Instructions for use of the MYPAN elicitation software

This software implements the methodology described in the paper:

Hampson LV, Whitehead J, Eleftheriou D, Brogan P. (2013) Bayesian methods for the design and interpretation of clinical trials in very rare diseases.

1. **General tips for running the software**

*Installing the Shiny package:* To use the software, the R package “shiny” must be installed. To install the latest version of the package, one can type the following commands into R:

>> install.packages(‘httpuv’, repos = c(RStudio = ‘http://rstudio.org/\_packages’, CRAN=’http://cran.rstudio.com’))

>> install.packages(‘devtools’)

>> devtools::install\_github(‘shiny’, ‘rstudio’)

*Running the web apps:* R files in the folders “public release Day 1 code” and “public release Day 2 code” run the software described in Sections 3 and 4 of the paper, respectively. To run the web app for use on Day 1, for example, open R and then change into the directory where the folder “elicitation\_Day1” has been saved: run the software by typing to the R command prompt:

>> library(shiny)

>> runApp()

This will cause a web browser to open. On the LHS panel of the web page, the prior elicitation questions are displayed; output is displayed on the tabs on the RHS of the webpage.

*Exiting the software:* To close the web application, go to the R console and then press the “Escape” button on your keyboard.

1. **Prior distributions based on expert opinion alone**

*Generating Prior Distributions:* When the Day 1 web app is first launched, answers to the four prior elicitation questions are initialized at pre-specified values; there will be a short wait while prior distributions corresponding to the starting values are calculated. After a few seconds, the tabs “Density: CYC Remission Rate”, “Density: MMF Remission Rate”, “Density: Log-odds Ratio” and “Summary” should be filled with plots of the prior distributions of

* pC  the 6-month remission rate on CYC,
* pE  the 6-month remission rate on MMF and
* θ = log{pE(1 pC)/[pC(1 pE)]}.

The “Summary” tab presents a more detailed summary of each of these prior distributions including 90% credibility intervals and effective prior sample sizes.

To modify answers to the prior elicitation questions simply move the sliders to the required values (sliders move in steps of 0.05). Once answers to the four questions have been updated, click on the “Update Bayesian Distributions” button which is located at the bottom of the LHS panel of the webpage. The output tabs on the RHS of the browser will then momentarily fade out while R code runs in the background to find prior distributions corresponding to the updated expert opinions.

The expert’s first name (or initials) should be entered into the field “Expert’s name:” found on the top of the LHS of the webpage. Once the expert’s prior distributions have been found, the R code will automatically output a pdf file containing their prior densities for pC, pE and θ. This pdf will be saved as “expertname-priorplot.pdf” in the directory form where the Day 1 software is called. Furthermore, titles of the plots will include the expert’s name so that once the pdfs have been printed it will be clear which priors correspond to which expert.

The last 2 questions on the Day 1 handout are not needed to identify the prior distribution of pE (under our assumed Bayesian model) but are included so that the adequacy of the fitted prior may be assessed. Answers to Q5 and Q6 for the fitted prior are listed on the “Summary” tab.

The software is programmed assuming that the non-inferiority margin for the trial is 0.1 (on the probability difference scale). To implement the methodology under a difference choice of non-inferiority margin, elicitation question (iv) would needed to be reworded in terms of this new margin, and the variable c2 in functions priorcall and postSumry saved in the script file “run\_analysis.R” would also need to be updated.

*Generating posterior distributions:*To examine how data from 40 patients would modify the elicited prior distributions, click on the option “Update prior distribution with observed data on 40 patients:”. Then click on the “Update Bayesian Distributions” button. Doing so reveals three additional input fields which ask the user to input: 1) the number of patients randomized to CYC in the hypothetical dataset (the remainder of the 40 patients are assumed to be allocated to MMF); 2) the number of patients on CYC who are in remission by 6-months (and adhere to the steroid dosing taper) in this hypothetical dataset; and 3) the number of patients on MMF who are in remission by 6 months (and adhere to the steroid taper).

These input fields are initialised at values of nC = 20 (and thus nE = 20), SC = 14 and SE = 14. When these fields are first revealed, R code is run to generate the posteriors corresponding to this initial data scenario: the RHS output tabs fade out and once the R code has finished running, prior and posterior distributions are displayed for pC, pE and θ, and a summary of the *posterior* distributions is displayed in the “Summary” tab. To examine what the posterior distributions would be for a different data scenario, modify the 3 input fields and then click on the “Update Bayesian Distributions” button.

Ticking the option “Summarise posterior distributions for a variety of data scenarios” generates three posterior distributions for the following scenarios (each of which assume nE = nC = 20): a) (SE = 14, SC = 14) – MMF and CYC equivalent with an observed response rate of 70% on each treatment arm; b) (SE = 12, SC = 14) – MMF non-inferior to CYC; c) (SE = 10, SC = 14) – MMF clearly inferior to CYC. Parameters specifying data scenarios a) – c) are saved in the file “data\_scenario.txt”. Different scenarios may be compared by modifying this file.

*Error messages:* Prior distributions cannot be found for every combination of answers to the four elicitation questions. Error messages will be displayed on the “Density: CYC Remission Rate” tab of the web browser in the following cases:

1. Error: Error in answers to elicitation questions Q1 and Q2: cannot determine Beta prior distribution for CYC remission rate.
2. Error: Given Beta prior for pC and answers to elicitation questions Q3 and Q4, cannot determine Normal prior distribution for log-odds ratio.

Errors 1 and 2 indicate that there is some inconsistency between answers to Q1 and Q2 (Error 1) or Q3 and Q4 (Error 2). To resolve these issues, ask the expert to modify their answers to either elicitation questions Q1 or Q2 (Error 1) or Q3 or Q4 (Error 2) and then click on the “Update Bayesian Distributions” button.

1. Error: Data on MMF: number of successes exceeds number randomized to MMF.
2. Total number randomized to MMF and CYC does not sum to 40
3. Total number randomized to MMF and CYC does not sum to 20

Error 3 indicates that SE > nE. Error 4 indicates that the sum of nE and nC does not equal 40 (if the user has declared that posterior inferences are to be based on data from 40 patients). Error 5 catches a similar error if the user has declared that posterior inferences are to be based on data from 20 patients. To resolve these errors, modify the inputs and click on the “Update Bayesian Distributions” button.

1. Error identifying CYC prior distribution: Stop because we cannot guarantee the accuracy of the numerical integration.
2. Prior density MMF remission rate is U (or L) shaped function of pE. Can't guarantee accuracy of numerical integration routines.

The prior probability that 0.001 ≤ pC ≤ 0.999 is less than the value which would be expected if a uniform distribution represented prior opinion for pC. If the prior density is U- or L-shaped (thus assigning high probabilities to values of pC close to 0 and/or 1) the accuracy of the numerical integration routines implemented in the R code cannot be guaranteed. Similar error messages will be outputted if prior (or posterior) opinion for pE is very vague.

1. **Prior distributions based on expert opinion and historical data**

The Day 2 software reads in the consensus priors for pC and θ agreed before the historical data are revealed, and updates these to incorporating data on from the MYCYC trial.

*Software inputs:* The results for the primary analysis of the MYCYC trial are summarised in “MYCYCdata.txt“. The results are listed (in order of appearance) as SCYC, FCYC, SMMF, FMMF, where S denotes the number of patients achieving the primary outcome and F denotes the number failing to achieve the primary outcome. The parameters of the consensus priors for pC and θ determined without knowledge of the MYCYC data should be recorded in “Day1Prior.txt”; specifically, the parameters listed (in order) are: a, b, μ and σ2, where pC ~Beta(a, b) and θ~N(μ, σ2). The answers to the Day 1 elicitation questions (asking for opinions about pC and θ before the MYCYC data are known) are listed in “D1Answers.txt”.

*Generating prior distributions:* When the Day 2 web app is first launched, answers to four prior elicitation questions are initialized at pre-specified values. Q1 and Q2 elicit opinion about the difference between the six-month remission rates on CYC in the MYCYC and MYPAN trial populations. Q3 and Q4 are similar in form to the first two questions and ask about the difference between the six-month remission rates on MMF in the MYCYC and MYPAN trial populations.

After the web app is launched, there will be a short wait while prior distributions corresponding to the starting values of Q1-4 are calculated. After a few seconds, the following tabs will be filled:

* “Summary: MYCYC CYC relevance” which plots the prior distribution of λC.
* “Summary: MYCYC MMF relevance” which plots the prior distribution of λE.
* “Summary: MYPAN treatments”, “MYPAN: CYC remission rate”, “MYPAN: MMF remission rate”, “MYPAN: CYC & MMF remission rate”, “MYPAN: log-odds ratio” summarise prior distributions for MMF and CYC remission rates in the MYPAN trial population.

Ticking the option “Discard MYCYC data from prior distributions” will mean that the R code will display and summarise the prior distributions for pC, pE and θ based on expert opinion without knowledge of the MYCYC data. The tabs “Summary: MYCYC CYC relevance” and “Summary: MYCYC MMF relevance” are left empty.

The software will save a file “expername-D2answer.txt” in the directory from which the R code is being called. This file will list the expert’s answers to the four Day 2 elicitation questions.

*Generating posterior distributions:* See Section 2 for instructions on how to use the software to generate posterior distributions for pC, pE and θ. The tabs “Summary: MYCYC CYC relevance” and “Summary: MYCYC MMF relevance” are not updated.

At the moment, in the interests of time the software does not compute credibility intervals or the values of P(θ >0) and P(pE > pC – 0.1) when prior distributions incorporate related data. However, functions to compute these quantities can be found in “D2design.R” and “D2prob\_calc.R”. It is straightforward to add calls to these functions in priorcall and postcall (saved in “D2run\_analysis.R”); to run the modified code, server.R would need to modified, specifically to add the lines

source(“D2design.R”)

source(“D2prob\_calc.R”)

to the top of server.R, and extra print statements would be needed in the reactive function assigning a value to output$mypan in order to print the summaries to the screen.

*Error messages:* Prior distributions cannot be found for every combination of answers to the four elicitation questions. Error messages will be displayed on the “Summary: MYCYC CYC relevance” tab of the web browser in the following cases:

1. Error: There is an inconsistency between the answers to elicitation questions Q1 and Q2: Q1 should be less than 1-Q2
2. Error: There is an inconsistency between the answers to elicitation questions Q3 and Q4: Q3 should be less than 1-Q4

Errors 1 and 2 indicate a logical inconsistency between the answers given to the elicitation questions. These should be revised and then click on the “Update Bayesian Distributions” button.

1. Error: Error in answers to Day 2 elicitation questions Q1 and Q2: cannot determine a Normal prior distribution linking CYC 6-month remission rates in MYCYC and MYPAN trial populations.
2. Error: Error in answers to elicitation questions Q3 and Q4: cannot determine a Normal prior distribution linking MMF 6-month remission rates in MYCYC and MYPAN trial populations.

Errors 3 and 4 indicate that there is some inconsistency between answers to Q1 and Q2 (Error 3) or Q3 and Q4 (Error 2). To resolve these issues, ask the expert to modify their answers to either elicitation questions Q1 or Q2 (Error 3) or Q3 or Q4 (Error 4) and then click on the “Update Bayesian Distributions” button.

1. Error: Data on MMF: number of successes exceeds number randomized to MMF.
2. Total number randomized to MMF and CYC does not sum to 40
3. Total number randomized to MMF and CYC does not sum to 20
4. Error 5 indicates that SE > nE. Error 6 indicates that the sum of nE and nC does not equal 40 (if the user has declared that posterior inferences are to be based on data from 40 patients). Error 7 catches a similar error if the user has declared that posterior inferences are to be based on data from 20 patients. To resolve these errors, modify the inputs and click on the “Update Bayesian Distributions” button.
5. Error evaluating prior density for CYC remission rate: can't guarantee accuracy of numerical integration routines when integrating U shaped function.

The prior probability that 0.001 ≤ pC ≤ 0.999 is less than the value which would be expected if a uniform distribution represented prior opinion for pC. If the prior density is U- or L-shaped (thus assigning high probabilities to values of pC close to 0 and/or 1) the accuracy of the numerical integration routines implemented in the R code cannot be guaranteed. Similar error messages may be outputted if prior opinion for pE (or posterior opinion for pE or pC) is very vague.

Lisa Hampson, December 2013.