

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



LSHTM Research Online

Thorpe, B; Carroll, O; Sharples, L; (2017) Methods for Handling Unobserved Covariates in a Bayesian Update of a Cost-Effectiveness Model. Medical decision making. 272989X17736780. ISSN 0272-989X
DOI: <https://doi.org/10.1177/0272989X17736780>

Downloaded from: <http://researchonline.lshtm.ac.uk/4645556/>

DOI: <https://doi.org/10.1177/0272989X17736780>

Usage Guidelines:

Please refer to usage guidelines at <https://researchonline.lshtm.ac.uk/policies.html> or alternatively contact researchonline@lshtm.ac.uk.

Available under license: <http://creativecommons.org/licenses/by-nc-nd/2.5/>

<https://researchonline.lshtm.ac.uk>

Supplementary material: OpenBUGS code

Benjamin Thorpe, PhD*, Orlagh Carroll, BSc†, Linda Sharples, PhD‡

September 12, 2017

We will describe the OpenBUGS code used to implement the model presented in Figure 8. Code for the models in Figures 2 to 7 can be obtained as simplifications of the code presented here.

Inputs to the OpenBUGS code, such as data and hyperparameters, were prepared using the R software and read into OpenBUGS using the package `R2OpenBUGS`. In particular, covariate data from both CE-MARC and CECaT were shifted and scaled in R using baseline summaries from the EUROPA study (except for symptomatic angina, which was not an input to either logistic regression). For example, 44% of EUROPA patients were on nitrates at baseline, so the binary covariate for nitrate use was shifted to take values -0.44 (not on nitrates) and 0.56 (on nitrates).

From CE-MARC (Study 2), each available variable was provided to OpenBUGS as a vector of length 272 (one element for each patient), including time to first CV event (`y2`), time of censoring (`c2`), and covariates such as use of nitrates (`nitrates2`) and age over 65 (`age2`). If patient i was censored then we set the i -th event time to be missing (`y2[i] <- NA`). If patient i experienced a CV event then we set the i -th censoring time to be zero, (`c2[i] <- 0`). Similarly, covariate data from CECaT (Study 3) were provided as vectors of length 217 (`nitrates3`, `age3`, etc.).

Hyperparameters describing the prior distribution of the parameter vector β also required attention. The first 17 elements of β were assigned a multivariate normal distribution (with mean and precision as estimated in EUROPA) and we specified an independent normal prior (mean 0 and standard deviation 0.97) for the 18th element, resulting in a multivariate normal prior for β overall (with known mean vector `mu_beta` and precision matrix `Tau_beta`). However,

*Leeds Institute of Clinical Trials Research, University of Leeds, UK

†Leeds Institute of Clinical Trials Research, University of Leeds, UK

‡London School of Hygiene and Tropical Medicine, UK

it was necessary to reparametrise due to the shifting and scaling of covariates. Again, this was performed in R (requiring only straightforward matrix multiplication), resulting in a multivariate normal distribution with mean vector `mu_beta_reparam` and precision matrix `Tau_beta_reparam` as inputs to OpenBUGS.

The following OpenBUGS code was used to sample (approximately) from the posterior distribution of β . Further manipulations of the OpenBUGS output, such as burn-in, thinning, reparametrisation and diagnostic checks, were performed in R.

```
model{

  # Prior for beta (reparametrised version):

  beta[1:18] ~ dmnorm(mu_beta_reparam[ ], Tau_beta_reparam[ , ])

  # Priors for logistic regression parameters:

  theta[1:11] ~ dmnorm(mu_theta[ ], Tau_theta[ , ])
  gamma[1:12] ~ dmnorm(mu_gamma[ ], Tau_gamma[ , ])

  # Loop over 217 CECaT patients:

  for(j in 1:217){

    # Imputing nitrate use for CECaT patients:
    # s[j] is the probability patient j used nitrates (x* = 0.56),
    # in the logistic regression p(x* | x'', theta).

    logit(s[j]) <- theta[1] +
      theta[2] * bloodpress3[j] +
      theta[3] * age3[j] +
      theta[4] * male3[j] +
      theta[5] * smoker3[j] +
      theta[6] * diabetes3[j] +
      theta[7] * familyhistory3[j] +
      theta[8] * lipid3[j] +
      theta[9] * obese3[j] +
      theta[10] * mi3[j] +
      theta[11] * stenosis3[j]
```

```
# Use of nitrates is generated from a Bernoulli distribution,
# giving values 0 and 1 which we shift to -0.44 and 0.56:

nitrates3_imp[j] ~ dbern(s[j])
nitrates3[j]     <- nitrates3_imp[j] - 0.44

# Fitting the model for angina:
# q[j] is the probability patient j had angina (x' = 1),
# in the logistic regression p(x' | x'', x*, gamma).

logit(q[j]) <- gamma[1]          +
               gamma[2] * bloodpress3[j] +
               gamma[3] * age3[j]      +
               gamma[4] * male3[j]     +
               gamma[5] * smoker3[j]   +
               gamma[6] * diabetes3[j] +
               gamma[7] * familyhistory3[j] +
               gamma[8] * lipid3[j]    +
               gamma[9] * obese3[j]    +
               gamma[10] * nitrates3[j] +
               gamma[11] * mi3[j]      +
               gamma[12] * stenosis3[j]

angina3[j] ~ dbern(q[j])

} # End of loop over CECaT patients.

# Loop over 272 CE-MARC patients:

for(i in 1:272){

  # Uncertainty in previous MI:
  # Previous acute coronary syndrome (acs2[i] = 0.35) is taken
  # to have probability 44% of being MI (mi2[i] = 0.35). Otherwise
  # (acs2[i] = -0.65) no previous MI was assumed (mi2[i] = -0.65).

  notmi[i] ~ dbern(0.56)
  mi2[i]    <- acs2[i] - notmi[i] * (acs2[i] + 0.65)
```

```

# Fitting the model for nitrate use:
# r[i] is the probability patient i used nitrates (x* = 0.56),
# in the logistic regression p(x* | x'', theta).

logit(r[i]) ~ theta[1] +
              theta[2] * bloodpress2[i] +
              theta[3] * age2[i] +
              theta[4] * male2[i] +
              theta[5] * smoker2[i] +
              theta[6] * diabetes2[i] +
              theta[7] * familyhistory2[i] +
              theta[8] * lipid2[i] +
              theta[9] * obese2[i] +
              theta[10] * mi2[i] +
              theta[11] * stenosis2[i]

# Shifting is required as OpenBUGS requires observations
# from a Bernoulli distribution to be 0 or 1:

nitrates2_shift[i] <- nitrates2[i] + 0.44
nitrates2_shift[i] ~ dbern(r[i])

# Imputing angina for CE-MARC patients:
# p[i] is the probability patient i had angina (x' = 1)
# in the logistic regression p(x' | x'', x*, gamma).

logit(p[i]) <- gamma[1] +
               gamma[2] * bloodpress2[i] +
               gamma[3] * age2[i] +
               gamma[4] * male2[i] +
               gamma[5] * smoker2[i] +
               gamma[6] * diabetes2[i] +
               gamma[7] * familyhistory2[i] +
               gamma[8] * lipid2[i] +
               gamma[9] * obese2[i] +
               gamma[10] * nitrates2[i] +
               gamma[11] * mi2[i] +
               gamma[12] * stenosis2[i]

angina2[i] ~ dbern(p[i])

```

```

# Fitting the model for time to CV event:
# lambda[i] is the CV event rate of patient i in the
# proportional hazards model p(y | x, beta).

log(lambda[i]) <- beta[1] +
  beta[2] * ace2[i] +
  beta[3] * age2[i] +
  beta[4] * male2[i] +
  beta[5] * smoker2[i] +
  beta[6] * tiapvdcva2[i] +
  beta[7] * diabetes2[i] +
  beta[8] * familyhistory2[i] +
  beta[9] * angina2[i] +
  beta[10] * bloodpress2[i] +
  beta[11] * clearance2[i] +
  beta[12] * obese2[i] +
  beta[13] * cholesterol2[i] +
  beta[14] * nitrates2[i] +
  beta[15] * mi2[i] +
  beta[16] * calcium2[i] +
  beta[17] * lipid2[i] +
  beta[18] * stenosis2[i]

# Each event time y2[i] is assumed to come from an
# exponential distribution with rate lambda[i].
# y2[i] was either observed or was censored at c2[i]
# (and therefore sampled assuming y2[i] > c2[i]):

y2[i] ~ dexp(lambda[i])I(c2[i], )

# Posterior predictive checks:
# We predict event times for each patient, using
# the cut function to prevent these from influencing
# the parameter distributions.

lambda_copy[i] <- cut(lambda[i])
y2_pred[i] ~ dexp(lambda_copy[i])

} # End of loop over CE-MARC patients.

} # End of Bayesian model specification.

```