WinBUGS : part 1

Bruno Boulanger
Jonathan Jaeger
Astrid Jullion
Philippe Lambert

Agenda

- Introduction to WinBUGS
- Exercice 1 : Normal with unknown mean and variance
  - Example description
  - Model specification in WinBUGS
  - Data and initial values in WinBUGS
  - Process in WinBUGS
  - Results in WinBUGS
- Exercice 2 : Comparison of 2 means-unequal variance

WinBUGS

- Free software :

  http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml

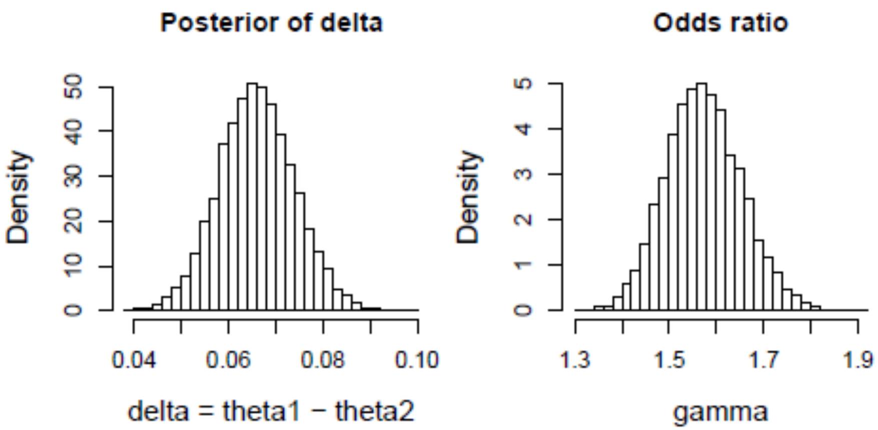
- WinBUGS implements various MCMC (Markov Chain Monte Carlo) algorithms to generate samples from posterior distributions.

- WinBUGS: Windows version of the BUGS program (Bayesian analysis using Gibbs Sampling).

WinBUGS


```
par(mfrow=c(1,2)) # Split window in 2
M = 10000 ; theta1 = rbeta(M,684,2538) ; theta2 = rbeta(M,1499,8748)
odds1 = theta1/(1-theta1) ; odds2 = theta2/(1-theta2)
delta = theta1-theta2 ; gam = odds1/odds2
hist(delta,breaks=25,freq=F,xlab="delta = theta1 - theta2",
     main="Posterior of delta")
round(mean (delta),3) ; round(quantile(delta,c(.025,.05,.5,.95,.975)),3)
hist(gam,breaks=25,freq=F,xlab="gamma",main="Odds ratio")
round(mean(gam),3) ; round(quantile(gam,c(.025,.05,.5,.95,.975)),3)
```

Comparison of 2 proportions



- Mean and quantiles of δ:

```
[1] 0.066
 2.5%    5%   50%   95% 97.5%
0.051 0.053 0.066 0.079 0.082
```

- Mean and quantiles of γ:

```
[1] 1.574
 2.5%    5%   50%   95% 97.5%
1.422 1.447 1.572 1.712 1.737
```

- Identified posterior distributions are not needed in WinBUGS.

Example description

- Assume that it is known from large studies that the mean cholesterol level in children aged 2-14 is 175 mg/dL. It is desirable to determine if there is familial aggregation of cholesterol levels.
- A group of fathers who had heart attack and elevated cholesterol levels (250 mg/dL) was identified and the cholesterol level of their offspring aged 2-14 was measured.
- The cholesterol levels of such **100** children have mean **207**mg/dL and standard deviation **30** mg/dL.
- Is there evidence that the cholesterol in that subpopulation is different from 175 mg/dL? (Rosner 2000, p.184).

Example description

- By experience, it is known that the cholesterol levels are approximately normally distributed.

- Therefore, we shall assume that $Y_i \sim N(\mu;\sigma^2)$ $(i = 1,\ldots,100)$ where Y_i denotes the cholesterol level of the ith child.

- Precision $\tau = 1/\sigma^2$

Example description

- **<u>Prior for μ</u>** : $\mu \sim N(\mu_0, \sigma^2_0)$

  A large variance σ^2_0 would translate a large uncertainty on the values of μ_0.

- **<u>Prior for τ</u>** : $\tau \sim G(a,b)$ with mean a/b and variance a/b^2

  A large prior variance would translate uncertainty on the plausible values for τ. A large prior variance can be obtained by taking a = b = 0.0001 (say).

Model specification in WinBUGS

Model specification in WinBUGS

To specify a Bayesian model, we need:

- **The likelihood:** $Y_i \sim N(\mu;\sigma^2)$ (i = 1,…,100)

  We use: $\tau = 1/\sigma^2$ (where τ is the precision).

- **Prior distribution of the model parameters:** μ; τ
  - We consider non-informative prior distributions:

    $\mu \sim N(200,10^4)$

    tau~ Gamma(0.0001,0.0001)

Model specification in WinBUGS

- Some common univariate distributions in WinBUGS:

dbin	binomial	r ~ dbin(p,n)
dnorm	normal	x ~ dnorm(mu, **tau**)
dlnorm	log-normal	x ~ dlnorm(mu, **tau**)
dpois	poisson	r ~ dpois(lambda)
dunif	uniform	x ~ dunif(a,b)
dgamma	gamma	x ~ dgamma(a,b)
dbeta	beta	x ~ dbeta(a,b)

Model specification in WinBUGS

Some rules :

- For loop: for (i in 1:n){ }
- Indice: y[i]
- Distribution: y[i] ~ dnorm(mu[i],tau)
- Assignment: variance <- 1/tau (computed in a deterministic way)

Model specification in WinBUGS

- **The likelihood:**

```
for (i in 1:n){
    y[i]~dnorm(mu,tau)}
```

- **Prior distribution of the model parameters:**

```
mu ~ dnorm (200 , 0.0001)
tau ~ dgamma (0.0001, 0.0001)
```

Model specification in WinBUGS


```
model{

#Likelihood
   for (i in 1:n){
      y[i]~dnorm(mu,tau)
                  }

#Prior
   mu~dnorm(200,0.0001)
   tau~dgamma(0.0001,0.0001)

#Computation of the standard deviation
   sd<-1/sqrt(tau)

}
```

Data and initial values in WinBUGS

Gabriele,
living with rheumatoid arthritis

Data and initial values

- **DATA**

- Make sure to provide all the data in the model specification.

```
for (i in 1:n){
    y[i]~dnorm(mu,tau)
            }
  #Prior
    mu~dnorm(200,0.0001)
    tau~dgamma(0.0001,0.0001)

  #Computation of the standard deviation
    sd<-1/sqrt(tau)
```

Data and initial values

DATA

```
list(y=c(2.51558E+02, 2.25099E+02, 1.66668E+02, 1.68022E+02, 2.19820E+02,
    1.77079E+02, 2.49650E+02, 1.87267E+02, 2.02742E+02, 2.20882E+02, 1.98411E
    +02, 1.65022E+02, 2.30645E+02, 2.32637E+02, 2.28971E+02, 2.11241E+02,
    1.54139E+02, 1.53285E+02, 1.96436E+02, 2.06299E+02, 1.93195E+02, 2.13863E
    +02, 1.47117E+02, 2.86469E+02, 1.91467E+02, 2.34219E+02, 1.71060E+02,
    1.72452E+02, 1.77047E+02, 2.02564E+02, 2.53186E+02, 2.24139E+02, 1.95245E
    +02, 1.50038E+02, 2.21593E+02, 2.05437E+02, 1.76135E+02, 1.92493E+02,
    2.09831E+02, 1.77484E+02, 1.99355E+02, 2.21080E+02, 2.15441E+02, 2.76884E
    +02, 2.10186E+02, 2.13041E+02, 2.17956E+02, 2.51468E+02, 2.25837E+02,
    2.23581E+02, 2.59690E+02, 1.96986E+02, 2.30439E+02, 2.13505E+02, 1.70554E
    +02, 2.04125E+02, 2.21069E+02, 2.29274E+02, 2.55735E+02, 2.21125E+02,
    2.28745E+02, 2.20769E+02, 2.12672E+02, 2.02715E+02, 2.01250E+02, 2.51791E
    +02, 2.17497E+02, 2.26841E+02, 1.95284E+02, 1.95802E+02, 1.89577E+02,
    2.36446E+02, 1.79112E+02, 2.12935E+02, 1.93293E+02, 1.84941E+02, 1.48437E
    +02, 2.22016E+02, 2.22240E+02, 2.10760E+02, 1.93071E+02, 1.90897E+02,
    1.69097E+02, 2.04470E+02, 1.68452E+02, 2.70150E+02, 2.18594E+02, 1.54905E
    +02, 2.16406E+02, 2.21260E+02, 2.06473E+02, 1.64567E+02, 2.61956E+02,
    1.71259E+02, 2.11077E+02, 2.35920E+02, 2.18069E+02, 1.57796E+02, 2.22255E
    +02, 2.05355E+02), n=1.00000E+02)
```

Data and initial values

- **Inital values**

- Make sure to provide initial values for all model parameters.

```
for (i in 1:n){
        y[i]~dnorm(mu,tau)
                }
#Prior
   mu~dnorm(200,0.0001)
   tau~dgamma(0.0001,0.0001)

#Computation of the standard deviation
   sd<-1/sqrt(tau)
```

Data and initial values

- **Inital values**

  list(tau=1.00000E+00, mu=2.00000E+02)

Process steps

Gabriele,
living with rheumatoid arthritis

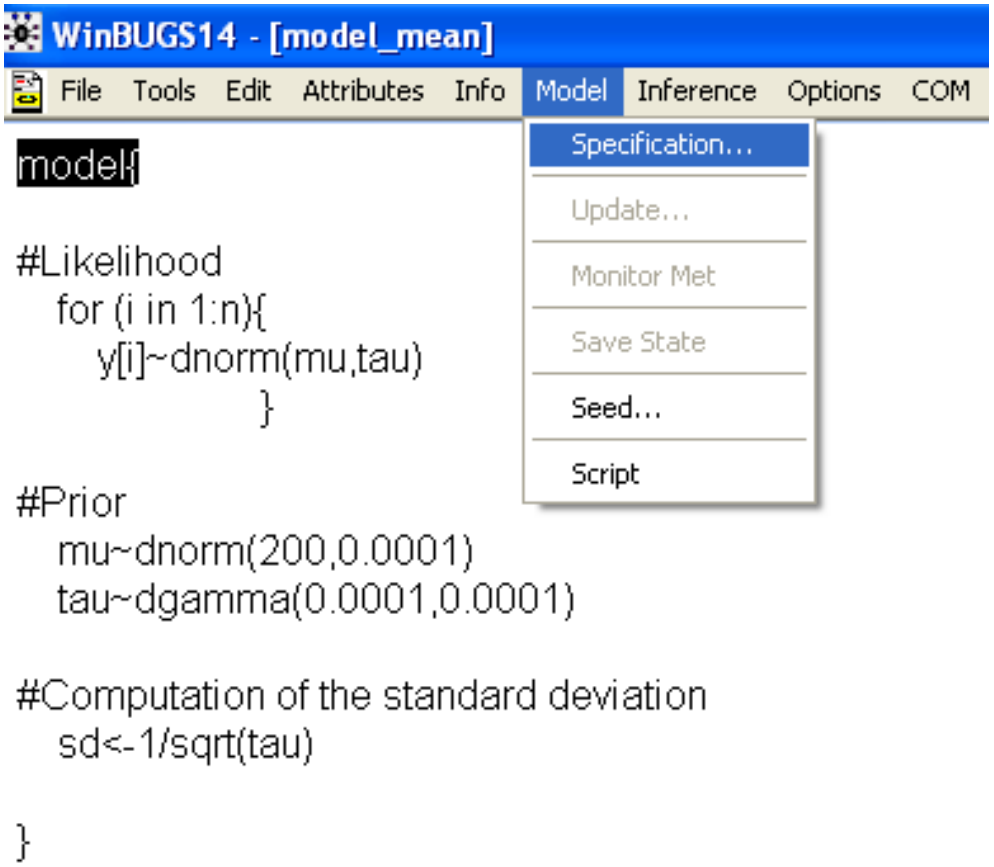
Process steps

- 1- check model
- 2- load data
- 3- compile model
- 4- load initial values
- 5- generate burn-in values
- 6- parameters to be monitored
- 7- perform the sampling to generate samples from the posteriors
- 8- display results

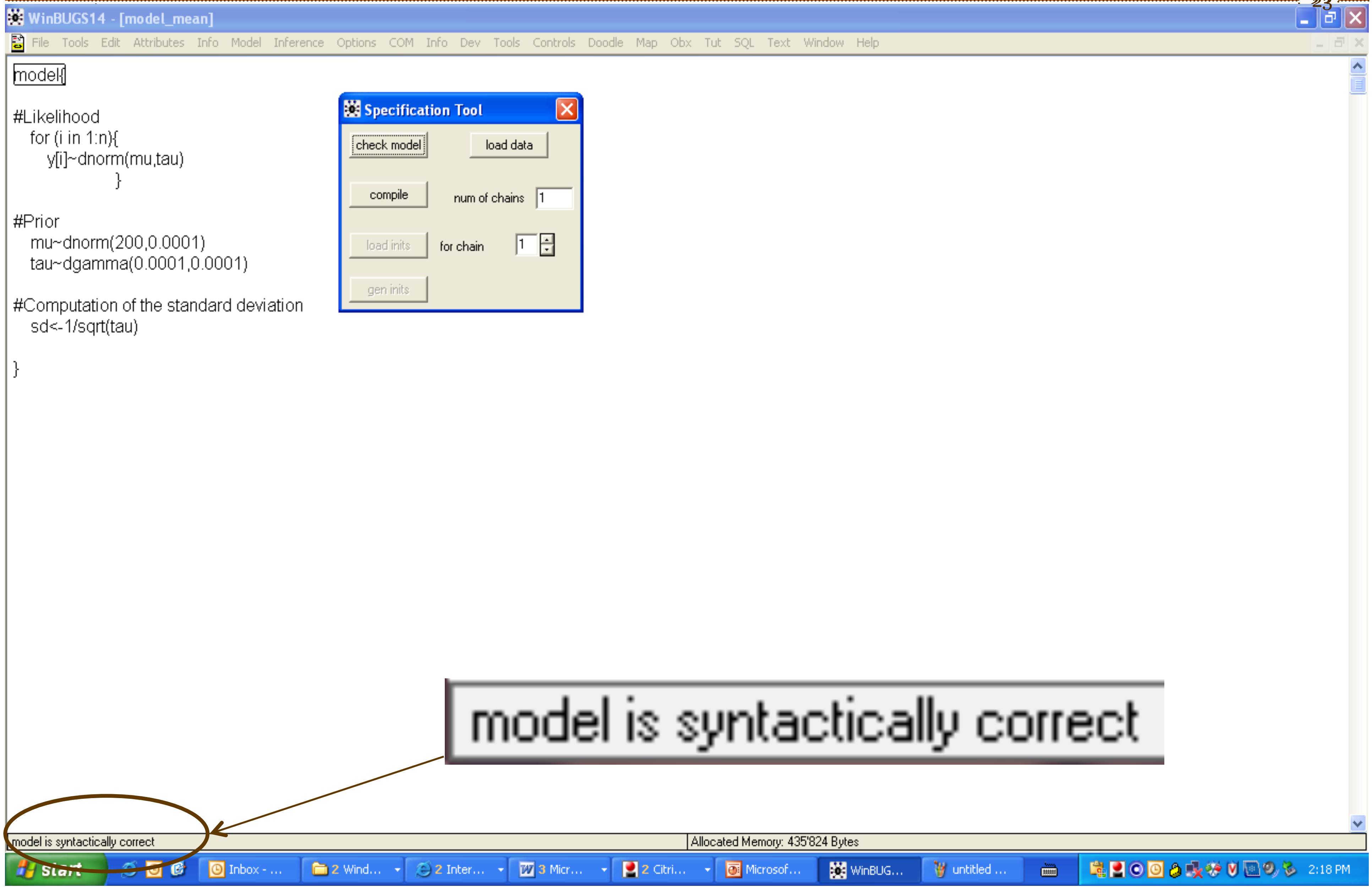
1. Check model

WinBUGS: step 1

- Highlight the word "model"
- Go to *Model -> Specification -> check model*



WinBUGS: step 1

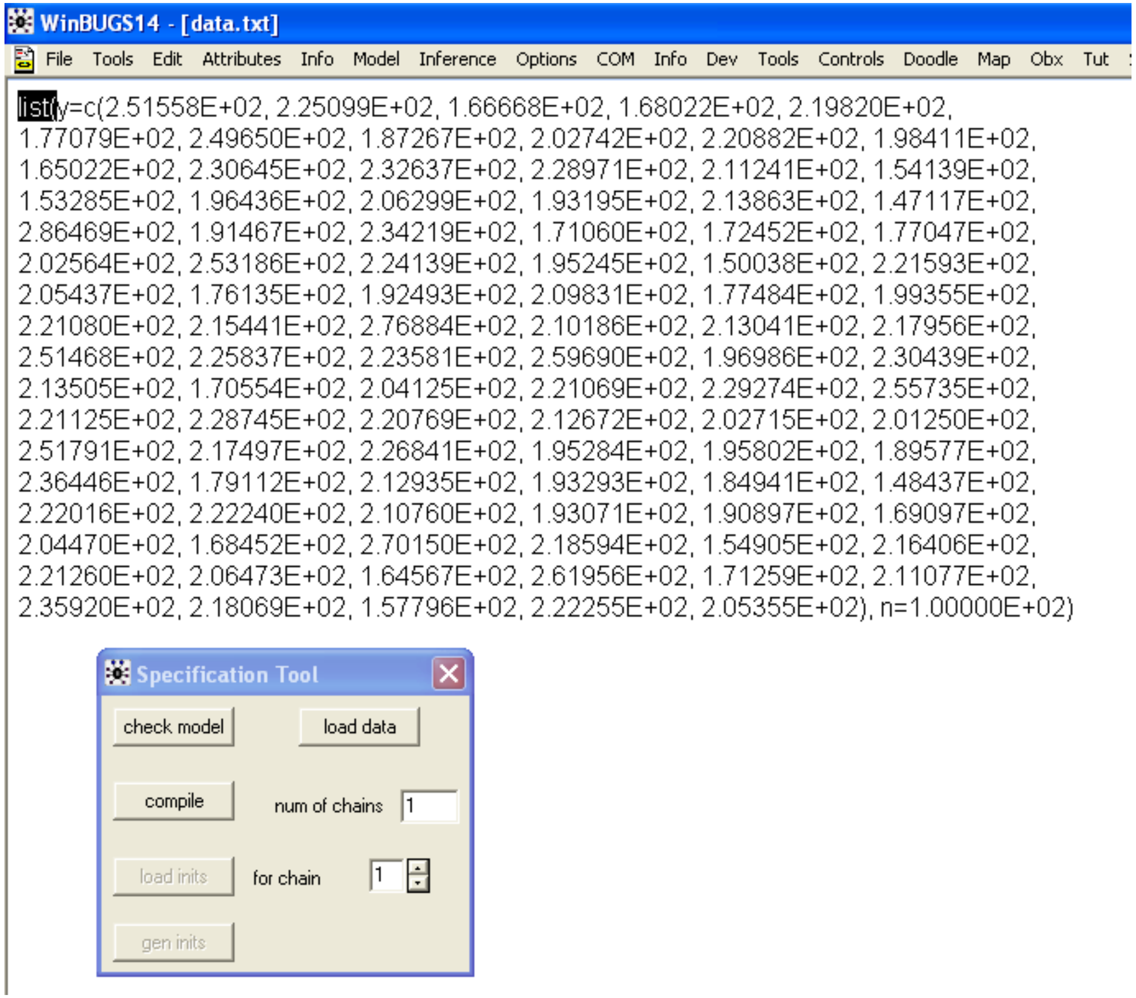


2. Load data

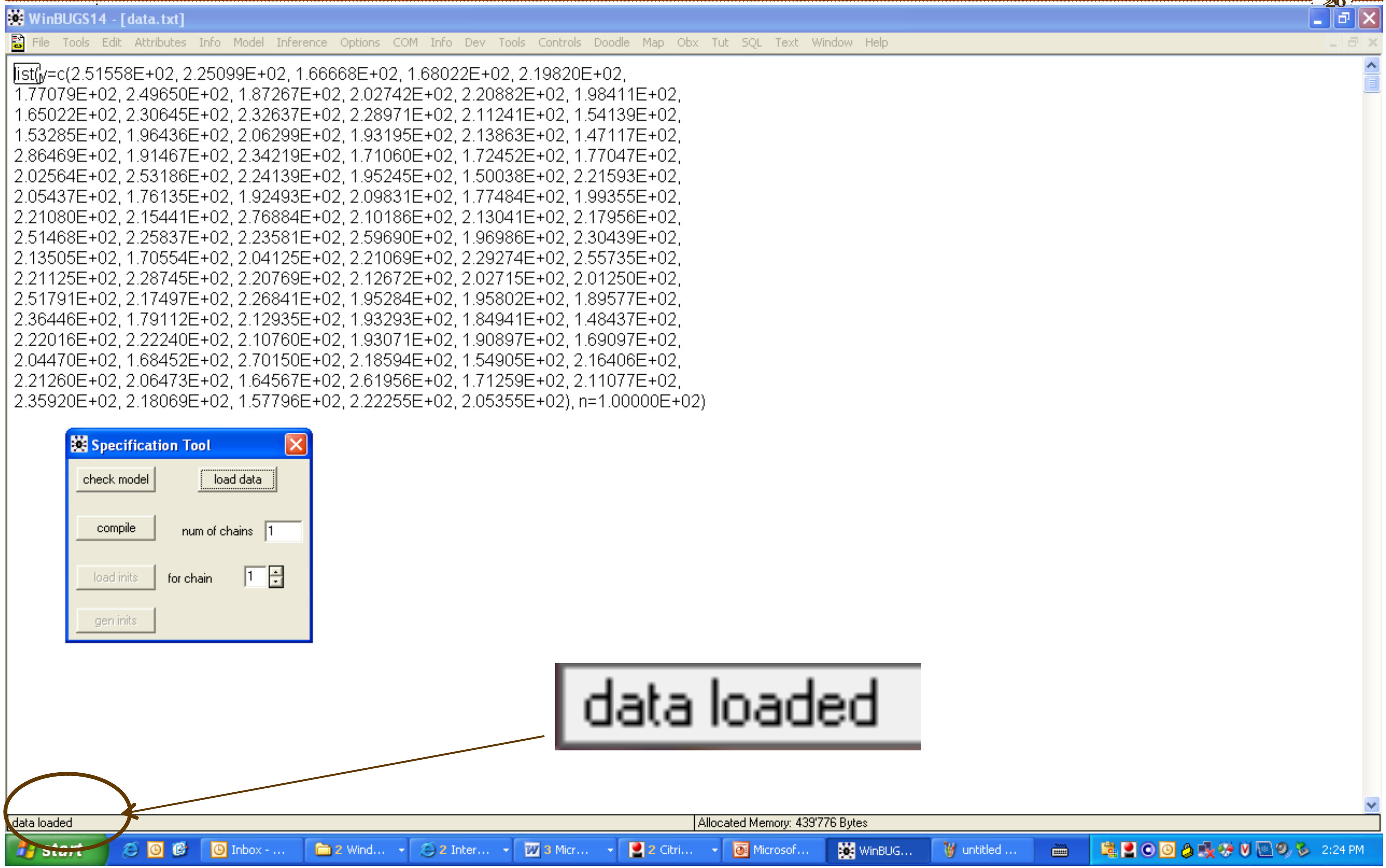
Gabriele,
living with rheumatoid arthritis

WinBUGS: step 2

- Open your data file (*file -> open*) [.txt file]
- Highlight the word "list" and click on "load data"



WinBUGS : step 2



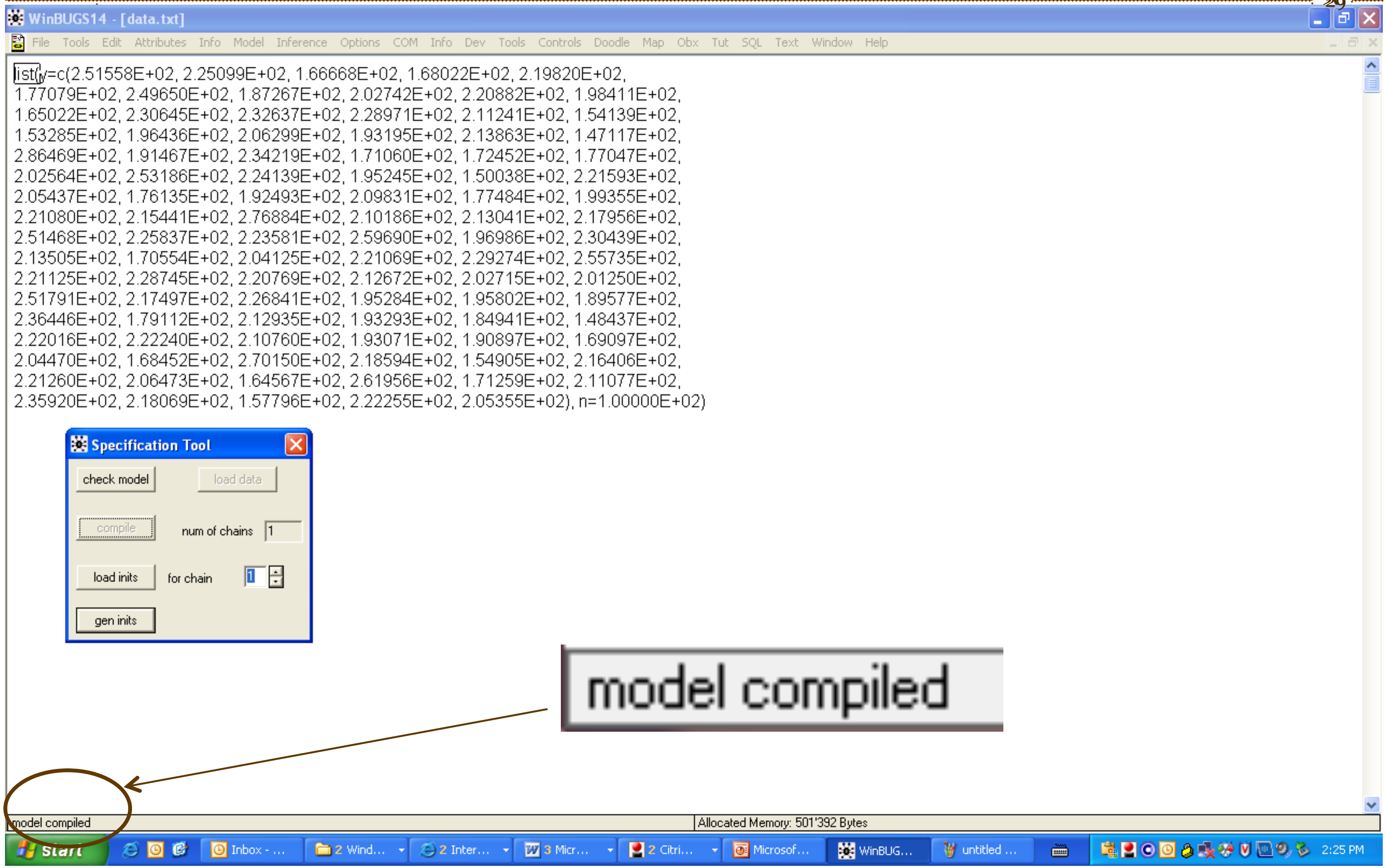
3. Compile

Gabriele,
living with rheumatoid arthritis

WinBUGS: step 3

- Click on “compile”

WinBUGS: step 3

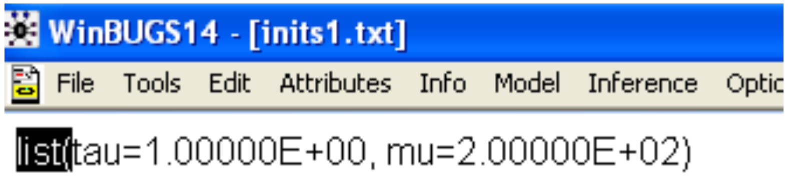


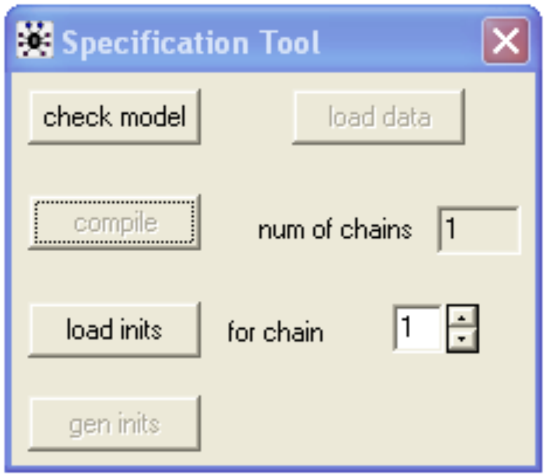
4. Load initial values

Gabriele,
living with rheumatoid arthritis

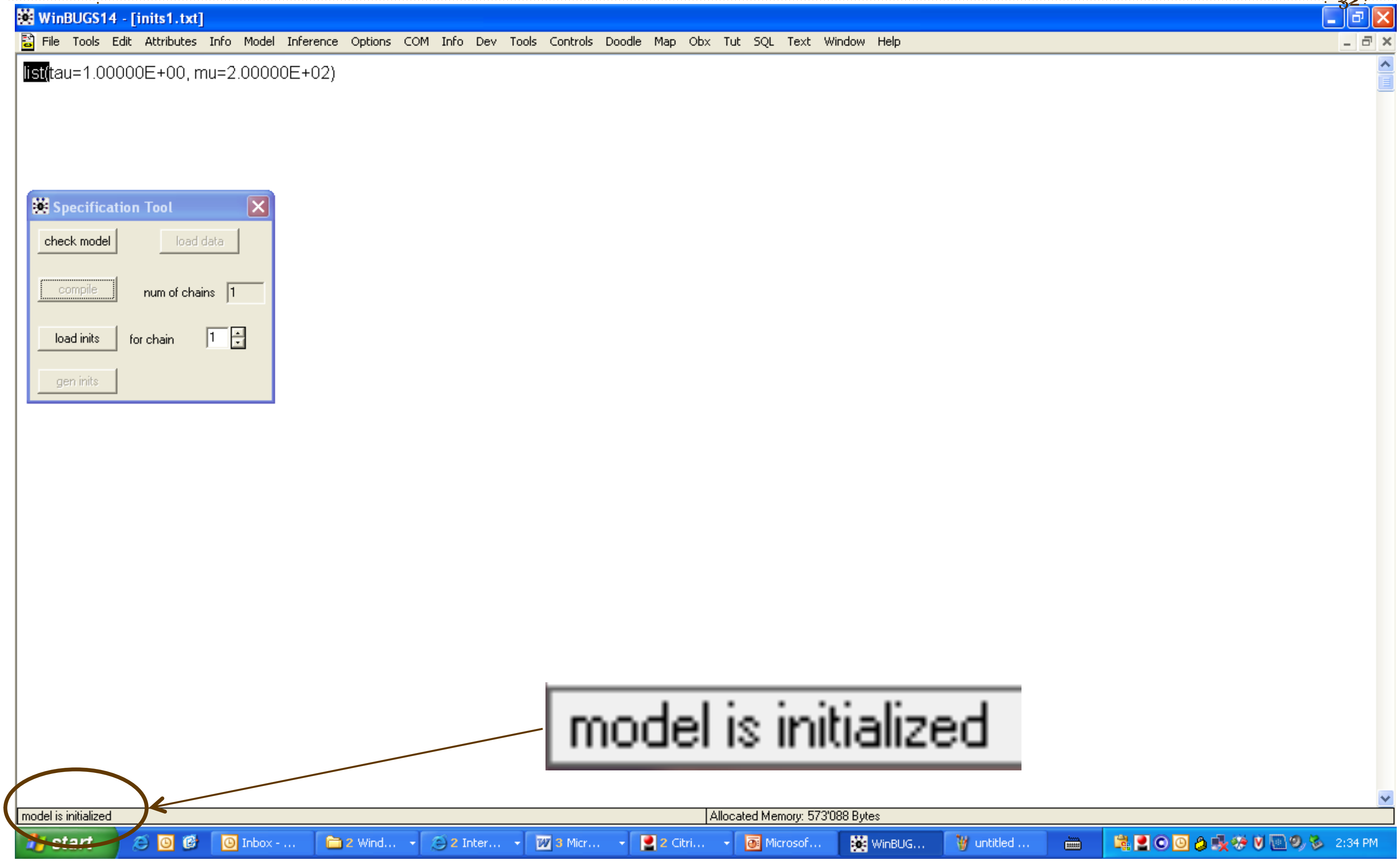
WinBUGS: step 4

- Open your init file (*file -> open*)   [.txt file]
- Highlight the word "list" and click on "load inits"





WinBUGS: step 4



WinBUGS: step 4

- If you want to run several chains, you have to specify initial values for each chain.

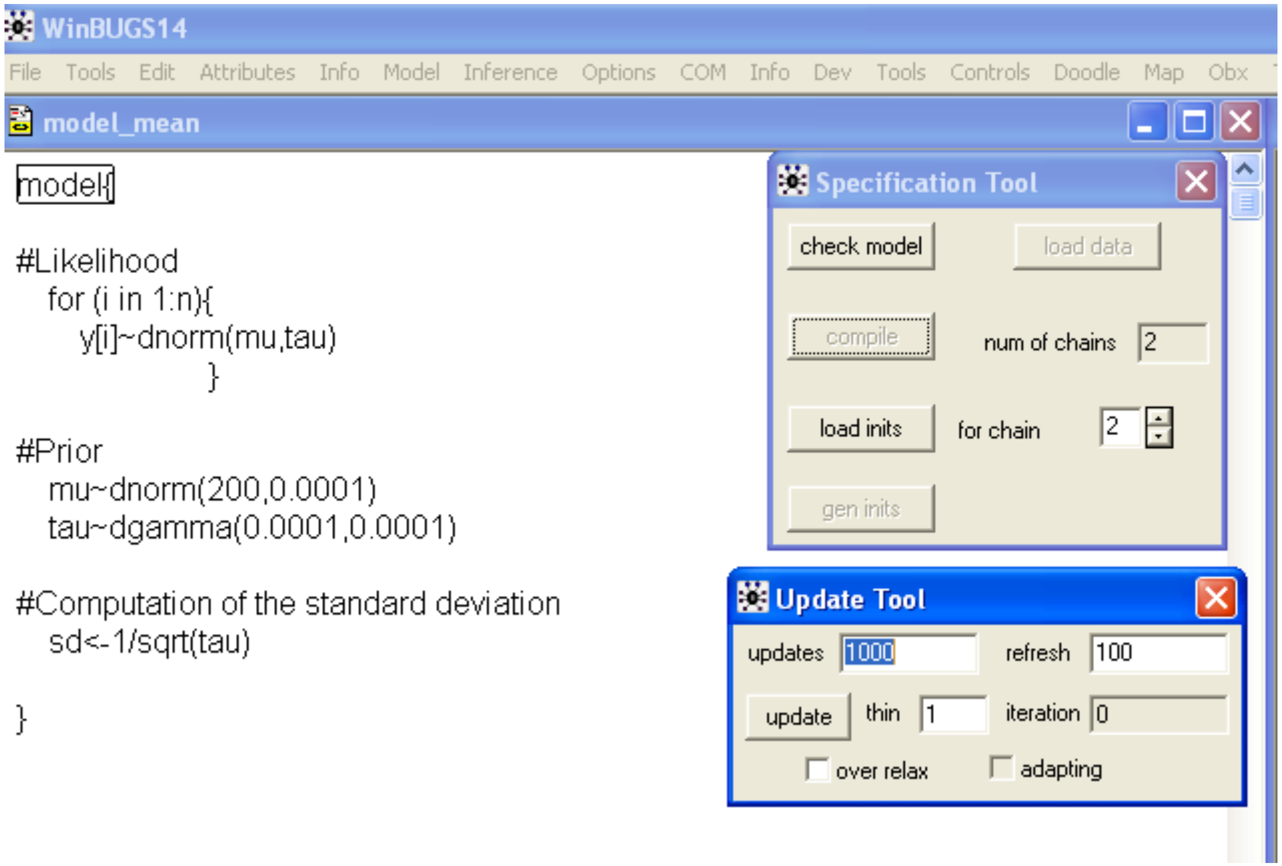
- In WinBUGS, give the number of chains after 'load data' and **before 'compile**'

5. Generate burn-in values


Gabriele,
living with rheumatoid arthritis

WinBUGS: step 5

- Go into *Model -> Update*
- Specify the burn-in
- Click on update

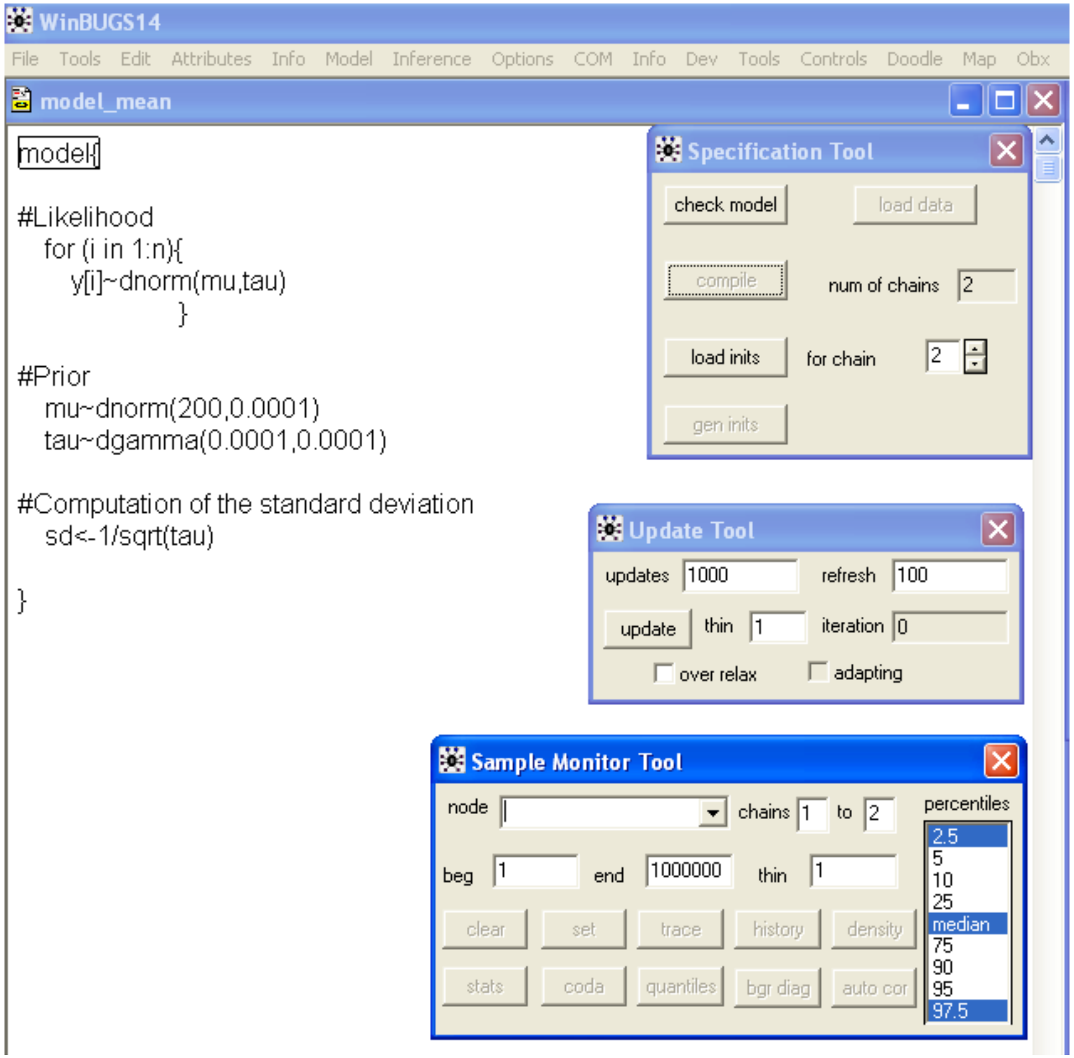


6. Parameters to monitor

Gabriele,
living with rheumatoid arthritis

WinBUGS: step 6

- Go into *Inference -> Samples*



WinBUGS: step 6

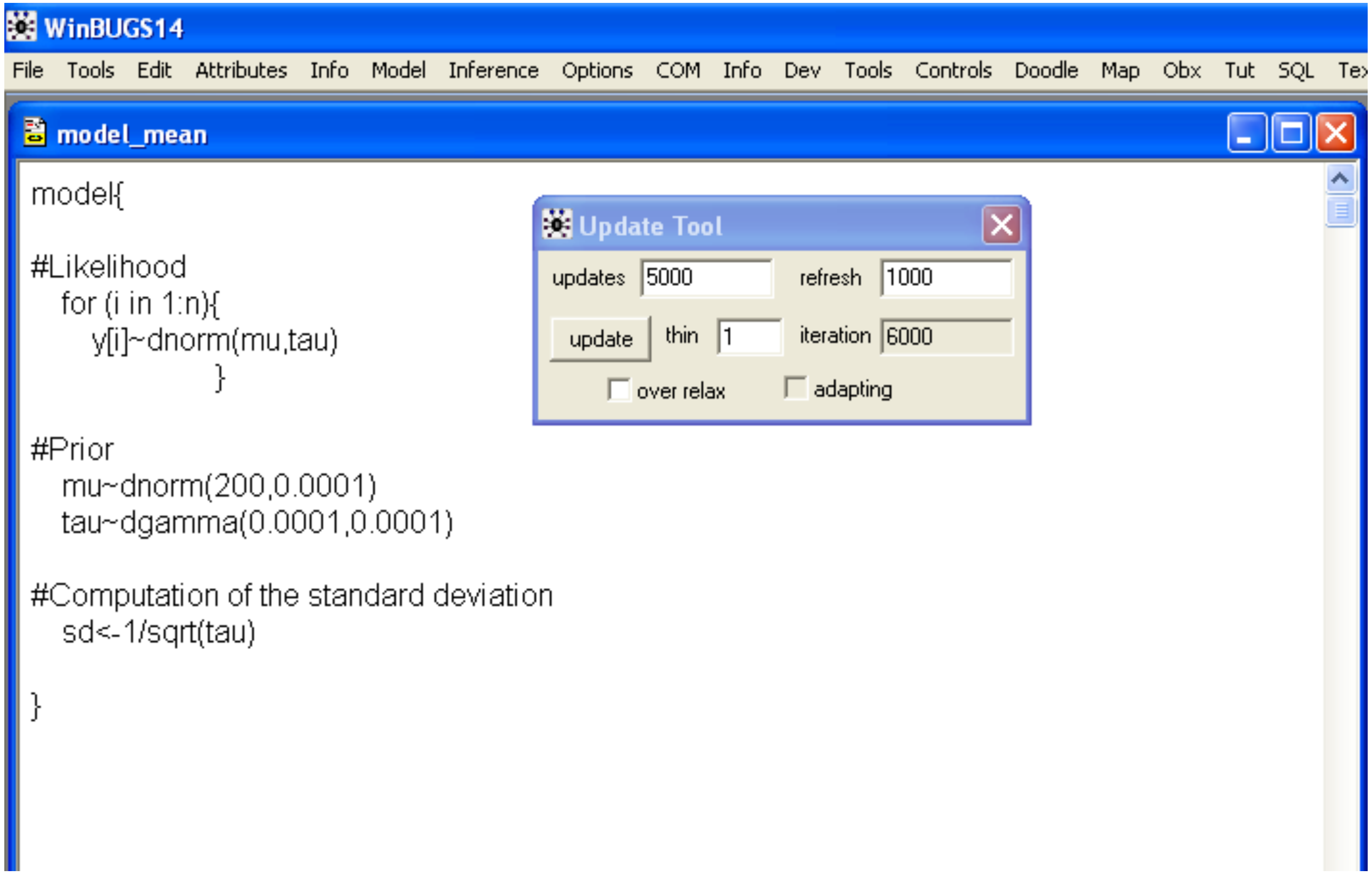
- In "node", write the parameter you want to monitor and click on "set".
- Repeat for all the parameters you want to monitor: mu, tau and sd.
- When done, write an asterisk

7. Perform the sampling to generate samples from posteriors


Gabriele,
living with rheumatoid arthritis

WinBUGS: step 7

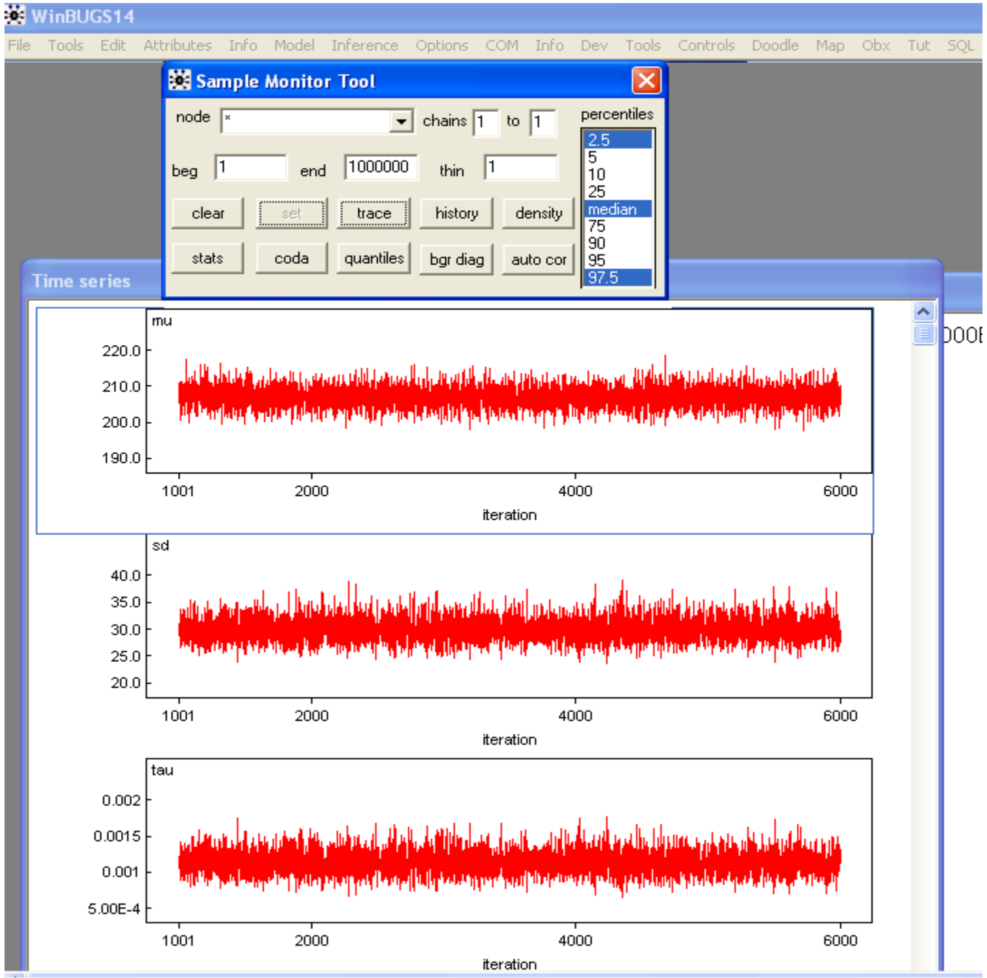
- In the "update tool", write the number of iterations and click on "update"



8. display the results

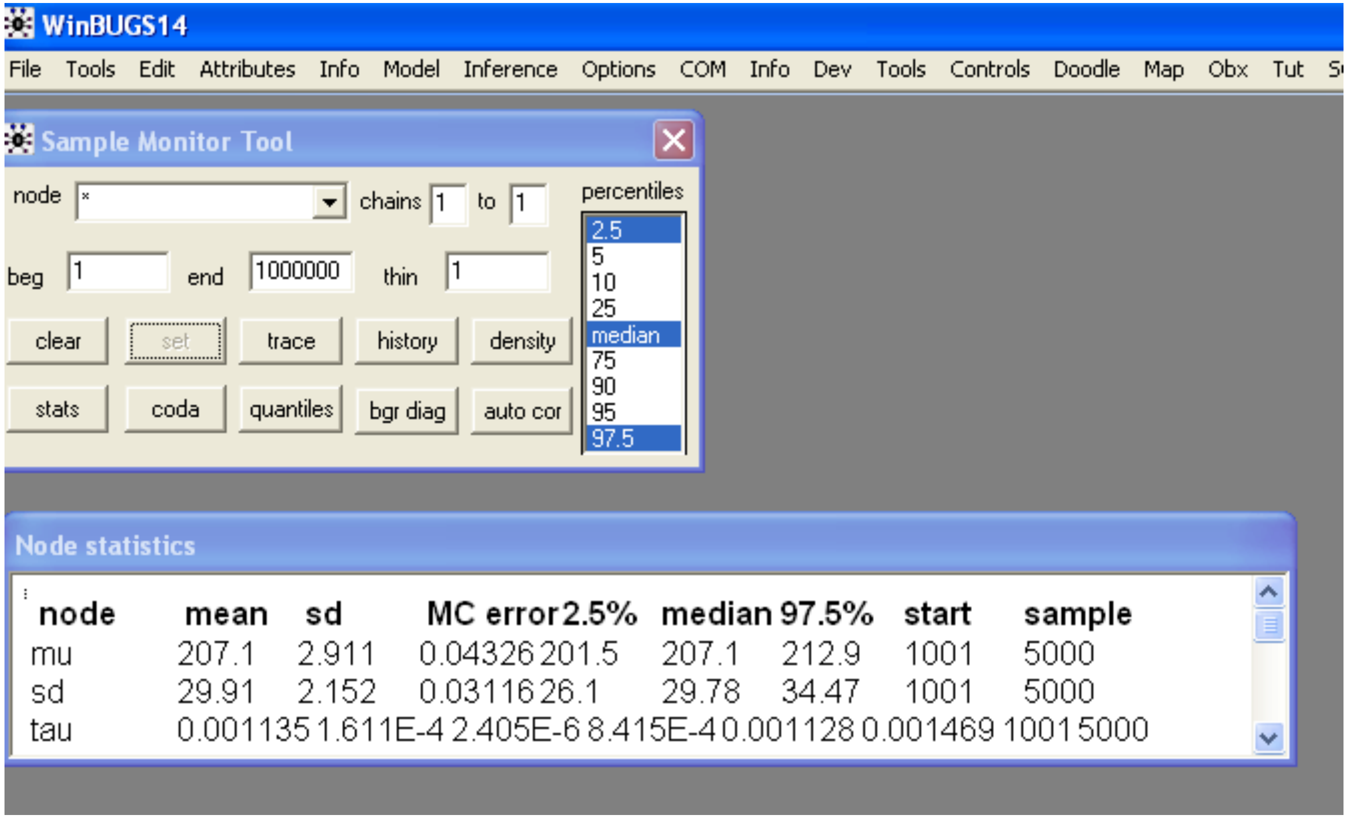
WinBUGS: step 8

- In the “sample monitor tool”, click on history to get the traces:



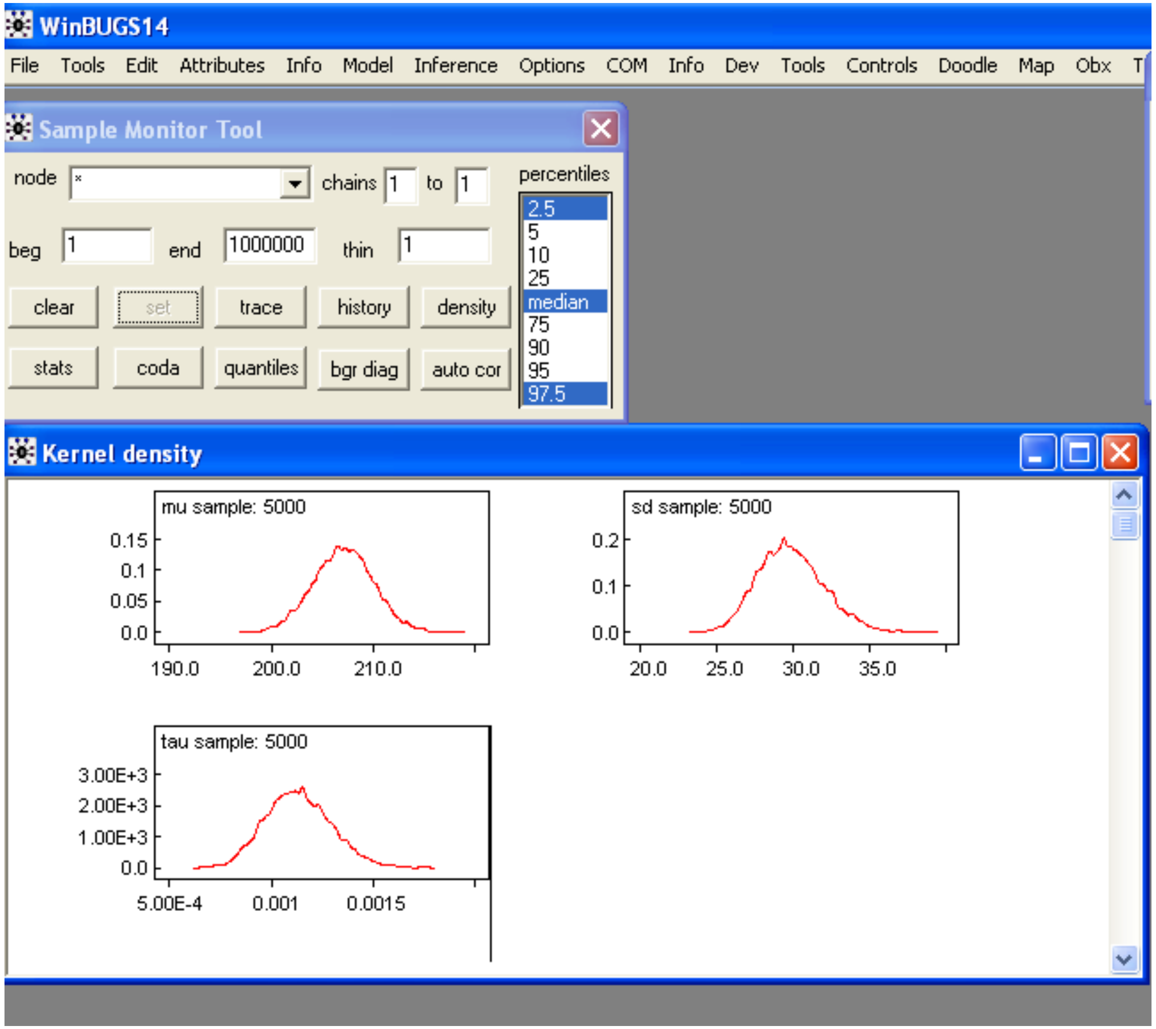
WinBUGS: step 8

- In the "sample monitor tool", click on "stats":



WinBUGS: step 8

- In the “sample monitor tool”, click on “density”:



WinBUGS: step 8

- Given that 175 mg/dL is not in the range of plausible values for μ , we conclude that this sub-group of children has a significantly larger mean cholesterol level than the corresponding overall population.

EXERCICE 2:

Comparison of 2 means-unequal variance

Gabriele,
living with rheumatoid arthritis

Exercice 2

- The cholesterol level in 2 groups of children has been measured in a case-control study where the case group consists of n_1 = 100 children whose father suffered from cardiovascular diseases.
- The mean cholesterol levels of the n_2 = 74 control kids is y_2 = 193.4 with standard deviation s_2 = 17.3. In the case group, one has y_1 = 207.3 and s_1 = 35.6. (Rosner 2000, p. 287).
- The normality hypothesis is reasonable:

  $Y_{1i} \sim N(\mu_1,\sigma^2{}_1)$ $(i = 1, \ldots ,n_1)$

  $Y_{2i} \sim N(\mu_{2,}\sigma^2{}_2)$ $(i = 1, \ldots ,n_2)$
- Estimate the difference between μ_1 and μ_2 : $\delta = \mu_1 - \mu_2$
- What do you conclude?

Exercice 2

- Non-informative priors :
  - $\mu_1 \sim N(180,10000)$
  - $\mu_2 \sim N(180,10000)$
  - $\tau_1 \sim Gamma(0.0001,0.0001)$
  - $\tau_2 \sim Gamma(0.0001,0.0001)$

with $\tau_1 = 1/\sigma_1^2$ and $\tau_2 = 1/\sigma_2^2$

- Data : data.txt
- Initial values : inits1.txt   inits2.txt