Bayesian prior elicitation: an application to the MYPAN trial in childhood polyarteritis nodosa

Lisa Hampson and John Whitehead
Lancaster University, UK

Despina Eleftheriou, Paul Brogan
University College London, UK

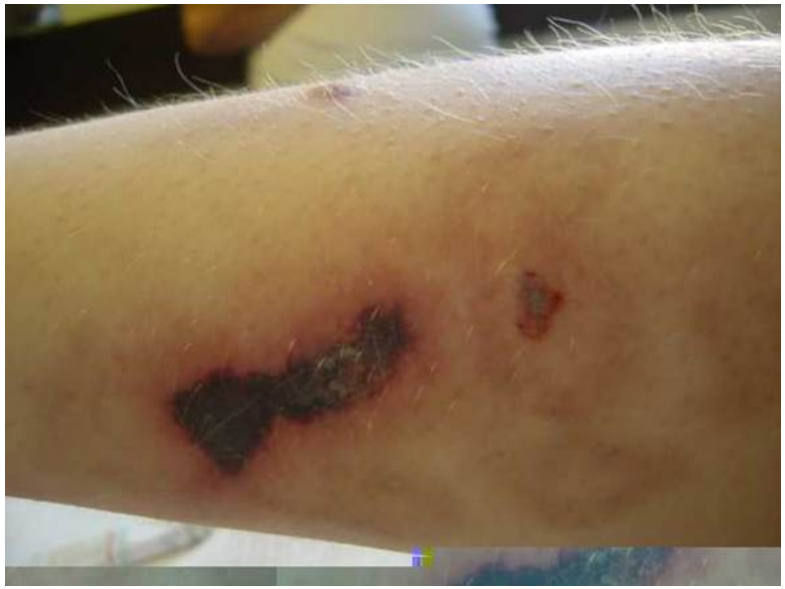
European Medicines Agency Workshop on Extrapolation,
London, 17th May 2016

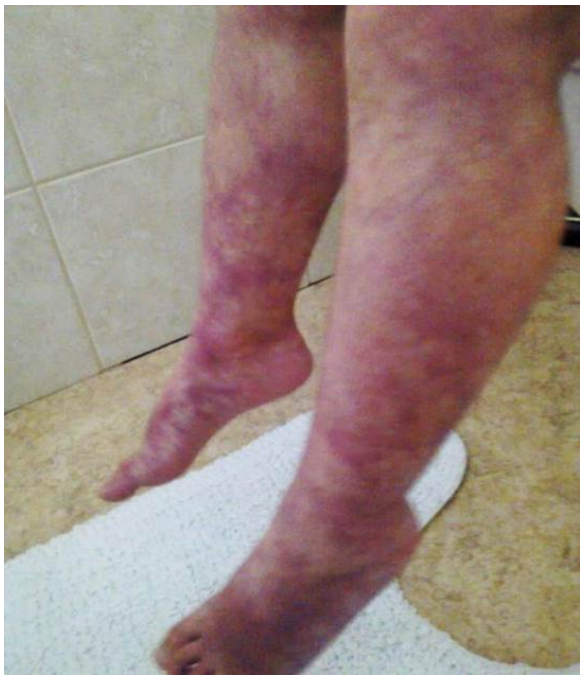
Acknowledgements

Grateful to acknowledge funding from the Medical Research Council and Arthritis Research UK.

The MYPAN trial

Childhood polyarteritis nodosa (PAN) is a serious inflammatory blood vessel disease affecting around 1 per million children.





The MYPAN trial

- Cyclophosphamide (CYC) has been standard treatment for past 35 years.
- Mycophenolate mofetil (MMF) is an immunosuppresant thought to have a lower risk of toxicity.
- MYPAN trial will compare MMF versus CYC for the treatment of childhood PAN.
- The primary endpoint is remission within 6-months. Probabilities of remission on MMF and CYC are p_E and p_C. MMF will be preferred to CYC if $p_E - p_C \geq -0.1$.
- A definitive trial would require 513 patients per arm when $p_E = p_C = 0.7$

The MYPAN trial

- Cyclophosphamide (CYC) has been standard treatment for past 35 years.
- Mycophenolate mofetil (MMF) is an immunosuppresant thought to have a lower risk of toxicity.
- MYPAN trial will compare MMF versus CYC for the treatment of childhood PAN.
- The primary endpoint is remission within 6-months. Probabilities of remission on MMF and CYC are p_E and p_C. MMF will be preferred to CYC if $p_E - p_C \geq -0.1$.
- A definitive trial would require 513 patients per arm when $p_E = p_C = 0.7$

PROBLEM: 20-30 European centres could recruit 40 patients over 4 years.

The MYPAN trial

- Cyclophosphamide (CYC) has been standard treatment for past 35 years.
- Mycophenolate mofetil (MMF) is an immunosuppresant thought to have a lower risk of toxicity.
- MYPAN trial will compare MMF versus CYC for the treatment of childhood PAN.
- The primary endpoint is remission within 6-months. Probabilities of remission on MMF and CYC are p_E and p_C. MMF will be preferred to CYC if $p_E - p_C \geq -0.1$.
- A definitive trial would require 513 patients per arm when $p_E = p_C = 0.7$

PROBLEM: 20-30 European centres could recruit 40 patients over 4 years.

SOLUTION: Aim for a more modest objective – to improve our understanding of treatment options for PAN.

MYPAN: A Bayesian RCT

Lilford et al. (1995); Tan et al. (2003); Johnson et al. (2010); Hampson et al. (2014); Hampson et al. (2015)

We will quantify prior uncertainty and the impact of new data using Bayesian methods.

MYPAN: A Bayesian RCT

Lilford et al. (1995); Tan et al. (2003); Johnson et al. (2010); Hampson et al. (2014); Hampson et al. (2015)

We will quantify prior uncertainty and the impact of new data using Bayesian methods.

Label MMF and CYC as treatments E and C, respectively. We represent the probability of remission on E and C as p_E and p_C.

Measure the advantage of E over C using the log-odds ratio

$$\theta = \log\left\{\frac{p_E(1-p_c)}{p_C(1-p_E)}\right\}.$$

MYPAN: A Bayesian RCT

Lilford et al. (1995); Tan et al. (2003); Johnson et al. (2010); Hampson et al. (2014); Hampson et al. (2015)

We will quantify prior uncertainty and the impact of new data using Bayesian methods.

Label MMF and CYC as treatments E and C, respectively. We represent the probability of remission on E and C as p_E and p_C.

Measure the advantage of E over C using the log-odds ratio

$$\theta = \log\left\{\frac{p_E(1-p_c)}{p_C(1-p_E)}\right\}.$$

No high quality data to base priors upon. Instead we elicit prior opinion on p_C and θ, modelling it as:

- $p_C \sim \text{Beta}(a, b)$
- $\theta \sim N(\mu, \sigma^2)$.

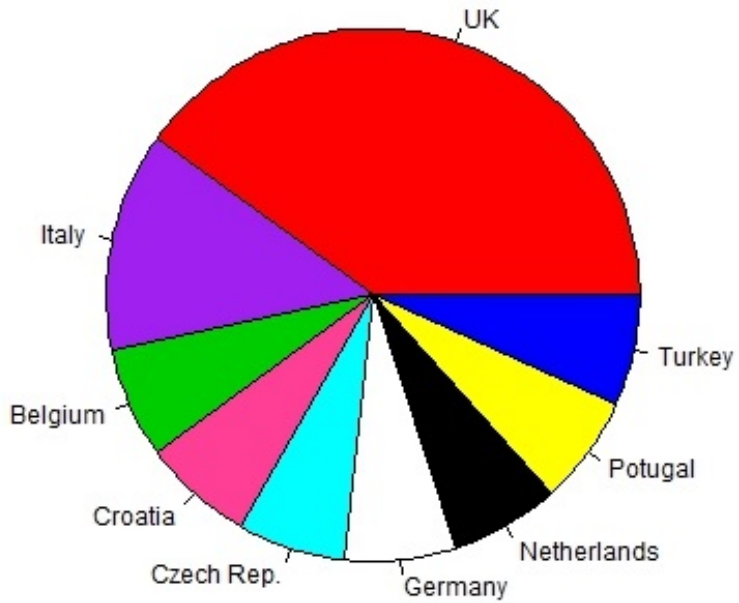
} independent

Identifying experts in childhood PAN

We defined an expert as a paediatric consultant

- Specialising in rheumatology, nephrology or immunology;
- With experience of treating children with PAN (on average 1 case every 2 years).

15 experts from across the EU and Turkey attended 2-day prior elicitation meeting.



Structure of the elicitation meeting

Day 1 objectives:

- Provide experts with relevant training;
- Elicit expert opinion about p_C and θ.
- Elicit individuals' opinions first . . .

Day 1: Eliciting individual opinion on p_C

Q1: What do you think the 6-month remission rate for children with PAN on CYC is?

- A1: Prior mode $= (a-1)/(a+b-2)$.

Day 1: Eliciting individual opinion on p_C

Q1: What do you think the 6-month remission rate for children with PAN on CYC is?

- A1: Prior mode $= (a-1)/(a+b-2)$.

Q2: Provide a proportion which you are 75% sure the true remission rate p_C exceeds.

- A2: $\pi_{0.25}$ satisfying $\Pr\{p_C < \pi_{0.25}; a, b\} = 0.25$.

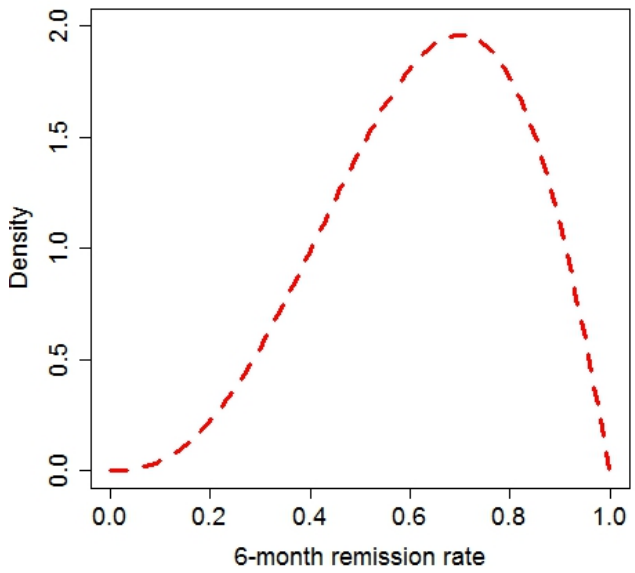
Day 1: Eliciting individual opinion on p_C

Q1: What do you think the 6-month remission rate for children with PAN on CYC is?

- A1: Prior mode $= (a-1)/(a+b-2)$.

Q2: Provide a proportion which you are 75% sure the true remission rate p_C exceeds.

- A2: $\pi_{0.25}$ satisfying $\Pr\{p_C < \pi_{0.25}; a, b\} = 0.25$.



Consensus: A1 = 0.7, A2 = 0.5 $\rightarrow p_C \sim$ Beta(3.6, 2.11)

Day 1: Eliciting individual opinion on θ

Q3: What is chance that the remission rate on MMF exceeds that on CYC?

- A3: $\Pr\{p_E > p_C\} = \Phi(\mu/\sigma)$.

Day 1: Eliciting individual opinion on θ

Q3: What is chance that the remission rate on MMF exceeds that on CYC?

- A3: $\Pr\{p_E > p_C\} = \Phi(\mu/\sigma)$.

Q4: What is chance that p_C exceeds p_E by more than 10%?

- A4: $\Pr\{p_E - p_C < -0.1\} = \int_0^1 \int_0^{\max\{p_C - 0.1, 0\}} g_0(p_C, p_E; a, b, \mu/\sigma, \sigma)\, \mathrm{d}p_E\, \mathrm{d}p_C$

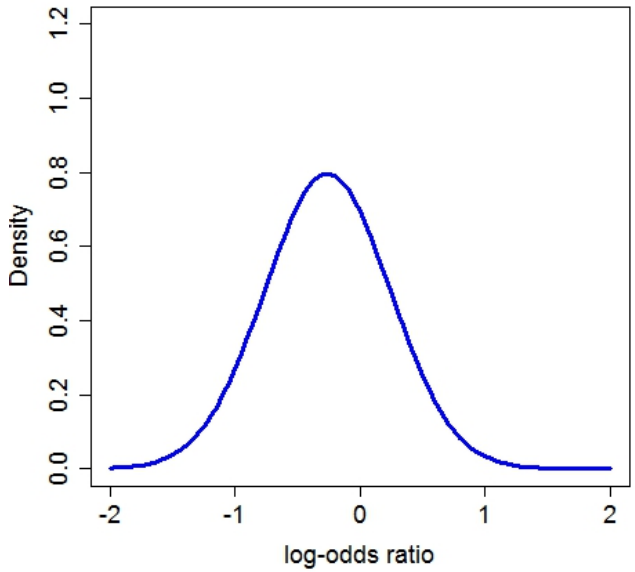
Day 1: Eliciting individual opinion on θ

Q3: What is chance that the remission rate on MMF exceeds that on CYC?

- A3: $\Pr\{p_E > p_C\} = \Phi(\mu/\sigma)$.

Q4: What is chance that p_C exceeds p_E by more than 10%?

- A4: $\Pr\{p_E - p_C < -0.1\} = \int_0^1 \int_0^{\max\{p_C - 0.1, 0\}} g_0(p_C, p_E; a, b, \mu/\sigma, \sigma)\, \mathrm{d}p_E\, \mathrm{d}p_C$



Consensus: A3 = 0.3, A4 = 0.3 $\rightarrow \theta \sim N(-0.26, 0.25)$

Day 1: Quantifying the strength of prior opinion

Morita et al. (2008), Neuenschwander.et al (2010)

A prior **Effective Sample Size** (ESS) characterises the strength of prior opinion:

- ESS is the number of observations needed to obtain the same amount of statistical information for a parameter as is represented by its prior distribution.

Day 1: Quantifying the strength of prior opinion

Morita et al. (2008), Neuenschwander.et al (2010)

A prior **Effective Sample Size** (ESS) characterises the strength of prior opinion:

- ESS is the number of observations needed to obtain the same amount of statistical information for a parameter as is represented by its prior distribution.

Calculate the ESS of $\omega = \log\{p_C/(1-p_C)\}$ as the sample size $n_C^\star$ satisfying

$$n_C^\star \int_{-\infty}^{\infty} \frac{1}{\mathrm{B}(a,b)} \frac{\exp\{(a+1)\omega\}}{\{1+\exp(\omega)\}^{a+b+2}} d\omega = \frac{1}{\mathrm{Var}_0(\omega)}$$

Interpretation: size of a single arm study evaluating CYC for which the expected Fisher's information for $\log\{p_C/(1-p_C)\}$ equals precision of elicited prior.

Day 1: Quantifying the strength of prior opinion

Calculate the ESS of θ as the sample size $n_\theta^\star$ satisfying

$$\int_0^1 \int_0^1 \frac{n_\theta^\star \bar{p}(1-\bar{p})}{4} g_0(p_E, p_C) dp_E dp_C = \sigma^{-2},$$

where $\bar{p} = (p_E + p_C)/2$.

Interpretation: sample size needed for an RCT allocating equal numbers to MMF and CYC to have expected Fisher's information for θ equal to precision of elicited prior.

Structure of the elicitation meeting

Day 1 objectives:

- Provide experts with relevant training;
- Elicit expert opinion about p_C and θ.
- Elicited individuals' opinions first . . .

Structure of the elicitation meeting

Day 1 objectives:

- Provide experts with relevant training;
- Elicit expert opinion about p_C and θ.
- Elicited individuals' opinions first . . .
- . . . before using behavioural aggregation to reach consensus prior distributions.

Structure of the elicitation meeting

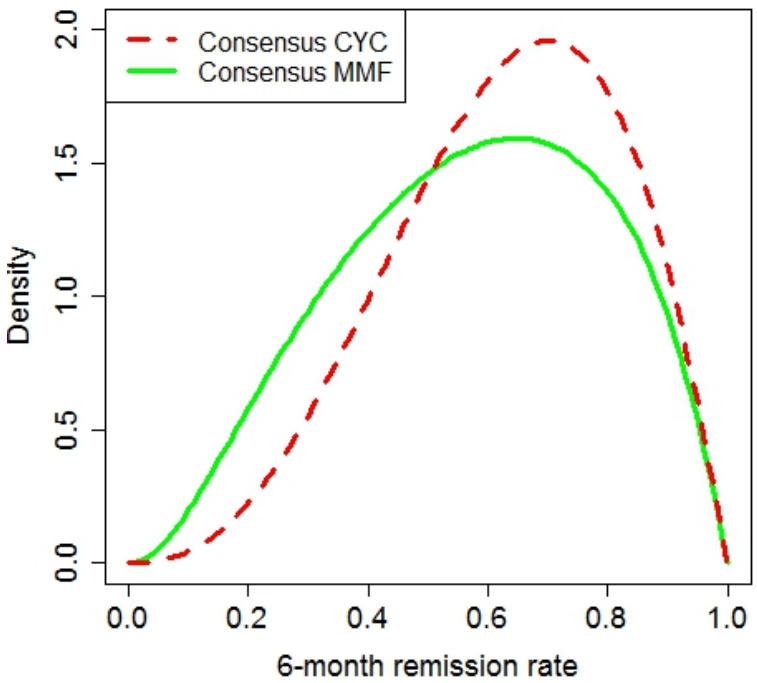
Day 1 objectives:

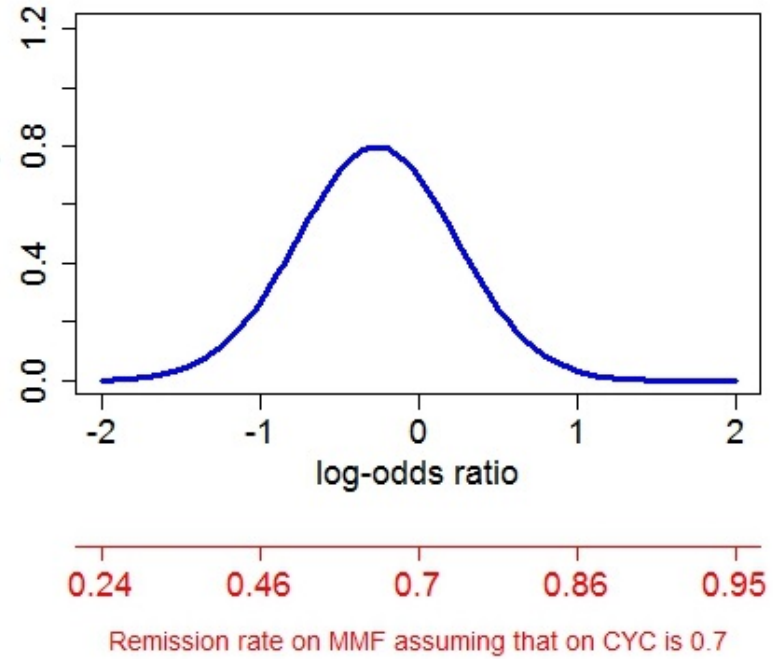
- Provide experts with relevant training;
- Elicit expert opinion about p_C and θ.
- Elicited individuals' opinions first . . .
- . . . before using behavioural aggregation to reach consensus prior distributions.

Reaching a consensus:

- A1-4 were displayed and discussed in a structured way.
- Mean and median answers were used as 'initial values' for consensus answers.
- Consensus answers determined by voting when choice was not unanimous.
- ESSs were influential in the group's final consensus decisions.

Day 1: Consensus prior distributions





- p_C: mode = 0.7, ESS = 5 patients on CYC
- θ : mode = -0.26, ESS = 39 patients per treatment
- p_E: mode = 0.65

Day 2: Extrapolating from adult data & across conditions

Expert opinion was combined with data from the MYCYC trial:

- MYCYC data were genuinely unknown to the experts on Day 1.
- MYCYC trial involved 132 adults and 8 children with a condition related to PAN.
- MYCYC compared MMF vs CYC.
- MYCYC primary endpoint was similar to MYPAN primary endpoint.

Day 2: Extrapolating from adult data & across conditions

Expert opinion was combined with data from the MYCYC trial:

- MYCYC data were genuinely unknown to the experts on Day 1.
- MYCYC trial involved 132 adults and 8 children with a condition related to PAN.
- MYCYC compared MMF vs CYC.
- MYCYC primary endpoint was similar to MYPAN primary endpoint.

Sought opinion on the relevance of MYCYC data **before** revealing the primary results.

Remission probabilities in the MYCYC and MYPAN trials linked via log-odds ratios

$$\lambda_C = \log\left\{\frac{p_{CR}(1-p_C)}{p_C(1-p_{CR})}\right\} \qquad \lambda_E = \log\left\{\frac{p_{ER}(1-p_E)}{p_E(1-p_{ER})}\right\}.$$

Day 2: Extrapolating from existing data

Prior opinion on the relevance of the MYCYC data was modelled as

$$\lambda_C \sim N(\alpha_C, \gamma_C^2) \quad \text{and} \quad \lambda_E \sim N(\alpha_E, \gamma_E^2).$$

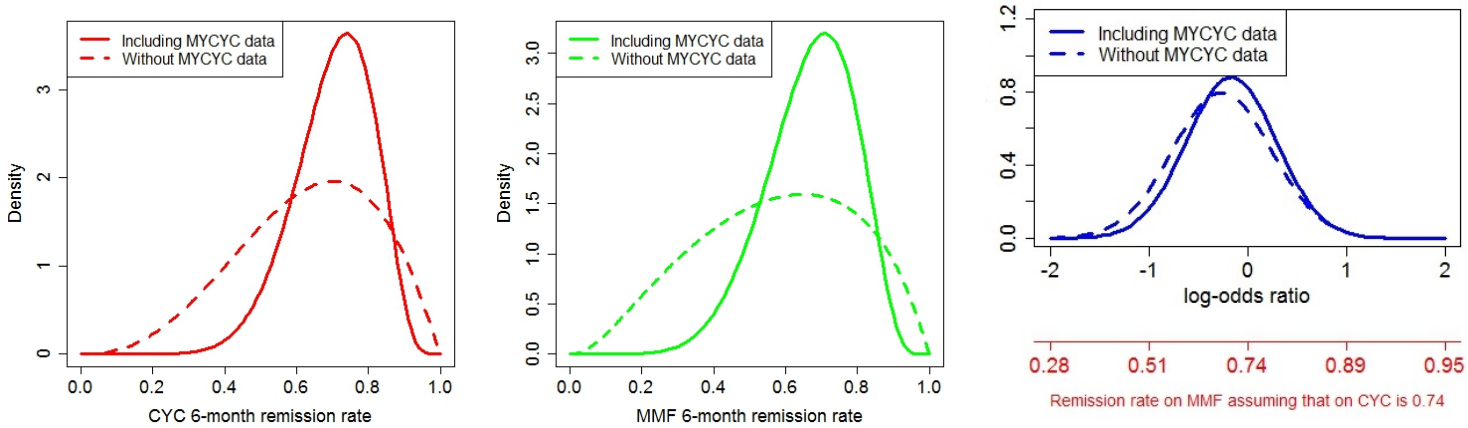
Experts were asked:

1. Q(a): What is the chance that the CYC remission rate in the MYCYC patient group exceeds that in the MYPAN patient group?
2. Q(b): What is the chance that the CYC remission rate in the MYPAN patient group exceeds that in the MYCYC patient group by more than 10%?
3. Q(c) - (d): Questions worded in terms of MMF.

We did not attempt to quantify the effective sample size of these priors.

Day 2: Extrapolating from existing data

MYCYC results: 74% successes on CYC; 73% successes on MMF.



Optimistic prior distribution with $\Pr\{p_E > p_C - 0.1\} = 0.77$.

Effective sample sizes were updated: 70 MYCYC patients per treatment increased the

- Effective Sample Size for p_C by 12;
- Effective Sample Size for θ by 9 per arm.

Summary
Schmidli et al. (2014)

MYPAN trial illustrates how a Bayesian approach can use existing information (expert opinion; adult data) to support paediatric drug development.

Bayesian methods could be used at different stages of the extrapolation process:

- Extrapolation concept: Quantify evidence supporting relevance of source data;
- Extrapolation plan: Determine expected value of new trial in target population;
- Extrapolation plan: Identify knowledge gaps – tailor studies to focus on parameters about which least is known or which have greatest impact on improving decision making (e.g., inform randomisation ratios, PK sampling times)
- Validation: Robust priors downweight source data when prior-data conflict.

Summary

Acceptability: The proposed approach would be used only if (after all reasonable efforts) a conventional trial in target population is deemed infeasible:

- Bayesian approach switches focus from hypothesis testing to estimation.
- A small Bayesian randomised trial may be ethically acceptable (if the data can shift the prior) because a more informative interpretation of the data is possible.

Challenges: Several potential contentious issues remain:

- Selection of experts.
- How to handle prior-data conflicts.

References

Billingham L, Malottki K, Steven N. Small sample sizes in clinical trials: a statistician's perspective. *Clinical Investigation* 2012; **2**:655.

Hampson LV, Whitehead J, Eleftheriou D, Brogan P. Bayesian methods for the design and interpretation of clinical trials in very rare diseases. *Statistics in Medicine* 2014; **33**:4186.

Hampson LV, Whitehead J, Eleftheriou D, et al. Elicitation of expert prior opinion: application to the MYPAN trial in childhood polyarteritis nodosa. *PLOS ONE* 2015;**10**:e0120981

Johnson SR, Tomlinson GA, Hawker GA, Granton JT, Feldman BM. Methods to elicit beliefs for Bayesian priors: a systematic review. *Journal of Clinical Epidemiology* 2010; **63**:355.

Lilford RJ, Thornton JG, Braunholtz D. Clinical trials and rare diseases: a way out of a conundrum. *British Medical Journal* 1995; **311**:1621.

References

Morita S, Thall PF, Müller P. Determining the effective sample size of a parametric prior. *Biometrics* 2008; **64**:595.

Neuenschwander B, Capkun-Niggli G, Branson M, Spiegelhalter DJ. Summarizing historical information on controls in clinical trials. *Clinical Trials* 2010; **7**:5.

Schmidli H, Gsteiger S, Roychoudhury S, et al. Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics* 2014; **70**:1023.

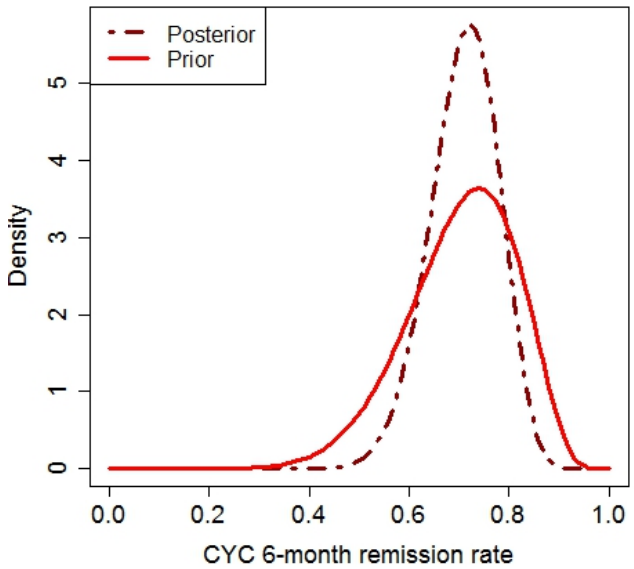
Tan S-B, Dear KBG, Bruzzi P, Machin D. Strategy for randomised clinical trials in rare cancers. *British Medical Journal* 2003; **327**:47.

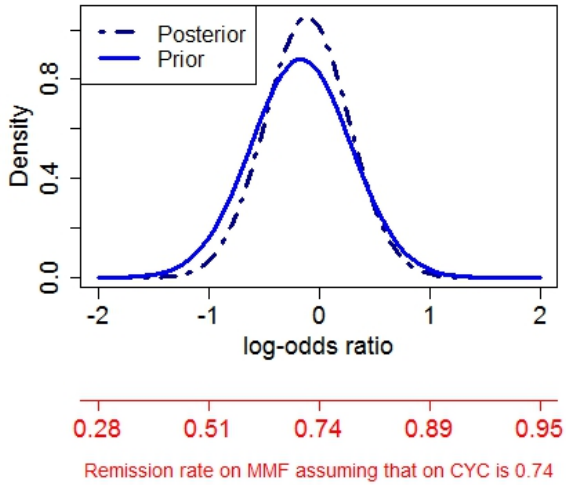
Tan S-B, Chung Y-F A, Tai B-C, Cheung Y-B, Machin D. Elicitation of prior distributions for a phase III randomized controlled trial of adjuvant therapy with surgery for hepatocellular carcinoma. *Controlled Clinical Trials* 2003; **24**:110.

Verifing consensus prior distributions

Idea: Present experts with hypothetical MYPAN datasets and ask whether they agree with 'their' posteriors derived using Bayes Theorem.

Example: Suppose we observed $n_E = 20$, $S_E = 14$, $n_C = 20$, $S_C = 14$.





- Posterior for p_C: mode = 0.72; Posterior for p_E: mode = 0.70
- $\Pr\{p_E > p_C - 0.1 \mid \text{data}\} = 0.84$.
- $\Pr\{p_E > p_C \mid \text{data}\} = 0.38$.