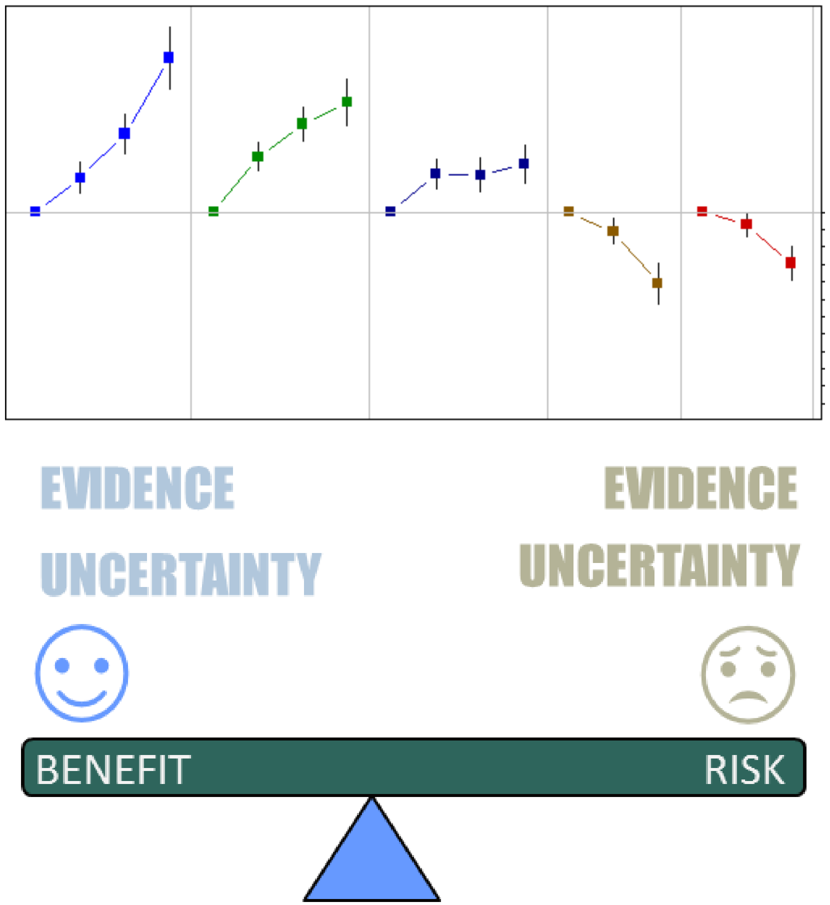
Benefit Risk Assessment Using Bayesian Discrete Choice Experiment

Saurabh Mukhopadhyay
Statistical Innovation Group
Data and Statistical Sciences, R&D
AbbVie Inc.
Email: saurabh.mukhopadhyay@abbvie.com

DIA Bayesian Scientific Working Group KOL Lecture Series

October 18, 2019



The support of this presentation was provided by AbbVie. AbbVie participated in the review and approval of the content.

Saurabh Mukhopadhyay is an employee of AbbVie.

Why Benefit-Risk Assessment?

Early assessment of benefit-risk (B-R) balance provides clarity on the treatment's utility in the population of interest

- Can help inform sponsors' decisions about the drug development programs

Due to increasing demand for evidence-based value judgments, B-R assessment of a treatment is very important throughout the drug lifecycle

- Sponsors generate B-R evidence to support their NDAs/BLAs
- Regulatory authorities use it to make decision on approvals and marketing authorizations

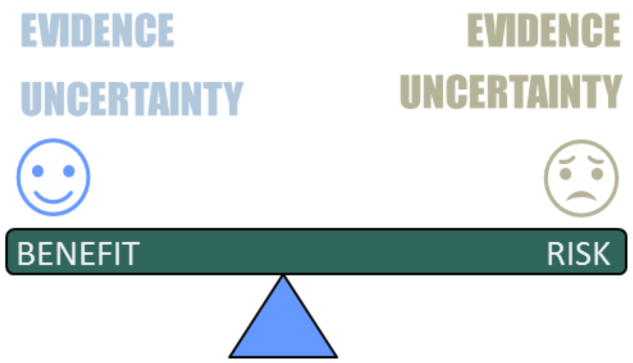
Patients and other stakeholders gain further insight on drug's benefit-risk balance and risk management

Some Challenges in B-R Assessment

- B-R assessment of a treatment is complex and involve confronting trade-offs between multiple, often conflicting features or attributes
- A large body of research shows that people are limited in the amount of information they can combine intuitively in balancing benefits and risks of a new treatment
- The problem is particularly acute for integrating the evidences across the attributes
- Also the decision on benefit-risk balance of a treatment may vary for different stakeholders, or for different subgroups in the patient population

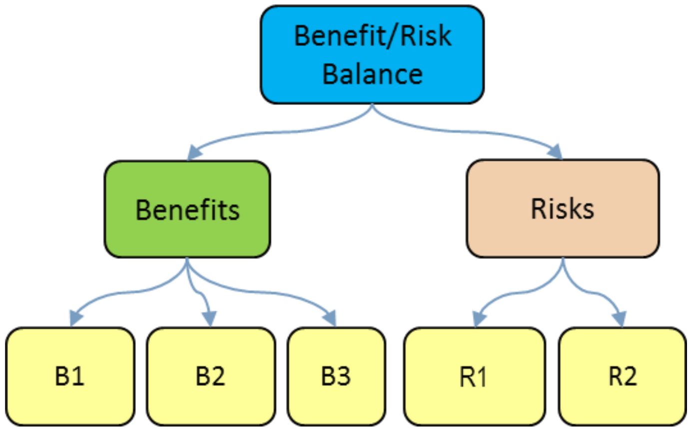
Structured Quantitative Framework for B-R Assessment

- A structured quantitative framework based on assessment of benefit and risk attributes can improve the quality of this important decision making
- Factors for Benefit-Risk Determination
  - Features or attributes
  - Magnitude, Severities, Probabilities – may be expressed as levels
  - Tradeoffs or relative importance of various benefit-risk attributes



Value Trees

- Value trees are a visual, hierarchical depiction of key aspects of a treatment that are of value to the decision-makers to understand which benefits and risks are pivotal to the benefit-risk balance.
  - A value tree provides a visual map to the research question
  - Important 1st step to identify attributes and levels



abbvie

DATA & STATISTICAL SCIENCES

- **Background**
  - Tradeoff Task
  - Discrete Choice Experiment
- Bayesian B-R Approach
  - DCE with Choice Pairs
  - Part Worth Utility
  - HBBR Utility Model
- Pilot Experiment
  - Model Fitting
  - Utility Scores
  - Overall B-R Balance
- R-Package for HBBR
- Augmented HBBR
- Summary

Selection of B-R Attributes and their Levels

Selection of attributes

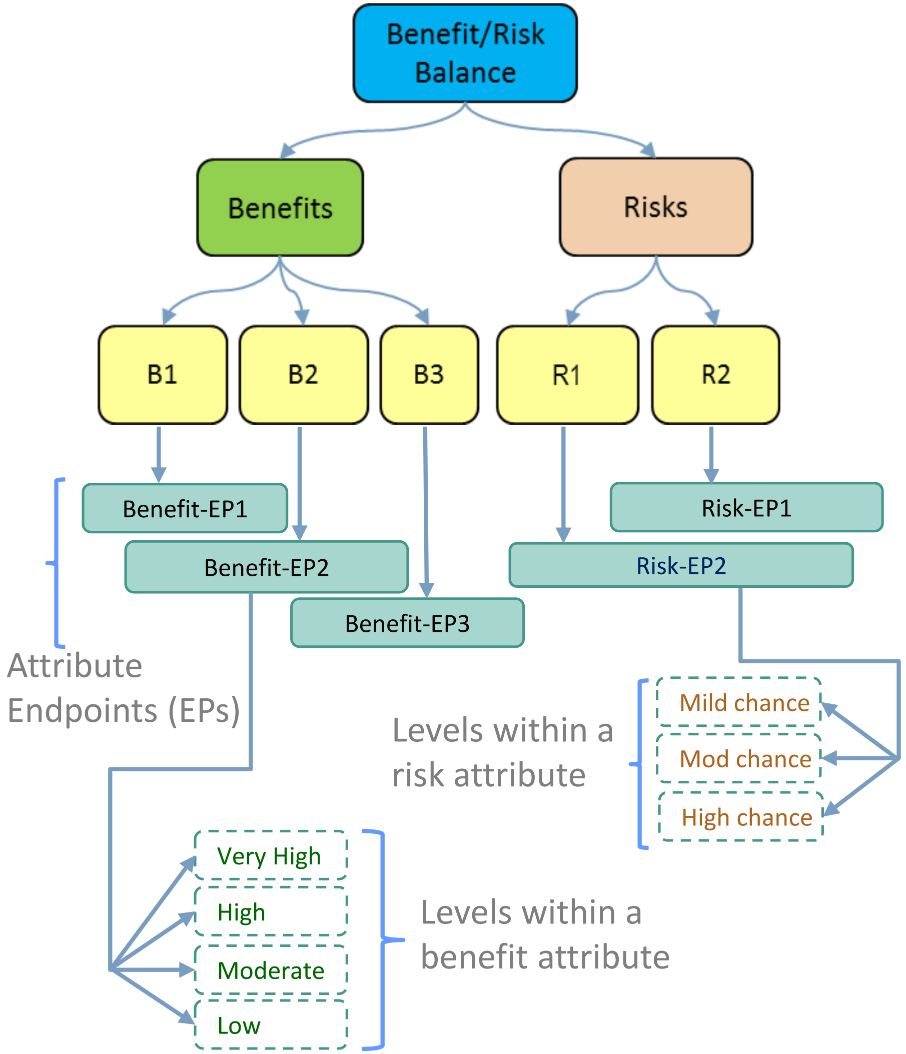
- Identify key efficacy and safety *attributes* that are relevant to research question
- Value-tree helps to connect fundamental objectives with attributes and endpoints

Identification of Levels

- Should avoid too many levels within each attribute
- General recommendation is to limit 3 to 4 levels per attribute
- Should avoid extreme values

Process should be transparent

- Assumptions should be stated
- Require literature review and discussion with medical experts and other stakeholders for identification of attributes/levels and calibration



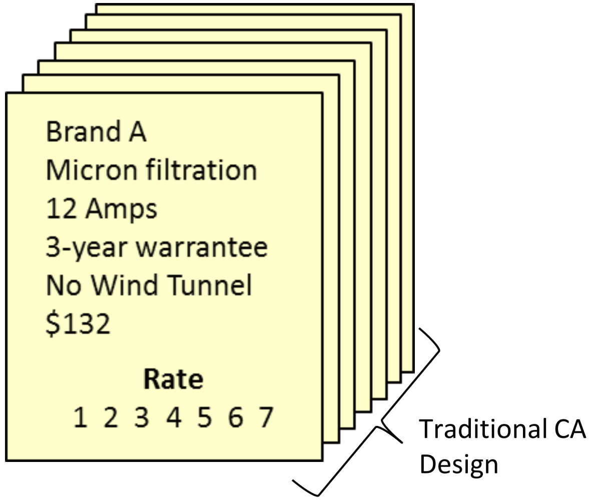
abbvie

DATA & STATISTICAL SCIENCES

Quantitative Assessment of Tradeoffs

Conjoint Analysis (CA) is the primary set of statistical technique to quantitatively assess tradeoffs among multi-attributed products or services

- An experimental method developed in the field of marketing research and has evolved over many decades
- Determines how people value different combinations of *attributes* and their *levels* that make up hypothetical product or service profiles
- In traditional CA, the profiles are presented to respondents for evaluation to express their underlying *tradeoffs*
  - Respondents rank or rate the profiles which are often very hard
- Conjoint experiments eliciting choice responses are known as discrete choice experiments (DCE)
  - Choosing among the profiles are easier than rating/ranking



abbvie

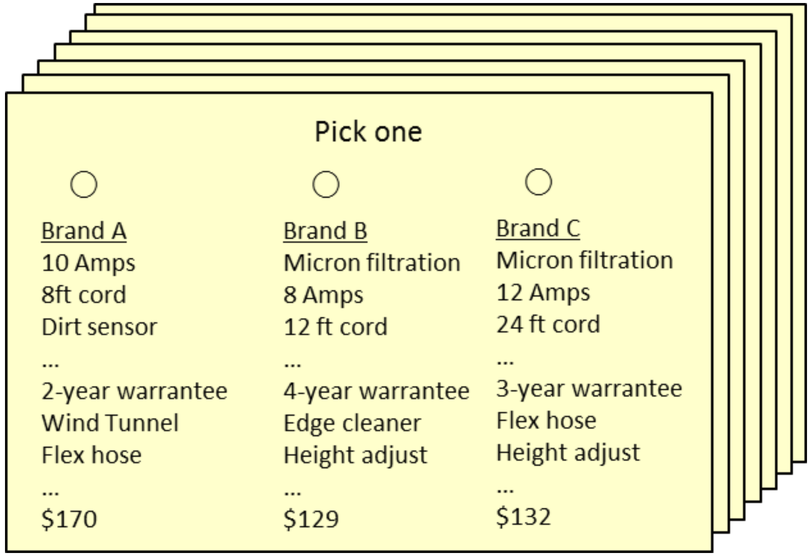
DATA & STATISTICAL SCIENCES

Discrete Choice Experiment

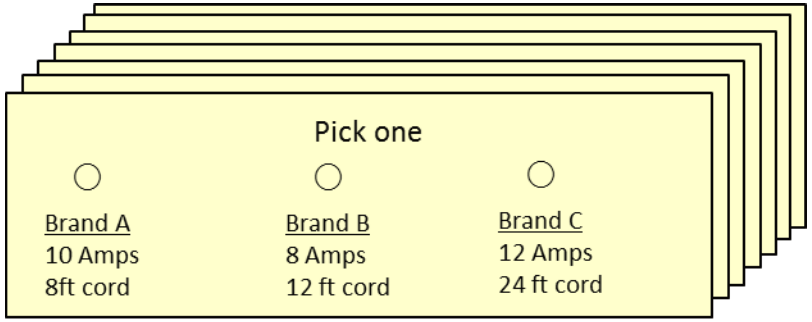
Discrete choice experiment (DCE) is also known as Choice Based Conjoint (CBC)

- In DCE respondents choose among sets of experimentally controlled sets of profiles
  - More discrimination power from tradeoff questions
- Determines how combination of attributes and levels can influence the overall choice or decision making
- Often recommended in heath outcome research
- Still poses a high cognitive burden to compare and choose from a set of full profiles
  - Even with a moderately large number of attributes produces a very large number of comparison each with high cognitive burden
- Partial Profiles look into a subset of attributes at a time
  - Less cognitive burden thus produces more quality responses
  - Also relatively fewer number of comparisons

Full Profiles



Partial Profiles



DCE with Partial Profiles in B-R Assessment

Even with partial profiles, usage of DCE in B-R assessment so far is very limited

- A large pool of respondents would be required to ensure proper estimation of underlying parameters using traditional frequentist methods
- Still requires each respondent to evaluate a large number of questions

Bayesian methods are ideally suited for such situation as they can leverage borrowing strength for analysis with limited data

abbvie
DATA & STATISTICAL SCIENCES

Bayesian Experiment and Modeling Framework

A Bayesian framework is proposed that *borrows strength* from respondents

- Allows to conduct the DCEs with only a limited number of respondents
- Respondents to choose only from a few pairs of profiles to state their preferences
  - Thus drastically reducing the cognitive burden

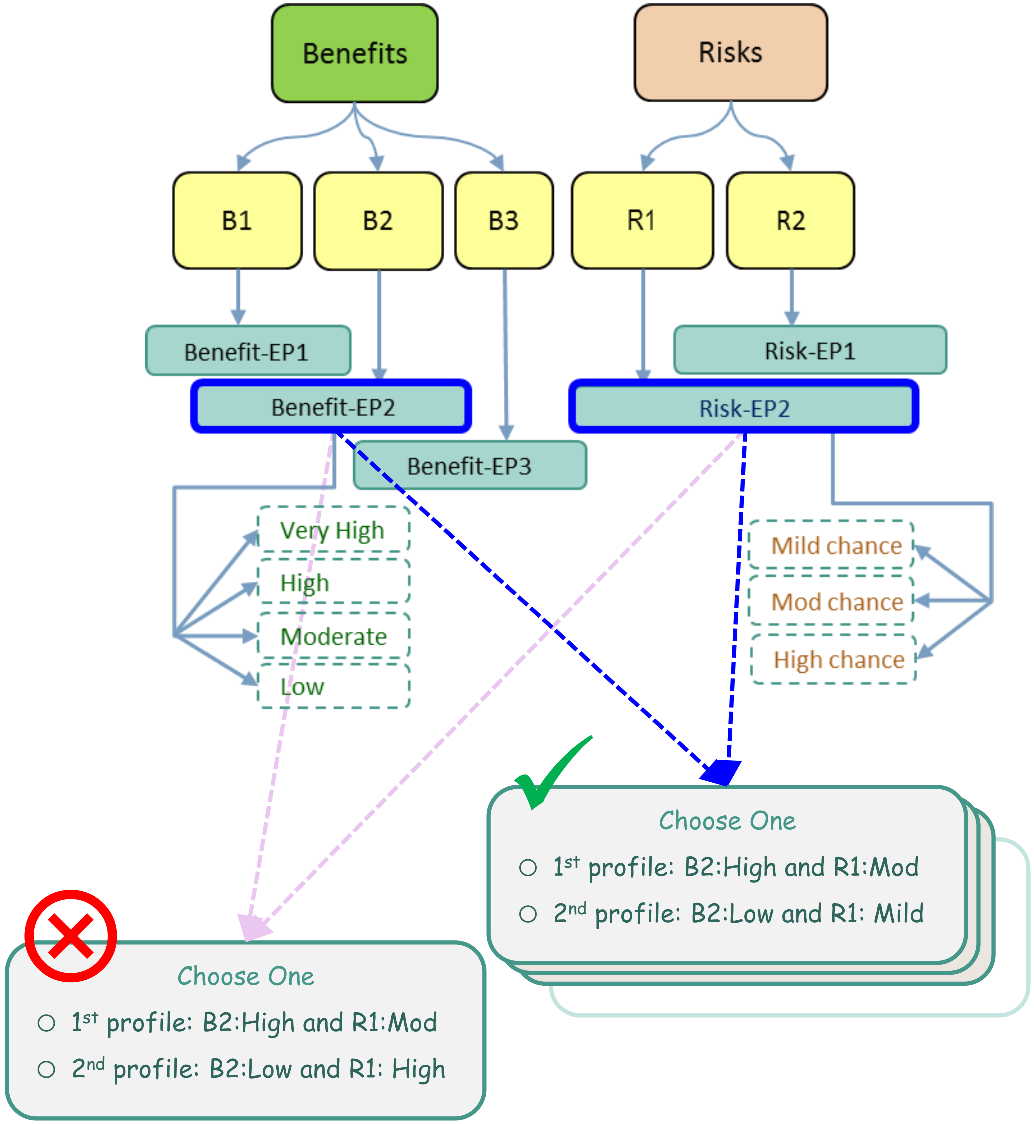
abbvie

DATA & STATISTICAL SCIENCES

Choice Pairs for the B-R Tradeoff Tasks

We propose a specific type of partial profile DCE based on 'choice pairs' questions:

- Each tradeoff task will consist of comparing *two* partial profiles – 'choice pair'
- One ***B*** attribute and one ***R*** attribute to be chosen at a time to prepare two partial profiles to construct a choice pair
- Not all paired alternatives will reflect real need for deliberation of tradeoffs - only *realistic (non-dominant) trade-off* tasks to be used
- Respondents will state their preferences by choosing one profile from each of the choice pairs presented to them

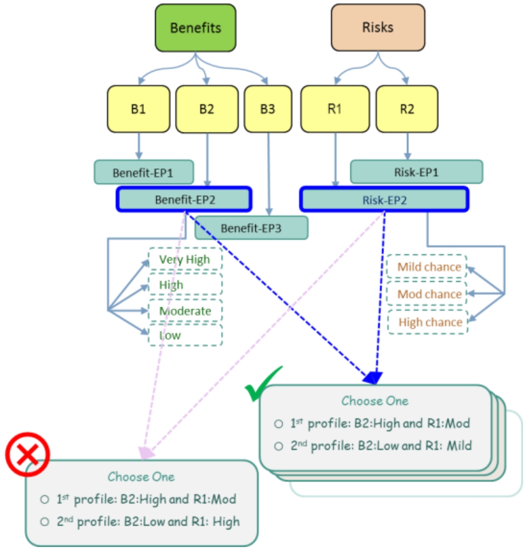


abbvie

DATA & STATISTICAL SCIENCES

Construction of Questionnaire

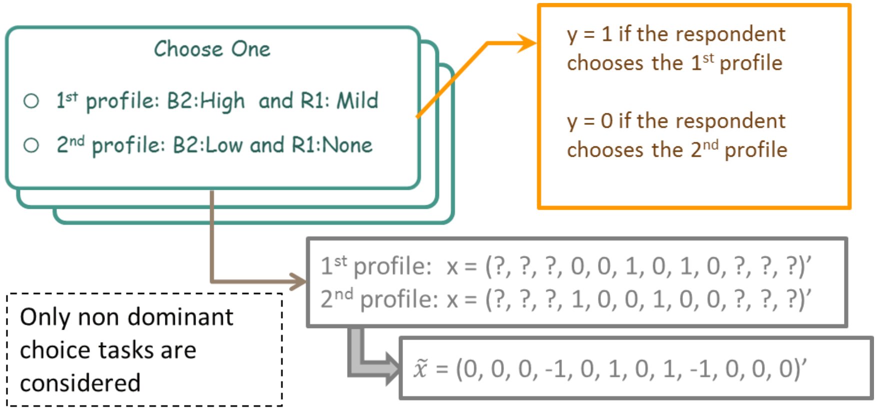
- Suppose there are *b* increasing levels in a Benefit attribute and *r* increasing levels in a Risk attribute:
  - There will be $\binom{b}{2}\binom{r}{2}$ non-dominant choice pairs for the particular combination of B and R attributes
- With S number of B attributes and L number R attributes, and j-th benefit attribute has b_j increasing levels and k-th risk attribute has and r_k increasing levels
  - There will be a total of $\mathrm{M} = \sum_{j=1}^{S}\sum_{k=1}^{L}\binom{b_j}{2}\binom{r_k}{2}$ non-dominant choice pairs in the experiment
  - If $b_j \equiv b$ and $r_k \equiv r$ then $M = S \cdot L \cdot \binom{b}{2}\binom{r}{2}$
- Questionnaire panels constructed by selecting a fixed (<<M) number of choice-pairs
  - Each respondent to evaluate one panel



Preference Data Format

Each respondent will be presented with a questionnaire panel consisting of a random subset of choice pairs

- Fixed number of trade-off tasks per respondent
- Respondent will state preferences by choosing one profile from each of the choice pairs in the panel
- Raw responses will then be processed
  - $y_{h,i}$ is the binary (1 or 0) response from h-th respondent for the i-th paired comparison task
  - $\tilde{x}_{h,i}$ is a vector of attribute differences – taking on values of 1, -1, or 0 based on whether the corresponding attribute level is in the 1st profile, 2nd profile or absent in both profiles, respectively

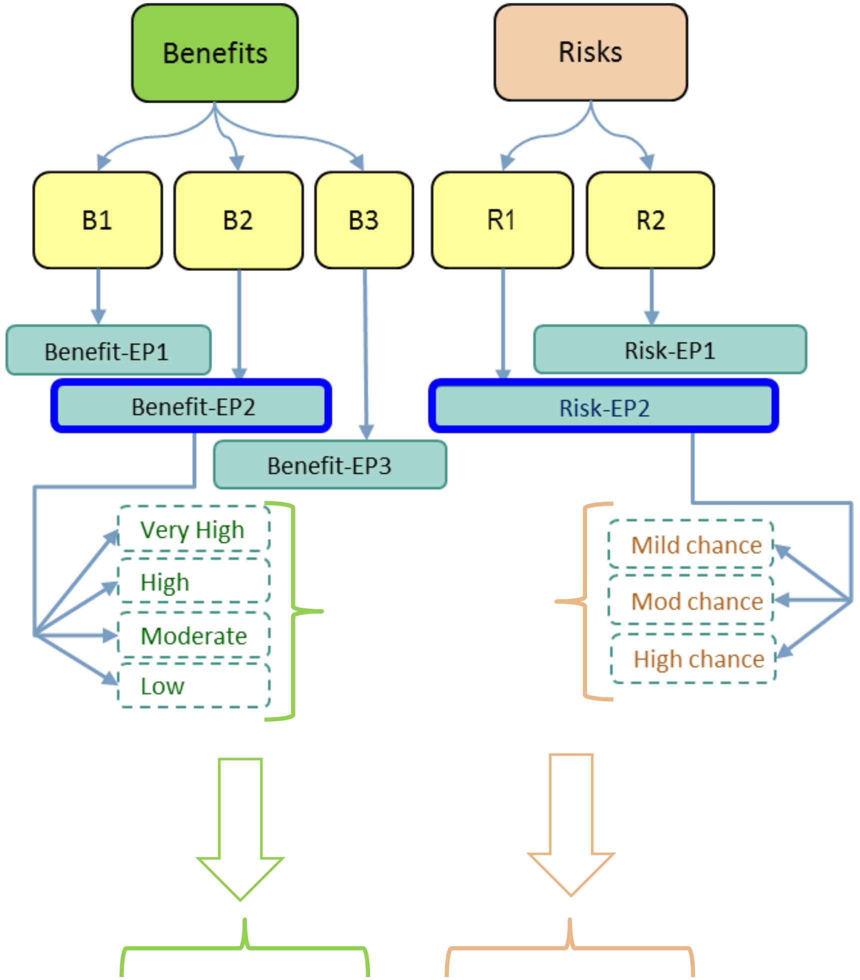


	id	y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
1	1	0	-1	1	0	0	0	0	0	-1	1	0	0	0
2	1	0	1	-1	0	0	0	0	1	-1	0	0	0	0
3	1	0	-1	1	0	0	0	0	-1	0	1	0	0	0
4	1	0	0	1	-1	0	0	0	0	0	0	0	1	-1
5	1	0	0	1	-1	0	0	0	0	0	0	1	0	-1
6	1	0	1	0	-1	0	0	0	0	0	0	0	1	-1
7	1	0	0	0	0	0	1	-1	1	-1	0	0	0	0

‘Part-worth’ measurements

A conjoint experiment decomposes customers’ preferences into “part-worth” measurements or preference scores with each level of each attribute

- We denote the part-worth vector from a respondent h as β_h
- Once the β_h are estimated, they can be recombined to estimate preferences for any possible combination of attribute levels



$$\beta_h = (\beta_{h,1}, \dots \beta_{h,i_1}, \beta_{h,i_1+1}, \dots, \beta_{h,i_2}, \beta_{h,i_2+1}, \dots, \beta_{h,k})$$

Hierarchical Bayes Benefit-Risk (HBBR) Utility Model

In the Hierarchical Bayes benefit risk (HBBR) approach we propose to use a random utility model to estimate the benefit-risk of a treatment

- Specifically, the overall B-R utility of a treatment profile for h^{th} respondent is modelled as

$$u_{\mathrm{h}} = x_{\mathrm{h}}' \beta_{\mathrm{h}}$$

  where, u_h is the overall B-R utility of a treatment profile from h^{th} respondent and x_h is a vector of 1's and 0's indicating whether or not the attribute levels are present in the treatment profile

We assume a hierarchical Bayes structure for the part-worth vectors β_h that borrows strength across and within respondents

- Will allow to work with only a limited number of respondents
- Also, each respondent needs to evaluate only a fraction of all choice pairs, thus respondents would not be fatigued from a long questionnaire

Hierarchical Prior Model for Part-Worth Vectors

Conjugate hierarchical priors are assumed for the part-worth vectors β_h

- Multivariate normal and inverse-Wishart priors are used

$$\beta_h \sim MVN(\bar{\beta}, V_\beta)$$
$$\bar{\beta} \sim MVN(\bar{\bar{\beta}}, B)$$
$$V_\beta \sim IW(\nu, V)$$

- Here $\bar{\beta}$ represents population level part worth utilities – parameter of interest

abbvie

DATA & STATISTICAL SCIENCES

Linking Stated Preference to the Utility Model

The modeling task is complicated as individual utility or preference is not directly observable in preference data

- If it is assumed that the preferred options will have higher utility it can be then derived (under some nominal assumption) that the preference probabilities can be expressed as multinomial logit (McFadden 1974)
- For our proposed DCE design with choice pairs, a (binomial) logit link connects the stated preferences to the utility model

$$P\left[y_{h,i} = 1\right] = logit\left(\tilde{x}_{h,i}'\beta_h\right) = \frac{\exp[\tilde{x}_{h,i}'\beta_h]}{1 + \exp[\tilde{x}_{h,i}'\beta_h]}$$

- Where $y_{h,i}$ is the response from hth respondent for the ith paired comparison and $\tilde{x}_{h,i}$ is a vector of attribute differences taking on values of 1, -1, or 0

Pilot implementation of HBBR approach for AML indication

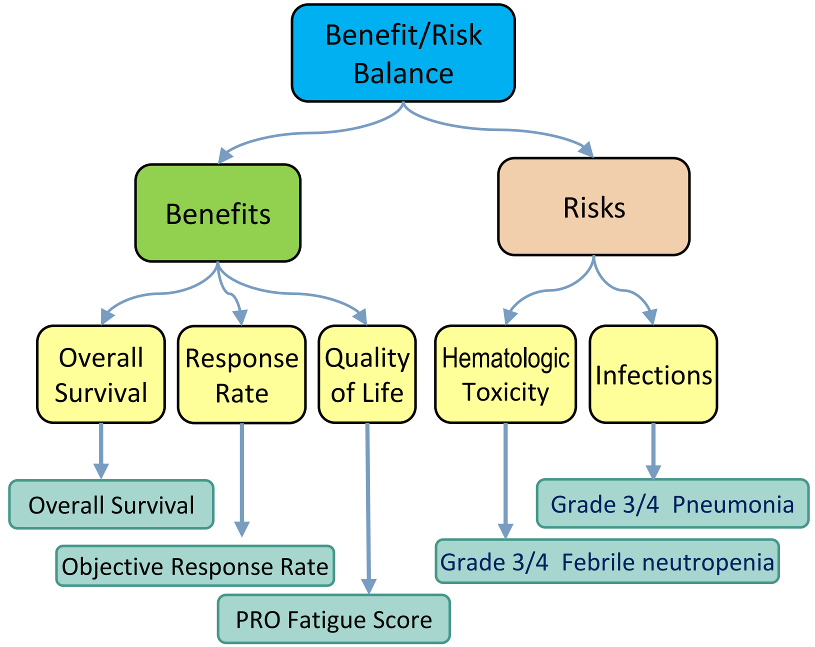
In this pilot experiment we identified

- 3 key benefit attributes each with 4 levels

Benefits	Less benefit than expected (E1)	Expected or most likely outcome (E2)	Better than expected, but reasonable (E3)	Way better, unlikely, but possible (E4)
Overall survival (months)	12	15	20	30
Objective response rate (%)	45	60	75	85
Fatigue reduction (%)	20	25	35	45

- 2 key risk attributes each with 3 levels

Risks	Better (less serious/severe or frequent) (H1)	Expected or most likely outcome (H2)	Worse (more serious/severe or frequent) (H3)
Hematologic toxicity (febrile neutropenia, %)	20	40	60
Infections (severe pneumonia, %)	5	10	20



OS: Overall Survival: 12, 15, 20, 30 months

ORR: Objective Response Rate : 45, 60, 75, 85%

FTG: Fatigue Improvement: 20, 25, 35, 45% chance

FebNeu: Febrile neutropenia: 20, 40, 60% chance

SevPNA: Pneumonia: 5, 10, 20% chance

Construction of Questionnaire Panels

Generating the Questionnaires:

- ALL non-dominant distinct choice pairs constructed
  - Each choice pair consists levels from one Benefit and one Risk attribute
  - There are a total of 108 choice pairs (108 = 3x2x6x3)
- Questionnaire panels were generated each with 18 choice pairs
  - Panels were constructed randomly with the aim of having 3 or more responses from each choice pair
  - 40 such panels were generated

1. Please select the most preferred of the following two options
   - ○ High (~20 months) OS and High (~60%) chance of febrile neutropenia
   - ○ Low (~12 months) OS and Moderate (~40%) chance of febrile neutropenia

2. Please select the most preferred of the following two options
   - ○ Low (~60%) ORR and Moderate (~10%) chance of severe pneumonia
   - ○ Very High (~85%) ORR and High (~20%) chance of severe pneumonia

abbvie

DATA & STATISTICAL SCIENCES

Stated Preference Data

Preference data were obtained from 23 SMEs

```
library(hbbr)
data(hbbrPilotResp)
brd = hbbrPilotResp$brdta
brd = data.frame(brd)
names(brd) = c("id", "y", "OS.L2","OS.L3","OS.L4", "ORR.L2", "ORR.L3","ORR.L4", "Ftg.L2","Ftg.L3","Ftg.L4",
                "Neut.L2","Neut.L3","Pneu.L2","Pneu.L3")
tail(brd)
```

	id	y	OS.L2	OS.L3	OS.L4	ORR.L2	ORR.L3	ORR.L4	Ftg.L2	Ftg.L3	Ftg.L4	Neut.L2	Neut.L3	Pneu.L2	Pneu.L3
409	23	1	0	0	0	0	0	0	-1	0	0	1	-1	0	0
410	23	0	0	0	0	0	0	0	0	0	1	-1	1	0	0
411	23	0	0	0	0	0	0	0	0	0	-1	-1	0	0	0
412	23	0	0	0	0	0	0	0	0	0	-1	0	0	0	-1
413	23	0	0	0	0	0	0	0	-1	1	0	0	0	-1	1
414	23	1	0	0	0	0	0	0	-1	0	0	0	0	-1	0

17. Please select the most preferred of the following two options
- ○ High (~35%) chance of FTG improvement and High (~20%) chance of severe pneumonia
- ● Moderate (~25%) chance of FTG improvement and Moderate (~10%) chance of severe pneumonia

HBBR Model to be used for the Preference Data

Recall the HBBR model:

$$P\left[y_{h,i} = 1\right] = \frac{\exp[\tilde{x}_{h,i}{}'\beta_h]}{1 + \exp[\tilde{x}_{h,i}{}'\beta_h]}$$

where

$$\beta_h \sim MVN(\bar{\beta}, V_\beta)$$
$$\bar{\beta} \sim MVN(\bar{\bar{\beta}}, B)$$
$$V_\beta \sim IW(\nu, V)$$

We specify hyper-parameters:

$$\bar{\bar{\beta}} = \mathbf{0}, \;\; B = 100 \cdot I$$

$$\nu = m + 2, \;\; V = \nu^{-1} \cdot I$$

We use an R-package `hbbr` to fit the model

Fitting HBBR Model using ‘hbbr’ Package

abbvie

DATA & STATISTICAL SCIENCES

Function `hbbr.Fit(…)` included in the `hbbr` package fits the model

```
hbfit = hbbr.Fit(brdta=brd, design=hbbrPilotResp$design,
                  mcmc=list(burnin=500, iter=10000, nc=2, thin=10))
```

Usage

```
hbbr.Fit(brdta, design, tune.param = list(tau = 0.01, eta = NULL, df.add
  = 2), mcmc = list(burnin = 5000, iter = 1e+05, nc = 2, thin = 20))
```

Arguments

`brdta`	processed and coded survey response data to be fitted to the hbbr model. It is a data frame in which 1st two columns indicate subject id and subject response (y = 0 or 1), and remaining columns contain information on design matrix (X). See Details below for more information.
`design`	design information of the experiment: design = list(b, r, bl, rl, blbls, rlbls) where, b is number of benefit attributes, r is number of risk attributes, bl and rl are vectors of integers of length b and r indicating number of levels in j-th benefit attribute and k-th risk attribute, respectively. blbls, rlbls consists of labels for benefit and risk attributes. When blbls is NULL, it uses "B1", "B2", ... and similarly for rlbls.
`tune.param`	a list of tuning hyper-parameters to be used; default tune.param=list(tau=0.01, eta=NULL). See Details below for more information.
`mcmc`	a list of mcmc parameters to be used in the Gibbs sampler to obtain posterior samples of the paramaters of interests; default: mcmc=list(burnin=5000, iter=100000, nc=2, thin=20). See Details below for more information.

```
print (hbbrPilotResp$design)
```

```
$b
[1] 3

$r
[1] 2

$bl
[1] 4 4 4

$rl
[1] 3 3
```

Summary of Average Part-Worth Utilities

Summary of average part-worth ($\bar{\beta}$) produced by `hbbr.Fit(…)`

```
Total Time Elapsed:  0.8


|**************************************************|
              summary of hbbr output
|**************************************************|
```

	mean	sd	2.5%	25%	50%	75%	97.5%	Rhat	n.eff
bbar[1]	1.9817	0.8892	0.2895	1.372	1.9627	2.6103	3.7095	1.002	760
bbar[2]	4.5013	1.1205	2.3341	3.785	4.4629	5.2202	6.7336	1.009	190
bbar[3]	8.8771	1.8127	5.6131	7.601	8.8069	10.0289	12.5728	1.011	150
bbar[4]	3.1914	0.8112	1.6472	2.637	3.1731	3.7273	4.8266	1.007	220
bbar[5]	5.0739	1.0076	3.1797	4.393	5.0509	5.7720	6.9876	1.024	79
bbar[6]	6.3152	1.3129	3.8110	5.445	6.3214	7.2037	8.9720	1.024	70
bbar[7]	2.1984	0.8726	0.5443	1.596	2.1728	2.7666	3.9564	1.000	1900
bbar[8]	2.1314	0.9639	0.2682	1.483	2.1087	2.7732	4.0346	1.003	520
bbar[9]	2.7781	1.1076	0.5593	2.080	2.8000	3.5012	4.8891	1.012	170
bbar[10]	-1.1079	0.7268	-2.5672	-1.592	-1.1099	-0.6336	0.3394	1.005	310
bbar[11]	-4.1297	1.1707	-6.5032	-4.874	-4.1270	-3.3276	-1.9442	1.008	190
bbar[12]	-0.7301	0.6365	-1.9416	-1.174	-0.7254	-0.3254	0.5764	1.011	140
bbar[13]	-2.9344	0.9736	-4.8351	-3.550	-2.9770	-2.2917	-0.9914	1.012	130
deviance	251.9389	20.6060	213.9608	238.774	251.4955	265.0703	292.2470	1.001	1900

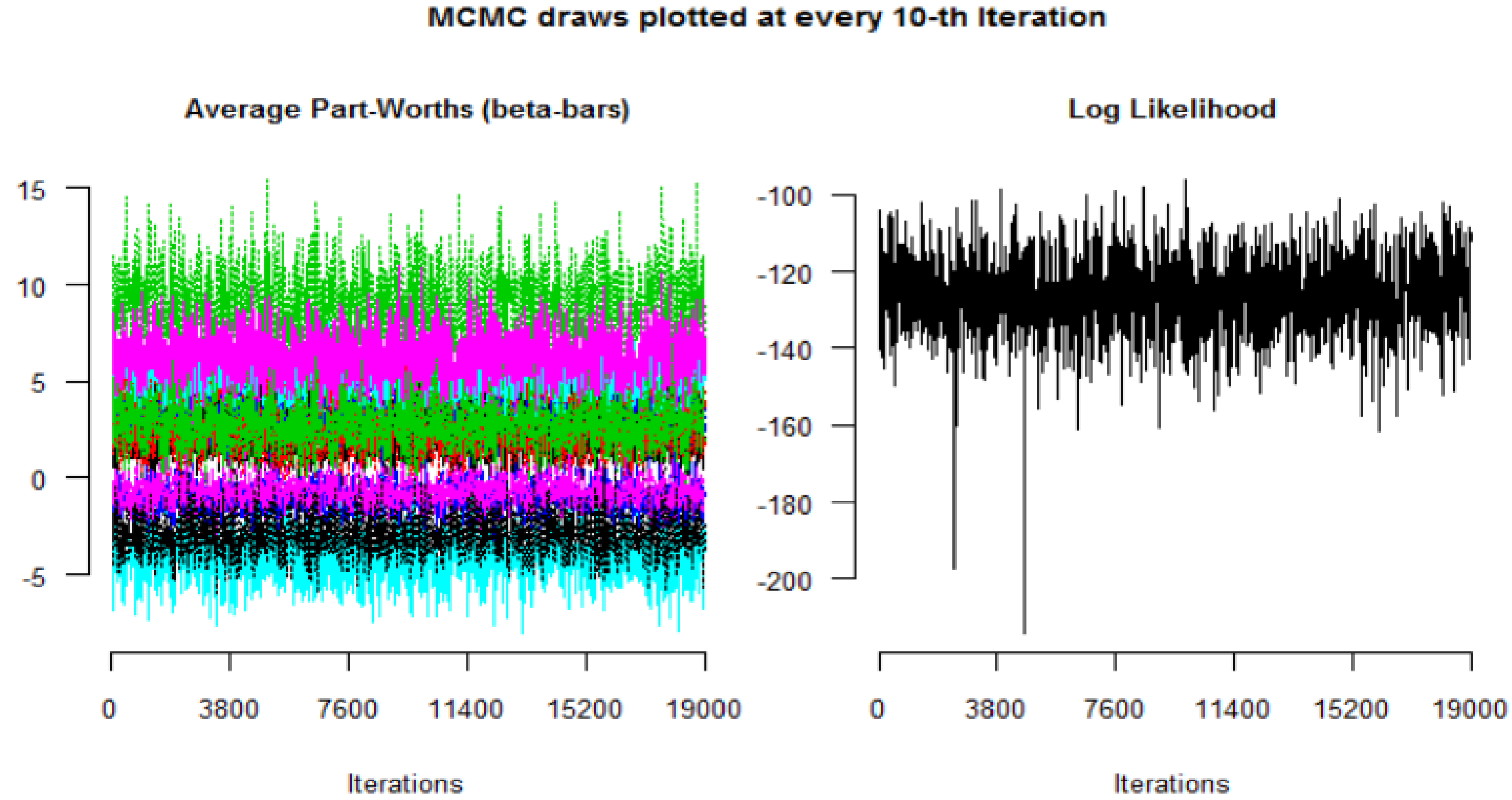
abbvie

DATA & STATISTICAL SCIENCES

Checking the MCMC Draws for $\bar{\beta}$

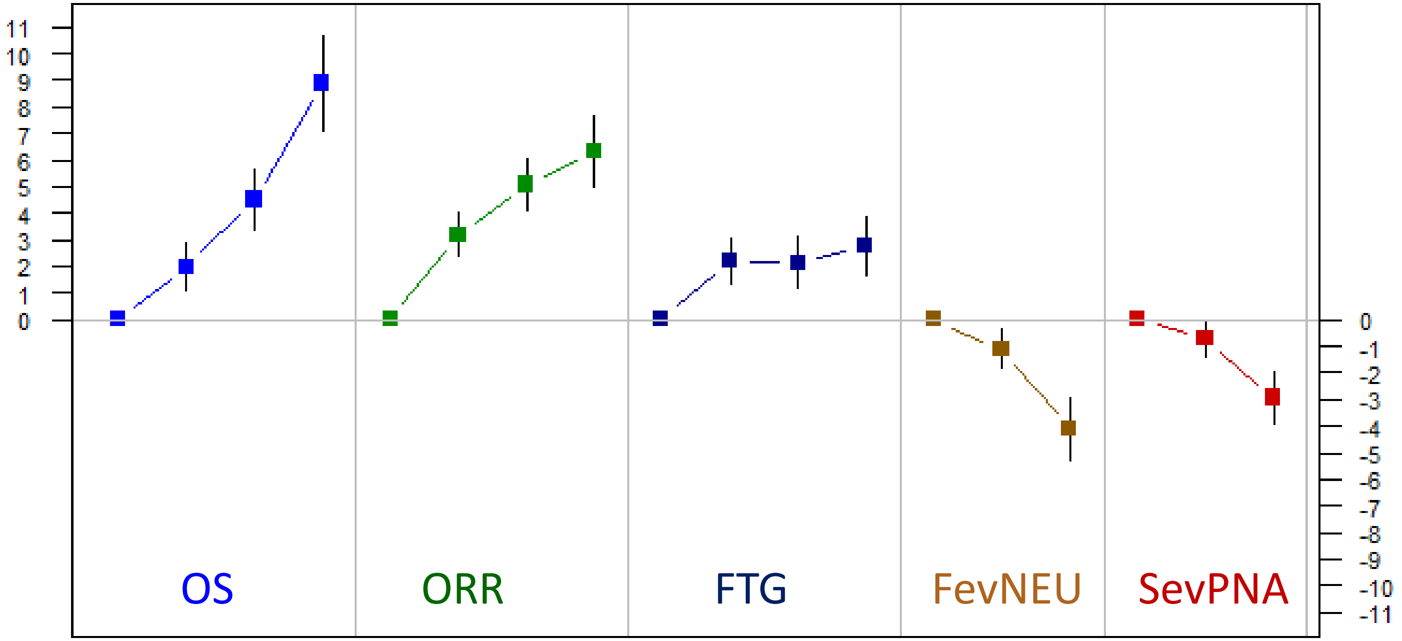
The HBBR model was fitted to the data using MCMC method

- The 1st plot shows traces of MCMC draws of $\bar{\beta}$
- The 2nd plot for the trace of log-likelihood ensures that the MCMC reached a stationary state



Summary of Average Part-Worth Utilities

Posterior estimates of average partworth utilities (mean ± SD) given the preference data



OS: Overall Survival: 12, 15, 20, 30 months

ORR: Objective Response Rate : 45, 60, 75, 85%

FTG: Fatigue Improvement: 20, 25, 35, 45% chance

FebNeu: Febrile neutropenia: 20, 40, 60% chance

SevPNA: Pneumonia: 5, 10, 20% chance

Some takeaways from the estimates:

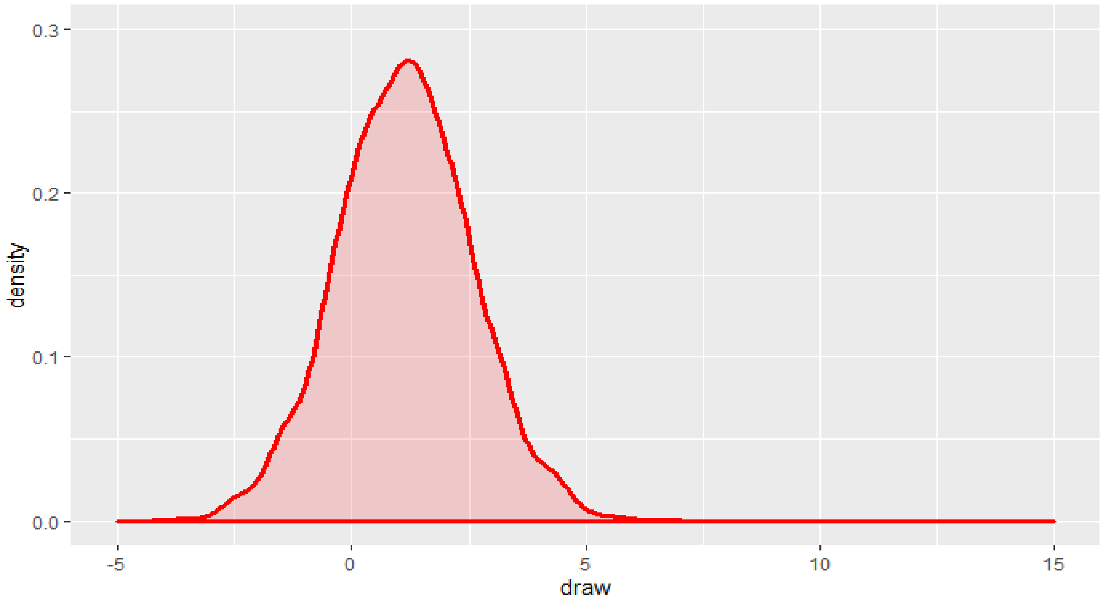
- If OS is very high (30 months) then average B-R is expected to be positive regardless of risk
- If risk of febrile neutropenia is high, OS and/or ORR must be high or very high for positive B-R
- Utility for fatigue improvement plateaus at 25% chance

Scoring a Treatment Profiles

Overall B-R utility score of a treatment profile can be obtained from the posterior distribution of $u = x'\bar{\beta}$

- Here x is the vector representing the treatment profile
- TP: OS 15 mo, ORR 60%, FTG 20% chance, Fev Neu 40% chance, and SevPNA 20% chance

$$x' = (0,1,0,0, \quad 0,1,0,0, \quad 1,0,0,0, \quad 0,1,0, \quad 0,0,1)$$



E[TP|Data] = 1.1

P(TP > 0|Data) = 0.79

abbvie

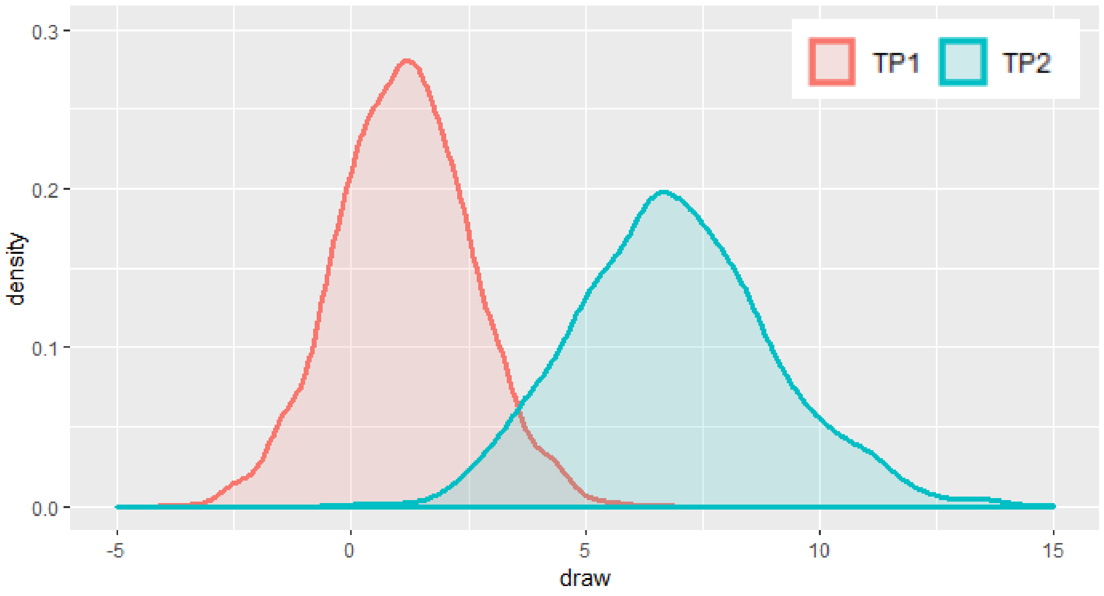
DATA & STATISTICAL SCIENCES

Comparing Treatment Profiles

We can easily compare overall B-R balance of two or more treatment profiles

- TP1: OS 15 mo, ORR 60%, FTG 20% chance, Fev Neu 40% chance, and SevPNA 20% chance
- TP2: OS 30 mo, ORR 75%, FTG 20% chance, Fev Neu 60% chance, and SevPNA 20% chance

$$x_1' = (0,1,0,0,\quad 0,1,0,0,\quad 1,0,0,0,\quad 0,1,0,\quad 0,0,1)$$
$$x_2' = (0,0,0,1,\quad 0,0,1,0,\quad 1,0,0,0,\quad 0,0,1,\quad 0,0,1)$$



E[TP2|Data) = 6.9

P(TP2 >0 |Data) > 0.999

E[TP2-TP1|Data) = 5.8

P(TP2 > TP1 |Data) > 0.997

Overall Assessment of Benefit-Risk Balance

For our pilot experiment there are 576 possible treatment profiles

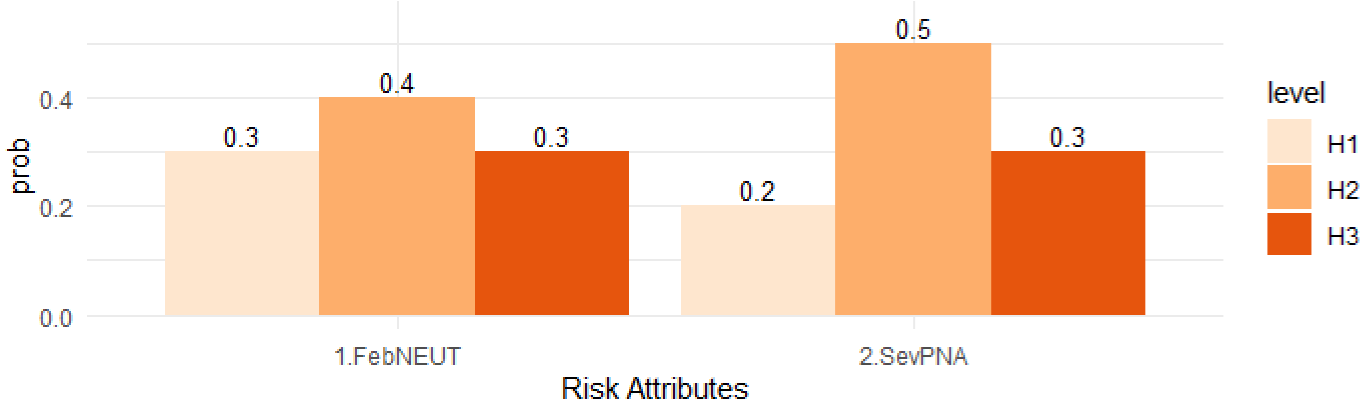
- We can study the overall utility score of all these profiles
- The posterior means of these utility score distributions range from -7 to 18

For a specific drug development program a large trial would provide us good estimates of frequencies of these treatment profiles

- We can then combine the various TPs using those frequencies to understand the distribution of patients experiencing various utility scores
- In absence of that information, we illustrate the steps by assuming known marginal proportion of levels within each attributes and if combination of attributes occur independently

abbvie
DATA & STATISTICAL SCIENCES

Assumed distribution of patients experiencing various attributes



- Consider e.g., TP2: OS E4, ORR E3, FTG E1, FevNeu H3, and SevPNA H3, then the proportion of patients with this treatment profile would be: 0.2 x 0.3 x 0.2 x 0.3 x 0.3 = 0.00108 = 0.108%
- We already know E[TP2|Data] = 6.9

abbvie

DATA & STATISTICAL SCIENCES

Proportion of patients on different treatment profiles

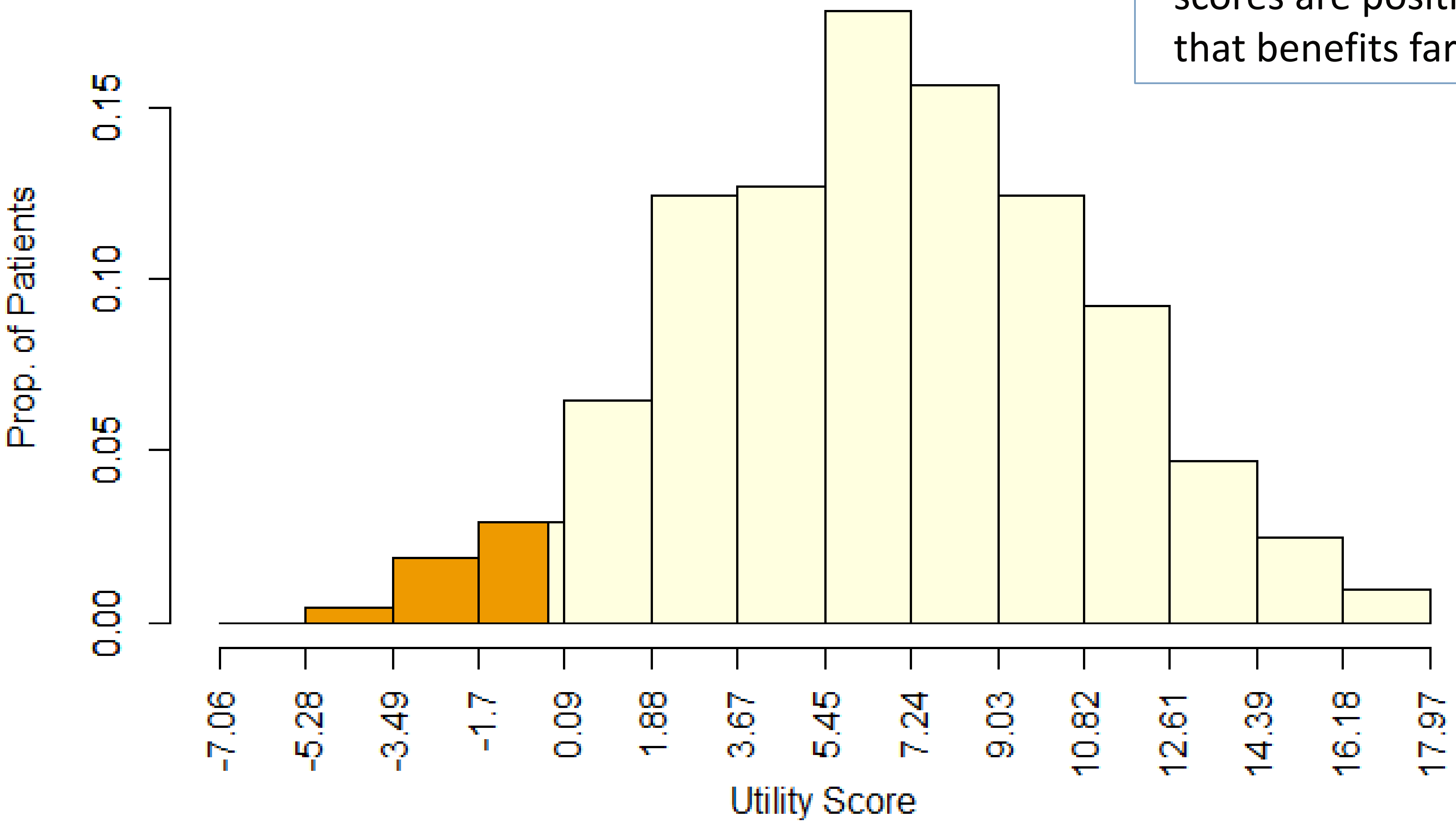
Similarly, we compute the proportions and corresponding posterior mean utility scores for all 576 profiles

```
head(profDistr, 10)
```

```
                        prfile       util    freq
1  b111 b211 b311 r111 r211  0.0000000 0.00024
2  b111 b211 b311 r111 r212 -0.7301269 0.00060
3  b111 b211 b311 r111 r213 -2.9343559 0.00036
4  b111 b211 b311 r112 r211 -1.1079205 0.00032
5  b111 b211 b311 r112 r212 -1.8380474 0.00080
6  b111 b211 b311 r112 r213 -4.0422763 0.00048
7  b111 b211 b311 r113 r211 -4.1297223 0.00024
8  b111 b211 b311 r113 r212 -4.8598493 0.00060
9  b111 b211 b311 r113 r213 -7.0640782 0.00036
10 b111 b211 b312 r111 r211  2.1984471 0.00036
```

abbvie
DATA & STATISTICAL SCIENCES

Overall distribution of patients experiencing various utility scores



Here over 95% of mean utility scores are positive, suggesting that benefits far outweigh risks

R Package and Codes

An R-package `hbbr` has been developed and submitted to CRAN

- `hbbr.Fit`: Fits response data to hbbr model
- `hbbrAug.Fit`: Fits the augmented hbbr model
- `hbbrPilotResp`: Contains response data from the pilot experiment and associated design information
- `simAugData`: Contains simulated data, design, baseline profiles, and true part-worth matrix for the Augmented HBBR model framework

The help files and vignettes include supplementary R codes

The package utilizes `R2jags` library

Hierarchical Bayesian Benefit-Risk Assessment Using Discrete Choice Experiment

Documentation for package 'hbbr' version 1.1.2

- DESCRIPTION file.

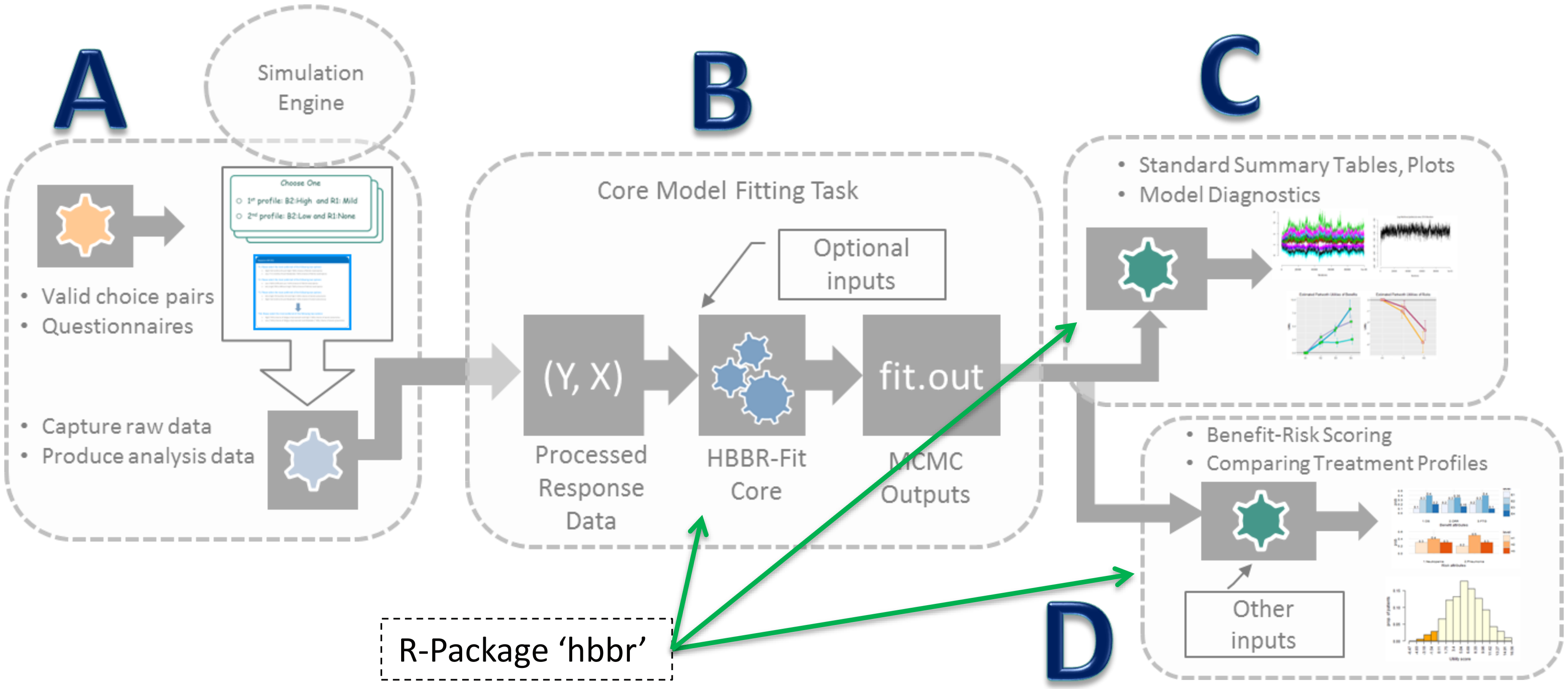
Help Pages

hbbr.Fit	hbbr.Fit (Fits processed response data to hbbr model)
hbbrAug.Fit	hbbrAug.Fit (Fits processed response data to the augmented hbbr model)
hbbrPilotResp	A list consisting of pilot data and associated discrete choice design information for the HBBR model framework.
simAugData	A list consisting of simulated data, design, baseline profiles, and true part-worth matrix for the Augmented HBBR model framework.

abbvie

DATA & STATISTICAL SCIENCES

End-to-end Implementation of HBBR



Augmented HBBR to include Patients' Characteristics

Ultimately, patients are the most important voice in the benefit–risk balance.

- So far, we did not assume that respondents' demographic or other characteristics could systematically influence their benefit–risk preferences
- In the real world however, it is likely that age, gender, disease status, and other baseline characteristics would affect the preferences that patients express
- Since patients are a key stakeholder for any benefit–risk assessment, it is extremely important to understand how those characteristics influence the benefit–risk preferences
- Furthermore, the ability to identify a subgroup of patients for whom benefit–risk conclusions might differ from the rest of the population could provide relevant information for indication and labeling claims as well as clinical guidance on most effective overall use of a new medication within a selected group of patients

Augmented HBBR

An augment HBBR model is proposed to implement the ability to examine contributing patients' characteristics

- Suppose that z_h represents the vector $(c \times 1)$ of observed baseline characteristics from h-th patient that can potential influence the benefit-risk preferences
- Recall that the B-R preferences are modelled through the part worth vector β_h $(m \times 1)$
- To incorporate z_h we now express

$$\beta_h = \Delta \cdot z_h + \varepsilon_h$$

where Δ is the matrix (of dimension $m \times c$) that incorporates the heterogeneity of regression coefficients due to baseline characteristics

- Conjugate priors specified to complete the model

$$\beta_h = \Delta \cdot z_h + \varepsilon_h, \quad \varepsilon_h \sim MVN\left(0, V_\beta\right)$$
$$\Delta = (\Delta_1, \Delta_2, \dots, \Delta_c), \qquad \Delta_j \sim MVN(0, Q^{-1})$$
$$V_\beta \sim IW(\nu, V)$$

Simulating Preference Data with Patients' Characteristics

An experiment with 2 benefit and 2 risk attributes was considered

- Each with 3 levels - thus m = (3-1)*2 + (3-1)*2 = 8
- N = 100 virtual patients
- Main effect and two baseline characteristics were used: age and disease status
  - Standardized age variable (z1) from standard normal was generated
  - Disease status (z2) 'yes' (1) and 'no' (-1) were assigned to 50:50 patients
- True Δ of dimension 8x3 was assumed:
  - The first column Δ_1 represents the overall part-worth effects; the second and third columns represent the additive effects of patient's age and disease status

```
# generate baseline characteristics:
set.seed(1234)
z1 = rnorm(N, 0,1)
z2 = rep(c(1,-1), N/2)
```

```
> Del
     [,1]  [,2]  [,3]
[1,]    3  0.70  1.20
[2,]   10  0.90  1.50
[3,]    2  0.01  0.05
[4,]    5  0.01  0.05
[5,]   -1  0.00 -0.90
[6,]   -7  0.00 -1.20
[7,]   -3 -0.50  0.00
[8,]   -5 -1.00  0.00
```

Simulating Preference Data with Patients' Characteristics

There are 36 total non-dominant choice pairs

- Questionnaire sets for 100 virtual patients were randomly generated each with 12 choice pairs

- Then virtual response were simulated using Bernoulli distribution where probability of choosing the 1st profile for each choice pair was computed from the logit link

$$P\left[y_{h,i} = 1\right] = \frac{\exp[\tilde{x}_{h,i}'\beta_h]}{1 + \exp[\tilde{x}_{h,i}'\beta_h]}$$

```
data("simAugData")
brdAug = simAugData$brdtaAug
brdAug = data.frame(brdAug)
names(brdAug) = c("id", "y", "B1.L2","B1.L3","B2.L2", "B2.L3", "R1.L2","R1.L3",
                  "R2.L2","R2.L3")
head(brdAug)
```

```
  id y B1.L2 B1.L3 B2.L2 B2.L3 R1.L2 R1.L3 R2.L2 R2.L3
1  1 0     1     0     0     0    -1     1     0     0
2  1 0    -1     0     0     0    -1     0     0     0
3  1 1     0     1     0     0     1     0     0     0
4  1 1     1     0     0     0     0     0     1     0
5  1 0     0    -1     0     0     0     0    -1     0
6  1 1     0     1     0     0     0     0     0     1
```

```
dim(brdAug)
```

```
[1] 1200   10
```

Fitting Augmented HBBR to the Simulating Preference Data

The R-package `hbbr` includes `hbbrAug.Fit()`

```
hbA = hbbrAug.Fit(brdta= brdAug, Z=simAugData$Z,
                  design=simAugData$design,
                  tune.param=list(tau=0.01, eta=NULL, df.add=2),
                  mcmc=list(burnin=500, iter=10000, nc=2, thin=10))
```

Augmented HBBR Model Specifications:

$$P\{y_{h,i} = 1\} = logit(X'_{h,i}\beta_h)$$

$$\beta_h = \Delta \cdot z_h + \varepsilon_h, \quad \varepsilon_h \sim MVN(0, V_\beta)$$

$$\Delta = (\Delta_1, \Delta_2, \Delta_3), \Delta_j \sim MVN(0, Q^{-1})$$

$$V_\beta \sim IW(\nu = m + 2 = 10, \Omega^{-1})$$

$$Q = 0.01 * I, \Omega = I$$

Note that we generate the posterior distributions of Δ matrix using the Gibbs chain; then obtain β_h estimates for any specified z_h as

$$\widehat{\beta_h} = \widehat{\Delta} \cdot z_h$$

where $\widehat{\Delta}$ is the posterior mean

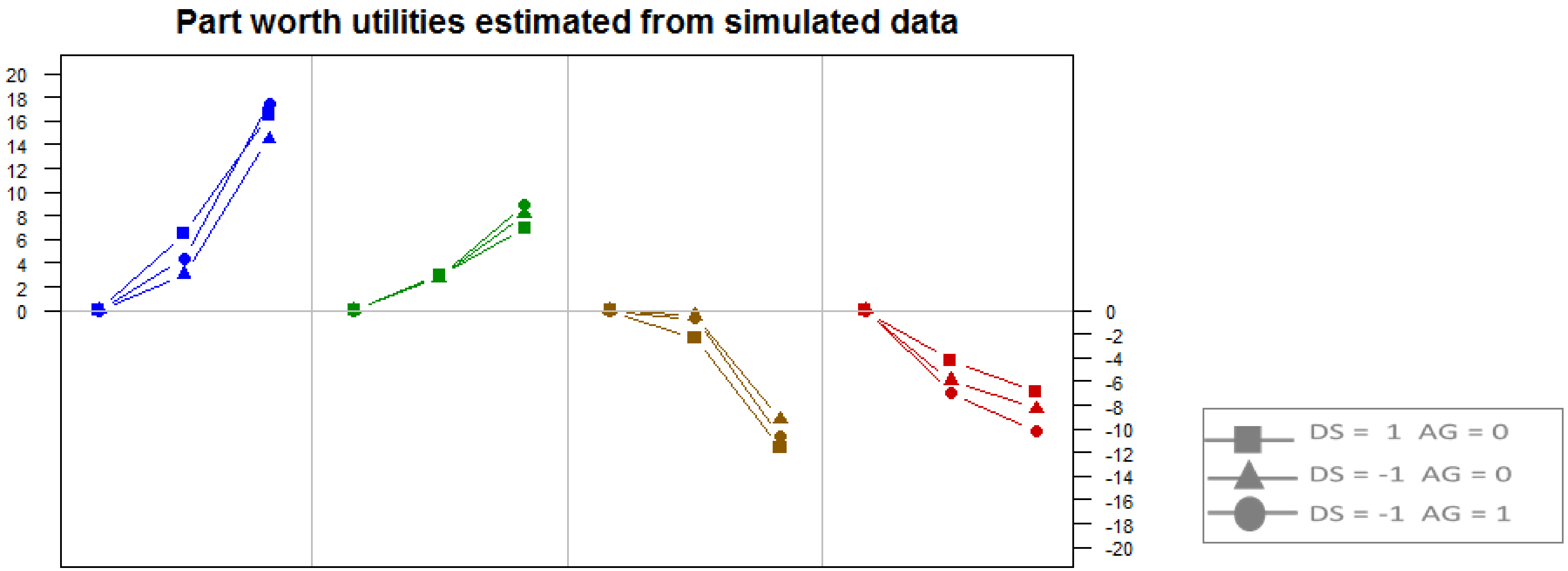
	mean	sd
Del[1,1]	4.77111	0.5862
Del[2,1]	15.51718	1.3229
Del[3,1]	2.84342	0.4305
Del[4,1]	7.55128	0.7528
Del[5,1]	-1.41375	0.4296
Del[6,1]	-10.44309	0.9436
Del[7,1]	-5.05170	0.5272
Del[8,1]	-7.64011	0.7597
Del[1,2]	1.30505	0.4831
Del[2,2]	2.99205	0.9307
Del[3,2]	0.09251	0.3842
Del[4,2]	0.80321	0.6013
Del[5,2]	-0.22883	0.4158
Del[6,2]	-1.34183	0.7141
Del[7,2]	-1.03237	0.4241
Del[8,2]	-1.91152	0.6028
Del[1,3]	1.73478	0.4729
Del[2,3]	1.05506	0.8896
Del[3,3]	0.07610	0.3700
Del[4,3]	-0.58933	0.5982
Del[5,3]	-0.94749	0.4057
Del[6,3]	-1.19701	0.7115
Del[7,3]	0.86696	0.4032
Del[8,3]	0.70273	0.5775
deviance	537.54853	30.9175

abbvie

DATA & STATISTICAL SCIENCES

Results from the Fitted Augmented HBBR Model

```
clrs = c("blue", "green4","orange4", "red3")
mns = hbA$del.means
betmn1 = mns %*% matrix(c(1, 0, 1), ncol=1)  # at mean age with disease staus=1
betmn2 = mns %*% matrix(c(1, 0, -1), ncol=1) # at mean age with disease staus=-1
betmn3 = mns %*% matrix(c(1, 1, -1), ncol=1) # at age = mean+1*SD, disease staus=-1
partworth.plot(attr.lvl = augattr.lvl, beta.mns = betmn1)
title("Part worth utilities estimated from simulated data")
partworth.plot(attr.lvl = augattr.lvl, beta.mns = betmn2, new=F, pnt=17)
partworth.plot(attr.lvl = augattr.lvl, beta.mns = betmn3, new=F, pnt=16)
```



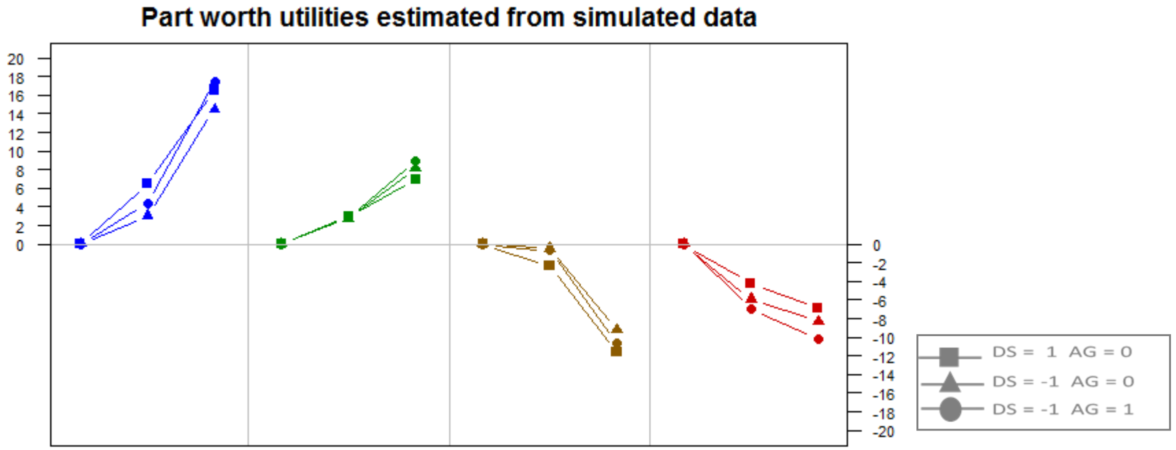
abbvie

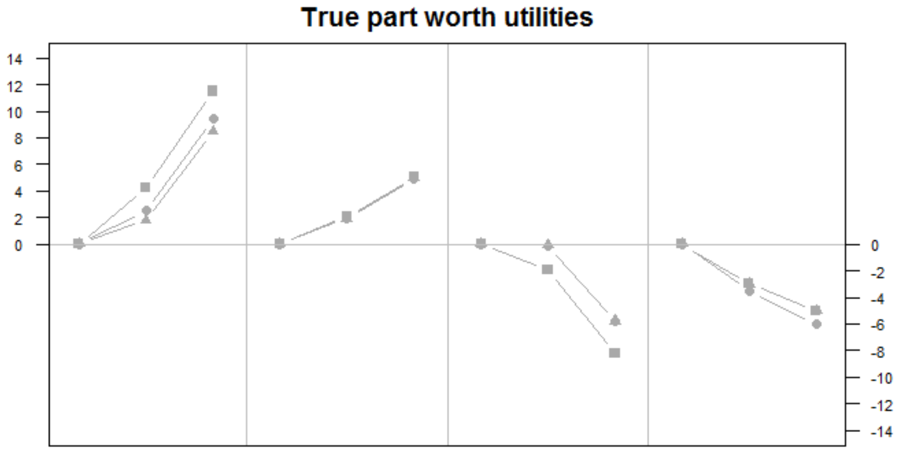
DATA & STATISTICAL SCIENCES

Comparing Fitted Results with True Values

We see an efficient recovery of part-worth utilities at various patient-level characteristics using the Augmented HBBR model

```
# Plotting true betas at those baseline characteristics
 Del = simAugData$Del
 clrs = rep("darkgrey", 4)
 # true part-worth values
 bmn1 = Del %*% matrix(c(1, 0, 1), ncol=1)  # at mean age with disease staus=1
 bmn2 = Del %*% matrix(c(1, 0, -1), ncol=1) # at mean age with disease staus=-1
 bmn3 = Del %*% matrix(c(1, 1, -1), ncol=1) # at age = mean+1*SD, disease staus=-1
 partworth.plot(attr.lvl = augattr.lvl, beta.mns = bmn1)
 title("True part worth utilities")
 partworth.plot(attr.lvl = augattr.lvl, beta.mns = bmn2, new=F, pnt=17)
 partworth.plot(attr.lvl = augattr.lvl, beta.mns = bmn3, new=F, pnt=16)
```





AbbVie
DATA & STATISTICAL SCIENCES

Key Steps for HBBR Approach

1 Identify attributes and assign levels

- Identify key efficacy and safety outcomes as attributes
- Assign levels based on
  - Strength of efficacy (or lack thereof)
  - Severity of safety outcomes
- Sources – consultation with medical experts in the TA, interviewing patients, literature review
  - Avoid too many levels
  - May require some validations

2 Generate Choice-pairs & Questionnaires

- Create *simple* and *realistic* hypothetical trade-off tasks
  - Each task involves two alternative treatment profiles
  - Profiles will differ only in one benefit and one risk attribute level
- Determine the number of trade-off tasks per respondent
- Produce randomized questionnaire panels
  - Preferably using online survey tools

3 Collect and Process Response Data

- Send (email) link to survey panels to study participants to respond within a short period of time
  - Study participants could be patients
  - OR medical professionals selected based on their experiences in patient care and/or expertise in the TA
- Raw responses to be coded according to the HBBR data coding instruction

4 Analyze & Report Benefit-Risk Assessment

- Fit the HBBR or Augmented HBBR model to the coded response data using `hbbr` package
  - The R-package fits the hierarchical Bayes model and estimates the part-worth utilities
  - Can be used to assess the utility of any full treatment profile or compare two or more profiles
  - Can be used to assess an overall B-R balance in a drug development program

Concluding remarks

- The proposed HBBR Framework consisted of a novel Bayesian approach for quantitatively assessing benefit-risk balance of a treatment
  - Borrows strength from respondents thus require a small number of respondents
  - Proposed DCE Design based on choice pairs is operationally efficient - consists of a modest number of easy-to-state-preference tasks per respondent
  - Expected to produce high-quality preference data as respondents would not become fatigued from a long questionnaire
  - Can be implemented at a very early stage of a drug development program – and can be updated as needed throughout the drug development lifecycle
  - Proposed augmented HBBR model allows to incorporate patients' characteristics to obtain a more precise estimate of benefit-risk balance
- Proper calibration of various attribute levels should be done in collaboration with experts in the therapeutic area using pilot experiments

DATA & STATISTICAL SCIENCES

Acknowledgement

I sincerely acknowledge Kimberley Dilley, Anthony Oladipo, and Jeremy Jokinen as coauthors of the manuscript referenced below from which part of the materials of this presentation were derived

Reference: Mukhopadhyay, S., Dilley, K., Oladipo, A. and Jokinen, J., Hierarchical Bayesian Benefit–Risk Modeling and Assessment Using Choice Based Conjoint. Statistics in Biopharmaceutical Research, (2019), 11(1), pp.52-60

Key References

1. Mukhopadhyay, S., Dilley, K., Oladipo, A., Jokinen, J. “Hierarchical Bayesian Benefit-Risk Modeling and Assessment Using Choice Based Conjoint” *Statistics in Biopharmaceutical Research* (2019), 11(1), pp. 52-60
2. Irony, T., Ho, M., Christopher, S., and Levitan, B. (2016), “Incorporating Patient Preferences Into Medical Device Benefit-Risk Assessments,” *Statistics in Biopharmaceutical Research*, 8, 230–236
3. Bridges, John FP, A. Brett Hauber, Deborah Marshall, Andrew Lloyd, Lisa A. Prosser, Dean A. Regier, F. Reed Johnson, and Josephine Mauskopf. "Conjoint analysis applications in health—a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force." *Value in health* (2011), 14(4), pp. 403-413
4. Ryan, M., Gerard, K. and Amaya-Amaya, M. eds., 2007. Using discrete choice experiments to value health and health care (Vol. 11). Springer Science & Business Media
5. McFadden, D., “Conditional logit analysis of qualitative choice behavior” (1973) in *Frontiers in Econometrics*, 105-142, P. Zarembka (ed.), Academic Press: New York
6. Baltas, G. and Doyle, P., Random utility models in marketing research: a survey. Journal of Business Research (2001), 51(2), pp.115-125

Thank you