(dkorange)) ///
, graphregion(color(white)) ///
xtitle("mean ARIA score", size(large)) ytitle("Predicted incidence rate*", size(large)) ///
legend(order(2 "STEMI") position(6) ring(0) row(1) col(2) region(lcolor(white) color(none))) ///
yscale(log range(2 80)) ylabel(2 "2" 5 "5" 10 "10" 20 "20" 50 "50" 100 "100", angle(0) format(%9.0f)
> ) xscale(range(0 5)) ///
title("`i´", placement(west) color(black) size(large))
graph save "Graph" fig_log`i´_1, replace
}
*Combine all graphs
{
graph combine ///
fig_logdeath_1.gph ///
fig_logdeath_0.gph ///
fig_logmace_1.gph ///
fig_logmace_0.gph ///
fig_logcvd_1.gph ///
fig_logcvd_0.gph ///
fig_logmi_1.gph ///
fig_logmi_0.gph ///
fig_logstroke_1.gph ///
fig_logstroke_0.gph ///
fig_loghf_1.gph ///
fig_loghf_0.gph ///
, altshrink cols(2) graphregion(color(white)) xsize(4.5) ysize(5)
graph export "G:\Adam\Project 2  - location and MI outcomes\Results\Fig_logoutcomes_STEMI_NSTEMI_FIR
> STMI.pdf", as(pdf) name("Graph") replace
}
```

```
}
```

Total cohort model

```
foreach i in death mace cvd mi stroke hf {
use FIRSTMI`i´_0_total, clear
append using FIRSTMI`i´_1_total
twoway ///
(rarea ul ll mean_ARIA if STEMI == 1, col(dkorange%30) fintensity(inten80) lwidth(none)) ///
(line A mean_ARIA if STEMI == 1 , col(dkorange)) ///
(rarea ul ll mean_ARIA if STEMI == 0, col(navy%30) fintensity(inten80) lwidth(none)) ///
(line A mean_ARIA if STEMI == 0, col(navy)) ///
, graphregion(color(white)) ///
xtitle("mean ARIA score") ytitle("Predicted incidence rate per 1000 person-years") ///
legend(order(4 "NSTEMI" 2 "STEMI") position(6) ring(0) row(1) col(2) region(lcolor(white) color(none
> ))) ///
yscale(log range(2 80)) ylabel(2 "2" 5 "5" 10 "10" 20 "20" 50 "50" 100 "100", angle(0) format(%9.0f)
> ) xscale(range(0 5)) ///
title("`i´", placement(west) color(black) size(medium))
graph save "Graph" fig_log`i´, replace
}
*Combine all graphs
{
graph combine ///
fig_logdeath.gph ///
fig_logmace.gph ///
fig_logcvd.gph ///
fig_logmi.gph ///
fig_logstroke.gph ///
fig_loghf.gph ///
, altshrink cols(2) graphregion(color(white)) xsize(4.5)
graph export "G:\Adam\Project 2  - location and MI outcomes\Results\Fig_logoutcomes_total_FIRSTMI.pd
> f", as(pdf) name("Graph") replace
}
```

References

[1] Salim S. Virani, Alvaro Alonso, Emelia J. Benjamin, Marcio S. Bittencourt, Clifton W. Callaway, April P. Carson, Alanna M. Chamberlain, Alexander R. Chang, Susan Cheng, Francesca N. Delling, Luc Djousse, Mitchell S.V. Elkind, Jane F. Ferguson, Myriam Fornage, Sadiya S. Khan, Brett M. Kissela, Kristen L. Knutson, Tak W. Kwan, Daniel T. Lackland, Tené T. Lewis, Judith H. Lichtman, Chris T. Longenecker, Matthew Shane Loop, Pamela L. Lutsey, Seth S. Martin, Kunihiro Matsushita, Andrew E. Moran, Michael E. Mussolino, Amanda Marma Perak, Wayne D. Rosamond, Gregory A. Roth, Uchechukwu K.A. Sampson, Gary M. Satou, Emily B. Schroeder, Svati H. Shah, Christina M. Shay, Nicole L. Spartano, Andrew Stokes, David L. Tirschwell, Lisa B. VanWagner, and Connie W. Tsao. Heart disease and stroke statistics—2020 update a report from the american heart association. *Circulation*, 141:E139–E596, 2020.

[2] Jane Jacobs, Karen Louise Peterson, Steven Allender, Laura Veronica Alston, and Melanie Nichols. Regional variation in cardiovascular mortality in australia 2009–2012: the impact of remoteness and socioeconomic status. *Australian and New Zealand Journal of Public Health*, 42:467–473, 10 2018.