

4. Individual level data in health economic evaluation

Anna Heath

Sick Kids Hospital & Dalla Lana School of Public Health | University of Toronto

- ✉ anna.heath@sickkids.ca
- 🌐 <https://sites.google.com/site/annaheathstats/>
- 🌐 <https://egon.stats.ucl.ac.uk/research/statistics-health-economics/>
- 🔗 <https://github.com/annaheath>
- 🔗 <https://github.com/StatisticsHealthEconomics>

Bayesian Methods in Health Economics, Lausanne

- Data & "standard" analysis
- Example: UK700
- Analysis of cost data
 - Alternative models
 - Goodness of fit
- Analysis of cost-effectiveness data
 - General structure
 - Alternative models

References

 *Bayesian Methods in Health Economics*, chapters 5.1, 5.2

 Book website (CRC)

Book website

Code

HTA alongside RCTs

- The available data usually look something like this

Demographics				Clinical outcomes				HRQL data				Resource use data				
ID	Trt	Sex	Age	...	y_0	y_1	...	y_J	u_0	u_1	...	u_J	c_0	c_1	...	c_J
1	1	M	23	...	y_{10}	y_{11}	...	y_{1J}	0.32	0.66	...	0.44	103	241	...	80
2	1	M	23	...	y_{20}	y_{21}	...	y_{2J}	0.12	0.16	...	0.38	1204	1808	...	877
3	2	F	19	...	y_{30}	y_{31}	...	y_{3J}	0.49	0.55	...	0.88	16	23	...	22
...

- y_{ij} = Survival time, event indicator (eg CVD), number of events, continuous measurement (eg blood pressure), ...
- u_{ij} = Utility-based score to value health (eg EQ-5D, SF-36, Hospital & Anxiety Depression Scale), ...
- c_{ij} = Use of resources (drugs, hospital, GP appointments), ...
- Usually aggregate longitudinal measurements into a cross-sectional summary and for each individual consider the pair (e_i, c_i)

HTA alongside RCTs

- The available data usually look something like this

Demographics				Clinical outcomes				HRQL data				Resource use data				
ID	Trt	Sex	Age	...	y_0	y_1	...	y_J	u_0	u_1	...	u_J	c_0	c_1	...	c_J
1	1	M	23	...	y_{10}	y_{11}	...	y_{1J}	0.32	0.66	...	0.44	103	241	...	80
2	1	M	23	...	y_{20}	y_{21}	...	y_{2J}	0.12	0.16	...	0.38	1204	1808	...	877
3	2	F	19	...	y_{30}	y_{31}	...	y_{3J}	0.49	0.55	...	0.88	16	23	...	22
...

- y_{ij} = Survival time, event indicator (eg CVD), number of events, continuous measurement (eg blood pressure), ...
- u_{ij} = Utility-based score to value health (eg EQ-5D, SF-36, Hospital & Anxiety Depression Scale), ...
- c_{ij} = Use of resources (drugs, hospital, GP appointments), ...
- Usually aggregate longitudinal measurements into a cross-sectional summary and for each individual consider the pair (e_i, c_i)
- HTA preferably based on **utility-based** measures of effectiveness
- Quality Adjusted Life Years (QALYs) are a measure of disease burden combining
 - Quantity** of life (the amount of time spent in a given health state)

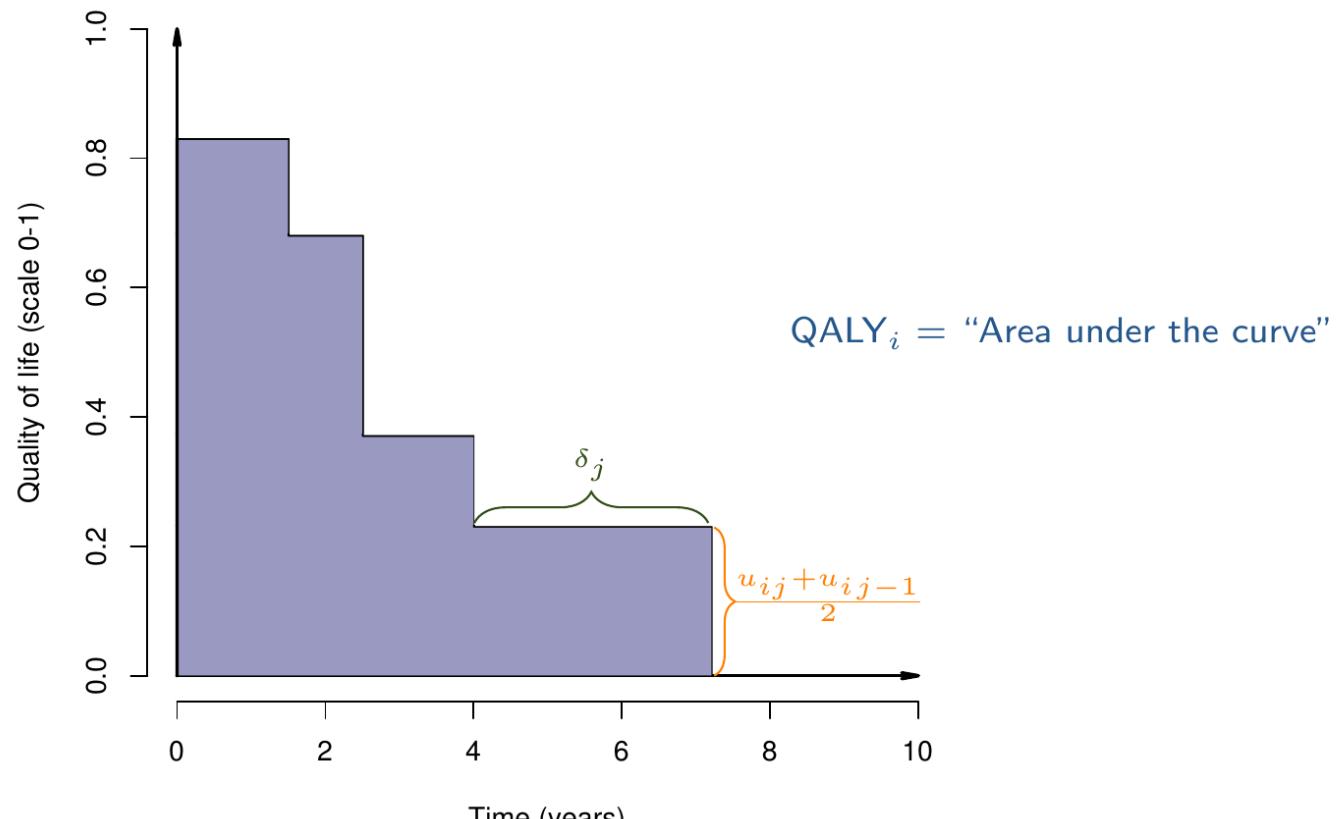
("Standard") Statistical modelling

See BMHE, pp. 23-25

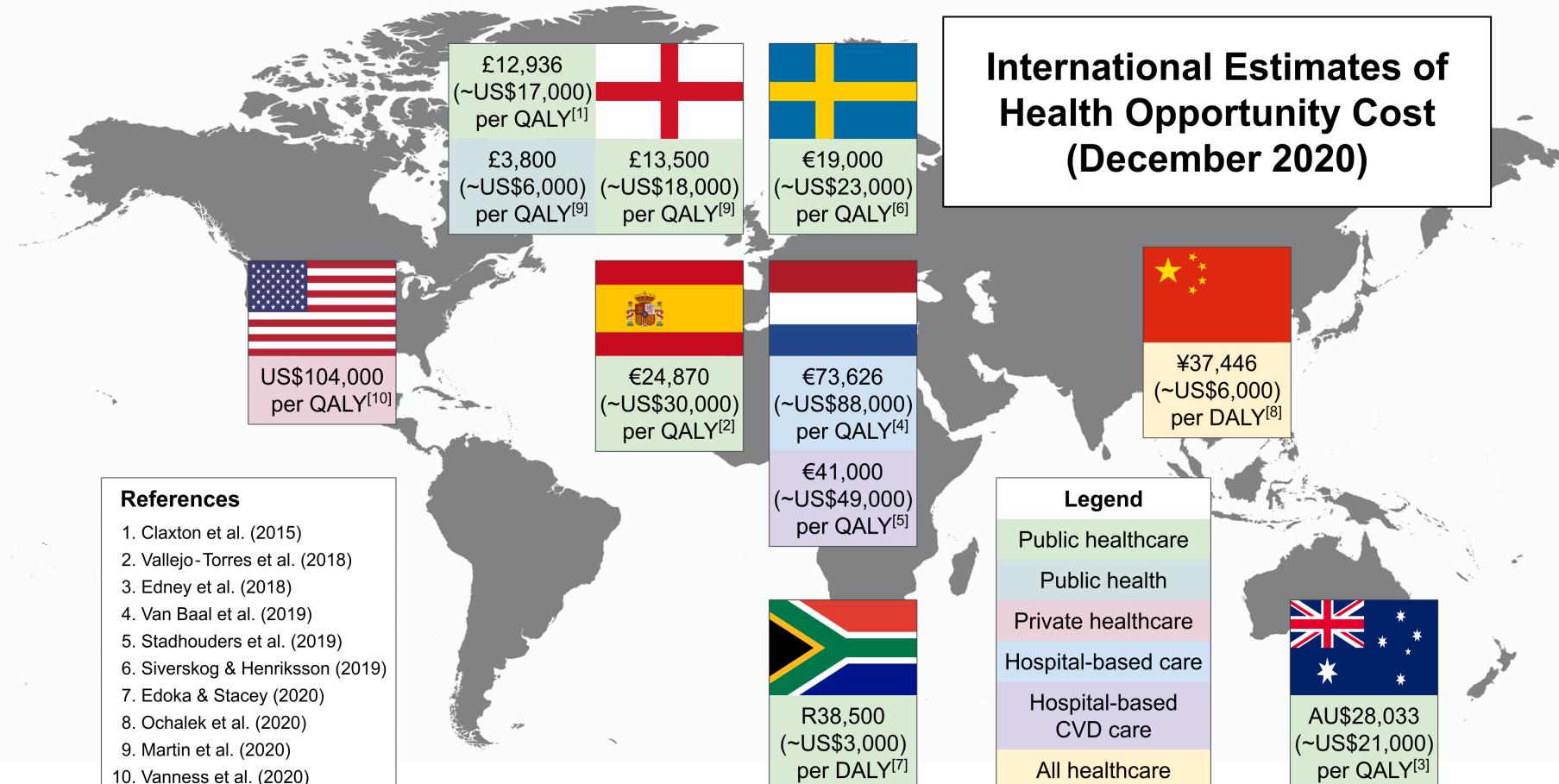


Compute individual QALYs and total costs as

$$e_i = \sum_{j=1}^J (u_{ij} + u_{ij-1}) \frac{\delta_j}{2} \quad \text{and} \quad c_i = \sum_{j=0}^J c_{ij} \quad \left[\text{with: } \delta_j = \frac{\text{Time}_j - \text{Time}_{j-1}}{\text{Unit of time}} \right]$$



A QALY is a QALY is a QALY(?)...



Dr Mike Paulden, School of Public Health, University of Alberta

paulden@ualberta.ca

@mikepaulden

mikepaulden.com

<https://twitter.com/mikepaulden/status/1333467554317156353>

("Standard") Statistical modelling

See BMHE, pp. 23-25

1

Compute individual QALYs and total costs as

$$e_i = \sum_{j=1}^J (u_{ij} + u_{ij-1}) \frac{\delta_j}{2} \quad \text{and} \quad c_i = \sum_{j=0}^J c_{ij} \quad \left[\text{with: } \delta_j = \frac{\text{Time}_j - \text{Time}_{j-1}}{\text{Unit of time}} \right]$$

2

(Often implicitly) assume normality and linearity and model **independently** individual QALYs and total costs by controlling for (**centered**) baseline values, eg $u^* = (u - \bar{u})$ and $c^* = (c - \bar{c})$

$$e_i = \alpha_{e0} + \alpha_{e1}u_{0i}^* + \alpha_{e2}\text{Trt}_i + \varepsilon_{ei} [+ \dots], \quad \varepsilon_{ei} \sim \text{Normal}(0, \sigma_e)$$
$$c_i = \alpha_{c0} + \alpha_{c1}c_{0i}^* + \alpha_{c2}\text{Trt}_i + \varepsilon_{ci} [+ \dots], \quad \varepsilon_{ci} \sim \text{Normal}(0, \sigma_c)$$

("Standard") Statistical modelling

See BMHE, pp. 23-25

- 1 Compute individual QALYs and total costs as

$$e_i = \sum_{j=1}^J (u_{ij} + u_{ij-1}) \frac{\delta_j}{2} \quad \text{and} \quad c_i = \sum_{j=0}^J c_{ij} \quad \left[\text{with: } \delta_j = \frac{\text{Time}_j - \text{Time}_{j-1}}{\text{Unit of time}} \right]$$

- 2 (Often implicitly) assume normality and linearity and model **independently** individual QALYs and total costs by controlling for (**centered**) baseline values, eg $u^* = (u - \bar{u})$ and $c^* = (c - \bar{c})$

$$e_i = \alpha_{e0} + \alpha_{e1}u_{0i}^* + \alpha_{e2}\text{Trt}_i + \varepsilon_{ei} [+ \dots], \quad \varepsilon_{ei} \sim \text{Normal}(0, \sigma_e)$$
$$c_i = \alpha_{c0} + \alpha_{c1}c_{0i}^* + \alpha_{c2}\text{Trt}_i + \varepsilon_{ci} [+ \dots], \quad \varepsilon_{ci} \sim \text{Normal}(0, \sigma_c)$$

- 3 Estimate population average cost and effectiveness differentials

- Under this model specification, these are $\Delta_e = \alpha_{e2}$ and $\Delta_c = \alpha_{c2}$

- 4 Quantify impact of uncertainty in model parameters on the decision making process

- In a fully frequentist analysis, this is done using resampling methods (eg bootstrap)

1 See discussion [here](#) for the interpretation of "centered" covariates + BMHE pp 68-73 for a discussion on the impact of centering for Bayesian modelling (autocorrelation)

What's wrong with this?...

- Potential correlation between costs & clinical benefits [Individual level + Aggregated level Data]
 - Strong positive correlation – effective treatments are innovative and result from intensive and lengthy research ⇒ are associated with higher unit costs
 - Negative correlation – more effective treatments may reduce total care pathway costs e.g. by reducing hospitalisations, side effects, etc.
 - Because of the way in which standard models are set up, bootstrapping generally only approximates the underlying level of correlation – MCMC does a better job!
- Joint/marginal normality not realistic [Mainly ILD]
 - Costs usually skewed and benefits may be bounded in $[0; 1]$
 - Can use transformation (e.g. logs) – but care is needed when back transforming to the natural scale
 - Should use more suitable models (e.g. Beta, Gamma or log-Normal) – **generally easier under a Bayesian framework**
 - Particularly relevant in presence of partially observed data – more on this later!
- Particularly as the focus is on decision-making (rather than just inference), we need to use all available evidence to fully characterise current uncertainty on the model parameters and outcomes [Mainly ALD]
 - A Bayesian approach is helpful in combining different sources of information
 - Propagating uncertainty is a fundamentally Bayesian operation!

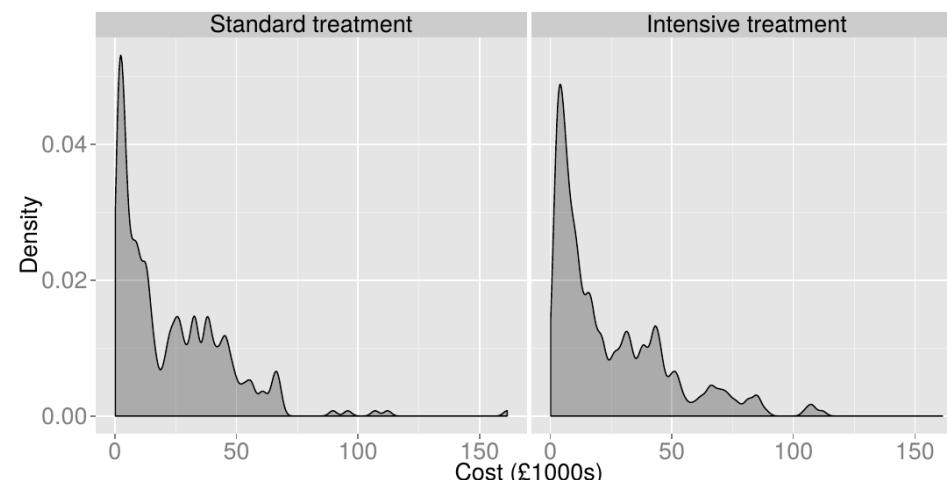
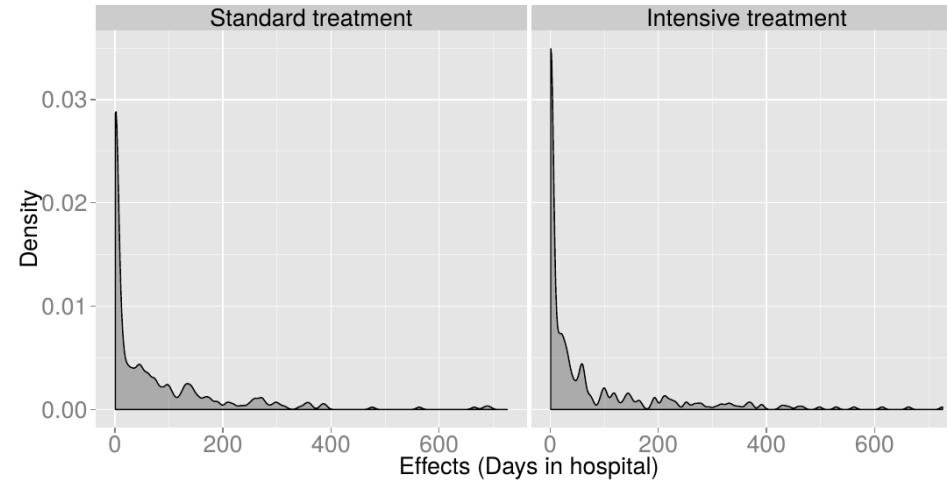
Example: UK700 data



- RCT of patients with severe psychosis
 - 281 standard case management
 - 272 intensive case management
- Effectiveness: Days in hospital for psychiatric problems
- Cost: Cost to society
- Baseline data: Days in hospital pre-intervention
- Subgroup data: Normal or borderline intelligence

Example UK700 data

Raw data plots



Dealing with skewness



1 Transform the scale of data (e.g. take logs)

- Models the re-scaled observed data
- **But:** simple re-scaling does not provide quantity of interest, if transformation is non-linear:

$$\text{median}(Y) = \exp(E[\log(Y)]) \neq E[\exp(\log(Y))] = \text{mean}[Y]$$

Dealing with skewness

1

Transform the scale of data (e.g. take logs)

- Models the re-scaled observed data
- **But:** simple re-scaling does not provide quantity of interest, if transformation is non-linear:

$$\text{median}(Y) = \exp(E[\log(Y)]) \neq E[\exp(\log(Y))] = \text{mean}[Y]$$

2

Use non-parametric methods (e.g. central limit theorem or bootstrap)

- Uses empirical distribution to make inferences about the uncertainty of the sample mean
- Theoretically, CLT holds for $n \rightarrow \infty$ and implies that the sample mean is **approximately** well described by a Normal distribution
- **But:** when data are very skewed in the first place, approximation is much rougher and unsatisfactory
- Bootstrap still relies on underlying assumptions based on some kind of Normality - may not be most efficient

Dealing with skewness

1 Transform the scale of data (e.g. take logs)

- Models the re-scaled observed data
- **But:** simple re-scaling does not provide quantity of interest, if transformation is non-linear:

$$\text{median}(Y) = \exp(E[\log(Y)]) \neq E[\exp(\log(Y))] = \text{mean}[Y]$$

2 Use non-parametric methods (e.g. central limit theorem or bootstrap)

- Uses empirical distribution to make inferences about the uncertainty of the sample mean
- Theoretically, CLT holds for $n \rightarrow \infty$ and implies that the sample mean is **approximately** well described by a Normal distribution
- **But:** when data are very skewed in the first place, approximation is much rougher and unsatisfactory
- Bootstrap still relies on underlying assumptions based on some kind of Normality - may not be most efficient

3 Parametric methods

- Builds a model for the observed data, in order to estimate **specifically** (directly or indirectly) the main parameter of interest

Fit parametric models

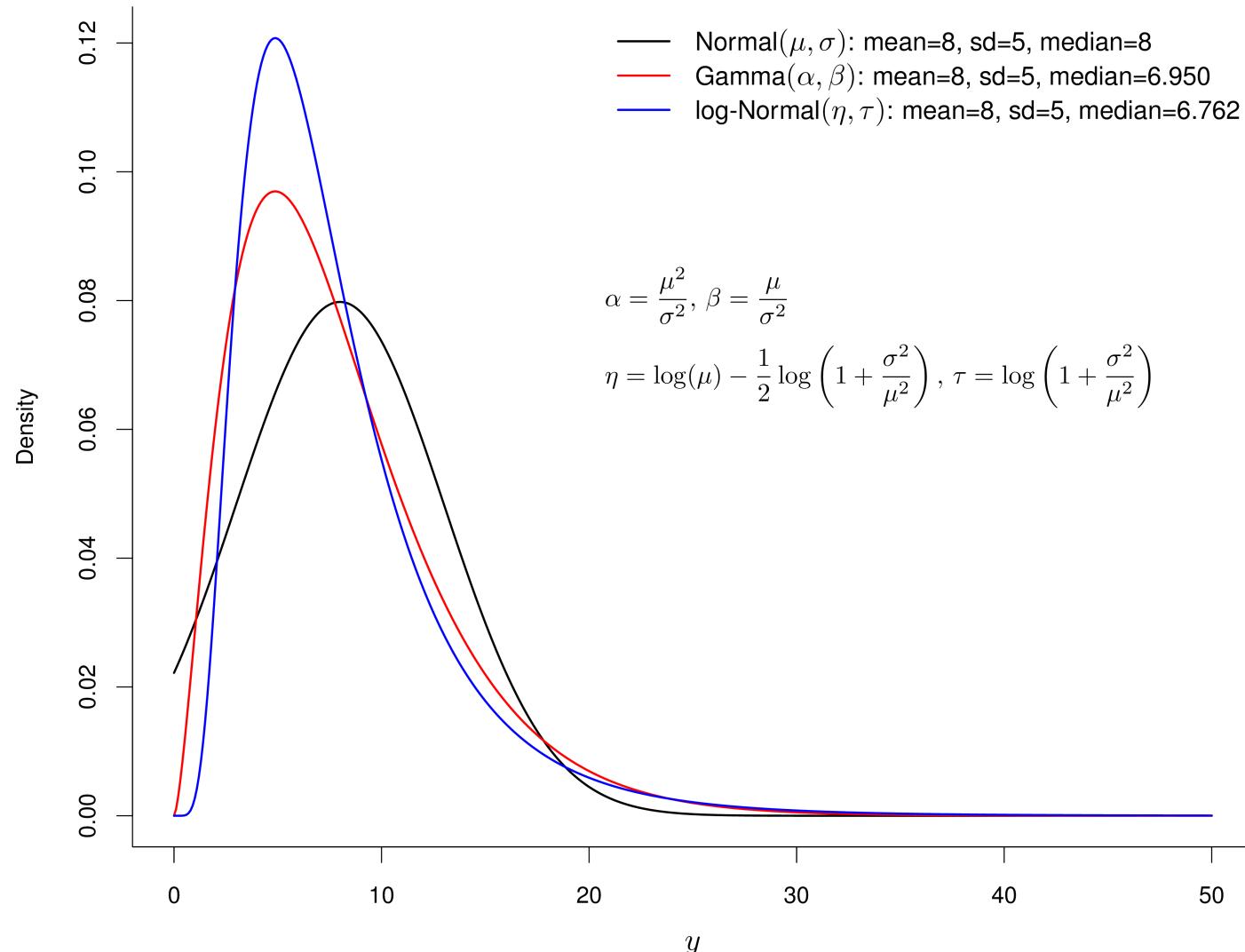
Data $y_i : i = 1, \dots, n$

Distribution	Density	Population average
Normal	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$	μ
Gamma	$\frac{1}{\Gamma(\rho)} \lambda^\rho y^{\rho-1} \exp(-\lambda y)$	$\frac{\rho}{\lambda}$
logNormal	$\frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\log y - \mu)^2}{2\sigma^2}\right)$	$\exp\left(\mu + \frac{\sigma^2}{2}\right)$

NB: In BUGS the Gamma distribution is parameterized in terms of **shape** ρ and **rate** λ where $\mu = \frac{\rho}{\lambda}$ (see [Lecture 1](#)). However, confusingly, in the OpenBUGS and WinBUGS manuals λ is called μ .

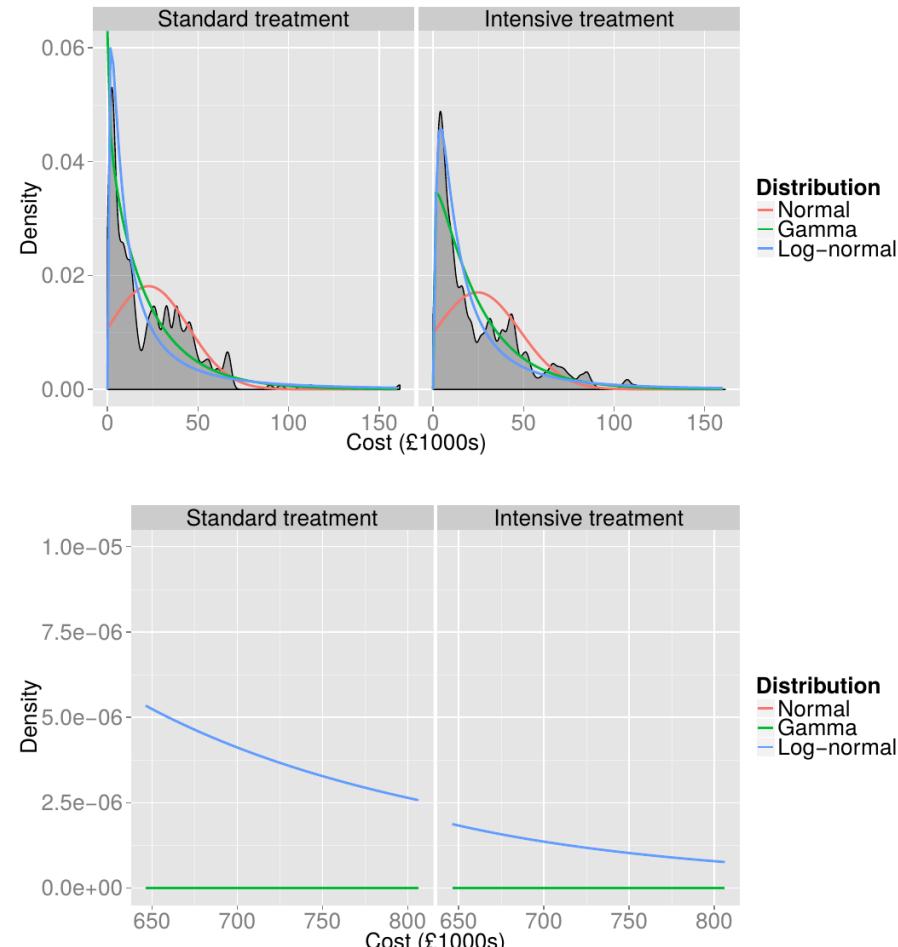
Also, other parameterisations exist, e.g. in terms of **shape** and **scale** = $1/\text{rate}$, so care is needed!

Fit parametric models



Example: UK700 data

Fit parametric model to costs



NB Tail of the log-Normal distribution is heavier than the Normal and Gamma!

Model selection: goodness of fit

See BMHE Section 3.6

- The **Deviance** is a measure of goodness of fit
 - It is defined as -2 log likelihood
- For example, if the data are associated with a Normal distribution, the likelihood function is

$$\mathcal{L}(\mu, \sigma) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - \mu)^2}{2\sigma^2}\right)$$

⚠ see <https://gianluca.statistica.it/teaching/intro-stats/estimation.html>

- So, for this model, the deviance is

$$D(\mu, \sigma) = -2 \sum_{i=1}^n \log \mathcal{L}(\mu, \sigma)$$

where $i = 1, \dots, n$ is the number of observations

- We measure goodness of fit by the posterior mean deviance $\bar{D}(\theta)$
- Lower deviance \Rightarrow better fit

(More on this in Lecture 5)

How to do this in BUGS

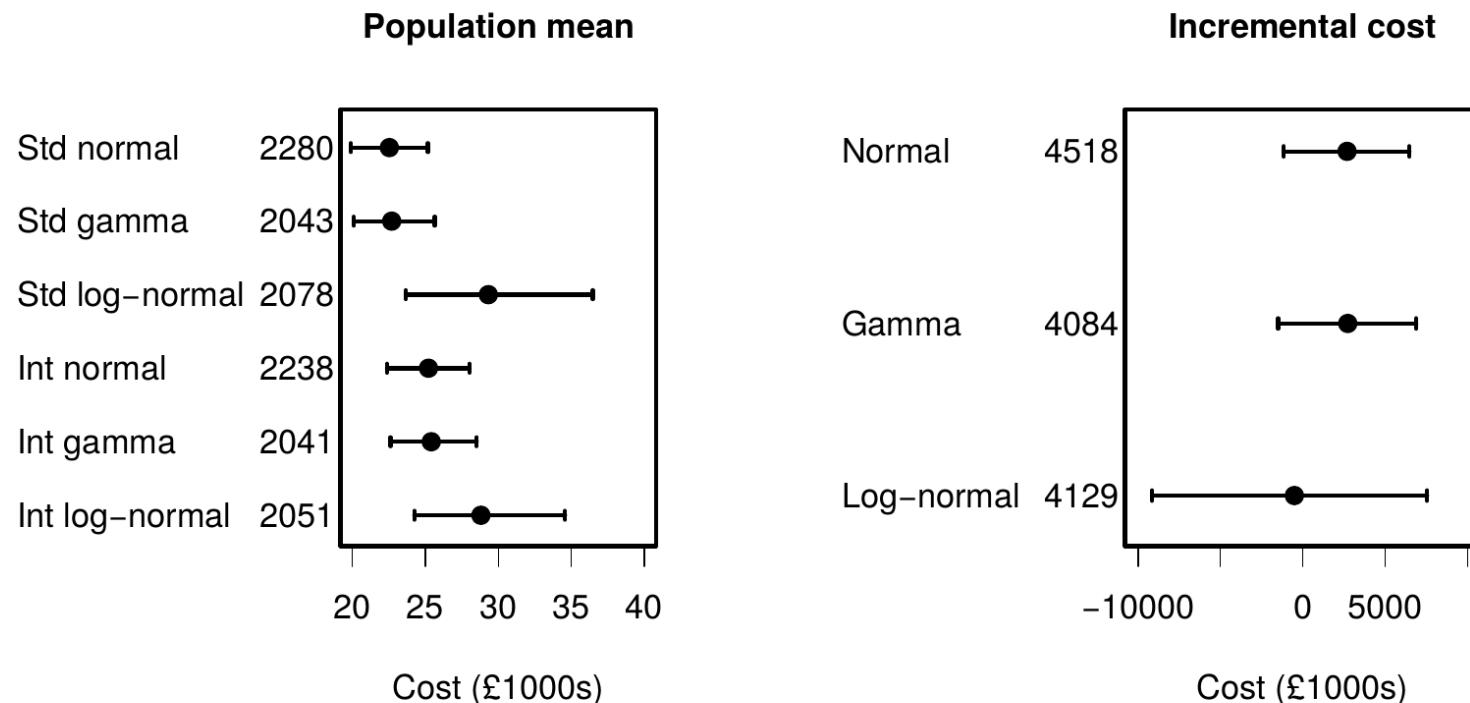
- Monitor the node deviance, which is automatically defined for each model (with observed data!)
- Press buttons
 - Update
 - Inference > DIC > set
 - Update
 - Inference > DIC > stats
- Or code the deviance explicitly in the model

```
for (i in 1:N1) {  
  logL[i] <- log(exp(-pow(cost1[i] - mu[1],2)/(2*ss[1])) /  
    (sqrt(2*3.1415927*ss[1])))  
}  
dev[1] <- -2*sum(logL[])
```

- **NB** When calling BUGS from R, assuming there are observed data, the deviance is monitored by default

Model selection: goodness of fit

Fit parametric models to costs



- Gamma model associated with smaller deviance for both intervention and control arms
- Gamma model best fitting also for the incremental costs (lower deviance)
- Normal model much worse fitting especially for incremental costs - but numerical results almost identical to Gamma!

What's wrong with this?...

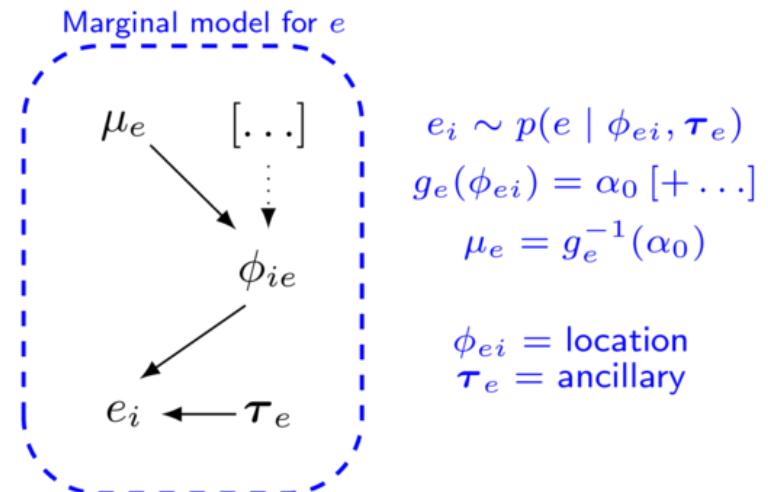
- Potential correlation between costs & clinical benefits [Individual level + Aggregated level Data]
 - Strong positive correlation - effective treatments are innovative and result from intensive and lengthy research ⇒ are associated with higher unit costs
 - Negative correlation - more effective treatments may reduce total care pathway costs e.g. by reducing hospitalisations, side effects, etc.
 - Because of the way in which standard models are set up, bootstrapping generally only approximates the underlying level of correlation - **MCMC does a better job!**
- Joint/marginal normality not realistic [Mainly ILD]
 - Costs usually skewed and benefits may be bounded in [0; 1]
 - Can use transformation (e.g. logs) - but care is needed when back transforming to the natural scale
 - Should use more suitable models (e.g. Beta, Gamma or log-Normal) - generally easier under a Bayesian framework
 - Particularly relevant in presence of partially observed data - more on this later!
- Particularly as the focus is on decision-making (rather than just inference), we need to use all available evidence to fully characterise current uncertainty on the model parameters and outcomes [Mainly ALD]
 - A Bayesian approach is helpful in combining different sources of information
 - Propagating uncertainty is a fundamentally Bayesian operation!

Bayesian HTA in action

- In general, can represent the joint distribution as $p(e, c) = p(e)p(c | e) = p(c)p(e | c)$

Bayesian HTA in action

- In general, can represent the joint distribution as $p(e, c) = p(e)p(c | e) = p(c)p(e | c)$



Bayesian HTA in action

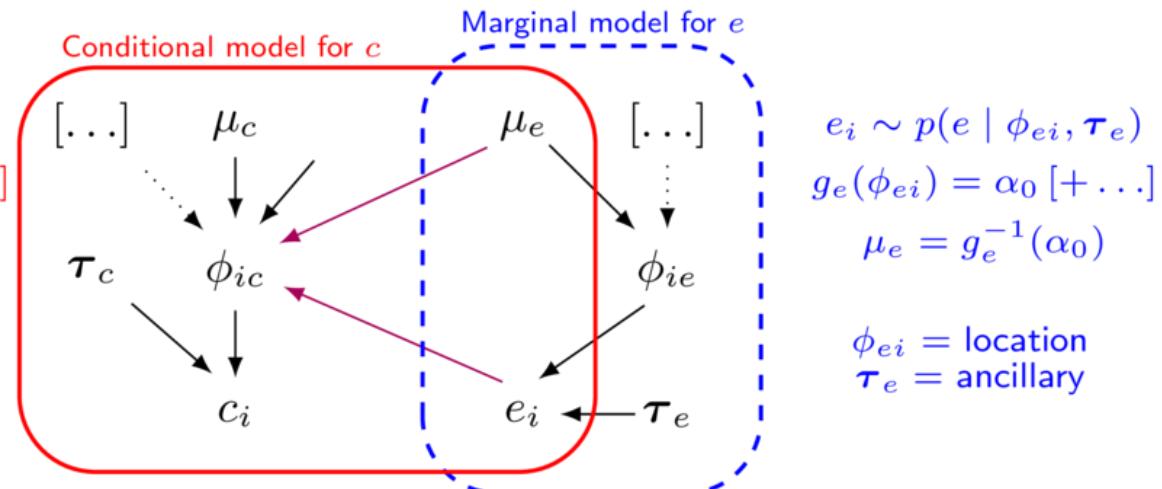
- In general, can represent the joint distribution as $p(e, c) = p(e)p(c | e) = p(c)p(e | c)$

$$c_i \sim p(c | e, \phi_{ci}, \tau_c)$$

$$g_c(\phi_{ci}) = \beta_0 + \beta_1(e_i - \mu_e) [+\dots]$$

$$\mu_c = g_c^{-1}(\beta_0)$$

ϕ_{ci} = location
 τ_c = ancillary



$$e_i \sim p(e | \phi_{ei}, \tau_e)$$

$$g_e(\phi_{ei}) = \alpha_0 [+\dots]$$

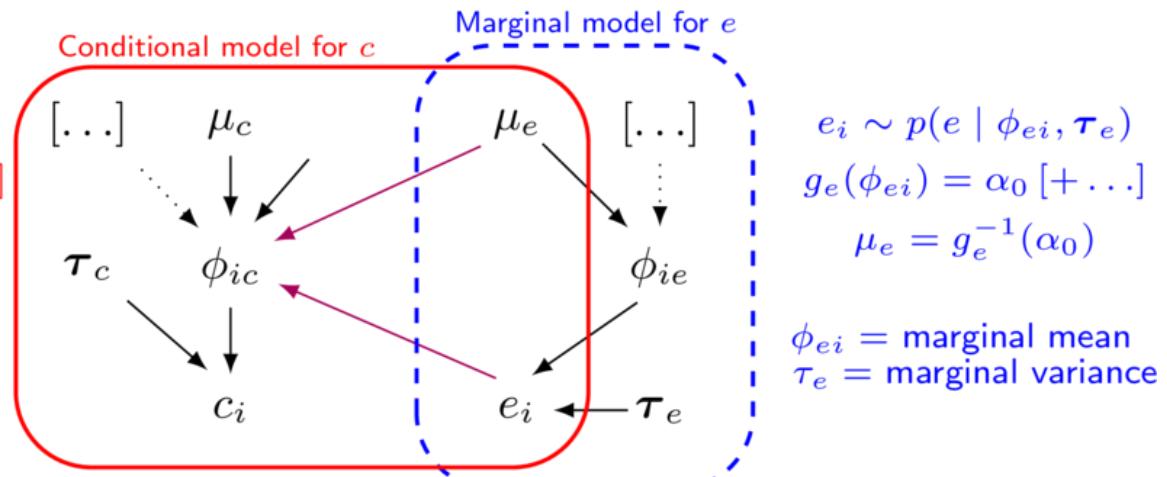
$$\mu_e = g_e^{-1}(\alpha_0)$$

ϕ_{ei} = location
 τ_e = ancillary

Bayesian HTA in action

- In general, can represent the joint distribution as $p(e, c) = p(e)p(c | e) = p(c)p(e | c)$

$$\begin{aligned}c_i &\sim p(c | e, \phi_{ci}, \tau_c) \\g_c(\phi_{ci}) &= \beta_0 + \beta_1(e_i - \mu_e) [+ \dots] \\\mu_c &= g_c^{-1}(\beta_0) \\\phi_{ci} &= \text{conditional mean} \\\tau_c &= \text{conditional variance}\end{aligned}$$



- For example:

$$\begin{array}{lll}e_i \sim \text{Normal}(\phi_{ei}, \tau_e) & \phi_{ei} = \alpha_0 [+ \dots] & \mu_e = \alpha_0 \\c_i | e_i \sim \text{Normal}(\phi_{ci}, \tau_c) & \phi_{ci} = \beta_0 + \beta_1(e_i - \mu_e) [+ \dots] & \mu_c = \beta_0\end{array}$$

Bayesian HTA in action

- In general, can represent the joint distribution as $p(e, c) = p(e)p(c | e) = p(c)p(e | c)$

$$c_i \sim p(c | e, \phi_{ci}, \tau_c)$$

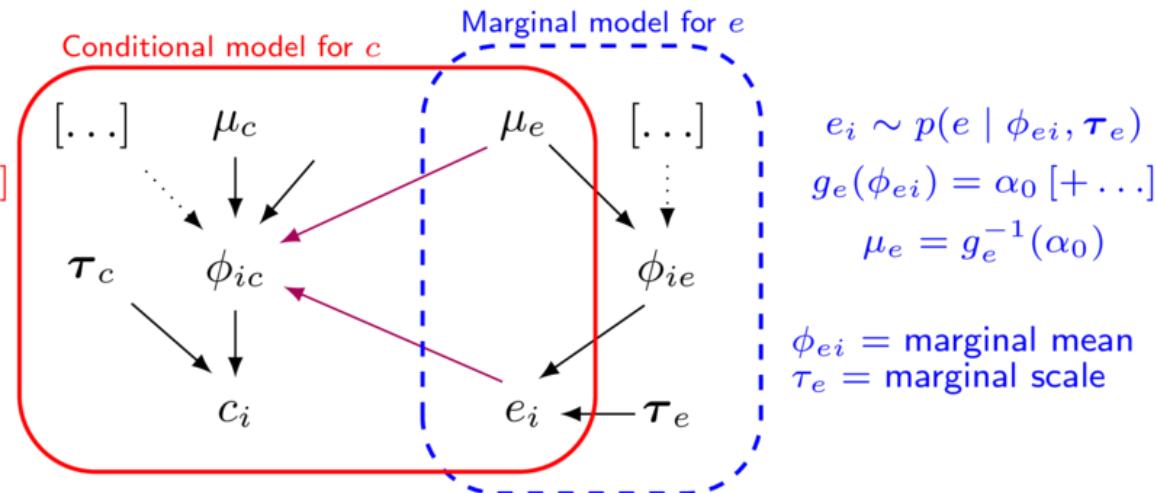
$$g_c(\phi_{ci}) = \beta_0 + \beta_1(e_i - \mu_e) [+ \dots]$$

$$\mu_c = g_c^{-1}(\beta_0)$$

ϕ_{ci} = conditional mean

τ_c = shape

τ_c/ϕ_{ci} = rate



$$e_i \sim p(e | \phi_{ei}, \tau_e)$$

$$g_e(\phi_{ei}) = \alpha_0 [+ \dots]$$

$$\mu_e = g_e^{-1}(\alpha_0)$$

ϕ_{ei} = marginal mean
 τ_e = marginal scale

- For example:

$$e_i \sim \text{Beta}(\phi_{ei}\tau_e, (1 - \phi_{ei})\tau_e)$$

$$\text{logit}(\phi_{ei}) = \alpha_0 [+ \dots]$$

$$\mu_e = \frac{\exp(\alpha_0)}{1 + \exp(\alpha_0)}$$

$$c_i | e_i \sim \text{Gamma}(\tau_c, \tau_c/\phi_{ci})$$

$$\log(\phi_{ci}) = \beta_0 + \beta_1(e_i - \mu_e) [+ \dots]$$

$$\mu_c = \exp(\beta_0)$$

Bayesian HTA in action

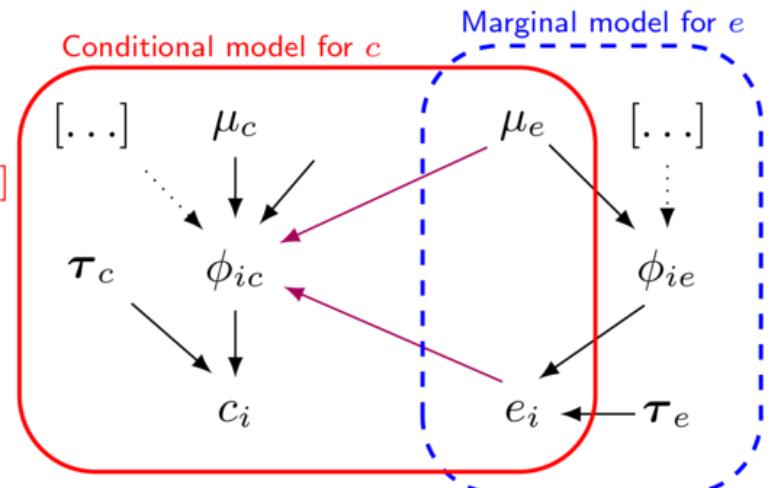
- In general, can represent the joint distribution as $p(e, c) = p(e)p(c | e) = p(c)p(e | c)$

$$c_i \sim p(c | e, \phi_{ci}, \tau_c)$$

$$g_c(\phi_{ci}) = \beta_0 + \beta_1(e_i - \mu_e) [+ \dots]$$

$$\mu_c = g_c^{-1}(\beta_0)$$

ϕ_{ci} = conditional mean
 τ_c = shape
 τ_c/ϕ_{ci} = rate



$$e_i \sim p(e | \phi_{ei}, \tau_e)$$

$$g_e(\phi_{ei}) = \alpha_0 [+ \dots]$$

$$\mu_e = g_e^{-1}(\alpha_0)$$

ϕ_{ei} = marginal mean
 τ_e = marginal scale

- For example:

$$e_i \sim \text{Beta}(\phi_{ei}\tau_e, (1 - \phi_{ei})\tau_e)$$

$$\text{logit}(\phi_{ei}) = \alpha_0 [+ \dots]$$

$$\mu_e = \frac{\exp(\alpha_0)}{1 + \exp(\alpha_0)}$$

$$c_i | e_i \sim \text{Gamma}(\tau_c, \tau_c/\phi_{ci})$$

$$\log(\phi_{ci}) = \beta_0 + \beta_1(e_i - \mu_e) [+ \dots]$$

$$\mu_c = \exp(\beta_0)$$

- Combining "modules" and fully characterising uncertainty about deterministic functions of random quantities is relatively straightforward using MCMC
- Prior information can help stabilise inference (especially with sparse data!), eg
 - Cancer patients are unlikely to survive as long as the general population
 - ORs are unlikely to be greater than ± 5

Model 1:

$$p(e) \sim \text{Normal}, p(c | e) \sim \text{Normal}$$

```
model {  
    for (i in 1:N1) {  
        eff1[i] ~ dnorm(phi.e1[i],tau.e[1]); phi.e1[i] <- mu.e[1]  
        cost1[i] ~ dnorm(phi.c1[i],tau.c[1]); phi.c1[i] <- mu.c[1]+beta[1]*(eff1[i]-mu.e[1])  
    }  
    for (i in 1:N2) {  
        eff2[i] ~ dnorm(phi.e2[i],tau.e[2]); phi.e2[i] <- mu.e[2]  
        cost2[i] ~ dnorm(phi.c2[i],tau.c[2]); phi.c2[i] <- mu.c[2]+beta[2]*(eff2[i]-mu.e[2])  
    }  
    # Node transformations for variances  
    for (t in 1:2) {  
        tau.e[t] <- 1/ss.e[t]; ss.e[t] <- s.e[t]*s.e[t]; s.e[t] <- exp(ls.e[t])  
        tau.c[t] <- 1/ss.c[t]; ss.c[t] <- s.c[t]*s.c[t]; s.c[t] <- exp(ls.c[t])  
        marg.ss.c[t] <- ss.c[t] + ss.e[t]*beta[t]*beta[t] # Marginal cost variance  
    }  
    # Prior distributions  
    ls.c[t] ~ dunif(-5, 10)  
    mu.c[t] ~ dnorm(0, 1.0E-6)  
    beta[t] ~ dunif(-5, 5)  
    ls.e[t] ~ dunif(-5, 10)  
    mu.e[t] ~ dnorm(0, 1.0E-6)  
}  
# HE outcomes  
delta.c <- mu.c[2]-mu.c[1]  
delta.e <- -(mu.e[2] - mu.e[1]) # NB: negative as days in hospital is a bad thing  
}
```

Model 2:

$p(e) \sim \text{Gamma}$, $p(c | e) \sim \text{Gamma}$ – (specification 1)

```
model {  
  for (i in 1:N1) {  
    eff1[i] ~ dgamma(shape.e[1],rate.e1[i]); log(phi.e1[i]) <- lm.e[1]  
    cost1[i] ~ dgamma(shape.c[1],rate.c1[i]); log(phi.c1[i]) <- lm.c[1]+beta[1]*(eff1[i]-mu.e[1])  
    rate.e1[i] <- shape.e[1]/phi.e1[i]  
    rate.c1[i] <- shape.c[1]/phi.c1[i]  
  }  
  for (i in 1:N2) {  
    eff2[i] ~ dgamma(shape.e[2],rate.e2[i]); log(phi.e2[i]) <- lm.e[2]  
    cost2[i] ~ dgamma(shape.c[2],rate.c2[i]); log(phi.c2[i]) <- lm.c[2]+beta[2]*(eff2[i]-mu.e[2])  
    rate.e2[i] <- shape.e[2]/phi.e2[i]  
    rate.c2[i] <- shape.c[2]/phi.c2[i]  
  }  
  # Prior distributions  
  for (t in 1:2) {  
    shape.e[t] ~ dunif(0,10)  
    shape.c[t] ~ dunif(0,10)  
    lm.e[t] ~ dnorm(0, 1.0E-6); mu.e[t] <- exp(lm.e[t])      # Prior on 'log(mu) and then rescale'  
    lm.c[t] ~ dnorm(0, 1.0E-6); mu.c[t] <- exp(lm.c[t])  
    beta[t] ~ dunif(-5, 5)  
  }  
  # HE outcomes  
  delta.c <- mu.c[2]-mu.c[1]  
  delta.e <- -(mu.e[2] - mu.e[1])   # NB: negative as days in hospital is a bad thing  
}
```

Model 2:

$p(e) \sim \text{Gamma}$, $p(c | e) \sim \text{Gamma}$ – (specification 2)

```
model {  
    for (i in 1:N1) {  
        eff1[i] ~ dgamma(shape.e[1],rate.e1[i]); phi.e1[i] <- mu.e[1]  
        cost1[i] ~ dgamma(shape.c[1],rate.c1[i]); phi.c1[i] <- mu.c[1]+beta[1]*(eff1[i]-mu.e[1])  
        rate.e1[i] <- shape.e[1]/phi.e1[i]  
        rate.c1[i] <- shape.c[1]/phi.c1[i]  
    }  
    for (i in 1:N2) {  
        eff2[i] ~ dgamma(shape.e[2],rate.e2[i]); phi.e2[i] <- mu.e[2]  
        cost2[i] ~ dgamma(shape.c[2],rate.c2[i]); phi.c2[i] <- mu.c[2]+beta[2]*(eff2[i]-mu.e[2])  
        rate.e2[i] <- shape.e[2]/phi.e2[i]  
        rate.c2[i] <- shape.c[2]/phi.c2[i]  
    }  
    # Prior distributions  
    for (t in 1:2) {  
        shape.e[t] ~ dunif(0,10)  
        shape.c[t] ~ dunif(0,10)  
        mu.e[t] ~ dunif(0, K)    # Prior 'directly on mu' (NB: needs to be positive!)  
        mu.c[t] ~ dunif(0, K)    # (for some large constant threshold K)  
        beta[t] ~ dunif(-5, 5)  
    }  
    # HE outcomes  
    delta.c <- mu.c[2]-mu.c[1]  
    delta.e <- -(mu.e[2] - mu.e[1])    # NB: negative as days in hospital is a bad thing  
}
```

Model 3:

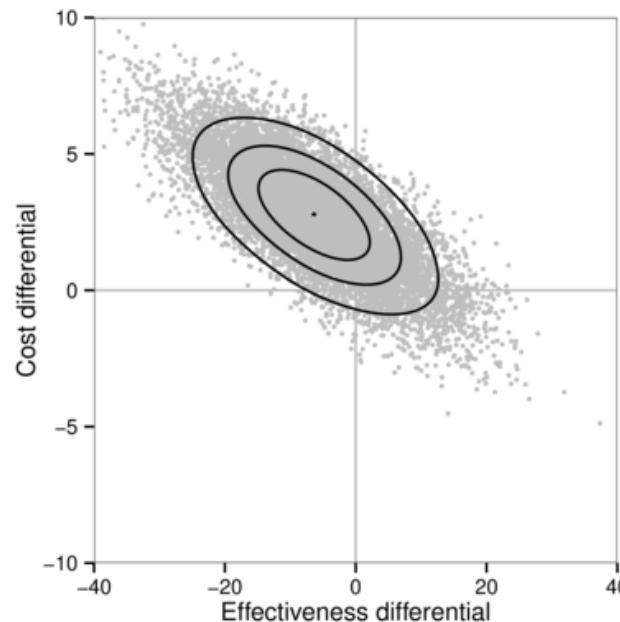
$p(c) \sim \text{logNormal}, p(e | c) \sim \text{Normal}$

$$c_i \sim \text{logNormal}(\phi_{ci}, \tau_c), \quad \exp\left(\phi_{ci} + \frac{\sigma_c^2}{2}\right) = \mu_c \Rightarrow \phi_{ci} = \log(\mu_c) - \frac{1}{2\tau_c}$$
$$e_i | c_i \sim \text{Normal}(\phi_{ei}, \tau_e), \quad \phi_{ei} = \mu_e + \beta(c_i - \mu_c)$$

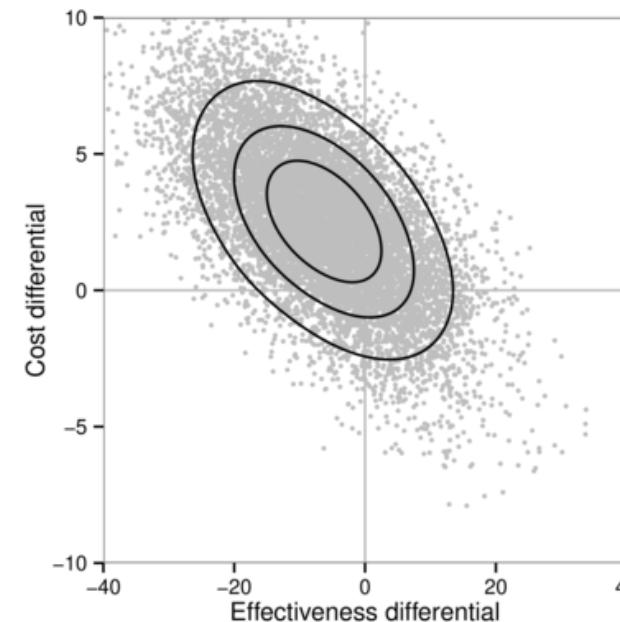
```
model {  
  for(i in 1:N1) {  
    eff1[i] ~ dnorm(phi.e1[i], tau.e[1]); phi.e1[i] <- mu.e[1]+beta[1]*(cost1[i]-mu.c[1])  
    cost1[i] ~ dlnorm(phi.c1[i], tau.c[1]); phi.c1[i] <- log(mu.c[1])-0.5*ss.c[1]  
  }  
  for(i in 1:N2) {  
    eff2[i] ~ dnorm(phi.e2[i], tau.e[2]); phi.e2[i] <- mu.e[1]+beta[1]*(cost2[i]-mu.c[2])  
    cost2[i] ~ dlnorm(phi.c2[i], tau.c[2]); phi.c2[i] <- log(mu.c[1])-0.5*ss.c[2]  
  }  
  # Node transformations for variances  
  for(t in 1:2) {  
    tau.e[t] <- 1/ss.e[t]; ss.e[t] <- s.e[t]*s.e[t]; s.e[t] <- exp(ls.e[t])  
    tau.c[t] <- 1/ss.c[t]; ss.c[t] <- s.c[t]*s.c[t]; s.c[t] <- exp(ls.c[t])  
  }  
  # Prior distributions  
  mu.e[t] ~ dnorm(0, 1.0E-6); mu.c[t] ~ dnorm(0, 1.0E-6); beta[t] ~ dunif(-5, 5)  
  ls.e[t] ~ dunif(-5, 10); ls.c[t] ~ dunif(-5, 10)  
}  
...  
}
```

Results

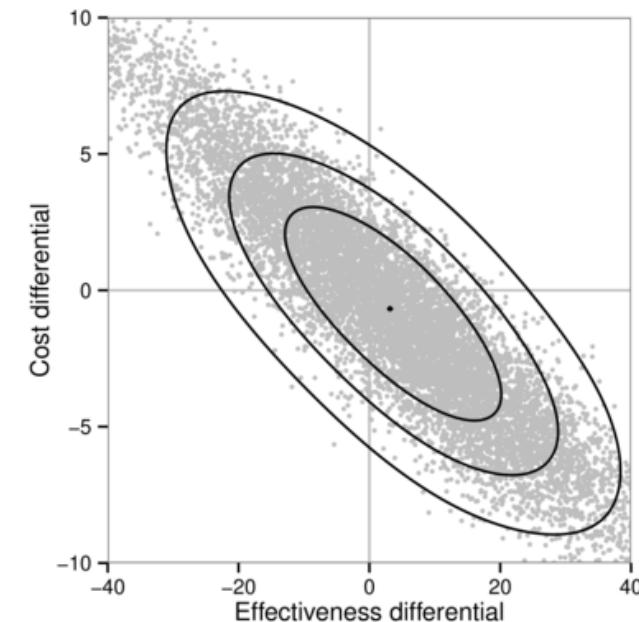
Costs and effects normal



Costs and effects gamma



Costs log-normal and effects normal



 [Next lecture](#)