Disease Model Test Harness

# Motivation

The disease risk calculator program calculates disease risk for pedigree members according to

* their personal attributes
* a model for the disease in question, incorporating personal attributes that are disease risk factors, heritability, prevalence etc. of the disease.

The disease model is specified in a disease model file. The disease risk calculator program interprets this file to obtain a disease model. This document describes testing done to ensure the correctness of the disease risk calculator program’s disease model file interpretation.

# Introduction

A set of disease model file test cases are created. They may be valid or invalid. A program known as a test harness uses the same procedure as the disease risk calculator program to interpret each of these disease model files. The output produced allows one to check whether the disease model file interpretation is working as expected. The test harness is called DisFileReaderTestHarness.R.

This is an R program, designed to be run from the command line.

It takes as its input the name of a disease model file.

It reads this disease model file to extract the disease model.

If any problems are detected, the program reports the problem and stops.

The test disease model files are held in directory – ValidateDiseaseModel

This contains two sub-directories

* good – contains disease model files that should run without any errors being reported
* bad – contains problem disease model files, a meaningful error should be reported

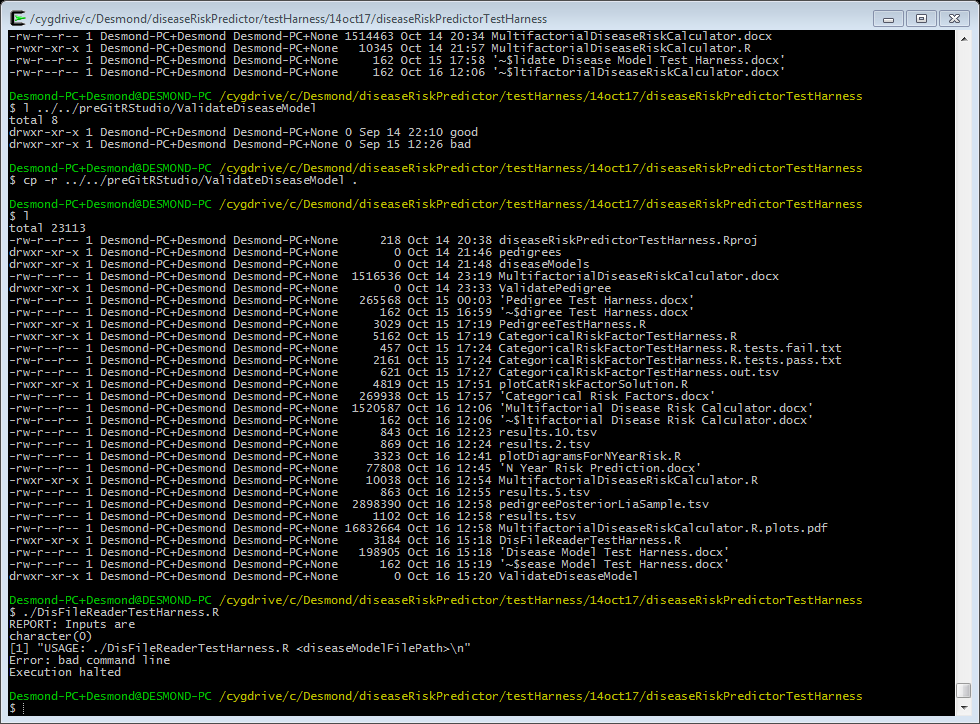
# Check the Tester Harness is Working

Open a command line prompt.

Go to the directory containing the test harness program

Enter the following at the command line prompt

./DisFileReaderTestHarness.R

The program should report a USAGE statement telling you what constitutes a valid command line for the program. This is what that looked like when run on my machine from a Cygwin prompt 

If you do not get a USAGE statement, then the program probably did not run. The following requirements must be met for the program to run

* the file DisFileReaderTestHarness.R needs execute permissions
* R must be on the PATH

# Testing a disease model file

First we are going to test a valid disease model file.

Enter the following at the command prompt

./DisFileReaderTestHarness.R ValidateDiseaseModel/good/dm.Con1.txt

You should see output like the following

$ ./DisFileReaderTestHarness.R ValidateDiseaseModel/good/dm.Con1.txt

REPORT: Inputs are

[1] "ValidateDiseaseModel/good/dm.Con1.txt"

REPORT: Read disease model from file - ValidateDiseaseModel/good/dm.Con1.txt

REPORT: fnReadFileDis()

...

REPORT: Disease model is

$heritability

[1] 0.6

$lifetimeRisk

[1] 0.05

$title

[1] "1. Single Stratum Congenital Disease"

$description

[1] "This single stratum disease is congenital (expressed from birth)."

[2] "Population variance in disease liability is explained by genes and non-shared environmental effects."

[3] "This is the simplest disease model possible in this program."

$liaThr

[1] 1.644854

$vsuAllRF

[1] "liability"

$aCov

, , Tot

liability

liability 1

, , Asq

liability

liability 0.6

, , Csq

liability

liability 0

, , Esq

liability

liability 0.4

REPORT: Completed successfully

The final line says the test harness program completed successfully.

The disease model, generated through interpretation of the disease model file, is reported in the lines following the line

REPORT: Disease model is

Compare this to the contents of the disease model file - ValidateDiseaseModel/good/dm.Con1.txt.

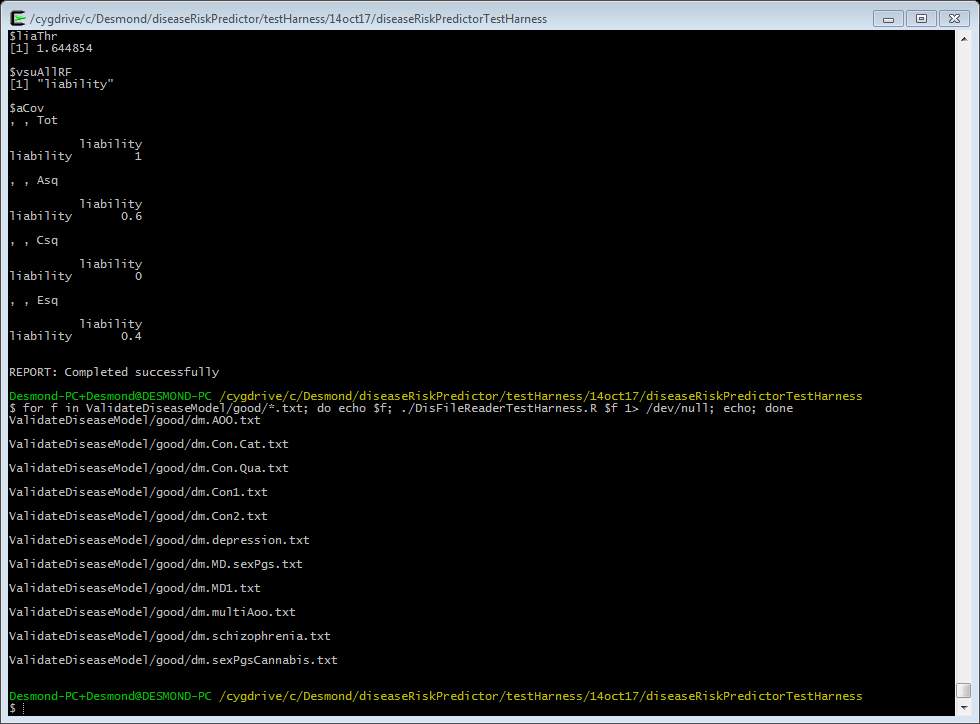
# Testing the good disease model files

Testing each disease model file individually is a bit tedious; let’s automate that. We will apply the following procedure to each disease model file in the ValidateDiseaseModel\good directory

* report file name
* invoke the test harness program on the disease model file
* discard all standard (i.e. non-error) output

Enter the following command at the command prompt.

for f in ValidateDiseaseModel/good/\*.txt; do echo $f; ./DisFileReaderTestHarness.R $f 1> /dev/null; echo; done

You should see output like 

You should see reported the name of each disease model file tested, and nothing else. All these disease model files should be valid. If anything else is reported there is a problem.

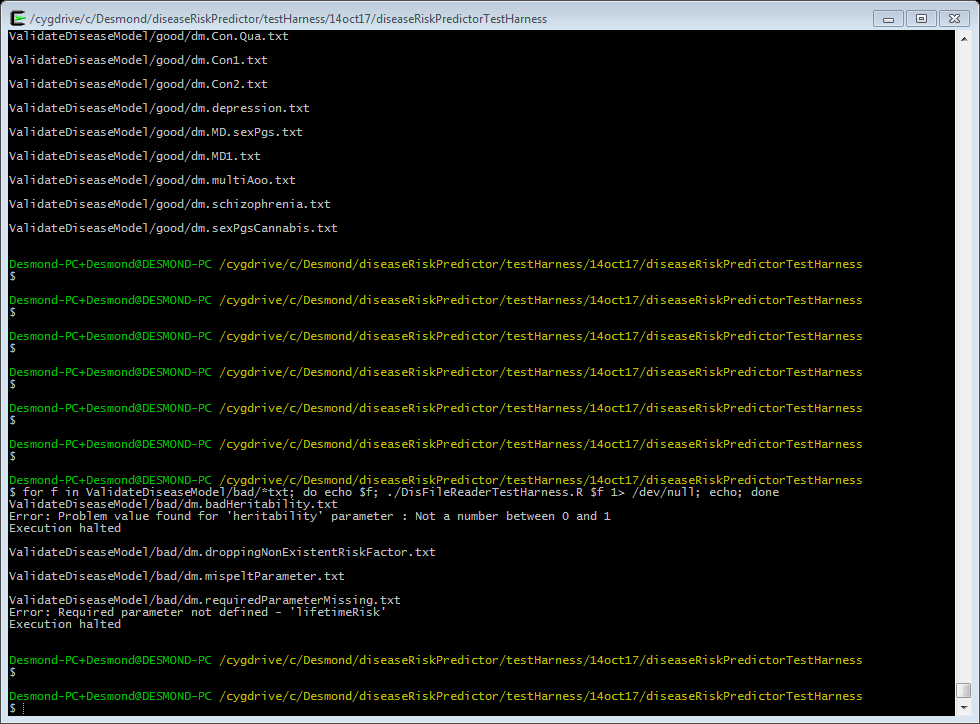
In fact all we have done is seen whether any of these files produce an error. Whether each disease model file has been correctly interpreted can only be done by comparing the test harness output with the original disease model file.

# Testing the bad disease model files

We can do the same for the bad disease model files. These are disease model files constructed so as to be invalid in some way.

Enter the following command at the command prompt.

for f in ValidateDiseaseModel/bad/\*txt; do echo $f; ./DisFileReaderTestHarness.R $f 1> /dev/null; echo; done

You should see output like 

The test harness should report the correct error for each disease model file. Check that the error reported makes sense given the name of the disease model file.