```
                                                      ___  ____  ____  ____  ____(R)
                                                     /__    /   ____/   /   ____/
                                                    ___/   /   /___/   /   /___/
                                                      Statistics/Data analysis

                                                         User: Yong Yi Lee
                                                      Project: Target_D
```

```
1 . codebook
```

---

study_id — **Study identifier**

```
                 Type: Numeric (int)

                Range: [1,4371]                      Units: 1
        Unique values: 1,868                         Missing .: 0/1,868

                 Mean: 2403.31
            Std. dev.: 1206.56

          Percentiles:     10%       25%       50%       75%       90%
                           681    1389.5      2445    3445.5      3992
```

NOTE: Follow-up periods are denoted by the following suffixes appended to each variable name.

_t = Baseline
_0 = Baseline
_1 = Baseline
_2 = 3-month follow-up
_3 = 12-month follow-up

---

group — **Randomisation group**

```
                 Type: Numeric (byte)
                Label: group_lbl

                Range: [0,1]                         Units: 1
        Unique values: 2                             Missing .: 0/1,868

           Tabulation: Freq.   Numeric  Label
                         935         0  Control
                         933         1  Intervention
```

---

phqdep_total_impute_t — **Depression - PHQ-9 score (baseline)**

```
                 Type: Numeric (float)

                Range: [0,27]                        Units: 1
        Unique values: 28                            Missing .: 0/1,868

                 Mean: 9.23233
            Std. dev.: 5.71362

          Percentiles:     10%       25%       50%       75%       90%
                             3         5         8        13        17
```

---

clinic_0 — **Clinic**

```
                 Type: Numeric (byte)

                Range: [1,14]                        Units: 1
        Unique values: 14                            Missing .: 0/1,868

                 Mean: 8.93737
            Std. dev.: 3.65385

          Percentiles:     10%       25%       50%       75%       90%
                             4         5        10        12        13
```

---

risk — **Prognostic group**

```
                 Type: Numeric (byte)
                Label: risk

                Range: [1,3]                         Units: 1
        Unique values: 3                             Missing .: 0/1,868
```

Tabulation:	Freq.	Numeric	Label
	1,357	**1**	Mild
	288	**2**	Moderate
	223	**3**	Severe

---

age_num_0 **Age**

Type: Numeric (**double**)

Range: [**18,65**] Units: **1**
Unique values: **48** Missing .: **0/1,868**

Mean: **35.531**
Std. dev.: **12.0869**

Percentiles:	10%	25%	50%	75%	90%
	22	**26**	**33**	**44**	**55**

---

gender_0 **Gender**

Type: Numeric (**byte**)
Label: **gender**

Range: [**0,2**] Units: **1**
Unique values: **3** Missing .: **0/1,868**

Tabulation:	Freq.	Numeric	Label
	590	**0**	Male
	1,270	**1**	Female
	8	**2**	Other

---

costintv **Intervention cost (applies to intervention group only)**

Type: Numeric (**float**)

Range: [**0,676.96**] Units: **.01**
Unique values: **4** Missing .: **0/1,868**

Tabulation:	Freq.	Value
	935	**0**
	679	**14.96**
	143	**132.96001**
	111	**676.96002**

SA1_costintv **Intervention cost under sensitivity analysis incorporating sunk costs for Target-D CPT research & development**

Type: Numeric (**float**)

Range: [**0,678.3**] Units: **.1**
Unique values: **4** Missing .: **0/1,868**

Tabulation:	Freq.	Value
	935	**0**
	679	**16.299999**
	143	**134.3**
	111	**678.29999**

SA2_costintv **Intervention cost under sensitivity analysis with higher unit cost for moderate prognostic group**

Type: Numeric (**float**)

Range: [**0,676.96**] Units: **.01**
Unique values: **4** Missing .: **0/1,868**

Tabulation: Freq. Value
935 0
679 14.96
143 222.96001
111 676.96002

---

costtothealthprof_2 **Total cost of health professional visits (3 months)**

Type: Numeric (**float**)

Range: [**0,4021.0701**] Units: **1.000e-06**
Unique values: **711** Missing .: **809/1,868**

Mean: **340.673**
Std. dev.: **418.347**

Percentiles:	10%	25%	50%	75%	90%
	0	**42.5987**	**228.087**	**491.278**	**806.205**

---

costtotervisit_2 **Total cost of emergency department visits (3 months)**

Type: Numeric (**float**)

Range: [**0,2236.36**] Units: **.01**
Unique values: **8** Missing .: **846/1,868**

Tabulation: Freq. Value
953 0
1 4.5
2 554.59003
53 559.09003
6 659.09003
5 1118.1801
1 1188.1801
1 2236.3601
846 .

---

costtotovnstay_2 **Total cost of overnight hospital stays (3 months)**

Type: Numeric (**float**)

Range: [**0,26772.9**] Units: **.1**
Unique values: **11** Missing .: **847/1,868**

Mean: **252.687**
Std. dev.: **1899.1**

Percentiles:	10%	25%	50%	75%	90%
	0	**0**	**0**	**0**	**0**

---

costtotdiagtest_2 **Total cost of diagnostic tests (3 months)**

Type: Numeric (**float**)

Range: [**0,2028.84**] Units: **.01**
Unique values: **130** Missing .: **835/1,868**

Mean: **67.4869**
Std. dev.: **165.32**

Percentiles:	10%	25%	50%	75%	90%
	0	**0**	**0**	**49.76**	**177.23**

---

costtotmed_2 — **Total cost of medications (3 months)**

Type: Numeric (**float**)

Range: [**0,488.25**] Units: **.01**
Unique values: **118** Missing .: **844/1,868**

Mean: **17.6252**
Std. dev.: **41.426**

Percentiles:	10%	25%	50%	75%	90%
	0	**0**	**0**	**23.17**	**55.93**

---

costtothealthprof_3 — **Total cost of health professional visits (12 months)**

Type: Numeric (**float**)

Range: [**0,6171.1445**] Units: **1.000e-06**
Unique values: **602** Missing .: **1,070/1,868**

Mean: **503.983**
Std. dev.: **586.04**

Percentiles:	10%	25%	50%	75%	90%
	0	**83.852**	**301.507**	**741.049**	**1280.5**

---

costtotervisit_3 — **Total cost of emergency department visits (12 months)**

Type: Numeric (**float**)

Range: [**0,1677.27**] Units: **.01**
Unique values: **9** Missing .: **1,113/1,868**

Tabulation:

Freq.	Value
687	**0**
1	**99**
4	**554.59003**
55	**559.09003**
1	**563.59003**
2	**659.09003**
3	**1118.1801**
1	**1218.1801**
1	**1677.27**
1,113	**.**

---

costtotovnstay_3 — **Total cost of overnight hospital stays (12 months)**

Type: Numeric (**float**)

Range: [**0,24339**] Units: **.1**
Unique values: **10** Missing .: **1,113/1,868**

Mean: **538.359**
Std. dev.: **2365.24**

Percentiles:	10%	25%	50%	75%	90%
	0	**0**	**0**	**0**	**0**

costtotdiagtest_3 — **Total cost of diagnostic tests (12 months)**

Type: Numeric (**float**)

Range: [**0,2419.07**] Units: **.01**
Unique values: **145** Missing .: **1,104/1,868**

Mean: **101.61**
Std. dev.: **217.935**

Percentiles:	10%	25%	50%	75%	90%
	0	**0**	**24.88**	**82.16**	**279.91**

costtotmed_3 — **Total cost of medications (12 months)**

Type: Numeric (**float**)

Range: [**0,995.05**] Units: **.01**
Unique values: **114** Missing .: **1,107/1,868**

Mean: **53.5464**
Std. dev.: **115.896**

Percentiles:	10%	25%	50%	75%	90%
	0	**0**	**0**	**79.02**	**153.03**

costtotprodloss_2 — **Total productivity losses (3 months)**

Type: Numeric (**float**)

Range: [**0,45090.672**] Units: **.0001**
Unique values: **238** Missing .: **841/1,868**

Mean: **4697.63**
Std. dev.: **5586.42**

Percentiles:	10%	25%	50%	75%	90%
	0	**0**	**3119.77**	**7737.8**	**11561.3**

costtotprodloss_3 — **Total productivity losses (12 months)**

Type: Numeric (**float**)

Range: [**0,82854.328**] Units: **.0001**
Unique values: **253** Missing .: **1,103/1,868**

Mean: **11459**
Std. dev.: **12577.1**

Percentiles:	10%	25%	50%	75%	90%
	0	**0**	**8438.29**	**18381.3**	**26852**

AQoL8DUtility_1 — **AQoL-8D utility score (baseline)**

Type: Numeric (**float**)

Range: [**.11649436,1**] Units: **1.000e-08**
Unique values: **1,684** Missing .: **184/1,868**

Mean: **.569382**
Std. dev.: **.196468**

Percentiles:	10%	25%	50%	75%	90%
	.3115	**.422853**	**.553744**	**.714437**	**.847783**

AQoL8DUtility_2 — **AQoL-8D utility score (3 months)**

Type: Numeric (**float**)

Range: [**.11332551,1**] Units: **1.000e-08**
Unique values: **1,012** Missing .: **856/1,868**

Mean: **.592451**
Std. dev.: **.203951**

Percentiles:	10%	25%	50%	75%	90%
	.321678	**.431112**	**.58164**	**.749469**	**.878988**

AQoL8DUtility_3 — **AQoL-8D utility score (12 months)**

Type: Numeric (**float**)

Range: [**.14523593,.99968475**] Units: **1.000e-08**
Unique values: **750** Missing .: **1,118/1,868**

Mean: **.609107**
Std. dev.: **.204689**

Percentiles:	10%	25%	50%	75%	90%
	.330536	**.45506**	**.607112**	**.782252**	**.888144**

---

phqdep_total_impute_2 — **Depression - PHQ-9 score (3 months)**

Type: Numeric (**float**)

Range: [**0,27**] Units: **1**
Unique values: **28** Missing .: **606/1,868**

Mean: **8.73296**
Std. dev.: **6.2971**

Percentiles:	10%	25%	50%	75%	90%
	1	**4**	**7**	**13**	**18**

---

phqdep_total_impute_3 — **Depression - PHQ-9 score (12 months)**

Type: Numeric (**float**)

Range: [**0,27**] Units: **1**
Unique values: **28** Missing .: **703/1,868**

Mean: **8.11674**
Std. dev.: **6.03413**

Percentiles:	10%	25%	50%	75%	90%
	1	**3**	**7**	**12**	**17**

---

education_rec_0 — **Highest level of education attained**

Type: Numeric (**float**)
Label: **education_rec**

```
               Range: [1,4]                          Units: 1
      Unique values: 4                           Missing .: 0/1,868

         Tabulation: Freq.   Numeric  Label
                       221         1  <Year 10/Year 10/equiv/Year
                                      11/equiv
                       282         2  Year 12/equiv
                       441         3  Certificate/diploma
                       924         4  Bachelor degree or higher
```

---

employment_0 **Employment status**

```
                Type: Numeric (byte)
               Label: employment

               Range: [0,3]                          Units: 1
      Unique values: 3                           Missing .: 0/1,868

         Tabulation: Freq.   Numeric  Label
                     1,353         0  Employed/working
                       211         2  Unemployed and looking for work
                       304         3  Neither working nor looking for
                                      work
```

---

card_none_0 **Health care card status**

```
                Type: Numeric (float)
               Label: yesno

               Range: [0,1]                          Units: 1
      Unique values: 2                           Missing .: 0/1,868

         Tabulation: Freq.   Numeric  Label
                       473         0  No
                     1,395         1  Yes
```

---

illness_t **Long-term illness status**

```
                Type: Numeric (byte)
               Label: yesno

               Range: [0,1]                          Units: 1
      Unique values: 2                           Missing .: 0/1,868

         Tabulation: Freq.   Numeric  Label
                     1,353         0  No
                       515         1  Yes
```

---

health_0 **Self-rated health**

```
                Type: Numeric (byte)
               Label: health

               Range: [1,5]                          Units: 1
      Unique values: 5                           Missing .: 0/1,868

         Tabulation: Freq.   Numeric  Label
                       102         1  Excellent
                       580         2  Very good
                       779         3  Good
                       319         4  Fair
                        88         5  Poor
```

live_alone_0 — **Live alone?**

Type: Numeric (**float**)
Label: **yesno**

Range: [**0**,**1**] Units: **1**
Unique values: **2** Missing .: **0**/**1,868**

Tabulation:	Freq.	Numeric	Label
	1,629	**0**	No
	239	**1**	Yes

incomerec2_t — **Ability to manage on available income**

Type: Numeric (**float**)
Label: **incomerec**

Range: [**1**,**2**] Units: **1**
Unique values: **2** Missing .: **0**/**1,868**

Tabulation:	Freq.	Numeric	Label
	1,649	**1**	Easily/Not too bad
	219	**2**	Difficult some of the time

psychpast_visited_0 — **Visited psychologist or counsellor in 12 months**

Type: Numeric (**float**)
Label: **yesno**

Range: [**0**,**1**] Units: **1**
Unique values: **2** Missing .: **0**/**1,868**

Tabulation:	Freq.	Numeric	Label
	1,078	**0**	No
	790	**1**	Yes

antidepressants_0 — **Currently using antidepressants?**

Type: Numeric (**float**)
Label: **yesno**

Range: [**0**,**1**] Units: **1**
Unique values: **2** Missing .: **0**/**1,868**

Tabulation:	Freq.	Numeric	Label
	1,452	**0**	No
	416	**1**	Yes

```
1
2
3   ********************************************************************************
4   ********************************************************************************
5   ***********                                                          ***********
6   ***********        STATA DO-FILE FOR TARGET-D ANALYSIS               ***********
7   ***********                                                          ***********
8   ********************************************************************************
9   ********************************************************************************
10
11
12  * NOTE: Use code below to conduct analysis of Target-D data set
13  *
14  * Code is split into 3 modules:
15  * (1) Complete case analysis;
16  * (2) Multiple imputation; and
17  * (3) Bootstrap analysis.
18  *
19  *
20
21
22
23  ******************
24  ***************   MODULE 1 - COMPLETE CASE ANALYSIS
25  ************
26
27  ****************************
28  ***   DATA PREPARATION   ***
29  ****************************
30
31  *** PREPARATORY CODE FOR COMPLETE CASE ANALYSIS
32
33  ** Set working directory
34
35  cd "C:\Stata"
36
37
38  ** Load original dataset, keep in-scope variables and save the revised dataset
39
40  clear all
41  use "C:\Stata\TargetD_MERGED_edited.dta", replace
42
43  global var_sid study_id
44
45  global var_reg group phqdep_total_impute_t clinic_0 risk /// *Regular_variables
46                 age_num_0 gender_0 costintv SA1_costintv SA2_costintv
47
48  global var_imp costtothealthprof_2 costtotervisit_2 /// *Imputed_variables
49                 costtotovnstay_2 costtotdiagtest_2 costtotmed_2 ///
50                 costtothealthprof_3 costtotervisit_3 costtotovnstay_3 ///
51                 costtotdiagtest_3 costtotmed_3 ///
52                 costtotprodloss_2 costtotprodloss_3 ///
53                 AQoL8DUtility_1 AQoL8DUtility_2 AQoL8DUtility_3 ///
54                 phqdep_total_impute_2 phqdep_total_impute_3
55
56  global var_adj education_rec_0 employment_0  /// *SA3_adjustment_variables
57                 card_none_0 illness_t  health_0  live_alone_0 ///
58                 incomerec2_t psychpast_visited_0  antidepressants_0
59
60  keep $var_sid $var_reg $var_imp $var_adj
61  order $var_sid $var_reg $var_imp $var_adj
62
63  save "C:\Stata\TargetD_MERGED_inscope.dta", replace
64
65
66  ** Load revised dataset containing only the in-scope variables
67
68  clear all
69  use "C:\Stata\TargetD_MERGED_inscope.dta", replace
70
71
72  ** Add aggregate cost variables by study perspective
73
74  * Health sector perspective
75
76  gen costhealthsector_2 = costtothealthprof_2 + costtotervisit_2 + ///
77                           costtotovnstay_2 + costtotdiagtest_2 + ///
78                           costtotmed_2 + costintv
79
```

```
gen costhealthsector_3 = costtothealthprof_3 + costtotervisit_3 + ///
                         costtotovnstay_3 + costtotdiagtest_3 + ///
                         costtotmed_3

gen costhealthsectortot = costhealthsector_2 + costhealthsector_3

* Societal perspective

gen costsocietal_2 = costhealthsector_2 + costtotprodloss_2

gen costsocietal_3 = costhealthsector_3 + costtotprodloss_3

gen costsocietaltot = costsocietal_2 + costsocietal_3


** Add aggregate QALY variables

gen qaly_2 = 0.5*(AQoL8DUtility_1 + AQoL8DUtility_2)*(3/12)
gen qaly_3 = 0.5*(AQoL8DUtility_2 + AQoL8DUtility_3)*(9/12)
gen qalytot = qaly_2 + qaly_3



*********************************************
***   TARGET_D - COMPLETE CASE ANALYSIS   ***
*********************************************

*** GLM ANALYSIS OF COST VARIABLES AT 3-MONTH FOLLOW-UP (T2)

** Analysis of health sector costs - 3-month follow-up (T2)

* Adjusted for baseline outcome and stratification factors

glm costhealthsector_2 i.group phqdep_total_impute_t i.clinic_0 i.risk, family(gamma) link(log) eform

* Subgroup analysis

glm costhealthsector_2 i.group phqdep_total_impute_t i.clinic_0 if risk==1, family(gamma) link(log) eform
glm costhealthsector_2 i.group phqdep_total_impute_t i.clinic_0 if risk==2, family(gamma) link(log) eform
glm costhealthsector_2 i.group phqdep_total_impute_t i.clinic_0 if risk==3, family(gamma) link(log) eform


** Analysis of societal costs - 3-month follow-up (T2)

* Adjusted for baseline outcome and stratification factors

glm costsocietal_2 i.group phqdep_total_impute_t i.clinic_0 i.risk, family(gamma) link(log) eform

* Subgroup analysis

glm costsocietal_2 i.group phqdep_total_impute_t i.clinic_0 if risk==1, family(gamma) link(log) eform
glm costsocietal_2 i.group phqdep_total_impute_t i.clinic_0 if risk==2, family(gamma) link(log) eform
glm costsocietal_2 i.group phqdep_total_impute_t i.clinic_0 if risk==3, family(gamma) link(log) eform



*** GLM ANALYSIS OF COST VARIABLES AT 12-MONTH FOLLOW-UP (T2+T3)

** Analysis of health sector costs - 12-month follow-up (T2+T3)

* Adjusted for baseline outcome and stratification factors

glm costhealthsectortot i.group phqdep_total_impute_t i.clinic_0 i.risk, family(gamma) link(log) eform

* Subgroup analysis

glm costhealthsectortot i.group phqdep_total_impute_t i.clinic_0 if risk==1, family(gamma) link(log) eform
glm costhealthsectortot i.group phqdep_total_impute_t i.clinic_0 if risk==2, family(gamma) link(log) eform
```

```
148  glm costhealthsectortot i.group phqdep_total_impute_t i.clinic_0 if risk==3, family(gamma)
      link(log) eform
149
150
151  ** Analysis of societal costs - 12-month follow-up (T2+T3)
152
153  * Adjusted for baseline outcome and stratification factors
154
155  glm costsocietaltot i.group phqdep_total_impute_t i.clinic_0 i.risk, family(gamma) link(
     log) eform
156
157  * Subgroup analysis
158
159  glm costsocietaltot i.group phqdep_total_impute_t i.clinic_0 if risk==1, family(gamma)
     link(log) eform
160  glm costsocietaltot i.group phqdep_total_impute_t i.clinic_0 if risk==2, family(gamma)
     link(log) eform
161  glm costsocietaltot i.group phqdep_total_impute_t i.clinic_0 if risk==3, family(gamma)
     link(log) eform
162
163
164
165  *** GLM ANALYSIS OF QALYS AT 3-MONTH FOLLOW-UP (T2)
166
167  ** Analysis of QALYs - 3-month follow-up (T2)
168
169  * Adjusted for baseline utility, baseline outcome and stratification factors
170
171  glm qaly_2 i.group AQoL8DUtility_1 phqdep_total_impute_t i.clinic_0 i.risk, family(
     gaussian) link(identity)
172
173  * Subgroup analysis
174
175  glm qaly_2 i.group AQoL8DUtility_1 phqdep_total_impute_t i.clinic_0 if risk==1, family(
     gaussian) link(identity)
176  glm qaly_2 i.group AQoL8DUtility_1 phqdep_total_impute_t i.clinic_0 if risk==2, family(
     gaussian) link(identity)
177  glm qaly_2 i.group AQoL8DUtility_1 phqdep_total_impute_t i.clinic_0 if risk==3, family(
     gaussian) link(identity)
178
179
180
181  *** GLM ANALYSIS OF QALYS AT 12-MONTH FOLLOW-UP (T2+T3)
182
183  ** Analysis of QALYs - 12-month follow-up (T2+T3)
184
185  * Adjusted for baseline utility, baseline outcome and stratification factors
186
187  glm qalytot i.group AQoL8DUtility_1 phqdep_total_impute_t i.clinic_0 i.risk, family(
     gaussian) link(identity)
188
189  * Subgroup analysis
190
191  glm qalytot i.group AQoL8DUtility_1 phqdep_total_impute_t i.clinic_0 if risk==1, family(
     gaussian) link(identity)
192  glm qalytot i.group AQoL8DUtility_1 phqdep_total_impute_t i.clinic_0 if risk==2, family(
     gaussian) link(identity)
193  glm qalytot i.group AQoL8DUtility_1 phqdep_total_impute_t i.clinic_0 if risk==3, family(
     gaussian) link(identity)
194
195
196
197  ******************
198  ***************    END OF MODULE 1 - COMPLETE CASE ANALYSIS
199  ************
200
201
202
203
204
205
206  ******************
207  ***************    START OF MODULE 2 - IMPUTATION ANALYSIS
208  ************
209
210
211  ******************************
212  ***    DATA PREPARATION    ***
213  ******************************
```

```
214
215  *** PREPARATORY CODE FOR MULTIPLE IMPUTATION ANALYSIS
216
217  ** Set working directory
218
219  cd "C:\Stata"
220
221
222  ** Load original dataset, keep in-scope variables and save the revised dataset
223
224  clear all
225  use "C:\Stata\TargetD_MERGED_edited.dta", replace
226
227  global var_sid study_id
228
229  global var_reg group phqdep_total_impute_t clinic_0 risk /// *Regular_variables
230                 age_num_0 gender_0 costintv SA1_costintv SA2_costintv
231
232  global var_imp costtothealthprof_2 costtotervisit_2 /// *Imputed_variables
233                 costtotovnstay_2 costtotdiagtest_2 costtotmed_2 ///
234                 costtothealthprof_3 costtotervisit_3 costtotovnstay_3 ///
235                 costtotdiagtest_3 costtotmed_3 ///
236                 costtotprodloss_2 costtotprodloss_3 ///
237                 AQoL8DUtility_1 AQoL8DUtility_2 AQoL8DUtility_3 ///
238                 phqdep_total_impute_2 phqdep_total_impute_3
239
240  global var_adj education_rec_0 employment_0  /// *SA3_adjustment_variables
241                 card_none_0 illness_t  health_0  live_alone_0 ///
242                 incomerec2_t psychpast_visited_0  antidepressants_0
243
244  keep $var_sid $var_reg $var_imp $var_adj
245  order $var_sid $var_reg $var_imp $var_adj
246
247  save "C:\Stata\TargetD_MERGED_inscope.dta", replace
248
249
250  ** Load revised dataset containing only the in-scope variables
251
252  clear all
253  use "C:\Stata\TargetD_MERGED_inscope.dta", replace
254
255
256
257  *** SET UP IMPUTATION DATA SET AND REGISTER VARIABLES
258
259  ** Duplicate existing variables for the multiple imputation analysis
260
261  gen costtothealthprof_2_imp = costtothealthprof_2
262  gen costtotervisit_2_imp = costtotervisit_2
263  gen costtotovnstay_2_imp = costtotovnstay_2
264  gen costtotdiagtest_2_imp = costtotdiagtest_2
265  gen costtotmed_2_imp = costtotmed_2
266
267  gen costtothealthprof_3_imp = costtothealthprof_3
268  gen costtotervisit_3_imp = costtotervisit_3
269  gen costtotovnstay_3_imp = costtotovnstay_3
270  gen costtotdiagtest_3_imp = costtotdiagtest_3
271  gen costtotmed_3_imp = costtotmed_3
272
273  gen costtotprodloss_2_imp = costtotprodloss_2
274  gen costtotprodloss_3_imp = costtotprodloss_3
275
276  gen AQoL8DUtility_1_imp = AQoL8DUtility_1
277  gen AQoL8DUtility_2_imp = AQoL8DUtility_2
278  gen AQoL8DUtility_3_imp = AQoL8DUtility_3
279
280  gen phqdep_total_impute_2_imp = phqdep_total_impute_2
281  gen phqdep_total_impute_3_imp = phqdep_total_impute_3
282
283
284  ** MI estimation commands
285
286  * Declare MI data set
287
288  mi set flong
289
290  * Register variables to be imputed
291
292  mi register imputed costtothealthprof_2_imp costtotervisit_2_imp costtotovnstay_2_imp ///
```

```
293                     costtotdiagtest_2_imp costtotmed_2_imp costtothealthprof_3_imp ///
294                     costtotervisit_3_imp costtotovnstay_3_imp costtotdiagtest_3_imp ///
295                     costtotmed_3_imp costtotprodloss_2_imp costtotprodloss_3_imp ///
296                     AQoL8DUtility_1_imp AQoL8DUtility_2_imp AQoL8DUtility_3_imp ///
297                     phqdep_total_impute_2_imp phqdep_total_impute_3_imp
298
299     mi register regular $var_reg $var_adj
300
301     * Perform multiple imputation
302
303     mi impute chained ///
304         (pmm, knn(10)) AQoL8DUtility_1_imp AQoL8DUtility_2_imp AQoL8DUtility_3_imp ///
305                     phqdep_total_impute_2_imp phqdep_total_impute_3_imp ///
306                     costtothealthprof_2_imp costtotervisit_2_imp costtotovnstay_2_imp ///
307                     costtotdiagtest_2_imp costtotmed_2_imp costtothealthprof_3_imp ///
308                     costtotervisit_3_imp costtotovnstay_3_imp costtotdiagtest_3_imp ///
309                     costtotmed_3_imp costtotprodloss_2_imp costtotprodloss_3_imp ///
310         = i.group i.gender_0 i.clinic_0 age_num_0 education_rec_0 psychpast_visited_0, ///
311         by(risk) add(100) burnin(100) rseed(2812281) dots
312
313     * Generate (and register) passive variables
314
315     mi passive: gen costhealthsector_2_imp = costtothealthprof_2_imp + costtotervisit_2_imp +
    ///
316                                             costtotovnstay_2_imp + costtotdiagtest_2_imp + ///
317                                             costtotmed_2_imp + costintv
318     mi passive: gen costhealthsector_3_imp = costtothealthprof_3_imp + costtotervisit_3_imp +
    ///
319                                             costtotovnstay_3_imp + costtotdiagtest_3_imp + ///
320                                             costtotmed_3_imp
321     mi passive: gen costhealthsectortot_imp = costhealthsector_2_imp + costhealthsector_3_imp
322
323     mi passive: gen costsocietal_2_imp = costhealthsector_2_imp + costtotprodloss_2_imp
324     mi passive: gen costsocietal_3_imp = costhealthsector_3_imp + costtotprodloss_3_imp
325     mi passive: gen costsocietaltot_imp = costsocietal_2_imp + costsocietal_3_imp
326
327     mi passive: gen qaly_2_imp = 0.5*(AQoL8DUtility_1_imp + AQoL8DUtility_2_imp)*(3/12)
328     mi passive: gen qaly_3_imp = 0.5*(AQoL8DUtility_2_imp + AQoL8DUtility_3_imp)*(9/12)
329     mi passive: gen qalytot_imp = qaly_2_imp + qaly_3_imp
330
331     mi passive: gen SA1_costhealthsector_2_imp = costtothealthprof_2_imp +
    costtotervisit_2_imp + ///
332                                             costtotovnstay_2_imp + costtotdiagtest_2_imp
    + ///
333                                             costtotmed_2_imp + SA1_costintv
334     mi passive: gen SA1_costhealthsectortot_imp = SA1_costhealthsector_2_imp +
    costhealthsector_3_imp
335
336     mi passive: gen SA1_costsocietal_2_imp = SA1_costhealthsector_2_imp + costtotprodloss_2_imp
337     mi passive: gen SA1_costsocietaltot_imp = SA1_costsocietal_2_imp + costsocietal_3_imp
338
339     mi passive: gen SA2_costhealthsector_2_imp = costtothealthprof_2_imp +
    costtotervisit_2_imp + ///
340                                             costtotovnstay_2_imp + costtotdiagtest_2_imp
    + ///
341                                             costtotmed_2_imp + SA2_costintv
342     mi passive: gen SA2_costhealthsectortot_imp = SA2_costhealthsector_2_imp +
    costhealthsector_3_imp
343
344     mi passive: gen SA2_costsocietal_2_imp = SA2_costhealthsector_2_imp + costtotprodloss_2_imp
345     mi passive: gen SA2_costsocietaltot_imp = SA2_costsocietal_2_imp + costsocietal_3_imp
346
347     * Save copy of the multiple imputation data set
348
349     describe, short
350     save "C:\Stata\TargetD_MERGED_MI_update.dta", replace
351
352     * Clear data set from memory
353
354     clear all
355
356
357
358     ********************************************
359     ***   TARGET_D - ANALYSIS OF IMPUTED DATA   ***
360     ********************************************
361
362     *** Set working directory
363
```

```
364  cd "C:\Stata"
365
366
367  *** LOAD IMPUTED DATA SET
368
369  use "C:\Stata\TargetD_MERGED_MI_update.dta", clear
370  describe, short
371
372
373  *** BASE CASE STATISTICAL ANALYSES USING IMPUTED DATA SET
374
375  ** GLM analysis of cost variables
376
377  * Analysis of health sector costs at 3-month follow-up (T2)
378
379  mi estimate, eform post: glm costhealthsector_2_imp i.group phqdep_total_impute_t i.
     clinic_0 i.risk, family(gamma) link(log)
380  mimrgns group, predict(xb) eform
381
382  * Subgroup analysis
383
384  mi estimate, eform post: glm costhealthsector_2_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==1, family(gamma) link(log)
385  mimrgns group, predict(xb) eform
386
387  mi estimate, eform post: glm costhealthsector_2_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==2, family(gamma) link(log)
388  mimrgns group, predict(xb) eform
389
390  mi estimate, eform post: glm costhealthsector_2_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==3, family(gamma) link(log)
391  mimrgns group, predict(xb) eform
392
393  * Analysis of health sector costs at 12-month follow-up (T2+T3)
394
395  mi estimate, eform post: glm costhealthsectortot_imp i.group phqdep_total_impute_t i.
     clinic_0 i.risk, family(gamma) link(log)
396  mimrgns group, predict(xb) eform
397
398  * Subgroup analysis
399
400  mi estimate, eform post: glm costhealthsectortot_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==1, family(gamma) link(log)
401  mimrgns group, predict(xb) eform
402
403  mi estimate, eform post: glm costhealthsectortot_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==2, family(gamma) link(log)
404  mimrgns group, predict(xb) eform
405
406  mi estimate, eform post: glm costhealthsectortot_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==3, family(gamma) link(log)
407  mimrgns group, predict(xb) eform
408
409  * Analysis of societal costs at 3-month follow-up (T2)
410
411  mi estimate, eform post: glm costsocietal_2_imp i.group phqdep_total_impute_t i.clinic_0 i
     .risk, family(gamma) link(log)
412  mimrgns group, predict(xb) eform
413
414  * Subgroup analysis
415
416  mi estimate, eform post: glm costsocietal_2_imp i.group phqdep_total_impute_t i.clinic_0
     if risk==1, family(gamma) link(log)
417  mimrgns group, predict(xb) eform
418
419  mi estimate, eform post: glm costsocietal_2_imp i.group phqdep_total_impute_t i.clinic_0
     if risk==2, family(gamma) link(log)
420  mimrgns group, predict(xb) eform
421
422  mi estimate, eform post: glm costsocietal_2_imp i.group phqdep_total_impute_t i.clinic_0
     if risk==3, family(gamma) link(log)
423  mimrgns group, predict(xb) eform
424
425  * Analysis of societal costs at 12-month follow-up (T2+T3)
426
427  mi estimate, eform post: glm costsocietaltot_imp i.group phqdep_total_impute_t i.clinic_0
     i.risk, family(gamma) link(log)
428  mimrgns group, predict(xb) eform
429
```

```
* Subgroup analysis

mi estimate, eform post: glm costsocietaltot_imp i.group phqdep_total_impute_t i.clinic_0 if risk==1, family(gamma) link(log)
mimrgns group, predict(xb) eform

mi estimate, eform post: glm costsocietaltot_imp i.group phqdep_total_impute_t i.clinic_0 if risk==2, family(gamma) link(log)
mimrgns group, predict(xb) eform

mi estimate, eform post: glm costsocietaltot_imp i.group phqdep_total_impute_t i.clinic_0 if risk==3, family(gamma) link(log)
mimrgns group, predict(xb) eform


** GLM analysis of QALY variables

* 3-month follow-up (T2)

mi estimate: glm qaly_2_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0 i.risk, family(gaussian) link(identity)
mimrgns group, predict(xb)

* Subgroup analysis

mi estimate: glm qaly_2_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0 if risk==1, family(gaussian) link(identity)
mimrgns group, predict(xb)

mi estimate: glm qaly_2_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0 if risk==2, family(gaussian) link(identity)
mimrgns group, predict(xb)

mi estimate: glm qaly_2_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0 if risk==3, family(gaussian) link(identity)
mimrgns group, predict(xb)

* 12-month follow-up (T2+T3)

mi estimate: glm qalytot_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0 i.risk, family(gaussian) link(identity)
mimrgns group, predict(xb)

* Subgroup analysis

mi estimate: glm qalytot_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0 if risk==1, family(gaussian) link(identity)
mimrgns group, predict(xb)

mi estimate: glm qalytot_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0 if risk==2, family(gaussian) link(identity)
mimrgns group, predict(xb)

mi estimate: glm qalytot_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0 if risk==3, family(gaussian) link(identity)
mimrgns group, predict(xb)


*** SENSITIVITY ANALYSES USING IMPUTED DATA SET

** SA1 -- GLM analysis of cost variables

* Analysis of health sector costs at 3-month follow-up (T2)

mi estimate, eform post: glm SA1_costhealthsector_2_imp i.group phqdep_total_impute_t i.clinic_0 i.risk, family(gamma) link(log)

* Subgroup analysis

mi estimate, eform post: glm SA1_costhealthsector_2_imp i.group phqdep_total_impute_t i.clinic_0 if risk==1, family(gamma) link(log)
mi estimate, eform post: glm SA1_costhealthsector_2_imp i.group phqdep_total_impute_t i.clinic_0 if risk==2, family(gamma) link(log)
mi estimate, eform post: glm SA1_costhealthsector_2_imp i.group phqdep_total_impute_t i.clinic_0 if risk==3, family(gamma) link(log)

* Analysis of health sector costs at 12-month follow-up (T2+T3)
```

NOTE: Sensitivity analysis SA1 involves the inclusion of sunk costs for the development of the Target-D CPT (cost of screening changed from $0.96 per person to $2.30 per person).

```
494  mi estimate, eform post: glm SA1_costhealthsectortot_imp i.group phqdep_total_impute_t i.
     clinic_0 i.risk, family(gamma) link(log)
495
496  * Subgroup analysis
497
498  mi estimate, eform post: glm SA1_costhealthsectortot_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==1, family(gamma) link(log)
499  mi estimate, eform post: glm SA1_costhealthsectortot_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==2, family(gamma) link(log)
500  mi estimate, eform post: glm SA1_costhealthsectortot_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==3, family(gamma) link(log)
501
502  * Analysis of societal costs at 3-month follow-up (T2)
503
504  mi estimate, eform post: glm SA1_costsocietal_2_imp i.group phqdep_total_impute_t i.
     clinic_0 i.risk, family(gamma) link(log)
505
506  * Subgroup analysis
507
508  mi estimate, eform post: glm SA1_costsocietal_2_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==1, family(gamma) link(log)
509  mi estimate, eform post: glm SA1_costsocietal_2_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==2, family(gamma) link(log)
510  mi estimate, eform post: glm SA1_costsocietal_2_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==3, family(gamma) link(log)
511
512  * Analysis of societal costs at 12-month follow-up (T2+T3)
513
514  mi estimate, eform post: glm SA1_costsocietaltot_imp i.group phqdep_total_impute_t i.
     clinic_0 i.risk, family(gamma) link(log)
515
516  * Subgroup analysis
517
518  mi estimate, eform post: glm SA1_costsocietaltot_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==1, family(gamma) link(log)
519  mi estimate, eform post: glm SA1_costsocietaltot_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==2, family(gamma) link(log)
520  mi estimate, eform post: glm SA1_costsocietaltot_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==3, family(gamma) link(log)
521                                                  NOTE: Sensitivity analysis SA2 involves a higher unit
522                                                  cost for the clinician-guided iCBT course delivered to
523  ** SA2 -- GLM analysis of cost variables          the moderate prognostic group (unit cost changed
524                                                  from $132 per person to $222 per person).
525  * Analysis of health sector costs at 3-month follow-up (T2)
526
527  mi estimate, eform post: glm SA2_costhealthsector_2_imp i.group phqdep_total_impute_t i.
     clinic_0 i.risk, family(gamma) link(log)
528
529  * Subgroup analysis
530
531  mi estimate, eform post: glm SA2_costhealthsector_2_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==1, family(gamma) link(log)
532  mi estimate, eform post: glm SA2_costhealthsector_2_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==2, family(gamma) link(log)
533  mi estimate, eform post: glm SA2_costhealthsector_2_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==3, family(gamma) link(log)
534
535  * Analysis of health sector costs at 12-month follow-up (T2+T3)
536
537  mi estimate, eform post: glm SA2_costhealthsectortot_imp i.group phqdep_total_impute_t i.
     clinic_0 i.risk, family(gamma) link(log)
538
539  * Subgroup analysis
540
541  mi estimate, eform post: glm SA2_costhealthsectortot_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==1, family(gamma) link(log)
542  mi estimate, eform post: glm SA2_costhealthsectortot_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==2, family(gamma) link(log)
543  mi estimate, eform post: glm SA2_costhealthsectortot_imp i.group phqdep_total_impute_t i.
     clinic_0 if risk==3, family(gamma) link(log)
544
545  * Analysis of societal costs at 3-month follow-up (T2)
546
547  mi estimate, eform post: glm SA2_costsocietal_2_imp i.group phqdep_total_impute_t i.
     clinic_0 i.risk, family(gamma) link(log)
548
549  * Subgroup analysis
550
551  mi estimate, eform post: glm SA2_costsocietal_2_imp i.group phqdep_total_impute_t i.
```

```
clinic_0 if risk==1, family(gamma) link(log)
mi estimate, eform post: glm SA2_costsocietal_2_imp i.group phqdep_total_impute_t i.
clinic_0 if risk==2, family(gamma) link(log)
mi estimate, eform post: glm SA2_costsocietal_2_imp i.group phqdep_total_impute_t i.
clinic_0 if risk==3, family(gamma) link(log)

* Analysis of societal costs at 12-month follow-up (T2+T3)

mi estimate, eform post: glm SA2_costsocietaltot_imp i.group phqdep_total_impute_t i.
clinic_0 i.risk, family(gamma) link(log)

* Subgroup analysis

mi estimate, eform post: glm SA2_costsocietaltot_imp i.group phqdep_total_impute_t i.
clinic_0 if risk==1, family(gamma) link(log)
mi estimate, eform post: glm SA2_costsocietaltot_imp i.group phqdep_total_impute_t i.
clinic_0 if risk==2, family(gamma) link(log)
mi estimate, eform post: glm SA2_costsocietaltot_imp i.group phqdep_total_impute_t i.
clinic_0 if risk==3, family(gamma) link(log)


** SA3 -- GLM analysis of cost and QALY variables after adjustment for variables
**        associated with non-response to the primary outcome measure (PHQ-9 score)

* Analysis of health sector costs at 3-month follow-up (T2)

mi estimate, eform post: glm costhealthsector_2_imp i.group phqdep_total_impute_t i.
clinic_0 i.risk age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.
illness_t  i.health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.
antidepressants_0, family(gamma) link(log)

* Subgroup analysis

mi estimate, eform post: glm costhealthsector_2_imp i.group phqdep_total_impute_t i.
clinic_0 age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t
  i.health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if
risk==1, family(gamma) link(log)
mi estimate, eform post: glm costhealthsector_2_imp i.group phqdep_total_impute_t i.
clinic_0 age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t
  i.health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if
risk==2, family(gamma) link(log)
mi estimate, eform post: glm costhealthsector_2_imp i.group phqdep_total_impute_t i.
clinic_0 age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t
  i.health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if
risk==3, family(gamma) link(log)

* Analysis of health sector costs at 12-month follow-up (T2+T3)

mi estimate, eform post: glm costhealthsectortot_imp i.group phqdep_total_impute_t i.
clinic_0 i.risk age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.
illness_t  i.health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.
antidepressants_0, family(gamma) link(log)

* Subgroup analysis

mi estimate, eform post: glm costhealthsectortot_imp i.group phqdep_total_impute_t i.
clinic_0 age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t
  i.health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if
risk==1, family(gamma) link(log)
mi estimate, eform post: glm costhealthsectortot_imp i.group phqdep_total_impute_t i.
clinic_0 age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t
  i.health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if
risk==2, family(gamma) link(log)
mi estimate, eform post: glm costhealthsectortot_imp i.group phqdep_total_impute_t i.
clinic_0 age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t
  i.health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if
risk==3, family(gamma) link(log)

* Analysis of societal costs at 3-month follow-up (T2)

mi estimate, eform post: glm costsocietal_2_imp i.group phqdep_total_impute_t i.clinic_0 i
.risk age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0, family
(gamma) link(log)

* Subgroup analysis

mi estimate, eform post: glm costsocietal_2_imp i.group phqdep_total_impute_t i.clinic_0
age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
```

```
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if risk
==1, family(gamma) link(log)
mi estimate, eform post: glm costsocietal_2_imp i.group phqdep_total_impute_t i.clinic_0
age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if risk
==2, family(gamma) link(log)
mi estimate, eform post: glm costsocietal_2_imp i.group phqdep_total_impute_t i.clinic_0
age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if risk
==3, family(gamma) link(log)

* Analysis of societal costs at 12-month follow-up (T2+T3)

mi estimate, eform post: glm costsocietaltot_imp i.group phqdep_total_impute_t i.clinic_0
i.risk age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i
.health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0,
family(gamma) link(log)

* Subgroup analysis

mi estimate, eform post: glm costsocietaltot_imp i.group phqdep_total_impute_t i.clinic_0
age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if risk
==1, family(gamma) link(log)
mi estimate, eform post: glm costsocietaltot_imp i.group phqdep_total_impute_t i.clinic_0
age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if risk
==2, family(gamma) link(log)
mi estimate, eform post: glm costsocietaltot_imp i.group phqdep_total_impute_t i.clinic_0
age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if risk
==3, family(gamma) link(log)

* Analysis of QALYS at 3-month follow-up (T2)

mi estimate: glm qaly_2_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0 i
.risk age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0, family
(gaussian) link(identity)

* Subgroup analysis

mi estimate: glm qaly_2_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0
age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if risk
==1, family(gaussian) link(identity)
mi estimate: glm qaly_2_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0
age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if risk
==2, family(gaussian) link(identity)
mi estimate: glm qaly_2_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0
age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if risk
==3, family(gaussian) link(identity)

* Analysis of QALYS at 12-month follow-up (T2+T3)

mi estimate: glm qalytot_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0
i.risk age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i
.health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0,
family(gaussian) link(identity)

* Subgroup analysis

mi estimate: glm qalytot_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0
age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if risk
==1, family(gaussian) link(identity)
mi estimate: glm qalytot_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0
age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if risk
==2, family(gaussian) link(identity)
mi estimate: glm qalytot_imp i.group AQoL8DUtility_1_imp phqdep_total_impute_t i.clinic_0
age_num_0 i.gender_0 i.education_rec_0 i.employment_0  i.card_none_0 i.illness_t  i.
health_0  i.live_alone_0 i.incomerec2_t i.psychpast_visited_0  i.antidepressants_0 if risk
==3, family(gaussian) link(identity)


** Clear data set from memory
```

```
clear all



******************
***************   END OF MODULE 2 - IMPUTATION ANALYSIS
************



******************
***************   START OF MODULE 3 - BOOTSTRAP ANALYSIS
************


*****************************************************************
***   TARGET_D - BOOTSTRAP WITH NESTED MULTIPLE IMPUTATION   ***
*****************************************************************

**
** The resampling method adopted in this code comprises the 'single imputation
** nested in bootstrapping' approach advocated by Brand et al., (2016) and other
** leading health economists, such as Prof Andrew Briggs. The code presented
** below involves using the MICE command to impute a singular complete dataset,
** then performing GLM analyses to examine differences in mean costs and QALYs.
**
** The margins command (dydx) is then used to estimate the marginal change in
** costs/QALYs due to a one-unit change in the intervention group (where
** 0 = control and 1 = intervention). ICERs are subsequently estimated as the
** difference in mean costs divided by the difference in mean QALYs.
**
** Brand et al., 2016 -- https://doi.org/10.1002/sim.7956
**
** NOTE: The nested bootstrap analysis has the advantage of enabling
**       the utilisation of the margins command. There is currently no
**       straightforward method by which to use the margins command alongside
**       the 'MI estimate' commands. See link below for further reading on
**       this issue.
**
** Link:
https://stats.idre.ucla.edu/stata/faq/how-can-i-get-margins-and-marginsplot-with-multiply-imputed-data/
**
**
** The bootstrap percentile method was used to estimate 95% confidence intervals
** (95% CIs) around the difference in mean costs, difference in mean QALYs and
** mean ICERs. This involves identifying the 2.5th and 97.5th percentiles of the
** 'm' bootstrap resamples that are generated. A total of 1,000 bootstrap resamples
** (m = 1,000) were produced.
**
**


*** PREPARATORY CODE

** Set working directory

cd "C:\Stata"


** Load dataset and keep in-scope variables

clear all
use "C:\Stata\TargetD_MERGED_inscope.dta", replace

global var_sid study_id

global var_reg group phqdep_total_impute_t clinic_0 risk /// *Regular_variables
               age_num_0 gender_0 costintv SA1_costintv SA2_costintv

global var_imp costtothealthprof_2 costtotervisit_2 /// *Imputed_variables
               costtotovnstay_2 costtotdiagtest_2 costtotmed_2 ///
               costtothealthprof_3 costtotervisit_3 costtotovnstay_3 ///
               costtotdiagtest_3 costtotmed_3 ///
               costtotprodloss_2 costtotprodloss_3 ///
```

```
708                 AQoL8DUtility_1 AQoL8DUtility_2 AQoL8DUtility_3 ///
709                 phqdep_total_impute_2 phqdep_total_impute_3
710
711 global var_adj education_rec_0 employment_0  /// *SA3_adjustment_variables
712                 card_none_0 illness_t  health_0  live_alone_0 ///
713                 incomerec2_t psychpast_visited_0  antidepressants_0
714
715 keep $var_sid $var_reg $var_imp $var_adj
716 order $var_sid $var_reg $var_imp $var_adj
717
718
719 ** Duplicate existing variables for the multiple imputation analysis
720
721 qui: gen costtothealthprof_2_imp = costtothealthprof_2
722 qui: gen costtotervisit_2_imp = costtotervisit_2
723 qui: gen costtotovnstay_2_imp = costtotovnstay_2
724 qui: gen costtotdiagtest_2_imp = costtotdiagtest_2
725 qui: gen costtotmed_2_imp = costtotmed_2
726
727 qui: gen costtothealthprof_3_imp = costtothealthprof_3
728 qui: gen costtotervisit_3_imp = costtotervisit_3
729 qui: gen costtotovnstay_3_imp = costtotovnstay_3
730 qui: gen costtotdiagtest_3_imp = costtotdiagtest_3
731 qui: gen costtotmed_3_imp = costtotmed_3
732
733 qui: gen costtotprodloss_2_imp = costtotprodloss_2
734 qui: gen costtotprodloss_3_imp = costtotprodloss_3
735
736 qui: gen AQoL8DUtility_1_imp = AQoL8DUtility_1
737 qui: gen AQoL8DUtility_2_imp = AQoL8DUtility_2
738 qui: gen AQoL8DUtility_3_imp = AQoL8DUtility_3
739
740 qui: gen phqdep_total_impute_2_imp = phqdep_total_impute_2
741 qui: gen phqdep_total_impute_3_imp = phqdep_total_impute_3
742
743
744 ** MI estimation commands
745
746 * Declare MI data set
747 *
748 * -------- IMPORTANT: DATA MUST BE SET IN THE WIDE FORMAT!! --------
749 * --------  BOOTSTRAP IS INCOMPATIBLE WITH THE LONG FORMAT  --------
750
751 mi set wide
752
753 * Register variables to be imputed
754
755 mi register imputed costtothealthprof_2_imp costtotervisit_2_imp costtotovnstay_2_imp ///
756                     costtotdiagtest_2_imp costtotmed_2_imp costtothealthprof_3_imp ///
757                     costtotervisit_3_imp costtotovnstay_3_imp costtotdiagtest_3_imp ///
758                     costtotmed_3_imp costtotprodloss_2_imp costtotprodloss_3_imp ///
759                     AQoL8DUtility_1_imp AQoL8DUtility_2_imp AQoL8DUtility_3_imp ///
760                     phqdep_total_impute_2_imp phqdep_total_impute_3_imp
761
762 mi register regular $var_reg $var_adj
763
764 * Generate (and register) passive variables
765
766 mi passive: gen costhealthsector_2_imp = costtothealthprof_2_imp + costtotervisit_2_imp +
    ///
767                                          costtotovnstay_2_imp + costtotdiagtest_2_imp + ///
768                                          costtotmed_2_imp + costintv
769 mi passive: gen costhealthsector_3_imp = costtothealthprof_3_imp + costtotervisit_3_imp +
    ///
770                                          costtotovnstay_3_imp + costtotdiagtest_3_imp + ///
771                                          costtotmed_3_imp
772 mi passive: gen costhealthsectortot_imp = costhealthsector_2_imp + costhealthsector_3_imp
773
774 mi passive: gen costsocietal_2_imp = costhealthsector_2_imp + costtotprodloss_2_imp
775 mi passive: gen costsocietal_3_imp = costhealthsector_3_imp + costtotprodloss_3_imp
776 mi passive: gen costsocietaltot_imp = costsocietal_2_imp + costsocietal_3_imp
777
778 mi passive: gen qaly_2_imp = 0.5*(AQoL8DUtility_1_imp + AQoL8DUtility_2_imp)*(3/12)
779 mi passive: gen qaly_3_imp = 0.5*(AQoL8DUtility_2_imp + AQoL8DUtility_3_imp)*(9/12)
780 mi passive: gen qalytot_imp = qaly_2_imp + qaly_3_imp
781
782
783 ** Save the declared multiple imputation dataset template
784
```

```
785 save "C:\Stata\TargetD_MERGED_inscope_MI_template.dta", replace
786
787
788
789 *** ESTIMATE ICERS AT 3-MONTH FOLLOW-UP (T2) AND 12-MONTH FOLLOW-UP (T2+T3)
790
791 ** BS10: Adjusted - Cost per QALY | Health sector & Societal | 3 mths & 12 mths
792
793 * Load multiple imputation dataset template
794
795 clear all
796 use "C:\Stata\TargetD_MERGED_inscope_MI_template.dta", clear
797
798
799 * Declare additional covariates for GLM analyses nested within bootstrap
800
801 global var_cov phqdep_total_impute_t i.clinic_0 i.risk
802
803
804 * Define program
805
806 program                 ,
807
808     ** Preserve data environment to be restored after program termination
809
810     preserve
811
812     ** Perform multiple imputation
813
814     mi impute chained ///
815     (pmm, knn(10)) AQoL8DUtility_1_imp AQoL8DUtility_2_imp AQoL8DUtility_3_imp ///
816                    phqdep_total_impute_2_imp phqdep_total_impute_3_imp ///
817                    costtothealthprof_2_imp costtotervisit_2_imp costtotovnstay_2_imp ///
818                    costtotdiagtest_2_imp costtotmed_2_imp costtothealthprof_3_imp ///
819                    costtotervisit_3_imp costtotovnstay_3_imp costtotdiagtest_3_imp ///
820                    costtotmed_3_imp costtotprodloss_2_imp costtotprodloss_3_imp ///
821     = i.group i.gender_0 i.clinic_0 age_num_0 education_rec_0 psychpast_visited_0, ///
822     by(risk) add(1) dots
823
824     * NOTE: Do not set the seed of the nested imputation within the bootstrap.
825     *       Also set the number of imputations to M=1.
826
827     ** Replace passive variables
828
829     mi passive: replace costhealthsector_2_imp = costtothealthprof_2_imp +
    costtotervisit_2_imp + ///
830                                          costtotovnstay_2_imp + costtotdiagtest_2_imp + ///
831                                          costtotmed_2_imp + costintv
832     mi passive: replace costhealthsector_3_imp = costtothealthprof_3_imp +
    costtotervisit_3_imp + ///
833                                          costtotovnstay_3_imp + costtotdiagtest_3_imp + ///
834                                          costtotmed_3_imp
835     mi passive: replace costhealthsectortot_imp = costhealthsector_2_imp +
    costhealthsector_3_imp
836
837     mi passive: replace costsocietal_2_imp = costhealthsector_2_imp + costtotprodloss_2_imp
838     mi passive: replace costsocietal_3_imp = costhealthsector_3_imp + costtotprodloss_3_imp
839     mi passive: replace costsocietaltot_imp = costsocietal_2_imp + costsocietal_3_imp
840
841     mi passive: replace qaly_2_imp = 0.5*(AQoL8DUtility_1_imp + AQoL8DUtility_2_imp)*(3/12)
842     mi passive: replace qaly_3_imp = 0.5*(AQoL8DUtility_2_imp + AQoL8DUtility_3_imp)*(9/12)
843     mi passive: replace qalytot_imp = qaly_2_imp + qaly_3_imp
844
845     ** Extract single imputed dataset
846
847     mi extract 1, clear
848
849
850     ** Conduct GLM analysis using single MI dataset in memory
851
852     * GLM analysis of QALY variables - 3 months
853
854     glm qaly_2_imp i.group AQoL8DUtility_1_imp $var_cov, family(gaussian) link(identity)
855
856     margins, dydx(group)
857     return scalar q1 = el(r(table),1,2)
858
859     * GLM analysis of QALY variables - 12 months
860
```

```
    glm qalytot_imp i.group AQoL8DUtility_1_imp $var_cov, family(gaussian) link(identity)

    margins, dydx(group)
    return scalar q6 = el(r(table),1,2)

    * GLM analysis of Cost variables - 3 months, Health sector

    glm costhealthsector_2_imp i.group $var_cov, family(gamma) link(log) eform

    margins, dydx(group)
    return scalar c1 = el(r(table),1,2)

    * GLM analysis of Cost variables - 3 months, Societal

    glm costsocietal_2_imp i.group $var_cov, family(gamma) link(log) eform

    margins, dydx(group)
    return scalar c2 = el(r(table),1,2)

    * GLM analysis of Cost variables - 12 months, Health sector

    glm costhealthsectortot_imp i.group $var_cov, family(gamma) link(log) eform

    margins, dydx(group)
    return scalar c6 = el(r(table),1,2)

    * GLM analysis of Cost variables - 12 months, Societal

    glm costsocietaltot_imp i.group $var_cov, family(gamma) link(log) eform

    margins, dydx(group)
    return scalar c7 = el(r(table),1,2)

end

* Perform bootstrap analysis

set seed 19

bootstrap inc_qalys_3m=r(q1)  inc_costs_hlth_3m=r(c1)  inc_costs_sctl_3m=r(c2)  ///
          inc_qalys_12m=r(q6) inc_costs_hlth_12m=r(c6) inc_costs_sctl_12m=r(c7), ///
          reps(10) strata(group) saving(TargetD_Bootstrap_GLM_BS10, replace) : myBS10

use TargetD_Bootstrap_GLM_BS10, clear
estat bootstrap, all



******************
***************    END OF MODULE 3 - BOOTSTRAP ANALYSIS
************
```