PseudoCons: Generating Case/Pseudocontrol Data from Pedigree Data!

Contents

1 Introduction

Do you desperately want to perform a case/control analysis but only have pedigree data? If each pedigree contains a case with parents then one option is to take the cases and create pseudocontrols using the non-transmitted alleles from the parents. The case/control analysis of interest can then be performed using the case/pseudocontrol data set. Unfortunately, creating the pseudocontrols can be a hassle, especially if the pedigree data does not consist of only case/parent trios. However, it is no longer a hassle as `PseudoCons` is here to help! Simply run `PseudoCons` on your pedigree data set to create a case/pseudocontrol data set.

Some theory behind case/pseudocontrol analysis can be found in the work by Cordell (2004) and Cordell et al. (2004).

2 Installation

Download an executable file from the download page for your system and off you go, or do the following.

1. Download the code from the download page.

2. Compile it by typing something like the following:

   ```
   g++ -O3 *.cpp -o pseudocons
   ```

3. Start creating case/control binary pedigree files with PseudoCons!

3 Using PseudoCons

3.1 Basic Usage

The program PseudoCons takes a PLINK binary pedigree file as input and outputs a PLINK binary pedigree file:

```
./pseudocons -i mydata.bed -o pseudoData.bed
```

This requires that the corresponding `.bim` and `.fam`, files are also available. A text PLINK pedigree file, `.ped`, with corresponding map file, `.map`, may be used to create a binary file using PLINK as follows:

```
plink --noweb --file mydata --make-bed --out mydata
```

This will create the binary pedigree file, `mydata.bed`, map file, `mydata.bim`, and family file, `mydata.fam` required for use with PseudoCons.

If for some reason you should wish to have a text PLINK pedigree file, this can be created using PLINK as follows:

```
plink --noweb --bfile pseudoData --make-bed --out pseudoData --recode
```

This will create the text PLINK pedigree file `pseudoData.ped` with map file `pseudoData.map`.

3.2 Options

Typing `pseudocons` with no options will output usage details:

```
PseudoCons: pseudocontrols from pedigree data v1.12
---------------------------------------------------
Copyright 2013-2014 Richard Howey, GNU General Public License, v3
Institute of Genetic Medicine, Newcastle University

Usage:
        ./pseudocons [options] inputFile
 or ./pseudocons -pf parameterfile

Options:
  -i filename        -- input filename
  -o filename        -- output filename
  -pc1               -- one pseudocontrol per trio
  -pc3               -- three pseudocontrols per trio
  -pc15              -- 15 pseudocontrols per trio (2 SNP interaction only)
  -snpnos snp1 snp2  -- SNP numbers of pair to create 15 pseudocontrols per trio
  -snpnames snp1 snp2 -- as above using SNP names
  -cepg              -- use CEPG giving 3, 7 and 31 pseudos for -pc1, 3 and 15
  -info info.dat     -- output info file
  -info-fa info.dat  -- output info file with father+child genotypes
  -info-ma info.dat  -- output info file with mother+child genotypes
  -info-fama info.dat -- output info file with father+mother+child genotypes
  -info-maxsnps n    -- max number of SNPs allowed in info file (default=20)
  -pro               -- proband filename
  -xtrio             -- allow extra trios
  -log               -- log filename
  -so                -- suppress output to screen

Default options:
  -pc1
```

3.3 Parameter file

A parameter file, `.pf`, may be used with PseudoCons instead of writing all of the options on the command line. To use a parameter file simply type:

`./pseudocons myparameters.pf`

The parameter file should be a text file with one option written on each line. For example, to perform the analysis above the file `myparameters.pf` would be as follows:

```
-i mydata.bed
-o myCasePseudoControlData.bed
```

```
-log myLog.log
-so
```

It is also possible to add comments to the file provided that the "-" character is not used, and to comment out any options by placing another character in front of any "-". For example, the above parameter file could be edited as follows to perform the next analysis given above:

```
This is in input file from my really great study
-i mydata.bed

This will be my case/pseudocontrol data file
-o myCasePseudoControlData.bed

Keep a log of the PseudoCons output here
-log myLog.log

Suppress the output from the screen
-so
```

3.4   Input

The following file types are input into PseudoCons:

File	Description
.bed	binary PLINK file
.bim	extended map file with allele names
.fam	family pedigree PLINK file

3.5   Output

The following file types are output from PseudoCons:

File	Description
.bed	binary PLINK file
.bim	extended map file with allele names
.fam	family pedigree PLINK file
.log	log file of PseudoCons

If the `-info info.dat` option is used then an additional information file is output, see section 4.6.

4   Data Processing

This section explains how PseudoCons processes the pedigree data to produce the case-control output data.

4.1 Trio Selection

By default one case/parent trio is taken from each pedigree and from this one case is taken and one pseudocontrol created. The trio chosen is simply decided by the first case in the pedigree file who also has two parents in the pedigree file.

Proband

It may be possible that there is a choice of case/parent trios from a pedigree to give the case and created pseudocontrol. For a pedigree file with many large pedigrees this could potentially alter the results of any subsequent analysis performed. For example, if pedigrees are ascertained on the basis of a particular affected child, but case/parent trios containing the parents and grandparents are chosen instead, this could then bias the analysis. With this in mind it is possible to supply an optional *proband* file containing a list of all the affected subjects that are of interest. The file is a list of subjects given by the pedigree name and subject name coresponding to the pedigree file given to PseudoCons. For example, a proband file may look as following:

```
1 4
2 5
3 2
5 12
7 3
9 3
10 2
```

The proband file is used in PseudoCons with the `-pro` option as follows:

```
./pseudocons -pro proband.dat -i mydata.bed -o mycasepscondata.bed
```

The name of the proband file should following immediately after the `-pb` option. The following points should be noted about proband files:

1. If a proband file is given it is not necessary to supply a subject for every pedigree. For example, for smaller pedigrees you may be happy to use the default setting.

2. The proband subjects do not need to appear in any particular order in the file.

3. If the proband subject is not affected a warning message will be displayed and the pedigree processed using the default settings.

4. If a proband subject does not exist in the pedigree file a warning message will be displayed and the pedigree file will be processed as normal.

Extra Trios

It is possible to use all possible case/parent trios from a pedigree, counting them as if they are independent, using the `-xtrio` option. The trios may overlap if a parent is also a case. Depending on the analysis you want to do, this assumption may be more or less valid.

4.2 One Pseudocontrol

The pseudocontrols are created using the non-transmitted alleles. For example, if the alleles of the case are A/A and the alleles of the parents are A/G and A/G, then the created pseudocontrol will have alleles G/G.

4.3 Three Pseudocontrols

The three pseudocontrols are created using any possible genotype from the parents that contains a non-transmitted alleles. For example, if the alleles of the case are A/A and the alleles of the parents are A/G and A/G, then the three created pseudocontrols will have alleles G/G, A/G and G/A.

4.4 Fifteen Pseudocontrols

Given two SNPs, the 15 pseudocontrols are created using any possible genotype pair from the parents that contains a non-transmitted allele.

4.5 CPG and CEPG

The standard procedure for `PseudoCons` is to assume Conditional on Parental Genotypes (CPG) rather than the Conditional on Exchangeable Parental Genotypes (CEPG) [Cordell (2004), Weinberg and Shi (2009)]. It is also possible to assume CEPG with the option `-cepg` which will create an additional set of pseudocontrols, resulting in 3, 7 and 31 pseudocontrols for options `-pc1`, `-pc3` and `-pc15` respectively.

Using options `-pc1 -cepg` will result in the following pseudocontrols: (i) the usual pseudocontrol with the non-transmitted genotype; and (ii) the parental genotypes swapped to give a pseudocontrol for the non-transmitted genotype and a pseudocontrol for the transmitted genotype.

Using options `-pc3 -cepg` will result in the following pseudocontrols: (i) the usual 3 pseudocontrols with a genotype containing a non-transmitted allele; and (ii) the parental genotypes swapped to give another 4 pseudocontrols given by the possible transmitted genotypes.

Using options `-pc15 -cepg` will result in the following pseudocontrols: (i) the usual 15 pseudocontrols given by any possible genotype pair from the parents that contains a non-transmitted allele; and (ii) the parental genotypes swapped to give another 16 pseudocontrols given by any possible genotype pair.

When the `-cepg` option is used the log file will include the following lines:

```
...
Number of pseudocontrols per trio: 1 + 2 = 3
Using Conditional on Exchangeable Parental Genotypes (CEPG)
...
```

showing the amended number of pseudocontrols and that CEPG has been assumed. If an infomation file is output when using the `-cepg` option then the extra pseudocontrols

created will have the maternal and paternal genotypes exchanged. So for options `-pc1`, `-pc3` and `-pc15` the pseudocontrols 2-3, 4-7 and 16-31 will have maternal and paternal genotypes exchanged respectively.

Warning: the options `-pc3` and `-pc1 -cepg` both result in 3 pseudocontrols but are different. Do not get the two mixed up!

4.6   Information File

It is possible to output an extra information file for follow up analysis. If the option `-info info.dat` is used then the text file `info.dat` is output with the following columns:

1. **Family ID**. The same family ID as in the `.bim` file.

2. **Individual ID**. The same individual ID as in the `.bim` file.

3. **Case ID**. The corresponding case ID (child ID) for the pseudocontrol (or case ID repeated).

4. **Set ID**. An ordinal number assigned to each set of pseudocons derived from the same case/parent trio.

If the option `-info-ma info.dat` is used then the text file `info.dat` also includes the following genotype information columns:

1. **Mother Genotypes**. Each of the SNPs will have 2 columns denoting the genotype given by the two allele names.

2. **Case ID**. Following the mother genotypes are the case genotypes.

If the option `-info-fa info.dat` is used then the text file `info.dat` also includes the following genotype information columns:

1. **Father Genotypes**. Each of the SNPs will have 2 columns denoting the genotype given by the two allele names.

2. **Case ID**. Following the father genotypes are the case genotypes.

If the option `-info-fama info.dat` is used then the text file `info.dat` also includes the following genotype information columns:

1. **Father Genotypes**. Each of the SNPs will have 2 columns denoting the genotype given by the two allele names.

2. **Mother Genotypes**. Following the father genotypes are the mother genotypes.

3. **Case ID**. Following the mother genotypes are the case genotypes.

The resultant text file may be very large if many SNPs are included, so a maximum number of 20 SNPs is imposed. This can be changed with the `-info-maxsnps` option. The log file includes information on the exact numbers of the columns for the genotypes. For example, if the `-info-ma info.dat` option were used with 10 SNPs the log file will include the lines:

```
...
Info file: info.dat
    Column 1: family ID
    Column 2: individual ID
    Column 3: corresponding case ID
    Column 4: pseudocontrol set ID
    Columns 5-24: mother genotype info
    Columns 25-44: child/pseudocontrol genotype info
...
```

The information file will then look something like the following (using `-pc1` with allele names 1 and 2):

```
10001 3 3 1 1 1 2 2 1 1 1 1 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 2 1 2 1 2 1 2 1 1 1 1 1 1
10001 3-pseudo-1 3 1 1 1 2 2 1 1 1 1 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 2 1 1
10002 3 3 2 1 1 2 2 1 1 1 1 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 2 1 2 1 2 1 2 1 1 1 1 1 1
10002 3-pseudo-1 3 2 1 1 2 2 1 1 1 1 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 2 1 2 1 2 1 2 1
10003 3 3 3 1 1 1 1 2 1 2 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
10003 3-pseudo-1 3 3 1 1 1 1 2 1 2 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 2 1 2 1 2 1
...
```

5 PseudoCons Examples

The different options of PseudoCons are demostrated using the example data set (included in the PseudoCons download) in the following examples. The example data set consists of 100 pedigrees where the first 50 are case/parent trios, the next 25 case/parent trios with an extra sibling, and the next 25 case/parent trios where the parents of the mother of the case is also included. There are 10 SNPs in the data set with allele names A and G.

5.1 One Pseudocontrol

To produce one pseudocontrol per case-parent trio type the following:

```
./pseudocons -i examplePsConsData.bed -o myCasePseudocontrols.bed
```

or

```
./pseudocons -pc1 -i examplePsConsData.bed -o myCasePseudocontrols.bed
```

This will create screen output similar to the following:

```
PseudoCons: pseudocontrols from pedigree data v1.0
--------------------------------------------------
Copyright 2013 Richard Howey, GNU General Public License, v3
Institute of Genetic Medicine, Newcastle University
```

```
Parameters:
Input file: examplePsConsData.bed
Output file: myCasePseudocontrols.bed
Log file: PseudoCons.log
Number of pseudocontrols per trio: 1


Number of subjects: 375
        Males: 188 (50.1333%)
        Females: 187 (49.8667%)
        Unknown sex: 0 (0%)
        Affected: 133 (35.4667%)
        Unaffected: 242 (64.5333%)
Number of SNPs: 10


Number of pedigrees: 100
Mean pedigree size: 3.75
Standard deviation of pedigree size: 0.833333


Number of trios used to create pseudocontrols: 100
Number of pedigrees with no pseudocontrols: 0


Number of cases: 100
        Males: 39 (39%)
        Females: 61 (61%)
        Unknown sex: 0 (0%)


Number of pseudocontrols: 100
        Males: 39 (39%)
        Females: 61 (61%)
        Unknown sex: 0 (0%)


Run time: less than one second
```

The screen output will also be saved to the log file, by default `PseudoCons.log`, but can be set using the `-log` option. The case/pseudocontrol files output are the binary pedigree plink files PLINK `myCasePseudocontrols.bed`, `myCasePseudocontrols.bim` and `myCasePseudocontrols.fam`. The created binary pedigree family file is as follows:

```
1 3 0 0 1 2
1 3-pseudo-1 0 0 1 1
2 3 0 0 2 2
2 3-pseudo-1 0 0 2 1
3 3 0 0 2 2
3 3-pseudo-1 0 0 2 1
4 3 0 0 1 2
4 3-pseudo-1 0 0 1 1
```

```
5 3 0 0 1 2
5 3-pseudo-1 0 0 1 1
6 3 0 0 1 2
6 3-pseudo-1 0 0 1 1
7 3 0 0 1 2
7 3-pseudo-1 0 0 1 1
...
```

The file consists of one case from each pedigree and one created pseudocontrol. The pedigree ID in the first column is repeated and the pseudocontrol subject ID is taken from the subject case ID with "pseudo-1" appended to it.

The created binary map file, `myCasePseudocontrols.bim`, is simply a repeat of the original binary map file since the used SNPs have not changed, which is:

```
1 rs1 0 1000 G A
1 rs2 0 2000 G A
1 rs3 0 3000 G A
1 rs4 0 4000 G A
2 rs5 0 10000 G A
2 rs6 0 20000 G A
2 rs7 0 30000 G A
3 rs8 0 16000 G A
3 rs9 0 32000 G A
3 rs10 0 48000 G A
```

5.2   Three Pseudocontrols

To produce three pseudocontrols per case-parent trio type the following:

```
./pseudocons -pc3 -i examplePsConsData.bed -o myCasePseudocontrols.bed
```

This will create screen output very similar to creating one pseudocontrol:

```
PseudoCons: pseudocontrols from pedigree data v1.0
--------------------------------------------------
Copyright 2013 Richard Howey, GNU General Public License, v3
Institute of Genetic Medicine, Newcastle University

Parameters:
Input file: releases/examplePsConsData.bed
Output file: myCasePseudocontrols3.bed
Log file: PseudoCons.log
Number of pseudocontrols per trio: 3

Number of subjects: 375
          Males: 188 (50.1333%)
          Females: 187 (49.8667%)
```

```
          Unknown sex: 0 (0%)
          Affected: 133 (35.4667%)
          Unaffected: 242 (64.5333%)
Number of SNPs: 10


Number of pedigrees: 100
Mean pedigree size: 3.75
Standard deviation of pedigree size: 0.833333


Number of trios used to create pseudocontrols: 100
Number of pedigrees with no pseudocontrols: 0


Number of cases: 100
          Males: 39 (39%)
          Females: 61 (61%)
          Unknown sex: 0 (0%)

Number of pseudocontrols: 300
          Males: 117 (39%)
          Females: 183 (61%)
          Unknown sex: 0 (0%)


Run time: less than one second
```

This time the created binary pedigree family file is as follows:

```
1 3 0 0 1 2
1 3-pseudo-1 0 0 1 1
1 3-pseudo-2 0 0 1 1
1 3-pseudo-3 0 0 1 1
2 3 0 0 2 2
2 3-pseudo-1 0 0 2 1
2 3-pseudo-2 0 0 2 1
2 3-pseudo-3 0 0 2 1
3 3 0 0 2 2
3 3-pseudo-1 0 0 2 1
3 3-pseudo-2 0 0 2 1
3 3-pseudo-3 0 0 2 1
4 3 0 0 1 2
4 3-pseudo-1 0 0 1 1
4 3-pseudo-2 0 0 1 1
4 3-pseudo-3 0 0 1 1
5 3 0 0 1 2
5 3-pseudo-1 0 0 1 1
5 3-pseudo-2 0 0 1 1
5 3-pseudo-3 0 0 1 1
```

. . .

The file consists of one case from each pedigree and three created pseudocontrols. The pedigree ID in the first column is repeated and the pseudocontrol subject IDs are taken from the subject case ID with "pseudo-1", "pseudo-2" and "pseudo-3" appended to it. Note that the sex of the case is repeated in the pseudocontrols.

As before the created binary map file, `.bim`, is the same.

5.3   Fifteen Pseudocontrols

To produce fifteen pseudocontrols per case-parent trio based on the non-transmitted allele combinations from two SNPs type the following:

`./pseudocons -pc15 -snpnames rs1 rs3 -i examplePsConsData.bed -o myCasePseudocontrols`

where the option `-snpnames rs1 rs3` picks the two SNPs to be consider using the SNP names. The SNPs can also be choosen by the order in which the SNPs appear in the file, so to choose the 1st and 3rd SNPs in the file type the following:

`./pseudocons -pc15 -snpnos 1 3 -i examplePsConsData.bed -o myCasePseudocontrols15.bed`

This will output to screen something similar to:

```
PseudoCons: pseudocontrols from pedigree data v1.0
--------------------------------------------------
Copyright 2013 Richard Howey, GNU General Public License, v3
Institute of Genetic Medicine, Newcastle University

Parameters:
Input file: examplePsConsData.bed
Output file: myCasePseudocontrols15.bed
Log file: PseudoCons.log
Interaction using SNP names rs1 and rs3
Number of pseudocontrols per trio: 15

Number of subjects: 375
        Males: 188 (50.1333%)
        Females: 187 (49.8667%)
        Unknown sex: 0 (0%)
        Affected: 133 (35.4667%)
        Unaffected: 242 (64.5333%)
Number of SNPs: 10

Number of pedigrees: 100
Mean pedigree size: 3.75
Standard deviation of pedigree size: 0.833333
```

```
Number of trios used to create pseudocontrols: 100
Number of pedigrees with no pseudocontrols: 0

Number of cases: 100
        Males: 39 (39%)
        Females: 61 (61%)
        Unknown sex: 0 (0%)

Number of pseudocontrols: 1500
        Males: 585 (39%)
        Females: 915 (61%)
        Unknown sex: 0 (0%)

Run time: less than one second
```

This time the created binary pedigree family file is as follows:

```
1 3 0 0 1 2
1 3-pseudo-1 0 0 1 1
1 3-pseudo-2 0 0 1 1
1 3-pseudo-3 0 0 1 1
1 3-pseudo-4 0 0 1 1
1 3-pseudo-5 0 0 1 1
1 3-pseudo-6 0 0 1 1
1 3-pseudo-7 0 0 1 1
1 3-pseudo-8 0 0 1 1
1 3-pseudo-9 0 0 1 1
1 3-pseudo-10 0 0 1 1
1 3-pseudo-11 0 0 1 1
1 3-pseudo-12 0 0 1 1
1 3-pseudo-13 0 0 1 1
1 3-pseudo-14 0 0 1 1
1 3-pseudo-15 0 0 1 1
2 3 0 0 2 2
2 3-pseudo-1 0 0 2 1
2 3-pseudo-2 0 0 2 1
2 3-pseudo-3 0 0 2 1
2 3-pseudo-4 0 0 2 1
2 3-pseudo-5 0 0 2 1
2 3-pseudo-6 0 0 2 1
2 3-pseudo-7 0 0 2 1
2 3-pseudo-8 0 0 2 1
2 3-pseudo-9 0 0 2 1
2 3-pseudo-10 0 0 2 1
2 3-pseudo-11 0 0 2 1
2 3-pseudo-12 0 0 2 1
```

```
2 3-pseudo-13 0 0 2 1
2 3-pseudo-14 0 0 2 1
2 3-pseudo-15 0 0 2 1
...
```

The file consists of one case from each pedigree and 15 created pseudocontrols, one for each pair of allele combinations from the two SNPs that were not transmitted. The pedigree ID in the first column is repeated and the pseudocontrol subject IDs are taken from the subject case ID with "pseudo-i" appended to it for the ith pseudocontrol. Note that the sex of the case is repeated in the pseudocontrols.

This time the created binary map file, `myCasePseudocontrols15.bim`, only consists of the two SNPs used to create the pseudocontrols.

```
1 rs1 0 1000 G A
1 rs3 0 3000 G A
```

The created binary pedigree, `myCasePseudocontrols15.bed`, also only consists of data with these two SNPs.

5.4   Proband

To choose which cases are chosen from a pedigree a proband file may be used as follows:

```
./pseudocons  -pro proband.dat -i examplePsConsData.bed -o myCasePseudocontrolsPro.be
```

The proband file is a list of pedigree IDs and subject IDs. The example proband file is as follows:

```
1 3
2 3
3 3
4 3
5 3
...
73 3
74 3
75 3
76 5
77 5
78 5
...
99 5
100 5
```

This will create screen output similar to the following:

```
PseudoCons: pseudocontrols from pedigree data v1.0
--------------------------------------------------
Copyright 2013 Richard Howey, GNU General Public License, v3
Institute of Genetic Medicine, Newcastle University

Parameters:
Input file: releases/examplePsConsData.bed
Output file: myCasePseudocontrolsPro.bed
Log file: PseudoCons.log
Number of pseudo controls per trio: 1
Proband file: releases/proband.dat

Number of subjects: 375
        Males: 188 (50.1333%)
        Females: 187 (49.8667%)
        Unknown sex: 0 (0%)
        Affected: 133 (35.4667%)
        Unaffected: 242 (64.5333%)
Number of SNPs: 10

Number of pedigrees: 100
Mean pedigree size: 3.75
Standard deviation of pedigree size: 0.833333

Number of trios used to create pseudo controls: 100
Number of pedigrees with no pseudo controls: 0

Number of cases: 100
        Males: 48 (48%)
        Females: 52 (52%)
        Unknown sex: 0 (0%)

Number of pseudo controls: 100
        Males: 48 (48%)
        Females: 52 (52%)
        Unknown sex: 0 (0%)

Run time: less than one second
```

Note that the number of males and females are different to previous due to different cases being chosen. The sex ratio is about 1 due to the proband file ensuring that affect offspring are chosen rather than affected mothers, which is poosible for the last group of pedigrees where the parents of the mother are also included.

For more on probands see section 4.1.

5.5 Extra Trios

It is possible to use more than one case/parent trio from each pedigree by using the
-xtrio as follows:

```
./pseudocons -xtrio -i examplePsConsData.bed -o myCasePseudocontrolsX.bed
```

This will create screen output similar to the following:

```
PseudoCons: pseudocontrols from pedigree data v1.0
--------------------------------------------------
Copyright 2013 Richard Howey, GNU General Public License, v3
Institute of Genetic Medicine, Newcastle University


Parameters:
Input file: releases/examplePsConsData.bed
Output file: myCasePseudocontrolsPro.bed
Log file: PseudoCons.log
Number of pseudocontrols per trio: 1
Allowing extra trios

Number of subjects: 375
        Males: 188 (50.1333%)
        Females: 187 (49.8667%)
        Unknown sex: 0 (0%)
        Affected: 133 (35.4667%)
        Unaffected: 242 (64.5333%)
Number of SNPs: 10

Number of pedigrees: 100
Mean pedigree size: 3.75
Standard deviation of pedigree size: 0.833333

Number of trios used to create pseudocontrols: 133
Number of pedigrees with no pseudocontrols: 0

Number of cases: 133
        Males: 58 (43.609%)
        Females: 75 (56.391%)
        Unknown sex: 0 (0%)

Number of pseudocontrols: 133
        Males: 58 (43.609%)
        Females: 75 (56.391%)
        Unknown sex: 0 (0%)

Run time: less than one second
```

Note that in this output there are 118 case/parent trios used to create the pseudocontrol data, but only 100 pedigrees. The extra 33 trios were taken from pedigrees containing more than one case/parent trio. This option will taken as many case/parent trios from one pedigree as possible, but for this example data set takes no more than 2 per pedigree. The number of pedigrees with no available case/parent trios are also reported, which for this data example data set is zero.

Extra care should be taken in interpreting any subsequent analysis using this option.

References

Cordell HJ. 2004. Properties of case/pseudocontrol analysis for genetic association studies: Effects of recombination, ascertainment, and multiple affected offspring. Genet Epidemiol 26:186–205.

Cordell HJ, Barratt BJ, Clayton DG. 2004. Case/pseudocontrol analysis in genetic association studies: a unified framework for detection of genotype and haplotype associations, gene-gene and gene-environment interactions and parent-of-origin effects. Genet Epidemiol 26:167–185.

Weinberg CR, Shi M. 2009. The genetics of preterm birth: Using what we know to design better association studies. Am J Epidemiol 170:1373–1381.