

```

1 *****
2 *
3 * This program was designed to assess *
4 * the data collected from a random *
5 * sample of 500 journals *
6 * Date: 10.08.15 *
7 *****;
8
9 /* 1: Create data trans in work library based on imported excel
10 document */
11
12 DATA trans;
13     SET TRANSR.Trans;
14
15 *** create new variables for use;
16     IF cit_art=0 THEN cit_art1=0;
17     ELSE IF cit_art=1 THEN cit_art1=1;
18
19 *** define studyfield for CLINICAL MEDICINE as 1 and all non-
20 clinical medicine as 0;
21     If studyfield='CLINICAL MEDICINE' then studyfield1=1;
22
23     else IF studyfield='NODOC' then studyfield1=.; /* Just incase
24 any NODOC or NESI remain */
25     else If studyfield='NOESI' then studyfield1=.;
26     else studyfield1=0;
27
28 *** create new funding variables - with funding1=5 for some
29 combination of 2-4;
30
31     *** 0=no mention, 1=no funding, 2=Public,
32 3=private, 4= other, 5= combination of funding 2&3,
33 6=combination 2&4, 7= combination 3&4,
34 8=combination 2-4, 88=not applicable;
35
36     if funding=0 then funding1=0;
37     else if funding =1 then funding1=1;
38     else if funding=2 then funding1=2;
39     else if funding=3 then funding1=3;
40     else if funding=4 then funding1=4;
41     else if funding=5 then funding1=5;
42     else if funding=6 then funding1=5;
43     else if funding=7 then funding1=5;
44     else if funding=8 then funding1=5;
45     else if funding=88 then funding1=88;
46
47
48 *** Citing article being replicated make sure those with no data are
49 not cited;
50     if cit_art=0 then cit_art1=0;
51     else if cit_art=1 then cit_art1=1;
52     else if cit_art=99 then cit_art1=0; *if no data available make
53 no cite;
54     else if cit_art=88 then cit_art1=88;

```

```

55
56 *** Citation by a systematic review;
57
58 ***0=no systematic review and/or meta-analysis has ever cited the
59 index paper,
60 1=at least one systematic review and/or meta-analysis has cited the
61 index paper
62 but none has included any of its data in quantitative syntheses for
63 any outcome,
64 2=at least one systematic review and/or meta-analysis has cited
65 the index paper and has included some of its data in quantitative
66 synthesis for at least one outcome 1.5=excluded data from review,
67 4= combination of 1 and 1.5, 5=combination of 1 and 2,
68 6=combination of 1.5 and 2, 7= combination of 1, 1.5 and 2,
69 88=Not Applicable, 99=No Data Available ;
70
71 ***For Cit_Sysrev =5 or 6 or 7, making it Cit_Sysrev 3: Any
72 combination of;
73
74     if Cit_Sysrev=0 then Cit_Sysrev1=0;
75     else if Cit_Sysrev=1 then Cit_Sysrev1=1;
76     else if Cit_Sysrev=1 then Cit_Sysrev1=2;
77     else if Cit_Sysrev=1.5 then Cit_Sysrev1=1.5;
78     else if Cit_Sysrev=4 then Cit_Sysrev1=4;
79     else if Cit_Sysrev=5 then Cit_Sysrev1=4;
80     else if Cit_Sysrev=6 then Cit_Sysrev1=4;
81     else if Cit_Sysrev=7 then Cit_Sysrev1=4;
82     else if Cit_Sysrev=88 then Cit_Sysrev1=88;
83     else if Cit_Sysrev=99 then Cit_Sysrev1=99;
84
85 *** Extra variable, incase needed in later analysis;
86
87     if Cit_Sysrev=0 then Cit_Sysrev2=0;
88     else if Cit_Sysrev=1 then Cit_Sysrev2=1;
89     else if Cit_Sysrev=2 then Cit_Sysrev2=2;
90     else if Cit_Sysrev=1.5 then Cit_Sysrev2=1.5;
91     else if Cit_Sysrev=88 then Cit_Sysrev2=88;
92     else if Cit_Sysrev=99 then Cit_Sysrev2=0;*if no data available
93 make no cite;
94
95
96 *** Replication of study
97 /* 0=based on the abstract and/or introduction,
98     the index paper claims that it presents some novel findings,
99     1=based on its abstract/intro, the index paper clearly claims
100 that
101     it is a replication effort trying to validate previous
102 knowledge,
103     4=unclear statement in the abstract, but based on its
104 introduction,
105     it is inferred that the index paper is a replication trying to
106 validate previous knowledge, 2=it claims to be both novel and
107 replicate previous findings, 3=no statement in the abstract
108 and/or

```

```

109     introduction about whether the index paper presents a novel
110 finding
111     or replication, 5=no abstract/no introduction, 88=not
112 applicable
113     if replication = 0 then replication1=0;
114
115     else if replication = 1 then replication1 = 1;
116     else if replication = 2 then replication1 = 2;
117     else if replication = 3 then replication1 = 3;
118     else if replication = 4 then replication1 = 1; *if unclear
119 abstract, but replication inferred by intro, make replication;
120     else if replication = 5 then replication1 = 3; *if there is no
121 intro abstract then make unclear;
122     else if replication = 88 then replication1 = 88;
123     else if replication = 99 then replication1 = 99;
124
125 *** Create new categorical variable for impact factor;
126
127     if 2 ge impact ge 0 then impact2a="0-2";
128     else if 4 ge impact gt 2 then impact2a=">2-4";
129     else if 6 ge impact gt 4 then impact2a=">4-6";
130     else if impact gt 6 then impact2a=">6";
131     else impact2a=" ";
132
133
134
135 *Create formats;
136
137 PROC FORMAT;
138
139 *Objective 1;
140
141
142     VALUE fundingf          0="No Mention"
143                           1="No Funding"
144                           2="Government"
145                           3="Private"
146                           4="Other"
147                           5="Both Public and Private"
148                           6="Both Public and Other"
149                           7="Both Private and Other"
150                           8="Public, Private and Other"
151                           88="Not Applicable"
152                           ;
153     VALUE fundinglf       0="No Mention"
154                           1="No Funding"
155                           2="Public"
156                           3="Private"
157                           4="Other"
158                           5="Combination"
159                           88="Not Applicable"
160                           ;
161
162     VALUE NIHFundf        1="True"

```

```

163             0="False"
164             88="Not Applicable"
165             ;
166     VALUE NSFFundf      1="True"
167             0="False"
168             88="Not Applicable"
169             ;
170     VALUE OtherFundf 1="True"
171             0="False"
172             88="Not Applicable"
173             ;
174     VALUE Cit_artf      0="No Citing Article"
175             1="At Least 1 Citing Article"
176             88="Not Applicable"
177             99="No Data Available"
178             ;
179
180     VALUE Cit_art1f     0="No Citing Article"
181             1="At Least 1 Citing Article"
182             88="Not Applicable"
183             ;
184
185     VALUE Cit_Sysrevf0="No Citing Article"
186             1="At Least 1 Citing Article, No
187 Data Included"
188             2="At Least 1 Citing Article, Data
189 Included"
190             1.5="At Least 1 Citing Article,
191 Data Excluded"
192             88="Not Applicable"
193             99="No Data Available"
194             ;
195
196     VALUE Cit_Sysrev2f 0="No Citing Article"
197             1="At Least 1 Citing Article, No
198 Data Included"
199             2="At Least 1 Citing Article, Data
200 Included"
201             1.5="At Least 1 Citing Article,
202 Data Excluded"
203             88="Not Applicable"
204             ;
205
206     VALUE Caserepf      1="True"
207             0="False"
208             ;
209     VALUE Cedaf         1="True"
210             0="False"
211             ;
212     VALUE Conflictsf 0="No Statement"
213             1="Statement Exists, Conflicts
214 Present"
215             2="Statement Exists, No Conflicts"
216             88="Not Applicable"

```

```

217                                     ;
218     VALUE Datasetf                 0="No Dataset"
219                                     1="Partial Coverage"
220                                     2="Full Coverage"
221                                     3="Cannot Be Determined"
222                                     88="Not Applicable"
223                                     ;
224     VALUE Modelf                   1="True"
225                                     0="False"
226                                     ;
227     VALUE Noresf                   1="True"
228                                     0="False"
229                                     ;
230     VALUE Otherf                   1="True"
231                                     0="False"
232                                     ;
233     VALUE RCTf                     1="True"
234                                     0="False"
235                                     ;
236     VALUE Sysmetf                  1="True"
237                                     0="False"
238                                     ;
239     VALUE protocolf                0="No Protocol"
240                                     1="Partial Coverage"
241                                     2="Full Coverage"
242                                     3="Cannot Be Determined"
243                                     88="Not Applicable"
244                                     ;
245
246     VALUE Replicationf             0="Novel"
247                                     1="Replication"
248                                     2="Novel and Replication"
249                                     3="No Statement on Novelty"
250                                     5="No Abstract and Introduction"
251                                     88="Not Applicable"
252                                     99="No Data Available"
253                                     ;
254
255     VALUE Replicationlf            0="Novel"
256                                     1="Replication"
257                                     2="Novel and Replication"
258                                     3="No Statement on Novelty"
259                                     88="Not Applicable"
260                                     99="No Data Available"
261                                     ;
262
263
264 *Objective 2;
265     VALUE Animalf                 1="True"
266                                     0="False"
267                                     77="No Abstract/Intro"
268                                     88="Not Applicable"
269                                     ;
270     VALUE Clinic_Trif              1="True"

```

```

271             0="False"
272             88="Not Applicable"
273             ;
274     VALUE Genesf             1="True"
275             0="False"
276             77="No Abstract/Intro"
277             88="Not Applicable"
278             ;
279     VALUE Humanf             1="True"
280             0="False"
281             77="No Abstract/Intro"
282             88="Not Applicable"
283             ;
284     VALUE PCMIDf             1="True"
285             0="False"
286             88="Not Applicable"
287             ;
288
289
290
291 RUN;
292
293 *Apply permanent Format;
294
295 DATA transf;
296     SET trans;
297
298     FORMAT Animal Animalf.;
299     FORMAT Caserep Caserepf.;
300     FORMAT Ceda Cedaf.;
301     FORMAT Clinic_Tri Clinic_Trif.;
302     FORMAT Conflicts Conflictsf.;
303     FORMAT Dataset Datasetf.;
304     FORMAT Funding Fundingf.;
305     FORMAT Funding1 Funding1f.;
306     FORMAT Genes Genesf.;
307     FORMAT Human Humanf.;
308     FORMAT NIHFund NIHFundf.;
309     FORMAT NSFFund NSFFundf.; /*And here*/
310     FORMAT OtherFund OtherFundf.;
311     FORMAT Model Modelf.;
312     FORMAT Nores Noresf.;
313     FORMAT Other Otherf.;
314     FORMAT PCMID PCMIDf.;
315     FORMAT Protocol Protocolf.;
316     FORMAT Replication Replicationf.;
317     FORMAT Replication1 Replication1f.;
318     FORMAT RCT RCTf.;
319     FORMAT Sysmet Sysmetf.;
320     FORMAT Cit_art Cit_artf.;
321     FORMAT Cit_art1 Cit_artf1.;
322     FORMAT Cit_Sysrev Cit_Sysrevf.;
323     FORMAT Cit_Sysrev2 Cit_Sysrevf2.;
324

```

```

325
326 RUN;
327
328
329
330
331 *****
332 * Descriptive Data Analysis *
333 * Make sure to remove all *
334 * non-medical and N/A articles*
335 *****;
336
337 *Number of Medical Journals v non-Medical;
338 PROC FREQ;
339     TABLES Medfield;
340 RUN;
341
342 *Distribution of type of research;
343 PROC FREQ;
344     TABLES Nores*Model*Caserep*RCT*sysmet*Ceda*Other/LIST;
345     WHERE Medfield=1;
346 RUN;
347
348 *Publically Available Protocols--WANT TO KEEP SYSMET, AND CEDA
349 HERE!!;
350 PROC FREQ;
351     TABLES protocol;
352     WHERE Medfield=1 and protocol NE 88;
353 RUN;
354
355 *Publically Available Datasets--WANT TO KEEP SYSMET AND CEDA HERE!!;
356 PROC FREQ;
357     TABLES dataset;
358     WHERE Medfield=1 and dataset NE 88 and caserep ne 1;
359 RUN;
360
361 *Funding;
362 PROC FREQ;
363     TABLES funding;
364     WHERE Medfield =1 and funding NE 88;
365 RUN;
366
367 *NIH Funding;
368 PROC FREQ;
369     TABLES NIHFund;
370     WHERE Medfield=1 and NIHFund NE 88;
371 RUN;
372
373
374 *NSF Funding;
375 PROC FREQ;
376     TABLES NSFFund;
377     WHERE Medfield=1 and NSFFund NE 88;
378 RUN;

```

```

379
380 *Other Funding;
381 PROC FREQ;
382     TABLES OtherFund;
383         WHERE Medfield=1 and OtherFund NE 88;
384 RUN;
385
386 *breakdown of government funding into exclusive funding categories;
387 PROC FREQ;
388     TABLES NIHFund*NSAfund*OtherFund/LIST;
389         WHERE Medfield=1 and NIHFund NE 88 and NSAfund NE 88 and
390 otherfund NE 88;
391 RUN;
392
393
394 *Trends in patterns of funding by year;
395 *Trends for funding over time;
396 PROC FREQ;
397     TABLES year*funding/LIST;
398         WHERE medfield=1;
399 RUN;
400
401 *Trends for public funding over time;
402 PROC FREQ;
403     TABLES year*NIHFund*NSAfund*OtherFund/LIST;
404         WHERE Medfield=1 and NIHFund NE 88 and NSAfund NE 88 and
405 otherfund NE 88;
406 RUN;
407
408
409
410 *****;
411 *                               ;
412 *   Empi Data Only comparison           ;
413 *                               ;
414 *   REPLICATION                       ;
415 *                               ;
416 *****;
417
418 PROC FREQ; *ensure new variable coded correctly;
419     TABLES replication*replication1/LIST MISSING;
420         WHERE Medfield =1 and replication NE 88 and sysmet ne 1
421 and ceda ne 1 and caserep ne 1;
422 RUN;
423
424 PROC FREQ; *only want Empirical data without casrep, ceda and
425 sysmets (total of 259);
426     TABLES replication1;
427         WHERE Medfield =1 and replication NE 88 and sysmet ne 1
428 and ceda ne 1 and caserep ne 1;
429 RUN;
430
431
432

```

```

433
434 *****;
435 *                                     ;
436 *   Empi Data Only comparison           ;
437 *                                     ;
438 *           CITING ARTICLE             ;
439 *                                     ;
440 *****;
441
442
443 PROC FREQ;*only want Empirical data without casrep, ceda and sysmets
444 (total of 259);
445     TABLES cit_art;
446     WHERE Medfield =1 and replication ne 88 and sysmet ne 1
447 and ceda ne 1 and caserep ne 1;
448 RUN;
449
450
451 *check to ensure that cited articles are correct with dataset;
452
453 PROC PRINT;
454     VAR PMID;
455     WHERE cit_art=1;
456 RUN;
457
458
459
460 *****;
461 *                                     ;
462 *   Empi Data Only comparison           ;
463 *                                     ;
464 *           CITING SYSMET             ;
465 *                                     ;
466 *****;
467
468 *Citing systematic review or meta analysis;
469 PROC FREQ;*only want Empirical data without casrep, ceda and sysmets
470 (total of 259);
471     TABLES cit_sysrev;
472     WHERE Medfield =1 and replication NE 88 and sysmet ne 1
473 and ceda ne 1 and caserep ne 1;
474 RUN;
475
476 PROC print;*ensure have correct articles with data included (16=n;
477     VAR PMID;
478     WHERE Medfield =1 and (cit_sysrev=2 or cit_sysrev=5 or
479 cit_sysrev=6 or cit_sysrev=7);
480 RUN;
481
482 PROC PRINT;*ensure have correct articles with data excluded (3=n;
483     VAR PMID;
484     WHERE Medfield =1 and (cit_sysrev=1.5 or cit_sysrev=4 or
485 cit_sysrev=6 or cit_sysrev=7);
486 RUN;

```

```

487
488 PROC PRINT;*ensure have correct articles with no data included but
489 cited (19=n;
490     VAR pmid;
491     WHERE Medfield =1 and (cit_sysrev=1 or cit_sysrev=4 or
492 cit_sysrev=5 or cit_sysrev=7);
493 RUN;
494
495
496
497 *****;
498 *                               ;
499 *   ALL Data (n=441)           ;
500 *                               ;
501 *           CONFLICTS         ;
502 *                               ;
503 *****;
504
505
506
507 *Conflicts of Interest all;
508
509 PROC FREQ;
510     TABLES conflicts;
511     WHERE Medfield =1 and conflicts NE 88;
512 RUN;
513
514 *Conflicts of Interest only RCTs (n=15);
515
516 PROC FREQ;
517     TABLES conflicts;
518     WHERE Medfield =1 and conflicts NE 88 and RCT=1;
519 RUN;
520
521 *Trends in Conflicts of Interest all (n=441);
522
523 *Conflicts by year;
524
525 PROC FREQ;
526     TABLES year*conflicts/LIST;
527     WHERE medfield=1;
528 RUN;
529
530
531 *****;
532 *                               ;
533 *   ALL Data (n=441)           ;
534 *                               ;
535 *           impact factor           ;
536 *                               ;
537 *****;
538 *Journal Impact Factor for 2013;
539 PROC UNIVARIATE;
540     VAR Impact2;

```

```

541             WHERE Medfield=1;
542 RUN;
543
544 *Journal impact factor categorized;
545 PROC FREQ; *ensure variable created correctly;
546     TABLES Impact2*impact2a/LIST MISSING;
547     WHERE Medfield=1;
548 RUN;
549
550 PROC FREQ; *ensure variable created correctly;
551     TABLES impact2a/LIST MISSING;
552     WHERE Medfield=1;
553 RUN;
554
555 *****;
556 *                               ;
557 *   ALL Data (n=441)           ;
558 *                               ;
559 *           PMCID              ;
560 *                               ;
561 *****;
562
563 PROC FREQ;
564     TABLES PMCID/LIST MISSING;
565     WHERE Medfield=1;
566 RUN;
567
568
569
570 *****;
571 *                               ;
572 *   ALL Data (n=441)           ;
573 *                               ;
574 *           studyfield         ;
575 *                               ;
576 *****;
577
578
579
580 *Study Field distribution for all data;
581 PROC FREQ;
582     TABLES studyfield/LIST MISSING;
583     WHERE Medfield=1;
584 RUN;
585
586
587
588
589
590 *****;
591 *                               ;
592 *   clinical medicine v other   ;
593 *                               ;
594 *                               ;

```

```

595 *
596 *****;
597
598 PROC FREQ;
599     TABLES studyfield*studyfield1/LIST MISSING;
600         WHERE medfield=1 ;
601 RUN;
602
603 *****;
604 *
605 *   clinical medicin v other
606 *                   n=441
607 *           FUNDING
608 *
609 *****;
610
611 PROC FREQ;
612     TABLES funding*funding1/LIST MISSING; *ensure new variable
613     that consolidates combo funding correct;
614 RUN;
615
616 PROC FREQ; *compare all types of research for funding, including
617     artiles without empirical data;
618     TABLES studyfield1*funding1/LIST MISSING;
619         WHERE medfield=1 and studyfield1 =1;
620 RUN;
621
622 PROC FREQ; *compare all types of research for funding, including
623     artiles without empirical data;
624     TABLES studyfield1*funding1/LIST MISSING;
625         WHERE medfield=1 and studyfield1 =0;
626 RUN;
627
628 PROC FREQ; *****WARNING: Computing exact p-values for this
629     problem may require much time and memory. Press the
630     system interrupt key to terminate exact computations.;
631     TABLES studyfield1*funding1/EXACT FISHER;
632         WHERE medfield=1;
633 RUN;
634
635
636
637 ***Run monte carlo approximation on fisher exact;
638 PROC FREQ;
639     TABLES studyfield1*funding1 / CHISQ EXPECTED;
640         EXACT FISHER / MC;
641 RUN;
642
643
644 *****;
645 *
646 *   clinical medicine v other
647 *                   n=441
648 *           FUNDING NIH

```

```

649 *
650 *****;
651
652
653 PROC FREQ; *compare all types of research for funding, including
654 artiles without empirical data;
655     TABLES studyfield1*NIHFUND/LIST MISSING;
656     WHERE medfield=1 and studyfield1 =1;
657 RUN;
658
659 PROC FREQ; *compare all types of research for funding, including
660 artiles without empirical data;
661     TABLES studyfield1*NIHFUND/LIST MISSING;
662     WHERE medfield=1 and studyfield1 =0;
663 RUN;
664
665 PROC FREQ;
666     TABLES studyfield1*NIHFUND/EXACT FISHER;
667     WHERE medfield=1;
668 RUN;
669
670
671
672
673 *****;
674 *
675 *   clinical medicine v other
676 *           n=259
677 *   REPLICATION
678 *
679 *****;
680 PROC FREQ; *only want Empirical data without casrep, ceda and
681 sysmets (total of 259);
682     TABLES studyfield1*replication1/LIST MISSING;
683     WHERE medfield=1 and replication ne 88 and studyfield1=0
684 and sysmet ne 1 and ceda ne 1 and caserep ne 1;
685 RUN; *clinical medicine studyfield1=1;
686
687 PROC FREQ; *only want Empirical data without casrep, ceda and
688 sysmets (total of 259);
689     TABLES studyfield1*replication1/LIST MISSING;
690     WHERE medfield=1 and replication ne 88 and studyfield1=1
691 and sysmet ne 1 and ceda ne 1 and caserep ne 1;
692 RUN; *clinical medicine studyfield1=1;
693
694 PROC FREQ; *only want Empirical data without casrep, ceda and
695 sysmets (total of 259);
696     TABLES studyfield1*replication1/EXACT FISHER;
697     WHERE medfield=1 and replication ne 88 and sysmet ne 1
698 and ceda ne 1 and caserep ne 1;
699 RUN; *clinical medicine studyfield1=1;
700
701
702 *****;

```

```

703 *
704 *   cli med v other comparison
705 *                               n=259
706 *   CITING ARTICLE
707 *
708 *****;
709 PROC FREQ;
710     TABLES cit_art*cit_art1/LIST MISSING; *make sure coded
711 correctly;
712     WHERE medfield=1 and replication ne 88 and sysmet ne 1
713 and ceda ne 1;
714 RUN;
715
716 PROC FREQ;
717     TABLES studyfield1*cit_art1/LIST MISSING;
718     WHERE medfield=1 and studyfield1 =1 and replication ne 88
719 and sysmet ne 1 and ceda ne 1;
720 RUN;
721
722
723
724 PROC FREQ;
725     TABLES studyfield1*cit_art1/LIST MISSING;
726     WHERE medfield=1 and studyfield1 =0 and replication ne 88
727 and sysmet ne 1 and ceda ne 1;
728 RUN;
729
730
731 PROC FREQ;
732     TABLES studyfield1*cit_art1/EXACT FISHER;
733     WHERE medfield=1 and replication ne 88 and sysmet ne 1
734 and ceda ne 1;
735 RUN;
736
737
738
739
740 *****;
741 *
742 *   cli med v other comparison
743 *                               n=259
744 *   CITING SYSMET
745 *
746 *****;
747 PROC FREQ;
748     TABLES cit_sysrev*cit_sysrev2/LIST;
749 RUN;
750 PROC FREQ;
751     TABLES studyfield1*Cit_Sysrev2/LIST MISSING;
752     WHERE medfield=1 and studyfield1 =1 and replication ne 88
753 and sysmet ne 1 and ceda ne 1;
754 RUN;
755
756

```



```

811         TABLES studyfield1*PMCID/LIST MISSING;
812             WHERE medfield=1 and studyfield1=1;
813 RUN; *clinical medicine studyfield1=1;
814
815 PROC FREQ;
816     TABLES studyfield1*PMCID/LIST MISSING;
817         WHERE medfield=1 and studyfield1=0;
818 RUN; *clinical medicine studyfield1=1;
819
820 PROC FREQ;
821     TABLES studyfield1*PMCID/EXACT FISHER;
822         WHERE medfield=1;
823 RUN; *clinical medicine studyfield1=1;
824
825
826 *****;
827 *                                     ;
828 *   additional code requested by reviewers   ;
829 *           to replicate table 1 with           ;
830 *   empirical studies only (n=304)           ;
831 *                                     ;
832 *****;
833
834
835 *****;
836 *                                     ;
837 *   additional code requested by reviewers   ;
838 *           to replicate table 1 with           ;
839 *   empirical studies only (n=304)           ;
840 *                                     ;
841 *****;
842
843 *Distribution of type of research for studies with empirical data
844 only n=304;
845 PROC FREQ;
846     TABLES Nores*Model*Caserep*RCT*sysmet*Ceda*Other/LIST;
847         WHERE Medfield=1 and nores ne 1 and model ne 1;
848 RUN;
849
850
851 *Distribution of PCMID for studies with empirical data
852 only n=304;
853
854 PROC FREQ;
855     TABLES PCMID/LIST MISSING;
856         WHERE Medfield=1 and nores ne 1 and model ne 1;
857 RUN;
858
859 *Distribution of impact factor for studies with empirical data
860 only n=304;
861
862 PROC FREQ;
863     TABLES impact2a/LIST MISSING;
864         WHERE Medfield=1 and nores ne 1 and model ne 1;

```

```

865
866 RUN;
867
868 PROC UNIVARIATE;
869     VAR Impact2;
870     WHERE Medfield=1 and nores ne 1 and model ne 1;
871
872 RUN;
873
874
875 *Distribution of field of research for studies with empirical data
876 only n=304;
877 PROC FREQ;
878     TABLES studyfield/LIST MISSING;
879     WHERE Medfield=1 and nores ne 1 and model ne 1;
880
881 RUN;
882
883
884 PROC FREQ;
885     TABLES impact2a/LIST MISSING;
886     WHERE Medfield=1 and nores ne 1 and model ne 1;
887
888 RUN;
889
890
891 /* PMCID among empirical data NON CLINICAL MEDICINE FIELD */
892 PROC FREQ;
893     TABLES studyfield1*PMCID/LIST MISSING;
894     WHERE medfield=1 and studyfield1=0 and nores ne 1 and
895 model ne 1;
896 RUN; *clinical medicine studyfield1=1;
897
898 /*PMCID among empirical data CLINICAL MEDICINE FIELD */
899 PROC FREQ;
900     TABLES studyfield1*PMCID*medfield/LIST MISSING;
901     WHERE medfield=1 and studyfield1=1 and nores ne 1 and
902 model ne 1;
903 RUN; *clinical medicine studyfield1=1;
904
905
906
907
908 /* Funding for Clinical Medicine Field with Empirical Data */
909 PROC FREQ;
910     TABLES studyfield1*funding1*medfield/LIST MISSING;
911     WHERE medfield=1 and studyfield1=1 and nores ne 1 and
912 model ne 1;
913 RUN; *clinical medicine studyfield1=1;
914
915 PROC FREQ; /*Funding for NON Clinical Medicine Field */
916     TABLES studyfield1*funding1*medfield/LIST MISSING;
917     WHERE medfield=1 and studyfield1=0 and nores ne 1 and
918 model ne 1;

```

```

919 RUN; *clinical medicine studyfield1=1;
920
921 PROC FREQ; /*Funding for empirical data studies among CLINICAL MED */
922     TABLES studyfield1*funding1*medfield/LIST MISSING;
923     WHERE medfield=1 and studyfield1=1 and nores ne 1 and
924 model ne 1;
925 RUN; *clinical medicine studyfield1=1;
926
927
928 PROC FREQ; /*Conflicts for empirical data studies among CLINICAL MED
929 */
930     TABLES studyfield1*conflicts*medfield/LIST MISSING;
931     WHERE medfield=1 and studyfield1=1 and nores ne 1 and
932 model ne 1;
933 RUN; *clinical medicine studyfield1=1;
934
935
936 PROC FREQ; /*Conflicts for empirical data among NON CLINICAL MED */
937     TABLES studyfield1*conflicts*medfield/LIST MISSING;
938     WHERE medfield=1 and studyfield1=0 and nores ne 1 and
939 model ne 1;
940 RUN; *clinical medicine studyfield1=1;
941
942
943
944
945
946
947
948 *This is the code for methods section to compare the proportion of
949 this study articles by year to articles by year in pubmed that are
950 English only published between 2000 and 2014;
951
952 DATA a;
953     INPUT year count expect;
954     CARDS;
955
956 2000 19 22.02537386
957 2001 30 22.65847202
958 2002 36 23.41544225
959 2003 30 24.6793644
960 2004 31 26.67482332
961 2005 37 29.34960617
962 2006 34 31.47882279
963 2007 23 33.26761167
964 2008 32 35.53346351
965 2009 31 37.44389672
966 2010 27 40.43980017
967 2011 49 43.95662537
968 2012 55 46.89548919
969 2013 34 49.62475904
970 2014 32 32.55644951
971 ;
972

```

```
973
974 PROC FREQ;
975     TABLES year / NOCUM CHISQ TESTP=(4.4050748 4.5316944 4.6830884
976 4.9358729 5.3349647 5.8699212 6.2957646 6.6535223 7.1066927
977 7.4887793
978 8.08796 8.7913251 9.3790978 9.9249518 6.5112899);
979     WEIGHT count;
980
981 RUN;
982
```