

```
1
2
3 //Simulation for Causal Pathway paper
4
5 //This runs the simulation and saves the output data in data files
6
7
8 //First, we set up the file
9
10 clear
11
12 clear matrix
13
14 set seed 1234
15
16 set scheme lean1
17
18 set more off
19
20 capture log close
21
22 global DateOfAnalysis 2016-2-1
23
24
25 log using "Log Simulation $DateOfAnalysis.log", replace
26
27
28 //Now we define all of the "magic numbers" of the simulation
29
30 global ProportionCD4Under200 0.3
31
32 global BaselineHazard 0.000125
33
34 global PropViroSuppIfCD4High 0.8
35
36 global PropViroSuppIfCD4Low 0.5
37
38 global ProbOfBactrimScript 0.8
39
40 global BactrimAdherenceIfVLLow 1
41
42 global BactrimAdherenceIfVLHigh 0.25
43
44 global CD4Under200HazardRatio 3
45
46 global ViralLoadUnder1000HazardRatio 0.5
47
48 global BactrimPrescriptionHazardRatio 0.7
49
50 global BactrimReceiptHazardRatio 0.7
51
52 global NumberOfReps 100
53
54 global SizeOfCohort 2000
55
56 global DaysOfFollowup 2192
57
58
59 //We now define all of the matrices which store the results of the simulation
60
61 matrix Model1=J($NumberOfReps , 6, .)
62
63 matrix colnames Model1 = CD4Under200Estimate CD4Under200SE ViralLoadUnder1000Estimate ViralLoadUnder1000SE BactrimReceiptEstimate
64 BactrimReceiptSE
65
66
67 matrix Model2=J($NumberOfReps , 2, .)
68
69 matrix colnames Model2 = CD4Under200Estimate CD4Under200SE
```

```

1
2
3
4
5 matrix Model3=J($NumberOfReps , 2, .)
6 matrix colnames Model3 = ViralLoadUnder1000Estimate ViralLoadUnder1000SE
7
8
9 matrix Model4=J($NumberOfReps , 2, .)
10 matrix colnames Model4 = BactrimReceiptEstimate BactrimReceiptSE
11
12
13 matrix Model5=J($NumberOfReps , 2, .)
14 matrix colnames Model5 = BactrimPrescriptionEstimate BactrimPrescriptionSE
15
16
17
18 matrix Model6=J($NumberOfReps , 4, .)
19 matrix colnames Model6 = CD4Under200Estimate CD4Under200SE ViralLoadUnder1000Estimate ViralLoadUnder1000SE
20
21
22
23 matrix Model7=J($NumberOfReps , 6, .)
24 matrix colnames Model7 = CD4Under200Estimate CD4Under200SE ViralLoadUnder1000Estimate ViralLoadUnder1000SE
25 BactrimPrescriptionEstimate BactrimPrescriptionSE
26
27
28 matrix Model8=J($NumberOfReps , 4, .)
29 matrix colnames Model8 = CD4Under200Estimate CD4Under200SE BactrimReceiptEstimate BactrimReceiptSE
30
31
32
33 matrix Model9=J($NumberOfReps , 4, .)
34 matrix colnames Model9 = CD4Under200Estimate CD4Under200SE BactrimPrescriptionEstimate BactrimPrescriptionSE
35
36
37
38 matrix Model10=J($NumberOfReps , 4, .)
39 matrix colnames Model10 = ViralLoadUnder1000Estimate ViralLoadUnder1000SE BactrimReceiptEstimate BactrimReceiptSE
40
41
42 matrix Model11=J($NumberOfReps , 4, .)
43 matrix colnames Model11 = ViralLoadUnder1000Estimate ViralLoadUnder1000SE BactrimPrescriptionEstimate BactrimPrescriptionSE
44
45
46
47 matrix Model12=J($NumberOfReps , 10, .)
48 matrix colnames Model12 = CD4Under200Est CD4Under200SE ViralLoadUnder1000Est ViralLoadUnder1000SE BactrimPrescriptionEst
49 BactrimPrescriptionSE ViralLoadBactrimRXInteractionEst ViralLoadBactrimRXInteractionSE ViralLoadBactrimRxLinComEst
50 ViralLoadBactrimRxLinComSE
51
52
53 matrix Model13=J($NumberOfReps , 10, .)
54 matrix colnames Model13 = CD4Under200Est CD4Under200SE ViralLoadUnder1000Est ViralLoadUnder1000SE BactrimReceiptEst
55 BactrimReceiptSE ViralLoadBactrimRecInterxnEst ViralLoadBactrimRecInterxnSE ViralLoadBactrimRecLinComEst
56 ViralLoadBactrimRecLinComSE
57
58
59
60

```

```
1
2
3 matrix NumFailures=J($NumberOfReps , 1, .)
4
5 matrix colnames NumFailures = NumFailures
6
7
8
9 //Now we run the simulation via loop:
10
11 forvalues i=1(1)$NumberOfReps {
12
13
14 clear
15
16 display "This is simulation #: " `i'
17
18 set obs $SizeOfCohort
19
20 generate PatientID=_n
21
22 generate CD4Under200=(runiform())<=$ProportionCD4Under200 )
23
24 generate ViralLoadUnder1000=.
25
26 replace ViralLoadUnder1000=(runiform())<=$PropViroSuppIfCD4High ) if CD4Under200==0
27
28 replace ViralLoadUnder1000=(runiform())<=$PropViroSuppIfCD4Low ) if CD4Under200==1
29
30
31 generate BactrimPrescription=0
32
33 replace BactrimPrescription=1 if ((runiform())<=$ProbOfBactrimScript )&(CD4Under200==1))
34
35
36 generate BactrimReceipt=.
37
38 replace BactrimReceipt=(runiform())<=$BactrimAdherenceIfVLLow )*BactrimPrescription if ViralLoadUnder1000==1
39
40 replace BactrimReceipt=(runiform())<=$BactrimAdherenceIfVLHigh )*BactrimPrescription if ViralLoadUnder1000==0
41
42
43 tab CD4Under200, missing
44
45 tab ViralLoadUnder1000, missing
46
47 tab BactrimPrescription, missing
48
49 tab BactrimReceipt, missing
50
51 bysort CD4Under200: tab ViralLoadUnder1000, missing
52
53 bysort CD4Under200: tab BactrimPrescription, missing
54
55 bysort ViralLoadUnder1000: tab BactrimReceipt, missing
56
57 bysort ViralLoadUnder1000: tab BactrimReceipt if CD4Under200, missing
58
59
60 expand $DaysOfFollowup
61
62 bysort PatientID: generate Day=_n
63
64 sort PatientID Day
```

```

1
2
3
4
5 generate RandomEvent=runiform()
6 generate CD4Hazard=1+($CD4Under200HazardRatio -1)*(CD4Under200==1)
7 generate ViralLoadHazard=1+($ViralLoadUnder1000HazardRatio -1)*(ViralLoadUnder1000==1)
8
9 generate BactrimHazard=1+($BactrimPrescriptionHazardRatio -1)*(BactrimReceipt==1)
10 generate CumulativeHazard= ($BaselineHazard )*CD4Hazard*ViralLoadHazard*BactrimHazard
11
12 tab CumulativeHazard
13
14 generate Event=( RandomEvent<=CumulativeHazard)
15
16 by PatientID: generate CumulativeEvent=sum(Event)
17
18 generate ST=(CumulativeEvent==0)|(CumulativeEvent==1 & Event==1)
19
20 //Now we identify who had an SSTI and drop observations after the SSTI
21 by PatientID: egen HadSSTI=max(Event)
22
23 drop if ST==0
24
25
26 //We identify the last date in the data set for each individual, which is either
27 //the date of SSTI or the date of exiting the cohort.
28
29 by PatientID: egen FinalDate=max(Day)
30
31
32 //We store the number of failures in the appropriate matrix
33
34 quietly summarize if (Event & ST)
35
36 matrix NumFailures[`i',1]=`r(N)'
37
38 display "Number of Failures is: " NumFailures[`i',1]
39
40
41 //We only keep the first observation since there is no time-varying data
42 //and running the Cox regressions on the large data set is simply way too slow.
43
44 keep if Day==1
45
46
47 //Now we stset the smaller data set:
48
49 stset FinalDate, failure( HadSSTI ==1)
50
51 //Now we run the Cox regressions and store the results in the appropriate matrices:
52
53 display ". stcox CD4Under200 ViralLoadUnder1000 BactrimReceipt, nolog nohr"
54
55 stcox CD4Under200 ViralLoadUnder1000 BactrimReceipt, nolog nohr
56
57 matrix b=e(b)
58
59 matrix V=e(V)
60

```

```
1
2
3     matrix Model1['i',1]=b[1,1]
4
5     matrix Model1['i',3]=b[1,2]
6
7     matrix Model1['i',5]=b[1,3]
8
9     matrix Model1['i',2]=(V[1,1])^0.5
10
11    matrix Model1['i',4]=(V[2,2])^0.5
12
13    matrix Model1['i',6]=(V[3,3])^0.5
14
15    display ". stcox CD4Under200, nolog nohr"
16
17    stcox CD4Under200, nolog nohr
18
19    matrix b=e(b)
20
21    matrix V=e(V)
22
23    matrix Model2['i',1]=b[1,1]
24
25    matrix Model2['i',2]=(V[1,1])^0.5
26
27    display ". stcox ViralLoadUnder1000, nolog nohr"
28
29    stcox ViralLoadUnder1000, nolog nohr
30
31    matrix b=e(b)
32
33    matrix V=e(V)
34
35    matrix Model3['i',1]=b[1,1]
36
37    matrix Model3['i',2]=(V[1,1])^0.5
38
39    display ". stcox BactrimReceipt, nolog nohr"
40
41    stcox BactrimReceipt, nolog nohr
42
43    matrix b=e(b)
44
45    matrix V=e(V)
46
47    matrix Model4['i',1]=b[1,1]
48
49    matrix Model4['i',2]=(V[1,1])^0.5
50
51    display ". stcox BactrimPrescription, nolog nohr"
52
53    stcox BactrimPrescription, nolog nohr
54
55    matrix b=e(b)
56
57    matrix V=e(V)
58
59    matrix Model5['i',1]=b[1,1]
60
61    matrix Model5['i',2]=(V[1,1])^0.5
```

```
1
2
3 display ". stcox CD4Under200 ViralLoadUnder1000, nolog nohr"
4
5 stcox CD4Under200 ViralLoadUnder1000, nolog nohr
6
7 matrix b=e(b)
8
9 matrix V=e(V)
10
11 matrix Model6[`i',1]=b[1,1]
12
13 matrix Model6[`i',3]=b[1,2]
14
15 matrix Model6[`i',2]=(V[1,1])^0.5
16
17 matrix Model6[`i',4]=(V[2,2])^0.5
18
19
20 display ". stcox CD4Under200 ViralLoadUnder1000 BactrimPrescription, nolog nohr"
21
22 stcox CD4Under200 ViralLoadUnder1000 BactrimPrescription, nolog nohr
23
24 matrix b=e(b)
25
26 matrix V=e(V)
27
28 matrix Model7[`i',1]=b[1,1]
29
30 matrix Model7[`i',3]=b[1,2]
31
32 matrix Model7[`i',5]=b[1,3]
33
34 matrix Model7[`i',2]=(V[1,1])^0.5
35
36 matrix Model7[`i',4]=(V[2,2])^0.5
37
38 matrix Model7[`i',6]=(V[3,3])^0.5
39
40
41 display ". stcox CD4Under200 BactrimReceipt, nolog nohr"
42
43 stcox CD4Under200 BactrimReceipt, nolog nohr
44
45 matrix b=e(b)
46
47 matrix V=e(V)
48
49 matrix Model8[`i',1]=b[1,1]
50
51 matrix Model8[`i',3]=b[1,2]
52
53 matrix Model8[`i',2]=(V[1,1])^0.5
54
55 matrix Model8[`i',4]=(V[2,2])^0.5
56
57
58 display ". stcox CD4Under200 BactrimPrescription, nolog nohr"
59
60 stcox CD4Under200 BactrimPrescription, nolog nohr
61
62 matrix b=e(b)
63
64 matrix V=e(V)
65
66 matrix Model9[`i',1]=b[1,1]
67
68 matrix Model9[`i',3]=b[1,2]
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

```
matrix Model9['i',2]=(V[1,1])^0.5
matrix Model9['i',4]=(V[2,2])^0.5

display ". stcox ViralLoadUnder1000 BactrimReceipt, nolog nohr"
stcox ViralLoadUnder1000 BactrimReceipt, nolog nohr

matrix b=e(b)
matrix V=e(V)

matrix Model10['i',1]=b[1,1]
matrix Model10['i',3]=b[1,2]
matrix Model10['i',2]=(V[1,1])^0.5
matrix Model10['i',4]=(V[2,2])^0.5

display ". stcox ViralLoadUnder1000 BactrimPrescription, nolog nohr"
stcox ViralLoadUnder1000 BactrimPrescription, nolog nohr

matrix b=e(b)
matrix V=e(V)

matrix Model11['i',1]=b[1,1]
matrix Model11['i',3]=b[1,2]
matrix Model11['i',2]=(V[1,1])^0.5
matrix Model11['i',4]=(V[2,2])^0.5

display ". stcox CD4Under200 i.ViralLoadUnder1000##i.BactrimPrescription, nolog nohr"
stcox CD4Under200 i.ViralLoadUnder1000##i.BactrimPrescription, nolog nohr

matrix b=e(b)
matrix V=e(V)

matrix Model12['i',1]=b[1,1]
matrix Model12['i',3]=b[1,3]
matrix Model12['i',5]=b[1,5]
matrix Model12['i',7]=b[1,9]
matrix Model12['i',2]=(V[1,1])^0.5
matrix Model12['i',4]=(V[3,3])^0.5
matrix Model12['i',6]=(V[5,5])^0.5
matrix Model12['i',8]=(V[9,9])^0.5
```

```
1
2
3
4
5 display ". lincom 1.BactrimPrescription+1.ViralLoadUnder1000#1.BactrimPrescription nohr"
6
7 lincom 1.BactrimPrescription+1.ViralLoadUnder1000#1.BactrimPrescription
8
9 matrix Model12[`i',9]=`r(estimate)'
10
11 matrix Model12[`i',10]=`r(se)'
12
13
14 display ". stcox CD4Under200 i.ViralLoadUnder1000##i.BactrimReceipt, nolog nohr"
15
16 stcox CD4Under200 i.ViralLoadUnder1000##i.BactrimReceipt, nolog nohr
17
18 matrix b=e(b)
19
20 matrix V=e(V)
21
22 matrix Model13[`i',1]=b[1,1]
23
24 matrix Model13[`i',3]=b[1,3]
25
26 matrix Model13[`i',5]=b[1,5]
27
28 matrix Model13[`i',7]=b[1,9]
29
30 matrix Model13[`i',2]=(V[1,1])^0.5
31
32 matrix Model13[`i',4]=(V[3,3])^0.5
33
34 matrix Model13[`i',6]=(V[5,5])^0.5
35
36 matrix Model13[`i',8]=(V[9,9])^0.5
37
38
39
40
41 display ". lincom 1.BactrimReceipt+1.ViralLoadUnder1000#1.BactrimReceipt nohr"
42
43 lincom 1.BactrimReceipt+1.ViralLoadUnder1000#1.BactrimReceipt
44
45 matrix Model13[`i',9]=`r(estimate)'
46
47 matrix Model13[`i',10]=`r(se)'
48
49 }
50
51 clear
52
53 //We list out the matrix values
54
55 forvalues i=1(1)13 {
56
57 matrix list Model`i'
58
59 }
60
61 matrix list NumFailures
62
63
64 //For each matrix, we save the data in a separate data set for later analysis:
65
66 svmat Model1, name(col)
67
68 save "Data Model1 $DateOfAnalysis.dta", replace
```



```
1
2
3      clear
4
5
6      svmat Model2, name(col)
7
8      save "Data Model2 $DateOfAnalysis.dta", replace
9
10     clear
11
12     svmat Model3, name(col)
13
14     save "Data Model3 $DateOfAnalysis.dta", replace
15
16     clear
17
18     svmat Model4, name(col)
19
20     save "Data Model4 $DateOfAnalysis.dta", replace
21
22     clear
23
24     svmat Model5, name(col)
25
26     save "Data Model5 $DateOfAnalysis.dta", replace
27
28     clear
29
30     svmat Model6, name(col)
31
32     save "Data Model6 $DateOfAnalysis.dta", replace
33
34     clear
35
36     svmat Model7, name(col)
37
38     save "Data Model7 $DateOfAnalysis.dta", replace
39
40     clear
41
42     svmat Model8, name(col)
43
44     save "Data Model8 $DateOfAnalysis.dta", replace
45
46     clear
47
48     svmat Model9, name(col)
49
50     save "Data Model9 $DateOfAnalysis.dta", replace
51
52     clear
53
54     svmat Model10, name(col)
55
56     save "Data Model10 $DateOfAnalysis.dta", replace
57
58
59
60
```

```
1
2
3 clear
4
5
6 svmat Model11, name(col)
7
8 save "Data Model11 $DateOfAnalysis.dta", replace
9
10 clear
11
12 svmat Model12, name(col)
13
14 save "Data Model12 $DateOfAnalysis.dta", replace
15
16 clear
17
18 svmat Model13, name(col)
19
20 save "Data Model13 $DateOfAnalysis.dta", replace
21
22 clear
23
24 svmat NumFailures, name(col)
25
26 save "Data NumFailures $DateOfAnalysis.dta", replace
27
28 clear
29
30
31 log close
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
```

```
1
2
3      clear
4
5      set scheme lean1
6
7      set more off
8
9      capture log close
10
11     log using "Log Output $DateOfAnalysis.log", replace
12
13
14     use "Data Model1 $DateOfAnalysis.dta", clear
15
16     display "Model1:"
17
18     foreach var in CD4Under200 ViralLoadUnder1000 BactrimReceipt {
19
20     sort `var'Estimate
21
22     generate SimNumber=_n
23
24     display ". summarize `var'Estimate"
25
26     summarize `var'Estimate
27
28     local `var'GeoMean=exp(`r(mean)')
29
30     display "Exponentiated mean of log HR of variable is: " `var'GeoMean'
31
32     histogram `var'Estimate
33
34     generate HR`var'Estimate=exp(`var'Estimate)
35
36     generate Lower`var'=exp(`var'Estimate+(invnorm(0.025))*`var'SE)
37
38     generate Upper`var'=exp(`var'Estimate+(invnorm(0.975))*`var'SE)
39
40     display ". summarize HR`var'"
41
42     summarize HR`var'Estimate
43
44     histogram HR`var'Estimate, xscale(log) frequency
45
46     graph export "Model1HR`var'Histogram $DateOfAnalysis.tif", replace
47
48     twoway (scatter HR`var'Estimate SimNumber, yscale(log) title("Forest plot of estimates of `var' in Model 1") ytitle(Hazard Ratio)
49     xtitle(Simulation Number))(rcap Upper`var' Lower`var' SimNumber, vertical), yline(`${var'HazardRatio}) yline(`var'GeoMean',
50     lpat(dash)) legend(off)
51
52     graph export "Model1HR`var'ForestPlot $DateOfAnalysis.tif", replace
53
54     drop SimNumber
55
56     }
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
```

```
1
2
3 summarize `var'Estimate
4
5 local `var'GeoMean=exp(`r(mean)')
6
7 display "Exponentiated mean of log HR of variable is: " `var'GeoMean'
8
9 histogram `var'Estimate
10
11 generate HR`var'Estimate=exp(`var'Estimate)
12
13 generate Lower`var'=exp(`var'Estimate+(invnorm(0.025))*`var'SE)
14
15 generate Upper`var'=exp(`var'Estimate+(invnorm(0.975))*`var'SE)
16
17 display ". summarize HR`var'"
18
19 summarize HR`var'Estimate
20
21 histogram HR`var'Estimate, xscale(log) frequency
22
23 graph export "Model2HR`var'Histogram $DateOfAnalysis.tif", replace
24
25 twoway (scatter HR`var'Estimate SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of `var' in Model 2")
26 xtitle(Simulation Number)) (rcap Upper`var' Lower`var' SimNumber, vertical), yline(`${var'HazardRatio}) yline(`${var'GeoMean',
27 lpat(dash)) legend(off)
28
29 graph export "Model2HR`var'ForestPlot $DateOfAnalysis.tif", replace
30
31 drop SimNumber
32
33 }
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
```

```
use "Data Model3 $DateOfAnalysis.dta", clear
display "Model3:"
foreach var in ViralLoadUnder1000 {
  sort `var'Estimate
  generate SimNumber=_n
  display ". summarize `var'Estimate"
  summarize `var'Estimate
  local `var'GeoMean=exp(`r(mean)')
  display "Exponentiated mean of log HR of variable is: " `var'GeoMean'
  histogram `var'Estimate
  generate HR`var'Estimate=exp(`var'Estimate)
  generate Lower`var'=exp(`var'Estimate+(invnorm(0.025))*`var'SE)
  generate Upper`var'=exp(`var'Estimate+(invnorm(0.975))*`var'SE)
  display ". summarize HR`var'"
  summarize HR`var'Estimate
  histogram HR`var'Estimate, xscale(log) frequency
  graph export "Model3HR`var'Histogram $DateOfAnalysis.tif", replace
```

```
1
2
3 twoway (scatter HR`var'Estimate SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of `var' in Model 3")
4 xtitle(Simulation Number))(rcap Upper`var' Lower`var' SimNumber, vertical), yline(`${var'HazardRatio}) yline(`${var'GeoMean',
5 lpat(dash)) legend(off)
6
7 graph export "Model3HR`var'ForestPlot $DateOfAnalysis.tif", replace
8
9 drop SimNumber
10
11 }
12
13 use "Data Model4 $DateOfAnalysis.dta", clear
14
15 display "Model4:"
16
17 foreach var in BactrimReceipt {
18
19 sort `var'Estimate
20
21 generate SimNumber=_n
22
23 display ". summarize `var'Estimate"
24
25 summarize `var'Estimate
26
27 local `var'GeoMean=exp(`r(mean)')
28
29 display "Exponentiated mean of log HR of variable is: " `var'GeoMean'
30
31 histogram `var'Estimate
32
33 generate HR`var'Estimate=exp(`var'Estimate)
34
35 generate Lower`var'=exp(`var'Estimate+(invnorm(0.025))*`var'SE)
36
37 generate Upper`var'=exp(`var'Estimate+(invnorm(0.975))*`var'SE)
38
39 display ". summarize HR`var'"
40
41 summarize HR`var'Estimate
42
43 histogram HR`var'Estimate, xscale(log) frequency
44
45 graph export "Model4HR`var'Histogram $DateOfAnalysis.tif", replace
46
47 twoway (scatter HR`var'Estimate SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of `var' in Model 4")
48 xtitle(Simulation Number))(rcap Upper`var' Lower`var' SimNumber, vertical), yline(`${var'HazardRatio}) yline(`${var'GeoMean',
49 lpat(dash)) legend(off)
50
51 graph export "Model4HR`var'ForestPlot $DateOfAnalysis.tif", replace
52
53 drop SimNumber
54
55 }
56
57 use "Data Model5 $DateOfAnalysis.dta", clear
58
59 display "Model5:"
60
61 foreach var in BactrimPrescription {
62
63 sort `var'Estimate
64
65 generate SimNumber=_n
66
67 display ". summarize `var'Estimate"
68
69 summarize `var'Estimate
70
71 local `var'GeoMean=exp(`r(mean)')
```

```

1
2
3 display "Exponentiated mean of log HR of variable is: " `var'GeoMean'
4
5 histogram `var'Estimate
6
7 generate HR`var'Estimate=exp(`var'Estimate)
8
9 generate Lower`var'=exp(`var'Estimate+(invnorm(0.025))*`var'SE)
10
11 generate Upper`var'=exp(`var'Estimate+(invnorm(0.975))*`var'SE)
12
13 display ". summarize HR`var'"
14
15 summarize HR`var'Estimate
16
17 histogram HR`var'Estimate, xscale(log) frequency
18
19 graph export "Model5HR`var'Histogram $DateOfAnalysis.tif", replace
20
21 twoway (scatter HR`var'Estimate SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of `var' in Model 5")
22 xtitle(Simulation Number)) (rcap Upper`var' Lower`var' SimNumber, vertical), yline(`${var'HazardRatio}) yline(`${var'GeoMean',
23 lpat(dash)) legend(off)
24
25 graph export "Model5HR`var'ForestPlot $DateOfAnalysis.tif", replace
26
27 drop SimNumber
28
29 }
30
31 use "Data Model6 $DateOfAnalysis.dta", clear
32
33 display "Model6:"
34
35 foreach var in CD4Under200 ViralLoadUnder1000 {
36
37 sort `var'Estimate
38
39 generate SimNumber=_n
40
41 display ". summarize `var'Estimate"
42
43 summarize `var'Estimate
44
45 local `var'GeoMean=exp(`r(mean)')
46
47 display "Exponentiated mean of log HR of variable is: " `var'GeoMean'
48
49 histogram `var'Estimate
50
51 generate HR`var'Estimate=exp(`var'Estimate)
52
53 generate Lower`var'=exp(`var'Estimate+(invnorm(0.025))*`var'SE)
54
55 generate Upper`var'=exp(`var'Estimate+(invnorm(0.975))*`var'SE)
56
57 display ". summarize HR`var'"
58
59 summarize HR`var'Estimate
60
61 histogram HR`var'Estimate, xscale(log) frequency
62
63 graph export "Model6HR`var'Histogram $DateOfAnalysis.tif", replace
64
65 twoway (scatter HR`var'Estimate SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of `var' in Model 6")
66 xtitle(Simulation Number)) (rcap Upper`var' Lower`var' SimNumber, vertical), yline(`${var'HazardRatio}) yline(`${var'GeoMean',
67 lpat(dash)) legend(off)
68
69 graph export "Model6HR`var'ForestPlot $DateOfAnalysis.tif", replace
70
71 drop SimNumber
72
73 }

```

```
1
2
3
4
5
6 use "Data Model7 $DateOfAnalysis.dta", clear
7
8 display "Model7:"
9
10 foreach var in CD4Under200 ViralLoadUnder1000 BactrimPrescription {
11
12     sort `var'Estimate
13
14     generate SimNumber=_n
15
16     display ". summarize `var'Estimate"
17
18     summarize `var'Estimate
19
20     local `var'GeoMean=exp(`r(mean)')
21
22     display "Exponentiated mean of log HR of variable is: " `var'GeoMean'
23
24     histogram `var'Estimate
25
26     generate HR`var'Estimate=exp(`var'Estimate)
27
28     generate Lower`var'=exp(`var'Estimate+(invnorm(0.025))*`var'SE)
29
30     generate Upper`var'=exp(`var'Estimate+(invnorm(0.975))*`var'SE)
31
32     display ". summarize HR`var'"
33
34     summarize HR`var'Estimate
35
36     histogram HR`var'Estimate, xscale(log) frequency
37
38     graph export "Model7HR`var'Histogram $DateOfAnalysis.tif", replace
39
40     twoway (scatter HR`var'Estimate SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of `var' in Model 7")
41     xtitle(Simulation Number)) (rcap Upper`var' Lower`var' SimNumber, vertical), yline(`${var'HazardRatio}) yline(`${var'GeoMean',
42     lpat(dash)) legend(off)
43
44     graph export "Model7HR`var'ForestPlot $DateOfAnalysis.tif", replace
45
46     drop SimNumber
47
48 }
49
50 use "Data Model8 $DateOfAnalysis.dta", clear
51
52 display "Model8:"
53
54 foreach var in CD4Under200 BactrimReceipt {
55
56     sort `var'Estimate
57
58     generate SimNumber=_n
59
60     display ". summarize `var'Estimate"
61
62     summarize `var'Estimate
63
64     local `var'GeoMean=exp(`r(mean)')
65
66     display "Exponentiated mean of log HR of variable is: " `var'GeoMean'
67
68     histogram `var'Estimate
69
70     generate HR`var'Estimate=exp(`var'Estimate)
71
72     generate Lower`var'=exp(`var'Estimate+(invnorm(0.025))*`var'SE)
```

```

1
2
3 generate Upper`var`=exp(`var`Estimate+(invnorm(0.975))*`var`SE)
4
5 display ". summarize HR`var`"
6
7 summarize HR`var`Estimate
8
9 histogram HR`var`Estimate, xscale(log) frequency
10
11 graph export "Model8HR`var`Histogram $DateOfAnalysis.tif", replace
12
13 twoway (scatter HR`var`Estimate SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of `var` in Model 8")
14 xtitle(Simulation Number))(rcap Upper`var` Lower`var` SimNumber, vertical), yline(`${var`HazardRatio}) yline(`var`GeoMean',
15 lpat(dash)) legend(off)
16
17 graph export "Model8HR`var`ForestPlot $DateOfAnalysis.tif", replace
18
19 drop SimNumber
20
21 }
22
23 use "Data Model9 $DateOfAnalysis.dta", clear
24
25 display "Model9:"
26
27 foreach var in CD4Under200 BactrimPrescription {
28
29 sort `var`Estimate
30
31 generate SimNumber=_n
32
33 display ". summarize `var`Estimate"
34
35 summarize `var`Estimate
36
37 local `var`GeoMean=exp(`r(mean)`)
38
39 display "Exponentiated mean of log HR of variable is: " `var`GeoMean'
40
41 histogram `var`Estimate
42
43 generate HR`var`Estimate=exp(`var`Estimate)
44
45 generate Lower`var`=exp(`var`Estimate+(invnorm(0.025))*`var`SE)
46
47 generate Upper`var`=exp(`var`Estimate+(invnorm(0.975))*`var`SE)
48
49 display ". summarize HR`var`"
50
51 summarize HR`var`Estimate
52
53 histogram HR`var`Estimate, xscale(log) frequency
54
55 graph export "Model9HR`var`Histogram $DateOfAnalysis.tif", replace
56
57 twoway (scatter HR`var`Estimate SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of `var` in Model 9")
58 xtitle(Simulation Number))(rcap Upper`var` Lower`var` SimNumber, vertical), yline(`${var`HazardRatio}) yline(`var`GeoMean',
59 lpat(dash)) legend(off)
60
61 graph export "Model9HR`var`ForestPlot $DateOfAnalysis.tif", replace
62
63 drop SimNumber
64
65 }
66
67 use "Data Model10 $DateOfAnalysis.dta", clear
68
69 display "Model10:"
70
71 foreach var in ViralLoadUnder1000 BactrimReceipt {

```



```

1
2
3     sort `var'Estimate
4
5     generate SimNumber=_n
6
7     display ". summarize `var'Estimate"
8
9     summarize `var'Estimate
10
11    local `var'GeoMean=exp(`r(mean)')
12
13    display "Exponentiated mean of log HR of variable is: " `var'GeoMean'
14
15    histogram `var'Estimate
16
17    generate HR`var'Estimate=exp(`var'Estimate)
18
19    generate Lower`var'=exp(`var'Estimate+(invnorm(0.025))*`var'SE)
20
21    generate Upper`var'=exp(`var'Estimate+(invnorm(0.975))*`var'SE)
22
23    display ". summarize HR`var'"
24
25    summarize HR`var'Estimate
26
27    histogram HR`var'Estimate, xscale(log) frequency
28
29    graph export "Modell0HR`var'Histogram $DateOfAnalysis.tif", replace
30
31    twoway (scatter HR`var'Estimate SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of `var' in Model
32    10") xtitle(Simulation Number)) (rcap Upper`var' Lower`var' SimNumber, vertical, yline(`${var'HazardRatio}) yline(`var'GeoMean',
33    lpat(dash)) legend(off)
34
35    graph export "Modell0HR`var'ForestPlot $DateOfAnalysis.tif", replace
36
37    drop SimNumber
38
39    }
40
41
42
43
44    use "Data Modell1 $DateOfAnalysis.dta", clear
45
46    display "Modell1:"
47
48    foreach var in ViralLoadUnder1000 BactrimPrescription {
49
50    sort `var'Estimate
51
52    generate SimNumber=_n
53
54    display ". summarize `var'Estimate"
55
56    summarize `var'Estimate
57
58    local `var'GeoMean=exp(`r(mean)')
59
60    display "Exponentiated mean of log HR of variable is: " `var'GeoMean'
61
62    histogram `var'Estimate
63
64    generate HR`var'Estimate=exp(`var'Estimate)
65
66    generate Lower`var'=exp(`var'Estimate+(invnorm(0.025))*`var'SE)
67
68    generate Upper`var'=exp(`var'Estimate+(invnorm(0.975))*`var'SE)
69
70    display ". summarize HR`var'"
71
72    summarize HR`var'Estimate
73
74    histogram HR`var'Estimate, xscale(log) frequency
75
76    graph export "Modell1HR`var'Histogram $DateOfAnalysis.tif", replace

```

```

1
2
3 twoway (scatter HR`var'Estimate SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of `var' in Model
4 11") xtitle(Simulation Number))(rcap Upper`var' Lower`var' SimNumber, vertical), yline(`${var'HazardRatio}) yline(`${var'GeoMean',
5 lpat(dash)) legend(off)
6
7 graph export "Modell1HR`var'ForestPlot $DateOfAnalysis.tif", replace
8
9 drop SimNumber
10
11 }
12
13 use "Data Modell12 $DateOfAnalysis.dta", clear
14
15 display "Modell12:"
16
17 foreach var in CD4Under200 ViralLoadUnder1000 BactrimPrescription {
18
19 sort `var'Est
20
21 generate SimNumber=_n
22
23 display ". summarize `var'Est"
24
25 summarize `var'Est
26
27 local `var'GeoMean=exp(`r(mean)')
28
29 display "Exponentiated mean of log HR of variable is: " `var'GeoMean'
30
31 histogram `var'Est
32
33 generate HR`var'Est=exp(`var'Est)
34
35 generate Lower`var'=exp(`var'Est+(invnorm(0.025))*`var'SE)
36
37 generate Upper`var'=exp(`var'Est+(invnorm(0.975))*`var'SE)
38
39 display ". summarize HR`var'"
40
41 summarize HR`var'Est
42
43 histogram HR`var'Est, xscale(log) frequency
44
45 graph export "Modell12HR`var'Histogram $DateOfAnalysis.tif", replace
46
47 twoway (scatter HR`var'Est SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of `var' in Model 12")
48 xtitle(Simulation Number))(rcap Upper`var' Lower`var' SimNumber, vertical), yline(`${var'HazardRatio}) yline(`${var'GeoMean',
49 lpat(dash)) legend(off)
50
51 graph export "Modell12HR`var'ForestPlot $DateOfAnalysis.tif", replace
52
53 drop SimNumber
54
55 }
56
57 sort ViralLoadBactrimRxLinComEst
58
59 generate SimNumber=_n
60
61 display ". summarize ViralLoadBactrimRxLinComEst"
62
63 summarize ViralLoadBactrimRxLinComEst
64
65 local ViralLoadBactrimRxLinComGeoMean=exp(`r(mean)')
66
67 display "Exponentiated mean of log HR of variable is: " `ViralLoadBactrimRxLinComGeoMean'
68
69 histogram ViralLoadBactrimRxLinComEst
70
71 generate HRViralLoadBactrimRxLinComEst=exp( ViralLoadBactrimRxLinComEst)

```

```

1
2
3 generate LowerViralLoadBactrimRxLinCom=exp(ViralLoadBactrimRxLinComEst+(invnorm(0.025))*ViralLoadBactrimRxLinComSE)
4 generate UpperViralLoadBactrimRxLinCom=exp(ViralLoadBactrimRxLinComEst+(invnorm(0.975))*ViralLoadBactrimRxLinComSE)
5
6
7
8 generate VLBactrimRxLinComEstZscore= ViralLoadBactrimRxLinComEst/ ViralLoadBactrimRxLinComSE
9
10 summarize VLBactrimRxLinComEstZscore
11
12 generate VLBactrimRxLinComEstSigPvalue=0
13
14 replace VLBactrimRxLinComEstSigPvalue=1 if (VLBactrimRxLinComEstZscore<invnorm(0.025))
15
16 replace VLBactrimRxLinComEstSigPvalue=1 if (VLBactrimRxLinComEstZscore>invnorm(0.975))
17
18 tab VLBactrimRxLinComEstSigPvalue
19
20
21 display ". summarize HRViralLoadBactrimRxLinCom"
22
23 summarize HRViralLoadBactrimRxLinComEst
24
25 histogram HRViralLoadBactrimRxLinComEst, xscale(log) frequency
26
27 graph export "Modell2HRViralLoadBactrimRxLinComHistogram $DateOfAnalysis.tif", replace
28
29 twoway (scatter HRViralLoadBactrimRxLinComEst SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of
30 ViralLoadBactrimRxLinCom in Model 12") xtitle(Simulation Number))(rcap UpperViralLoadBactrimRxLinCom
31 LowerViralLoadBactrimRxLinCom SimNumber, vertical), yline($BactrimReceiptHazardRatio) yline(`ViralLoadBactrimRxLinComGeoMean',
32 lpat(dash)) legend(off)
33
34 graph export "Modell2HRViralLoadBactrimRxLinComForestPlot $DateOfAnalysis.tif", replace
35
36 drop SimNumber
37
38
39 //We now see in what proportion of the simulations is the P-value for the interaction term significant:
40
41 generate VLBactrimRXInteractionZscore= ViralLoadBactrimRXInteractionEst/ ViralLoadBactrimRXInteractionSE
42
43 summarize VLBactrimRXInteractionZscore
44
45 generate VLBactrimRXInteractionSigPvalue=0
46
47 replace VLBactrimRXInteractionSigPvalue=1 if (VLBactrimRXInteractionZscore<invnorm(0.025))
48
49 replace VLBactrimRXInteractionSigPvalue=1 if (VLBactrimRXInteractionZscore>invnorm(0.975))
50
51 tab VLBactrimRXInteractionSigPvalue
52
53
54 //We now test the bias in the alternative strategy of only keeping the interaction term if significant:
55
56 generate AlternateVLBactrimEst=(VLBactrimRXInteractionSigPvalue)*(ViralLoadBactrimRxLinComEst)+(1-
57 VLBactrimRXInteractionSigPvalue)*(BactrimPrescriptionEst)
58
59 sort AlternateVLBactrimEst
60
61 generate SimNumber=_n
62
63 generate HRAAlternateVLBactrimEst=exp(AlternateVLBactrimEst)
64
65 generate AlternateVLBactrimSE=(VLBactrimRXInteractionSigPvalue)*(ViralLoadBactrimRxLinComSE)+(1-
66 VLBactrimRXInteractionSigPvalue)*(BactrimPrescriptionSE)
67
68 generate LowerAlternateVLBactrim=exp(AlternateVLBactrimEst+(invnorm(0.025))*AlternateVLBactrimSE)

```

```

1
2
3 generate UpperAlternateVLBactrim=exp(AlternateVLBactrimEst+(invnorm(0.975))*AlternateVLBactrimSE)
4
5 summarize AlternateVLBactrimEst
6
7 local AlternateVLBactrimGeoMean=exp(`r(mean)`)
8
9 display "Exponentiated mean of log HR of variable is: " `AlternateVLBactrimGeoMean'
10
11 twoway (scatter HRAlternateVLBactrimEst SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of
12 AlternateVLBactrim in Model 12") xtitle(Simulation Number)) (rcap UpperAlternateVLBactrim LowerAlternateVLBactrim SimNumber,
13 vertical), yline($BactrimReceiptHazardRatio) yline('AlternateVLBactrimGeoMean', lpattern(dash)) legend(off)
14
15 graph export "Model12HRAlternateVLBactrimForestPlot $DateOfAnalysis.tif", replace
16
17 drop SimNumber
18
19
20 use "Data Model13 $DateOfAnalysis.dta", clear
21
22 display "Model13:"
23
24 foreach var in CD4Under200 ViralLoadUnder1000 BactrimReceipt {
25
26 sort `var'Est
27
28 generate SimNumber=_n
29
30 display ". summarize `var'Est"
31
32 summarize `var'Est
33
34 local `var'GeoMean=exp(`r(mean)`)
35
36 display "Exponentiated mean of log HR of variable is: " ``var'GeoMean'
37
38 histogram `var'Est
39
40 generate HR`var'Est=exp(`var'Est)
41
42 generate Lower`var'=exp(`var'Est+(invnorm(0.025))*`var'SE)
43
44 generate Upper`var'=exp(`var'Est+(invnorm(0.975))*`var'SE)
45
46 display ". summarize HR`var'"
47
48 summarize HR`var'Est
49
50 histogram HR`var'Est, xscale(log) frequency
51
52 graph export "Model13HR`var'Histogram $DateOfAnalysis.tif", replace
53
54 twoway (scatter HR`var'Est SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of `var' in Model 13")
55 xtitle(Simulation Number)) (rcap Upper`var' Lower`var' SimNumber, vertical), yline(`${var'HazardRatio}) yline(`var'GeoMean',
56 lpat(dash)) legend(off)
57
58 graph export "Model13HR`var'ForestPlot $DateOfAnalysis.tif", replace
59
60 drop SimNumber
61
62 }
63
64 display ". summarize ViralLoadBactrimRecLinComEst"
65
66 sort ViralLoadBactrimRecLinComEst
67
68 generate SimNumber=_n
69
70 summarize ViralLoadBactrimRecLinComEst

```

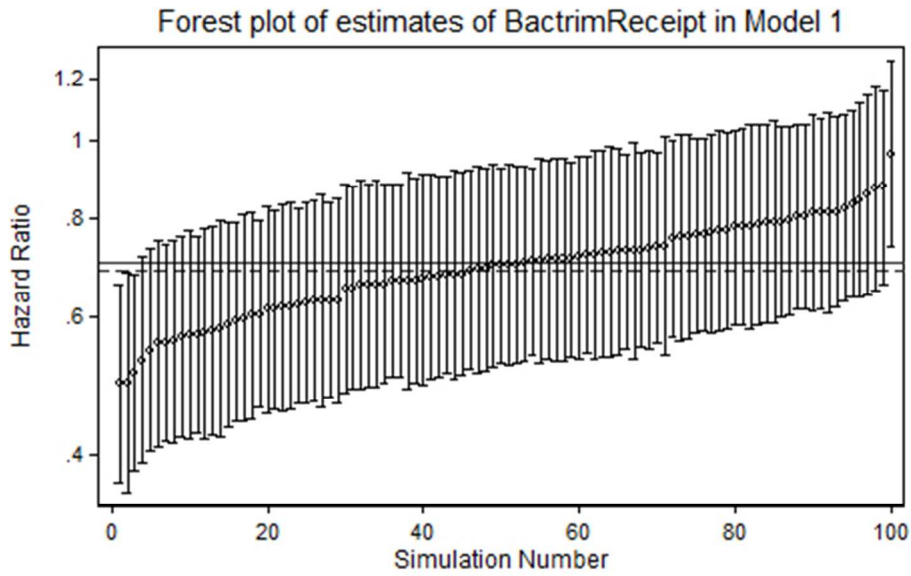
```

1
2
3 local VLBactrimRecLinComGeoMean=exp(`r(mean)`)
4
5 display "Exponentiated mean of log HR of variable is: " `VLBactrimRecLinComGeoMean'
6
7 histogram ViralLoadBactrimRecLinComEst
8
9 generate HRViralLoadBactrimRecLinComEst=exp( ViralLoadBactrimRecLinComEst)
10
11 generate LowerViralLoadBactrimRecLinComEst=exp(ViralLoadBactrimRecLinComEst+(invnorm(0.025))*ViralLoadBactrimRecLinComSE)
12
13 generate UpperViralLoadBactrimRecLinComEst=exp(ViralLoadBactrimRecLinComEst+(invnorm(0.975))*ViralLoadBactrimRecLinComSE)
14
15 generate VLBactrimRecLinComEstZscore= ViralLoadBactrimRecLinComEst/ ViralLoadBactrimRecLinComSE
16
17 summarize VLBactrimRecLinComEstZscore
18
19 generate VLBactrimRecLinComEstSigPvalue=0
20
21 replace VLBactrimRecLinComEstSigPvalue=1 if (VLBactrimRecLinComEstZscore<invnorm(0.025))
22
23 replace VLBactrimRecLinComEstSigPvalue=1 if (VLBactrimRecLinComEstZscore>invnorm(0.975))
24
25 tab VLBactrimRecLinComEstSigPvalue
26
27
28 display ". summarize HRViralLoadBactrimRecLinCom"
29
30 summarize HRViralLoadBactrimRecLinComEst
31
32 histogram HRViralLoadBactrimRecLinComEst, xscale(log) frequency
33
34 graph export "Modell3HRViralLoadBactrimRecLinComHistogram $DateOfAnalysis.tif", replace
35
36 twoway (scatter HRViralLoadBactrimRecLinComEst SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of
37 BactrimRecLinComEst in Model 13") xtitle(Simulation Number)) (rcap UpperViralLoadBactrimRecLinCom LowerViralLoadBactrimRecLinCom
38 SimNumber, vertical), yline($BactrimReceiptHazardRatio) yline(`VLBactrimRecLinComGeoMean', lpat(dash)) legend(off)
39
40 graph export "Modell3HRViralLoadBactrimRecLinComForestPlot $DateOfAnalysis.tif", replace
41
42 drop SimNumber
43
44
45 //We now see in what proportion of the simulations is the P-value for the interaction term significant:
46
47 generate VLBactrimRecInteractionZscore= ViralLoadBactrimRecInterxnEst/ ViralLoadBactrimRecInterxnSE
48
49 summarize VLBactrimRecInteractionZscore
50
51 generate VLBactrimRecInteractionSigPvalue=0
52
53 replace VLBactrimRecInteractionSigPvalue=1 if ( VLBactrimRecInteractionZscore<invnorm(0.025))
54
55 replace VLBactrimRecInteractionSigPvalue=1 if (VLBactrimRecInteractionZscore>invnorm(0.975))
56
57 tab VLBactrimRecInteractionSigPvalue
58
59
60 //We now test the bias in the alternative strategy of only keeping the interaction term if significant:
61
62 generate AlternateVLBactrimEst=( VLBactrimRecInteractionSigPvalue)*( ViralLoadBactrimRecLinComEst)+(1-
63 VLBactrimRecInteractionSigPvalue)*( BactrimReceiptEst)
64
65 sort AlternateVLBactrimEst
66
67 generate SimNumber=_n
68
69 generate HRAlternateVLBactrimEst=exp( AlternateVLBactrimEst)

```

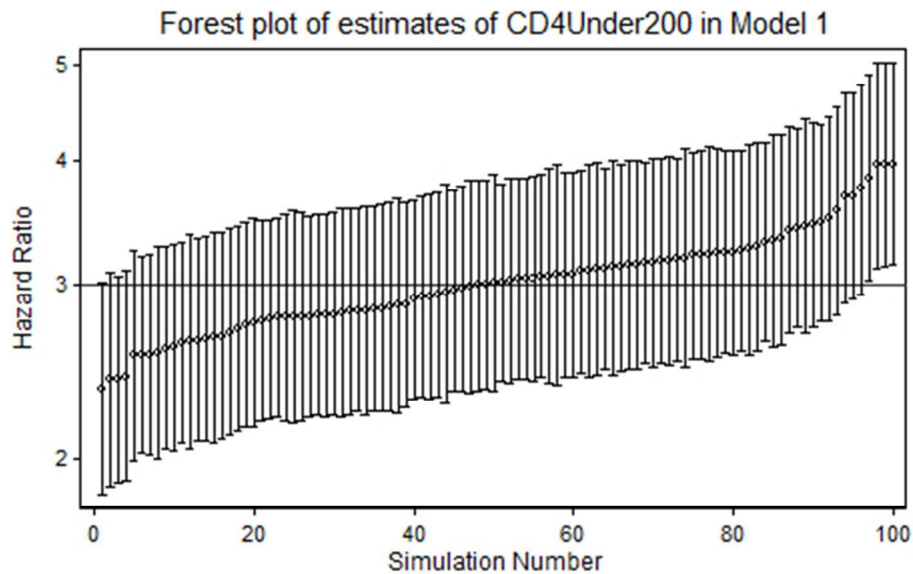
```
1
2
3 generate AlternateVLBactrimSE=( VLBactrimRecInteractionSigPvalue)*( ViralLoadBactrimRecLinComSE)+(1-
4 VLBactrimRecInteractionSigPvalue)*( BactrimReceiptSE)
5
6 generate LowerAlternateVLBactrim=exp(AlternateVLBactrimEst+(invnorm(0.025))*AlternateVLBactrimSE)
7
8 generate UpperAlternateVLBactrim=exp(AlternateVLBactrimEst+(invnorm(0.975))*AlternateVLBactrimSE)
9
10 summarize AlternateVLBactrimEst
11
12 local AlternateVLBactrimGeoMean=exp(`r(mean)')
13
14 display "Exponentiated mean of log HR of variable is: " `AlternateVLBactrimGeoMean'
15
16 twoway (scatter HRAlternateVLBactrimEst SimNumber, yscale(log) ytitle(Hazard Ratio) title("Forest plot of estimates of
17 AlternateVLBactrimEst in Model 13") xtitle(Simulation Number))(rcap UpperAlternateVLBactrim LowerAlternateVLBactrim SimNumber,
18 vertical), yline($BactrimReceiptHazardRatio) yline(`AlternateVLBactrimGeoMean', lpattern(dash)) legend(off)
19
20 graph export "Model13HRAAlternateVLBactrimForestPlot $DateOfAnalysis.tif", replace
21
22 drop SimNumber
23
24
25 use "Data NumFailures $DateOfAnalysis.dta", clear
26
27 summarize NumFailures
28
29 display "Average proportion with SSTI over five years: " `r(mean)'/2000
30
31 histogram NumFailures, frequency
32
33 graph export "NumFailuresHistogram $DateOfAnalysis.tif", replace
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with TMP-SMX use in each simulation under Model 1. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

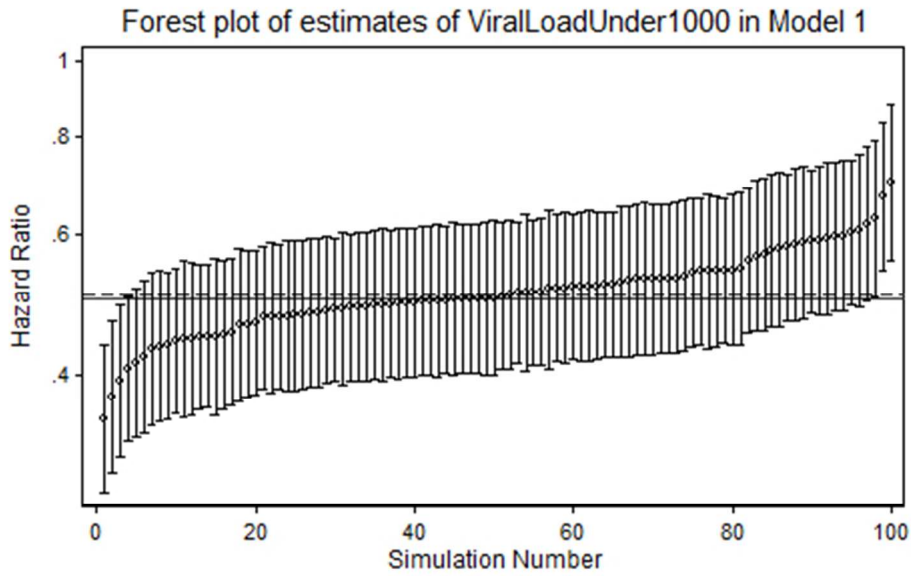


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with CD4 count under 200 cells/mL in each simulation under Model 1. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

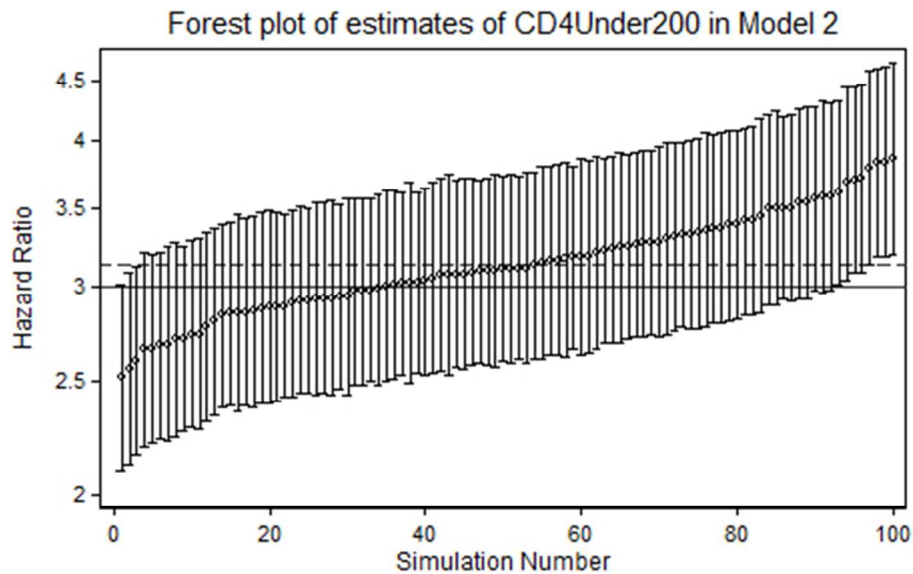
167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with viral load under 1000 copies/mL in each simulation under Model 1. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

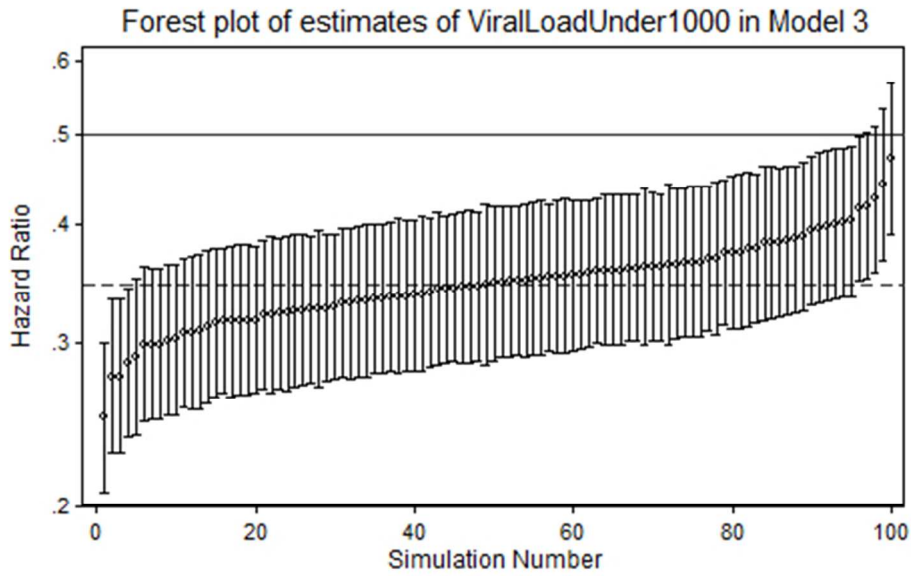
167x111mm (72 x 72 DPI)



28 Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with CD4 count
29 under 200 cells/mL in each simulation under Model 2. Estimates were sorted by size. The dashed line
30 represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the
31 solid line represents the true population value.

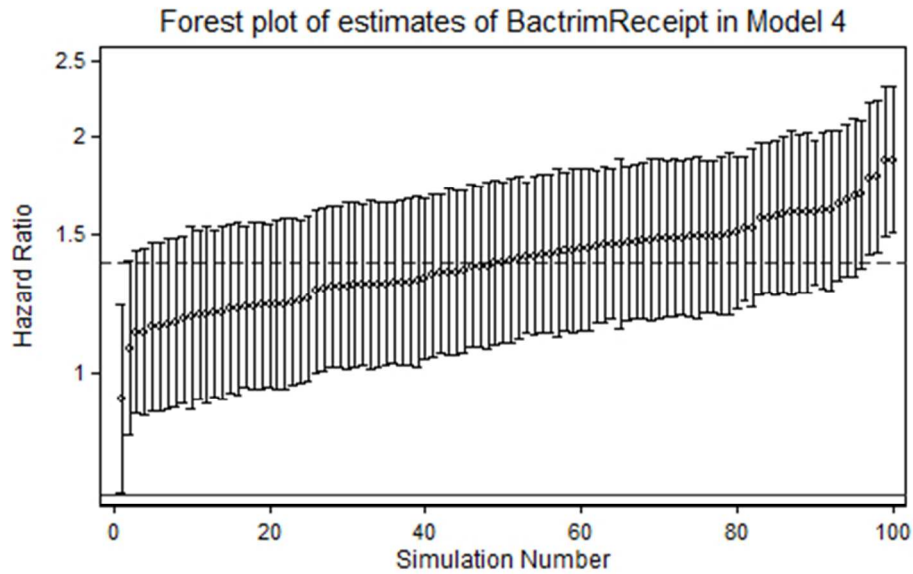
32 167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with viral load under 1000 copies/mL in each simulation under Model 3. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

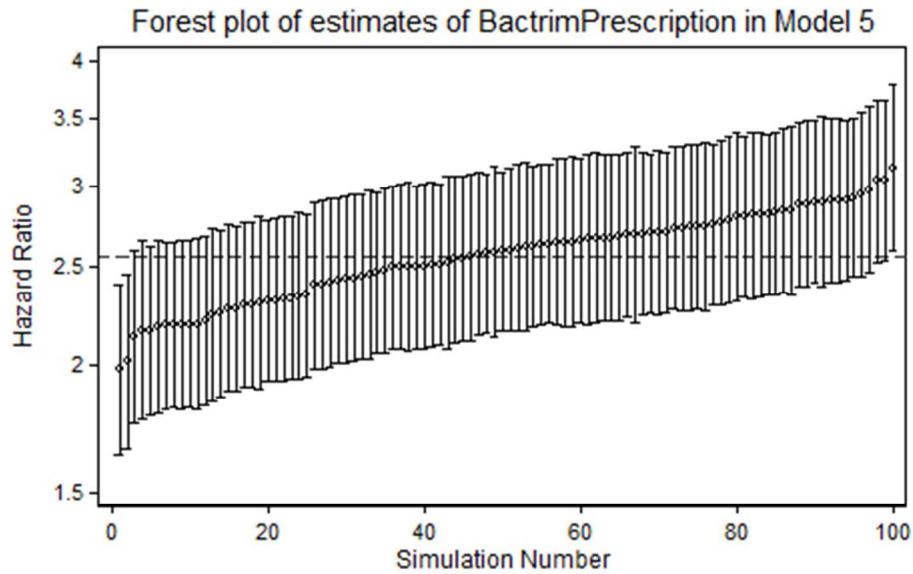


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with TMP-SMX use in each simulation under Model 4. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

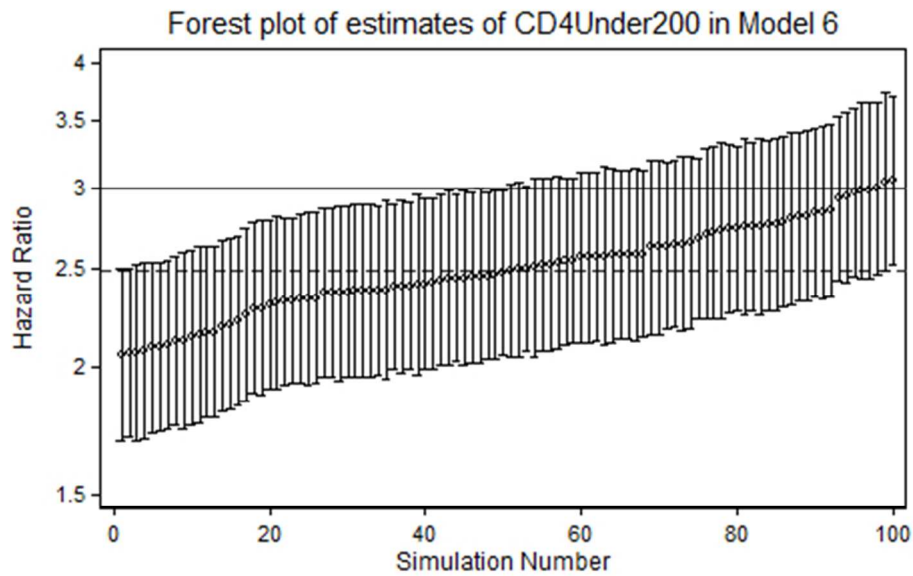
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



28 Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with TMP-SMX
29 prescription in each simulation under Model 5. Estimates were sorted by size. The dashed line represents the
30 geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line
31 represents the true population value.

32 167x111mm (72 x 72 DPI)

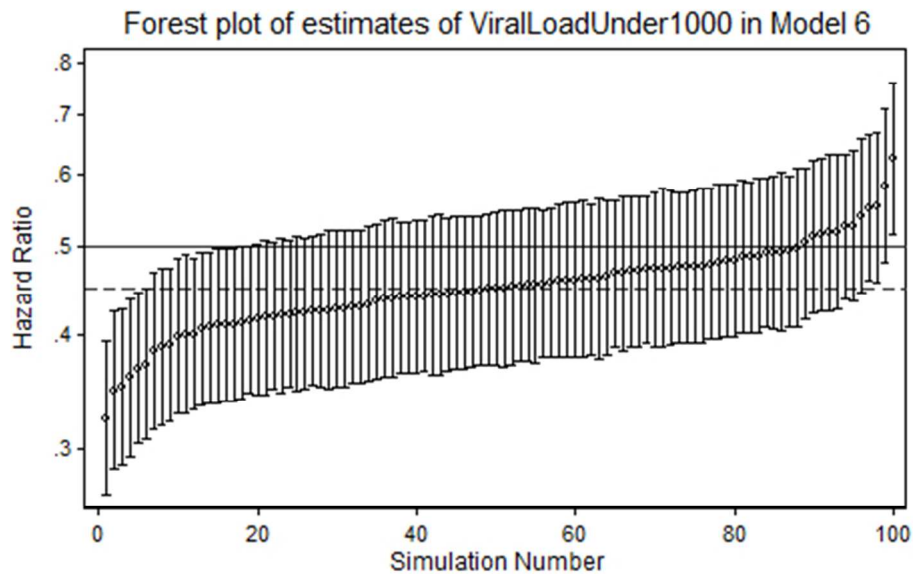
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



28 Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with CD4 count
29 under 200 cells/mL in each simulation under Model 6. Estimates were sorted by size. The dashed line
30 represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the
31 solid line represents the true population value.

32 167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

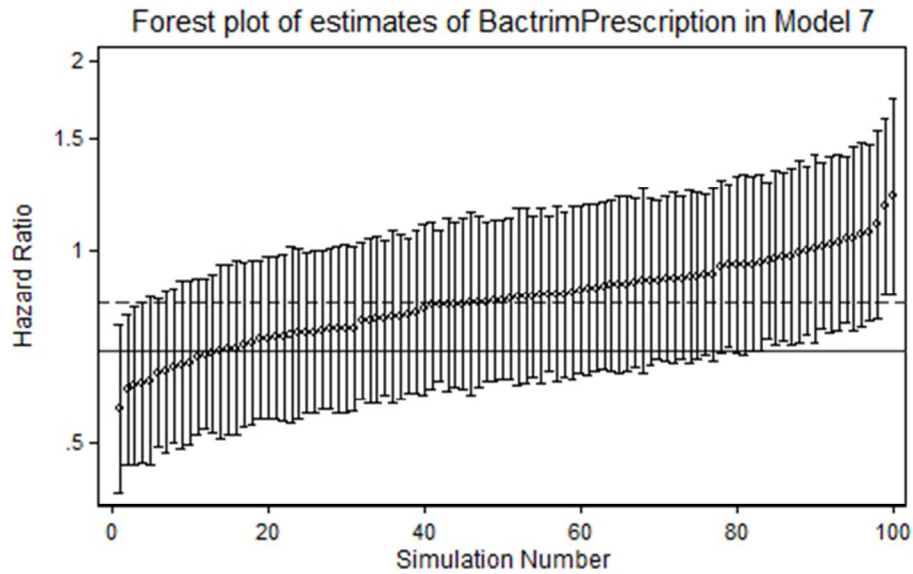


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with viral load under 1000 copies/mL in each simulation under Model 6. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

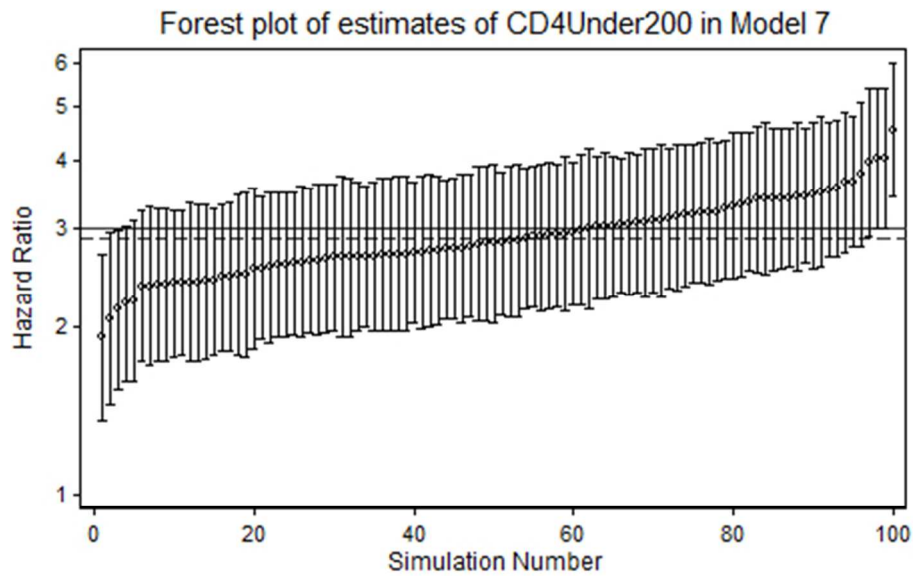


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with TMP-SMX prescription in each simulation under Model 7. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

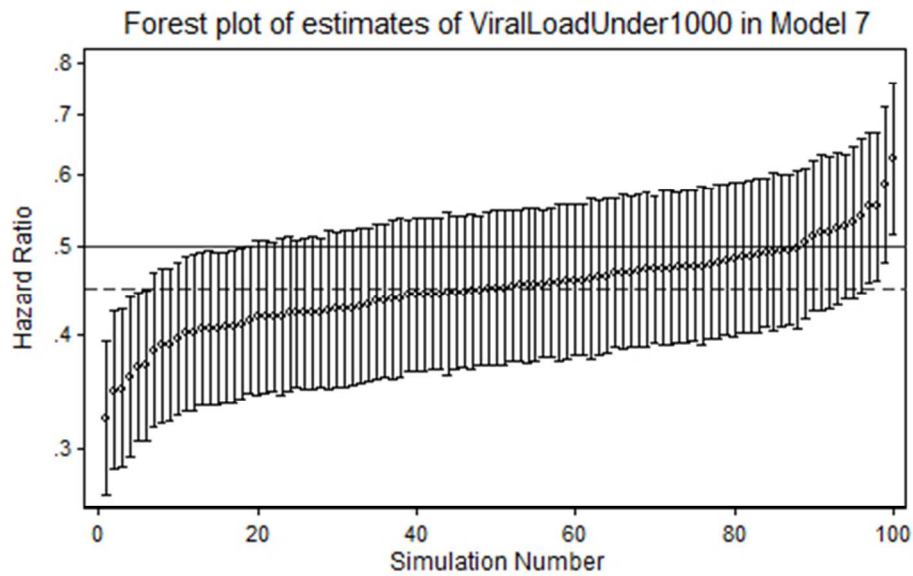
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



28 Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with CD4 count
29 under 200 cells/mL in each simulation under Model 7. Estimates were sorted by size. The dashed line
30 represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the
31 solid line represents the true population value.

32 167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

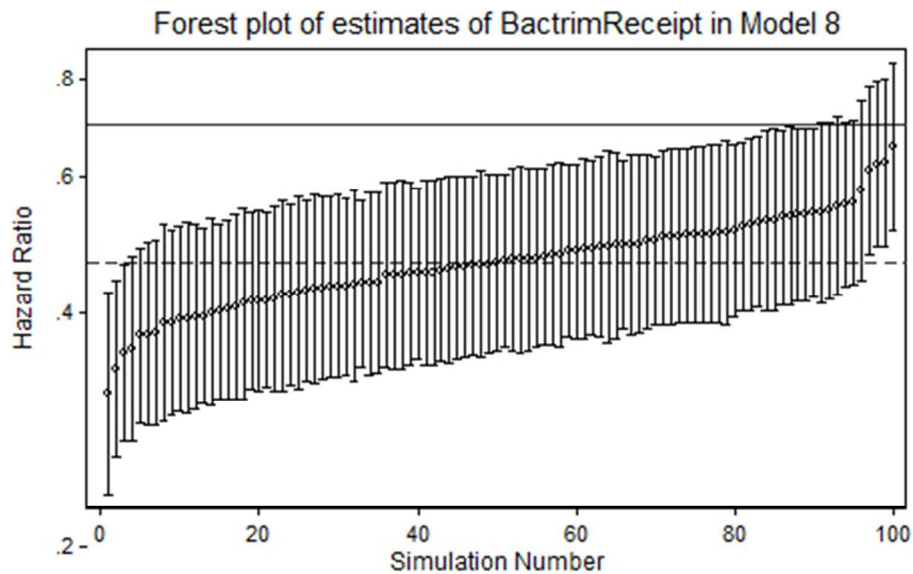


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with viral load under 1000 copies/mL in each simulation under Model 7. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

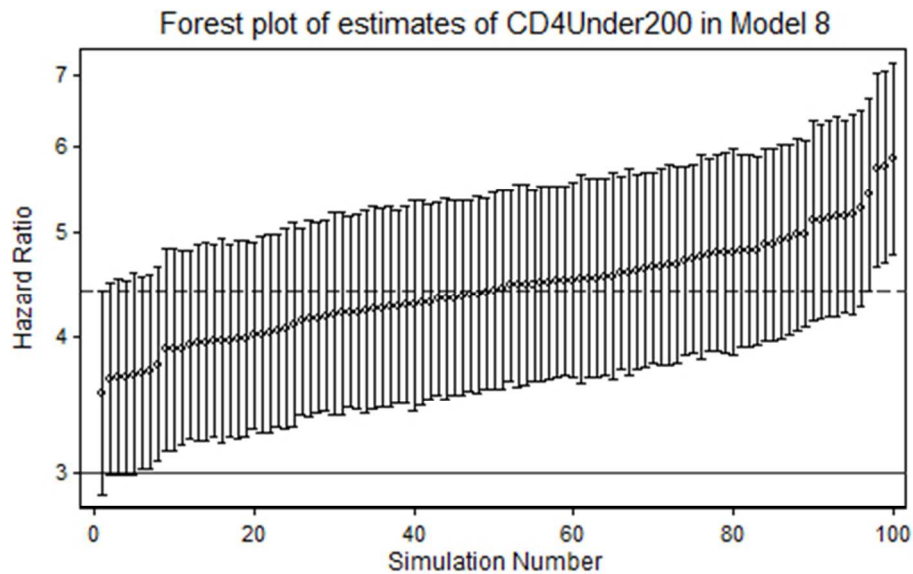
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



28 Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with TMP-SMX
29 use in each simulation under Model 8. Estimates were sorted by size. The dashed line represents the
30 geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line
31 represents the true population value.

32 167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

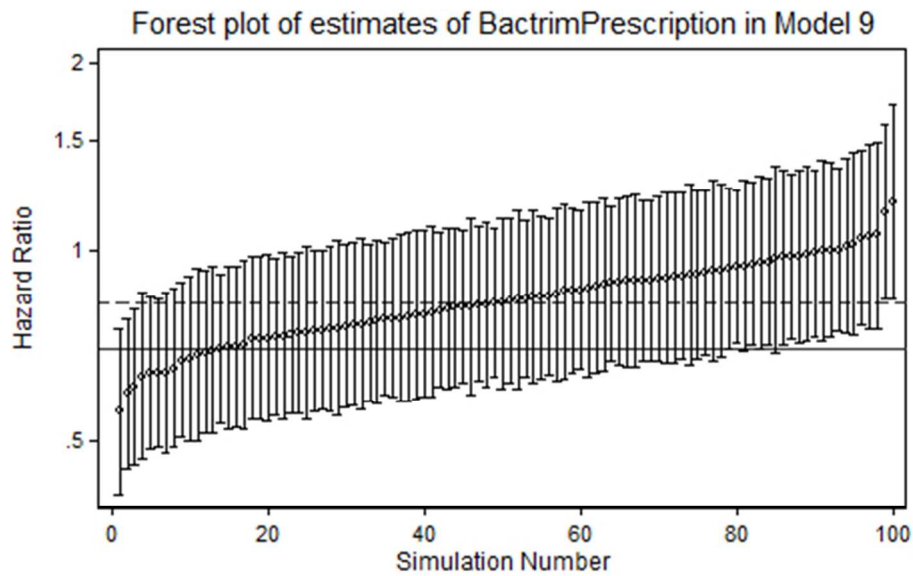


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with CD4 count under 200 cells/mL in each simulation under Model 8. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

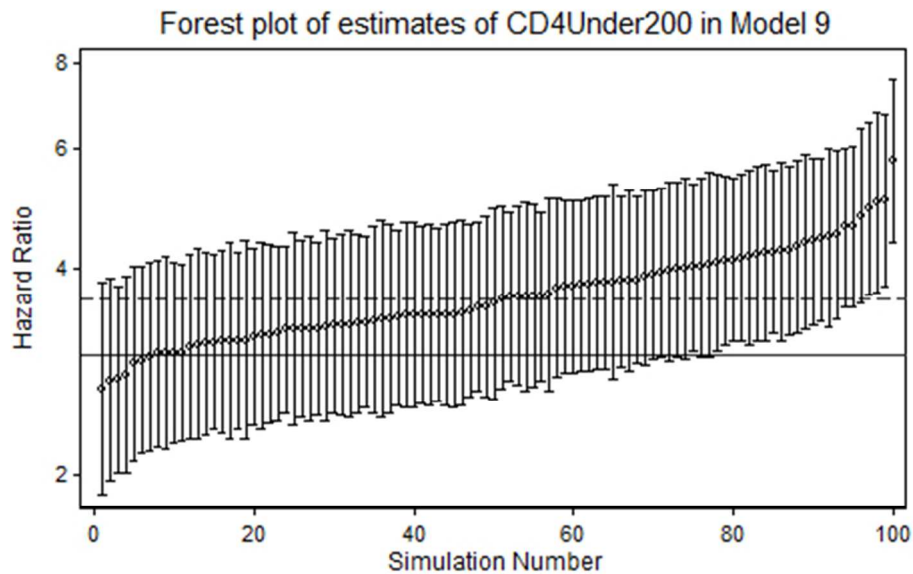


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with TMP-SMX prescription in each simulation under Model 9. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



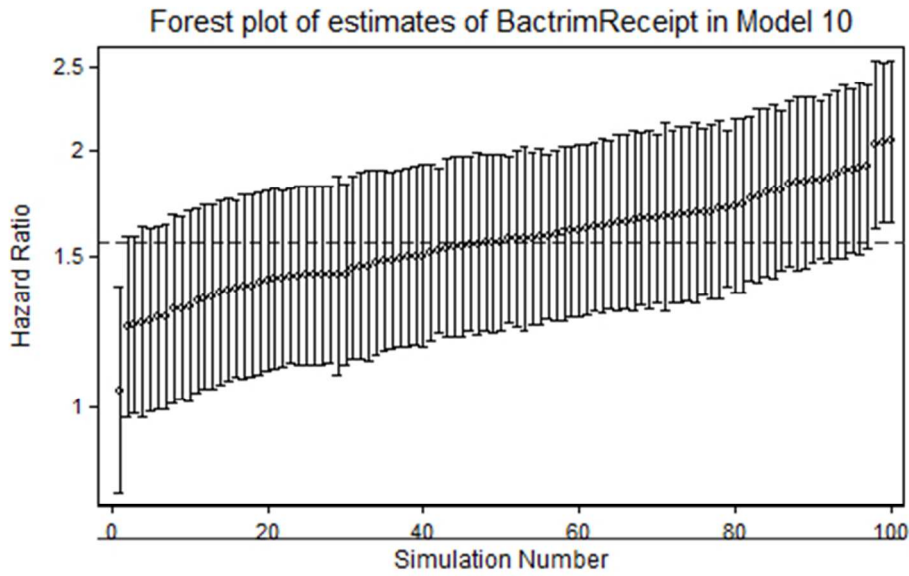
28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with CD4 count under 200 cells/mL in each simulation under Model 9. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

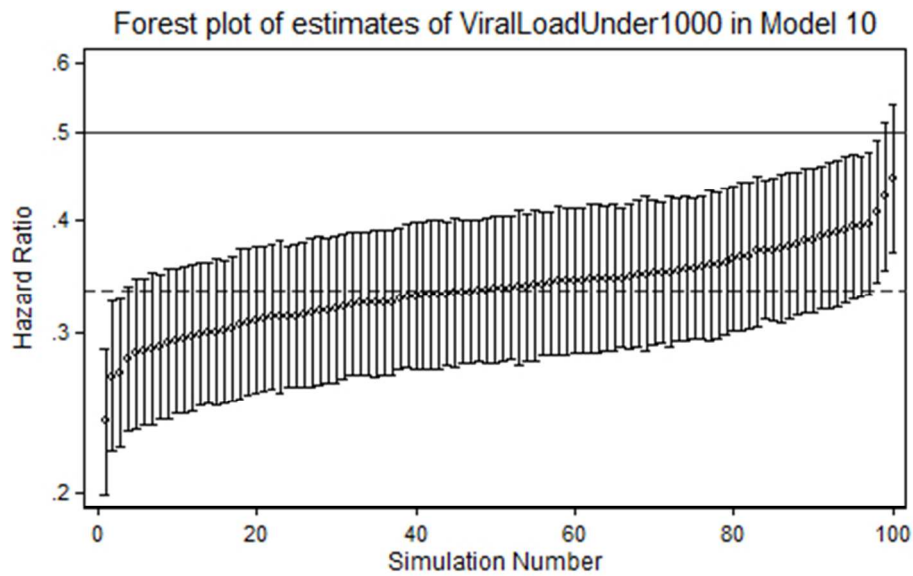
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with TMP-SMX use in each simulation under Model 10. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

View Only

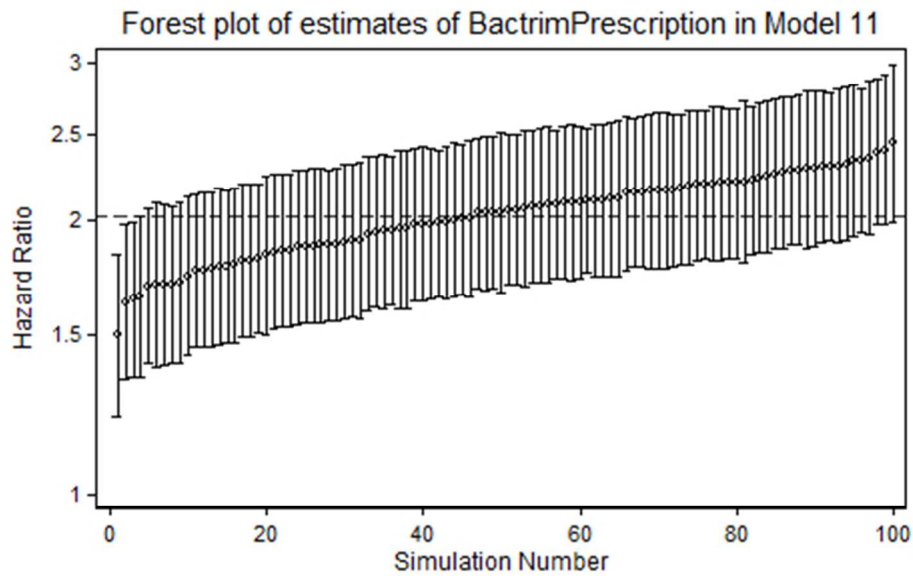


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with viral load under 1000 copies/mL in each simulation under Model 10. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

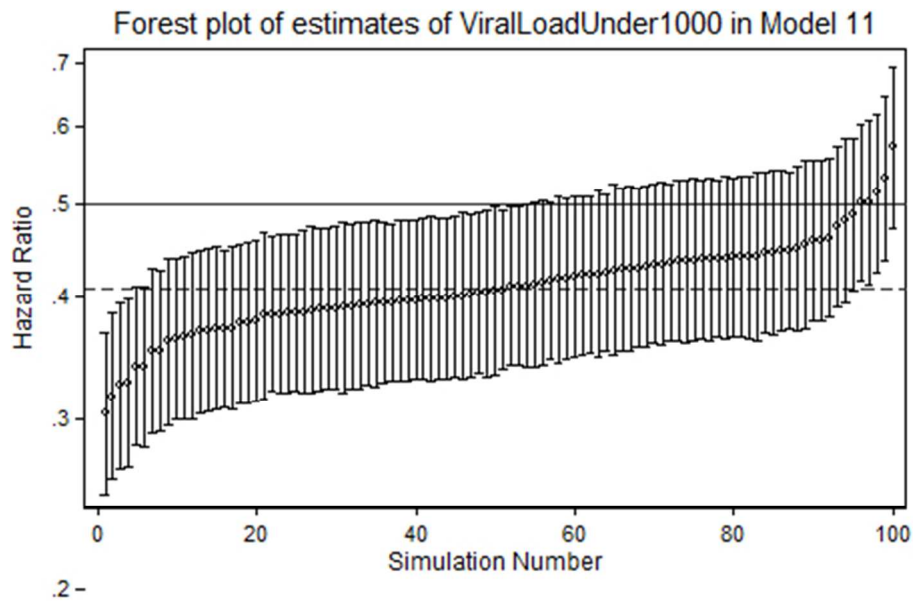
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



28 Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with TMP-SMX
29 prescription in each simulation under Model 11. Estimates were sorted by size. The dashed line represents
30 the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line
31 represents the true population value.

32 167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

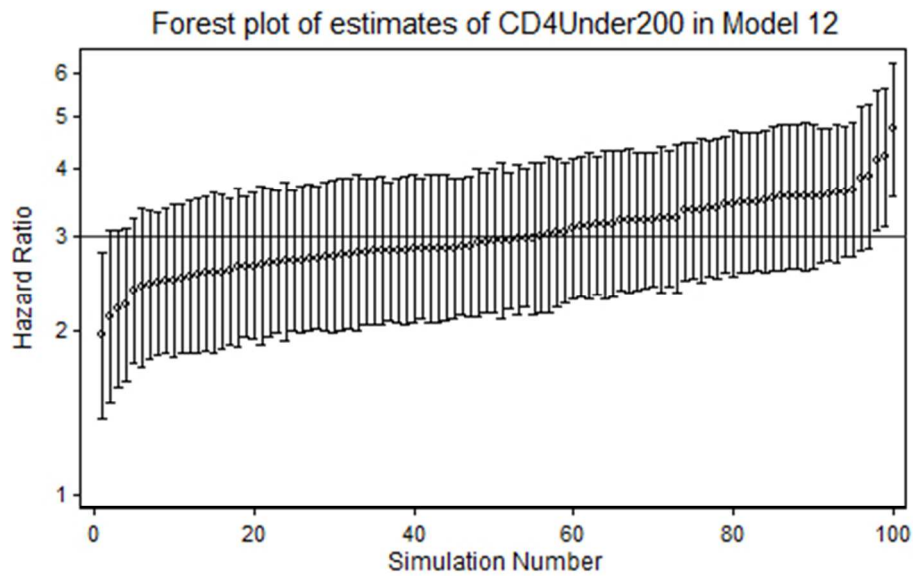


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with viral load under 1000 copies/mL in each simulation under Model 11. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

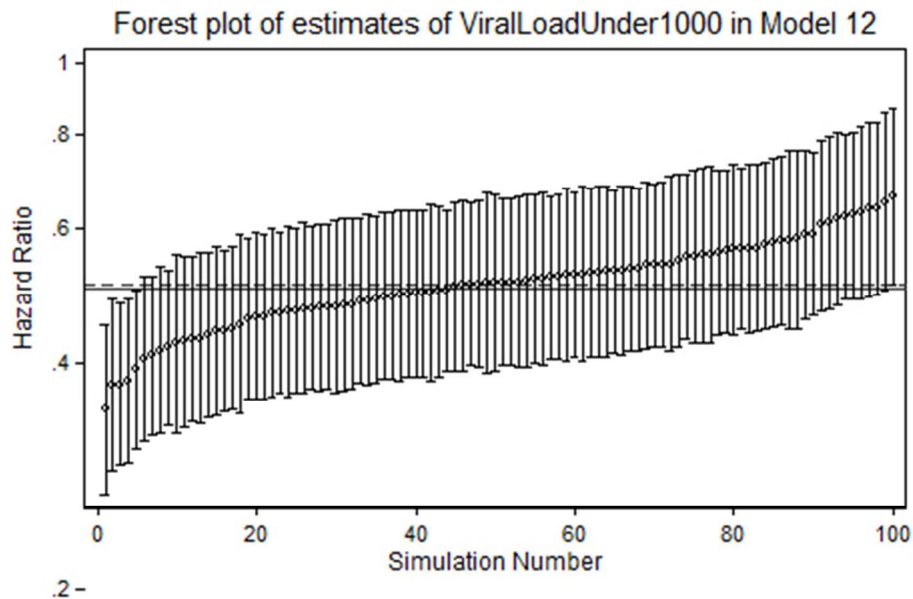


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with CD4 count under 200 cells/mL in each simulation under Model 12. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

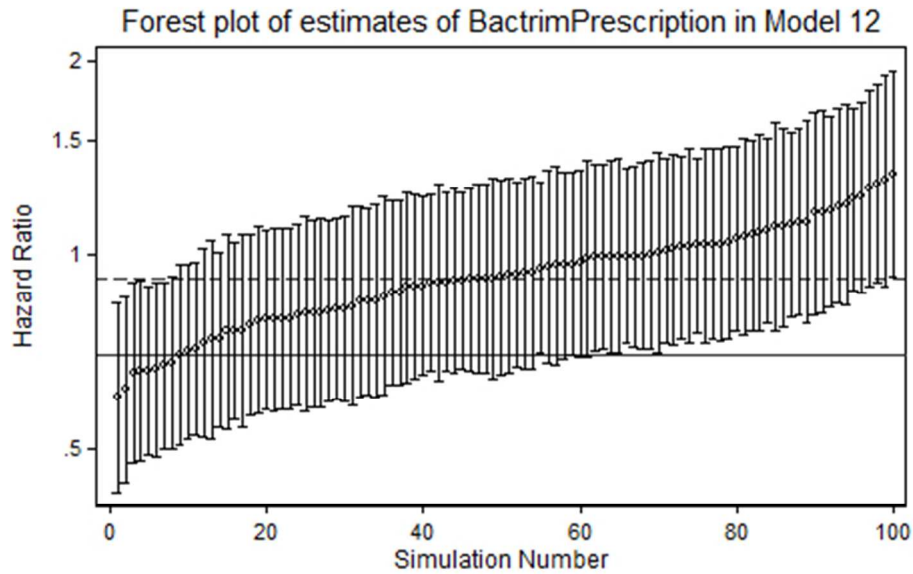


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with viral load under 1000 copies/mL for subjects not prescribed TMP-SMX in each simulation under Model 12. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

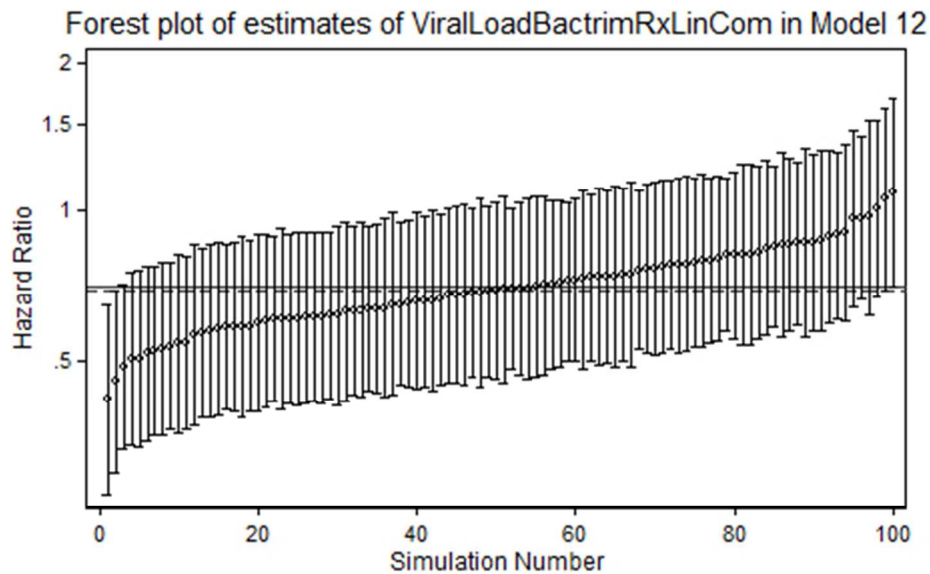


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with TMP-SMX prescription for subjects with high viral load in each simulation under Model 12. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

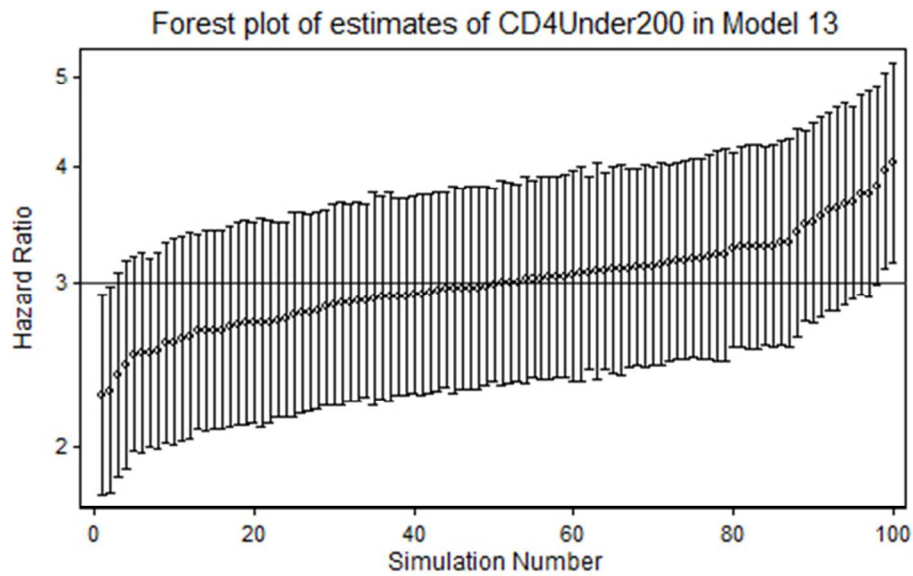


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with TMP-SMX prescription in subjects with viral load under 1000 copies/mL in each simulation under Model 12. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

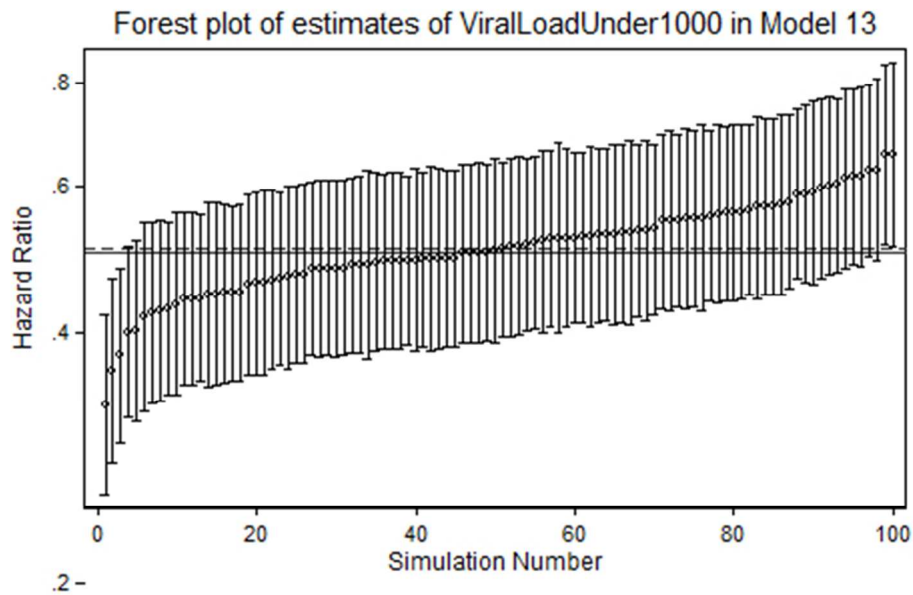
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



28 Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with CD4 count
29 under 200 cells/mL in each simulation under Model 13. Estimates were sorted by size. The dashed line
30 represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the
31 solid line represents the true population value.

32 167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

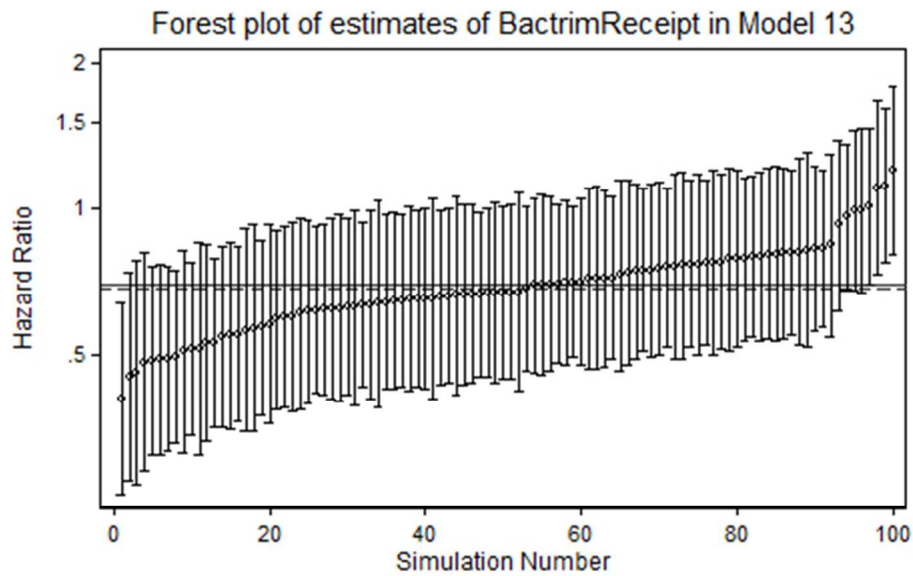


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with viral load under 1000 copies/mL for subjects not using TMP-SMX in each simulation under Model 13. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

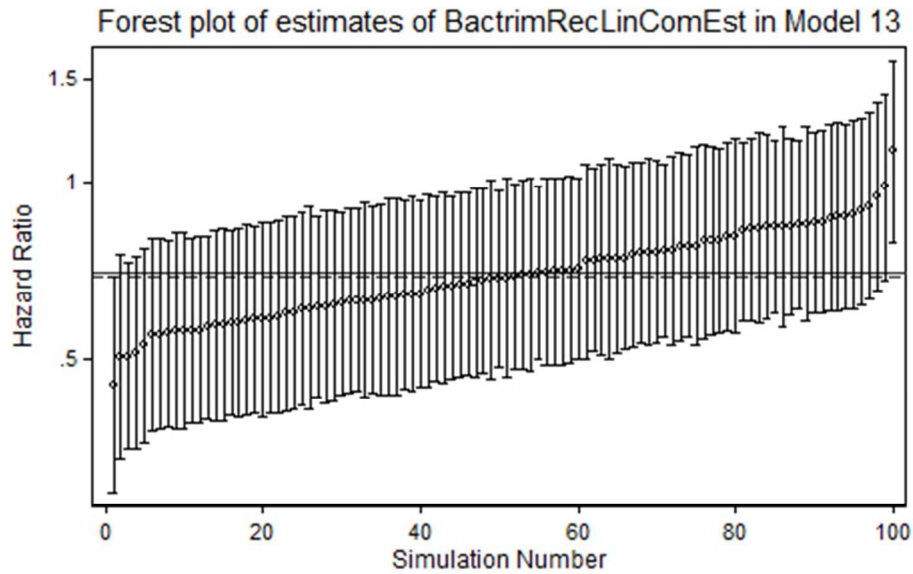


28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with TMP-SMX use for subjects with high viral load in each simulation under Model 13. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



28
29
30
31
32

Forest plot with point estimates and 95% confidence intervals for the hazard ratio associated with TMP-SMX use in subjects with viral load under 1000 copies/mL in each simulation under Model 13. Estimates were sorted by size. The dashed line represents the geometric mean of the hazard ratios obtained in the 100 simulations under the model; the solid line represents the true population value.

167x111mm (72 x 72 DPI)

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60