Incorporating prior beliefs into meta-analyses of health state utility values using the Bayesian Power Prior

Anthony Hatswell (ahatswell@deltahat.com)
Director, Delta Hat
Honorary Senior Research Fellow, Department of Statistical Science, UCL

Delta Hat

Background: Health state utilities

- Models for HTA generally use utilities for a health state, with the time spent in each health state used to generate QALYs

- There are many instruments for giving health state utilities
  - These can be generic: EQ-5D, SF-36, etc.
  - Or disease specific: EORTC QLQ-C30, CTCAE, etc.

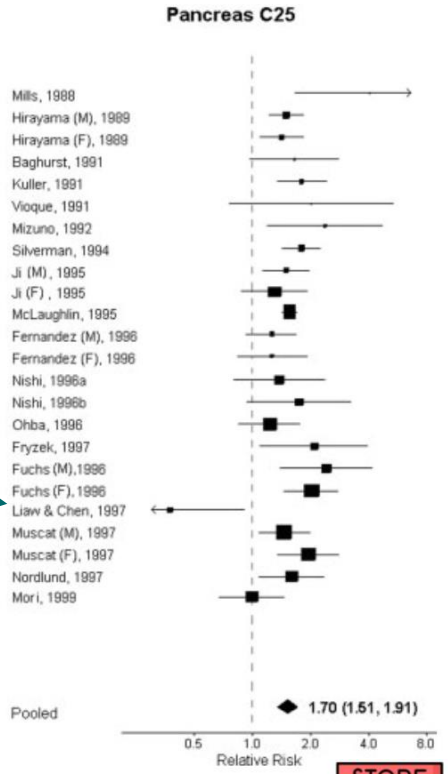
- HTA agencies also have a preference; NICE is EQ-5D, Canada is standard gamble based (i.e., SF-36), etc.

- The health state utility values (HSUVs) matter
  - In some diseases quality of life is the only driver of QALYs
  - In others it is quality of life *and* duration
  - It is rare (cost-minimisation?) that the utilities are not critical

- There is usually a large amount of uncertainty in estimated HSUVs
  - Patient responses have high variability
  - Often we have few observations

Big Tony's Sneaky Models LLC

- We are building a model, there are a few trials
- We want to follow 'Best Practice'[1]
- Does smoking cause pancreatic cancer?
  - No! (Liaw & Chen, 1997)
- Nailed it!



[1] Beca J, Husereau D, Chan KKW, Hawkins N, Hoch JS. Oncology Modeling for Fun and Profit! Key Steps for Busy Analysts in Health Technology Assessment. PharmacoEconomics. 2018 Jan 1;36(1):7-15

[2] Gandini S, Botteri E, Iodice S, Boniol M, Lowenfels AB, Maisonneuve P, et al. Tobacco smoking and cancer: A meta-analysis. Int J Cancer. 2008;122(1):155-64.

Background: Meta-analysis

- When we have multiple sets of clinical data, we meta-analyse to reflect the fact that each contributes to our overall knowledge

- With HSUVs, we as a field (generally) do not do this!
  - Instead, usually a single value is used as being the 'best'
  - Sensitivity analysis is then done with other values for comparison (Single Preferred Value, SPV)

- Meta-analysis is possible, but not widely used
  - There are a number of systematic review & meta-analyses (about 20) indexed in PubMed, generally using Random Effects Meta Analysis (REMA)
  - There is guidance from Petrou *et al.*, (2018) which is helpful, covering the main meta-analysis techniques (Fixed Effects, Random Effects, and meta-regression)

- One issue with meta-analysis however, is that all studies are assumed to have equal applicability

Petrou S, Kwon J, Madan J. A Practical Guide to Conducting a Systematic Review and Meta-analysis of Health State Utility Values. PharmacoEconomics. 2018 May 10; Available from: http://link.springer.com/10.1007/s40273-018-0670-1

Background: Bayesian power prior

- The Bayesian Power Prior (BPP) is a simple concept; historical data should count, but be downweighted (Duan *et al.*, 2006)
  - It was conceptualised for sequential water quality testing, with the previous values informing the latest test, but not overriding it

- It has been used in medical research to:
  - Allow historical controls to be used to supplement contemporary patients, and
  - Increase the power of basket trials (which enrol patients with different tumour sites)

Duan, Y., Ye, K. and Smith, E.P. (2006), Evaluating water quality using power priors to incorporate historical information. Environmetrics, 17: 95-106. https://doi.org/10.1002/env.752

Here: Application (/co-option) of the BPP to HSUVs

- With utilities is there are many more reasons we might prefer one estimate over another:
  - Interventions (e.g., drug of interest, vs drugs with different mechanisms and/or safety profiles)
  - Preferred vs non preferred instrument (e.g., SF-6D vs EQ-5D)
  - Value set (e.g., Spanish rather than UK)
  - Setting (e.g., US ‘health care’ vs European, trial vs observation setting)
  - Study age (e.g., data collected in 2001 vs 2021)

- Again, in a meta-analysis, all studies are treated as equally applicable and weights are derived solely from standard deviations / standard errors

- This presentation applies the BPP to utility data, incorporating downweighting for dissimilar results
  - Note: results are downweighted, but not adulterated in any way

What is presented here?

- 4 case studies:
  - Non-small cell lung caner (NSCLC)
  - Dialysis
  - Cirrhosis
  - blindness

- HSUVs are meta-analysed using the mean and SEM

- 4 meta-analysis methods compared:
  - [1] SPV i.e., picking the 'best'
  - [2] Fixed-Effects Meta Analysis & [3] Random Effects Meta-Analysis
    - Implemented using the '*metafor*' package
  - [4] Bayesian Power Prior
    - Implemented using '*rstan*' & custom code based on examples using the Gelman '8 schools' example
    - Weights given by myself, and the author of a systematic review of blindness

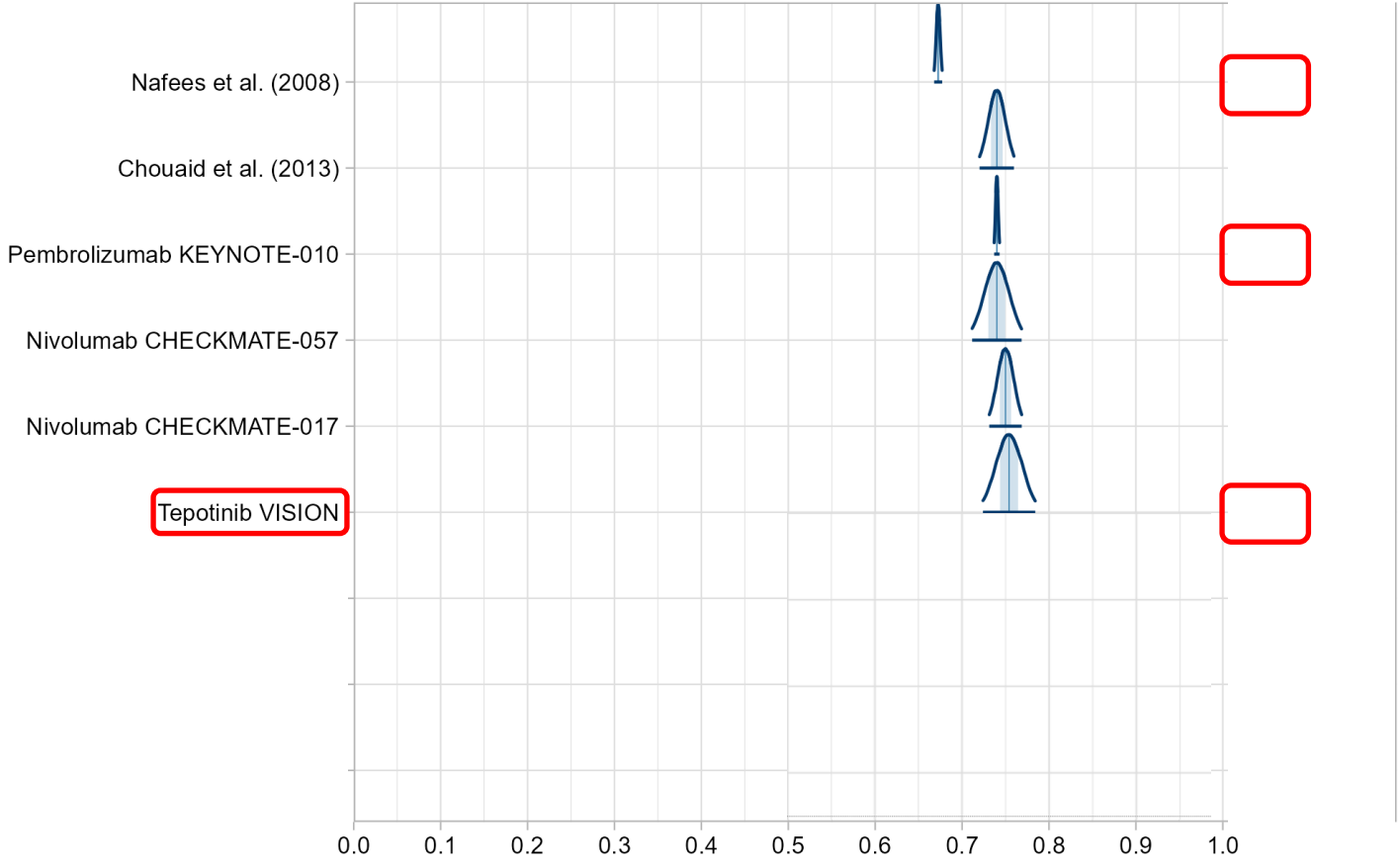
```
data {
  int<lower=0> J;             // number of studies
  real y[J];                  // estimated treatment effects
  real<lower=0> sigma[J];     // s.e. of effect estimates
  real<lower=0> weight[J];    // weight
}
parameters {
  real mu;
  real theta[J];
  real<lower=0> tau;
}

model {
  for (n in 1:J) {
    tau ~ normal(0, 0.5);
    theta[n] ~ normal(mu, tau);
    target += weight[n] * normal_lpdf(y | theta, sigma[n]);
  }
}
```

Delta Hat

NSCLC: What value to use for previously treated tepotinib patients?

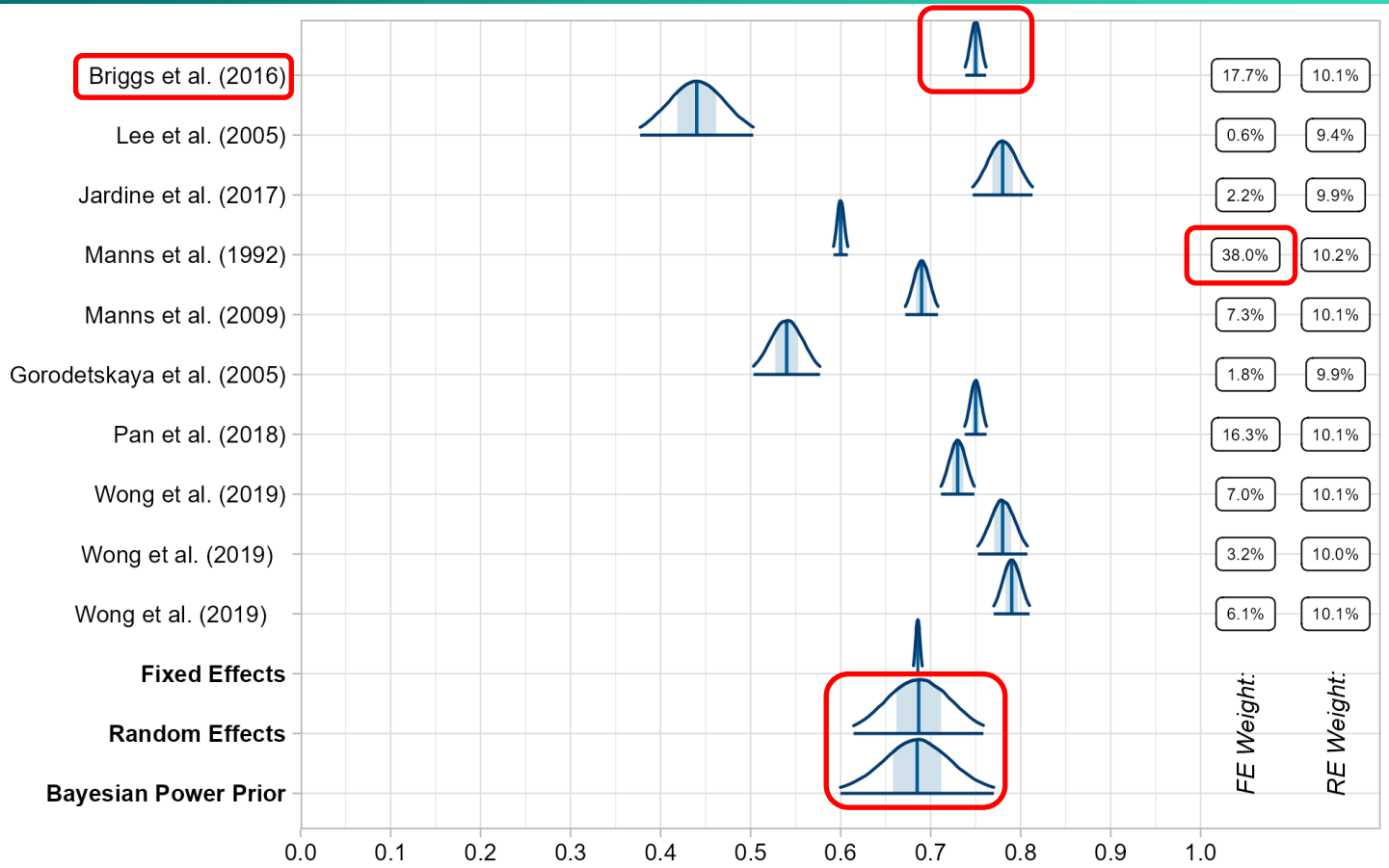
Study	HSUV (SD)	Sample size	Power Prior weight	Justification of weight used
Teptotinib VISION study	0.754 (0.261)	290	1	The contemporary study of patients treated with tepotinib
Nafees (2008)	0.674 (0.008)	10	0.05	Vignette study of 10 healthcare practitioners, approximately 15 years ago
Chouaid (2013)	0.74 (0.180)	319	0.5	Patients were treated in a non-trial setting, with a variety of therapies which pre-date current standard of care
NICE TA428L Pembrolizumab	0.74 (0.051)	1034	0.9	Although treated in a recent study, patients received immunotherapy, not targeted therapy
NICE TA484: Nivolumab, squamous	0.74 (0.23)	252	0.9	Although treated in a recent study, patients received immunotherapy, not targeted therapy
NICE TA655: Nivolumab, non-squamous	0.75 (0.23)	582	0.9	Although treated in a recent study, patients received immunotherapy, not targeted therapy

Results: NSCLC



Delta Hat

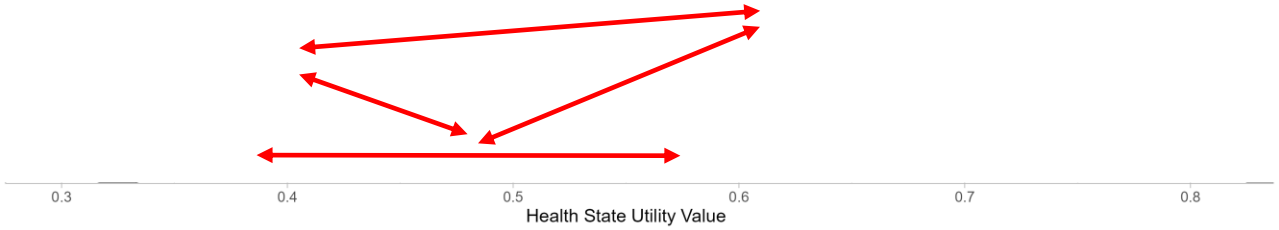
Results: Dialysis



A comparison of results

N = 9

N = 14



Delta Hat

Implications

- Firstly, Fixed Effects Meta Analysis has real face validity issues – the assumption of a fixed effect does not seem justifiable for HSUVs
- SPV might be widely used, but seems to have (unnecessarily) large confidence intervals and/or different values
  - Surely other studies tell us *something*?
- The BPP appears to work, and gives a method for differential weighting
- However, more work is needed if it is to be widely used:
  - Hierarchical models to account for correlation at the study level
  - Inclusion of multiple health states in the same model e.g., pre & post-progression
  - How do we set weights?

More broadly

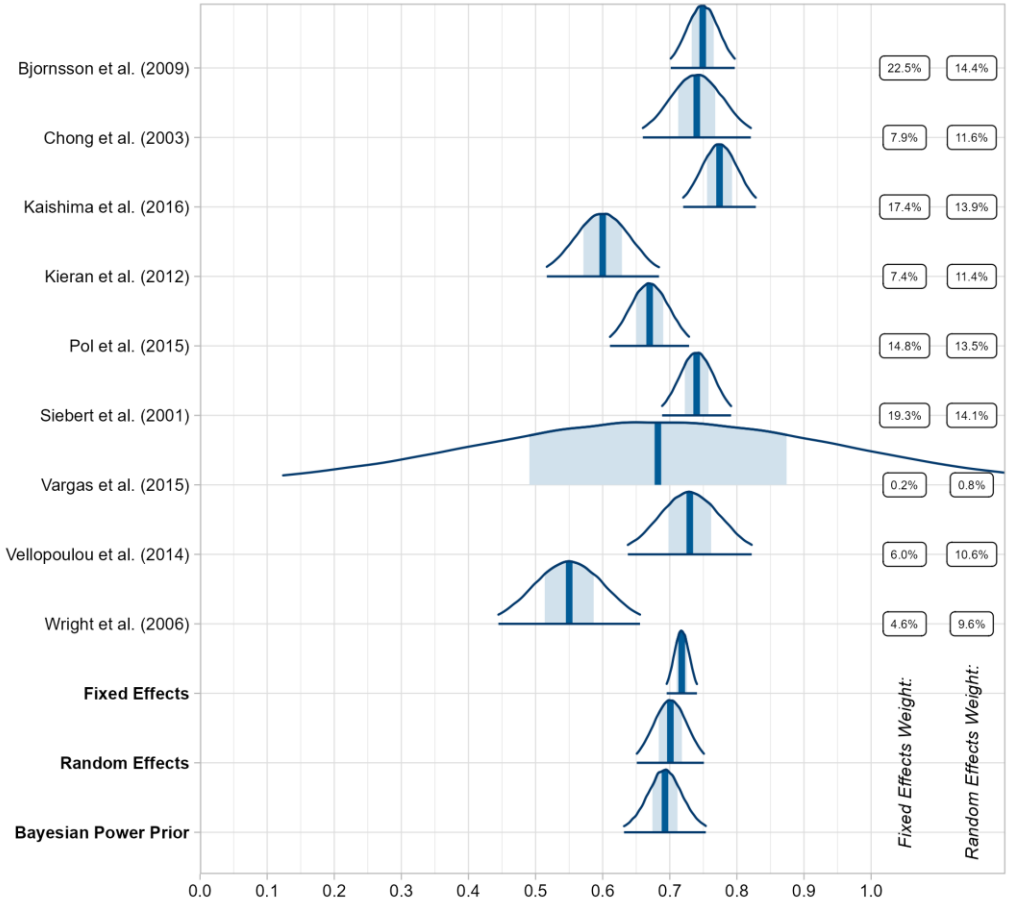
- Reporting needs to improve for HSUVs
  - Methods are seldom clear
  - Measures of dispersion are rarely included – please always include a SD or SE as a minimum!
- More thought is needed to how HSUVs are used
  - The automatic use of an SPV needs to stop
  - Sometimes meta-analysis will be appropriate
  - … but values are not always equally appropriate, careful thought and judgement can help here
- Getting it right matters
  - HSUVs power models
  - Models affect adoption decisions

Make sense? Or  ?

Anthony Hatswell (ahatswell@deltahat.com)

Delta Hat

Backup: Results, Cirrhosis



Backup: Results, Diabetic retinopathy