Enhancement of an IDEAL framework for the assessment of surgical interventions: understanding the complexities and constructing solutions

Olympia Papachristofi[1]
with Prof Linda Sharples[2]
[1]MRC Biostatistics Unit, Cambridge, UK
[2]University of Leeds, UK
olympia.papachristofi@mrc-bsu.cam.ac.uk

May 2014, SCT Philadelphia

Outline

Background

* **Complex Interventions**

Interventions composed by several interacting components.

Background

* **Complex Interventions**

Interventions composed by several interacting components.

IDEAL framework

↪ Five-stage paradigm delineating the development of innovative surgical procedures.

↪ Acknowledges that the defining features of complex interventions naturally characterise surgical interventions.

Background

↪ Correspondence with the Phases of pharmaceutical development.

↪ Highlights study designs and reporting standards most useful at each stage of this more complex setting.

Figure: IDEAL framework

Aims

- Consider a view of surgical interventions within the IDEAL evaluative framework in order to:

Aims

- Consider a view of surgical interventions within the IDEAL evaluative framework in order to:

  - ascertain methodological issues arising due to these complexities.

  - implement methods to resolve these within the 5-stage study design concept.

Multidisciplinary team and interactions

∗ Minimal attention during trial design on potential co-intervention effects.

Surgical interventions

↪ Examine the extent to which patients treated by the *same medical team* are more likely to have the same outcome.

↪ Establish the individual surgeon and anaesthetist effects on outcome.

↪ Investigate potential anaesthetist-surgeon collaborations.

Multidisciplinary team and interactions

* Minimal attention during trial design on potential co-intervention effects.

Surgical interventions

↪ Examine the extent to which patients treated by the *same medical team* are more likely to have the same outcome.

↪ Establish the individual surgeon and anaesthetist effects on outcome.

↪ Investigate potential anaesthetist-surgeon collaborations.

* Obtain a measure of the degree of correlation amongst observations within a cluster.
↪ *Intra-Class Correlation coefficient* (ICC).

Random effects models

- Models:

1. Surgeon random effects - two-level random intercept model.
2. Anaesthetist random effects - two-level random intercept model.

Random effects models

- Models:

1. Surgeon random effects - two-level random intercept model.
2. Anaesthetist random effects - two-level random intercept model.

↪ Cannot be extended to 3-level hierarchical models as data structure not fully nested.

3. *Cross-classified model* accounting for lower level units belonging to more than one clusters which are not nested.

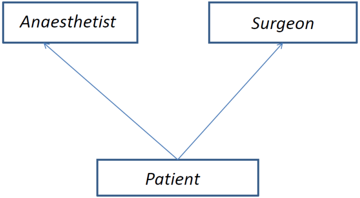
Anaesthetist performance

- Anaesthetists: "members" of more than one surgeon clusters.

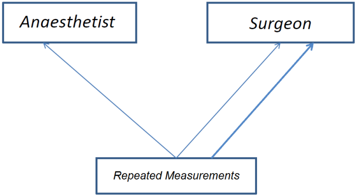
- Interest: variation in the anaesthetists' performance attributable to surgeons and its dependence on the partition of the anaesthetists' caseload across surgeons.
$\hookrightarrow$ *Variance Partition coefficient* (VPC).

- *ICC*: correlation between two anaesthetists depending on their surgeon collaboration profiles.

Classification diagrams



(a) Cross-classified model



(b) Multiple Membership Multiple classification model

Application

- Large case series study (>18000) on heterogeneous cardiac surgery patients.
- 18426 operations conducted by 18 surgeons and 24 anaesthetists.
- Binary response variable: in-hospital death.
- First level covariate: the logistic-EuroSCORE of each patient.

Application

General two-level hierarchical model

$$logit(\pi_{ijk}) = \beta_{0ij} + \sum_{w} \beta_w (x_{ijkw} - \overline{x_w}) \qquad (1)$$

where $y_{ijk} | \pi_{ijk} \sim Binomial(1, \pi_{ijk})$.

π_{ijk} the probability of an in hospital death for the k^{th} patient treated by the i^{th} surgeon and j^{th} anaesthetist.

$\beta_{0ij} = \alpha + u_i + v_j$, $u_i \sim N(0, \sigma_u^2)$ the i^{th} surgeon random intercept
$v_j \sim N(0, \sigma_v^2)$ respectively for the j^{th} anaesthetist.

$e_{ijk} \sim Logistic(0, 1) \Rightarrow \sigma_e^2 = \pi^2/3$

x_{ijkw} the w^{th} covariate for the k^{th} patient treated by the i^{th} surgeon and j^{th} anaesthetist.

Anaesthetist Multiple Membership Multiple classification model

$$logit(\pi_i) = \beta_{0i} + \sum_w \beta_w (x_{iw} - \overline{x_w}) \qquad (2)$$

where $y_i | \pi_i \sim Binomial(1, \pi_i)$.

$\beta_{0i} = \alpha + v_{A_{(i)}} + \sum_{j \in Sur_{(i)}} w_{i,j} u_j$.

π_i the probability of an in hospital death for the i^{th} patient treated by surgeon $Sur_{(i)} \subset (1, .., n_i)$ and anaesthetist $A_{(i)} \subset (1, .., I)$

Random intercepts' distributions: $u_j \sim N(0, \sigma_u^2)$ for the surgeon
$v_{A_{(i)}} \sim N(0, \sigma_v^2)$ for the anaesthetist.

$e_i \sim Logistic(0, 1) \Rightarrow \sigma_e^2 = \pi^2/3$

x_{iw} the w^{th} covariate for the i^{th} patient treated by the surgeon $Sur_{(i)}$ and anaesthetist $A_{(i)}$.

Choice of weights and priors

Weights

$$w_{i,j} = \frac{operations\ done\ by\ A_{(i)}\ with\ Sur_{(i),j}}{total\ operations\ done\ by\ A_{(i)}}$$

$\hookrightarrow$ Fitted using MCMC methods in package MCMCglmm in R.

Binary response model

- Residual variance fully specified by mean $\rightarrow$ non-identifiable.
- Fix it to some positive value $\rightarrow$ 1 recommended.

Choice of Priors

- Random effects variance

$*$ $\Gamma^{-1}(\epsilon, \epsilon)$, $\epsilon \to 0$

$**$ Improper uniform on σ_u^2.

Drawbacks:

$*$ bad behaviour for variances close to 0.

$**$ miscalibration towards higher σ_u values and need more than 4 clusters to get a proper posterior.

Choice of Priors

- Random effects variance

$*$ $\Gamma^{-1}(\epsilon, \epsilon)$, $\epsilon \to 0$

$**$ Improper uniform on σ_u^2.

Drawbacks:

$*$ bad behaviour for variances close to 0.

$**$ miscalibration towards higher σ_u values and need more than 4 clusters to get a proper posterior.

$\hookrightarrow$ **Suggest**: Improper Uniform on σ_u.

$\hookrightarrow$ Parameter-expanded priors and slice sampling to ensure effective mixing.

Results

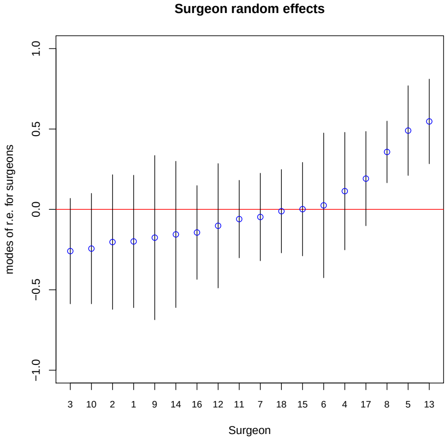


Figure: Surgeon random effects

Fixed effects	Estimate	Standard Error	$p - value$
Intercept $\hat{\alpha}$	-4.06	0.10	0.00
logit(EuroSCORE) $\hat{\beta}$	0.90	0.03	0.00
Random effects	Groups	$\hat{\sigma}_u^2$	$\hat{\sigma}_u$
	Surgeon	0.093981	0.30656
Diagnostics	*AIC*	4319	
	BIC	4343	

Table: Fixed and random effects estimates for surgeon random intercept model

Results

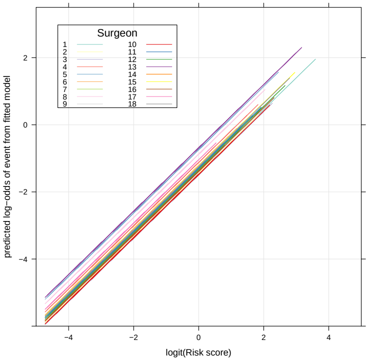
Fixed effects	Estimate	Standard Error	$p - value$
Intercept $\hat{\alpha}$	-4.06	0.10	0.00
logit(EuroSCORE) $\hat{\beta}$	0.90	0.03	0.00
Random effects	Groups	$\hat{\sigma}_u^2$	$\hat{\sigma}_u$
	Surgeon	0.093981	0.30656
Diagnostics	AIC	4319	
	BIC	4343	

Table: Fixed and random effects estimates for surgeon random intercept model

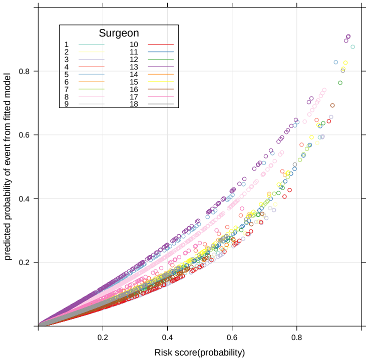
Main interest

$$VPC \equiv ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} \Longrightarrow \mathbf{ICC_{surgeon} = 0.0278}\,(3s.f.)$$

Results

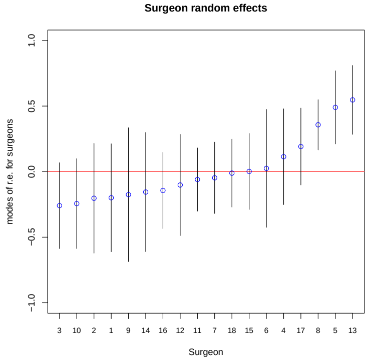


(a) Predicted log-odds of event per surgeon for different risk scores

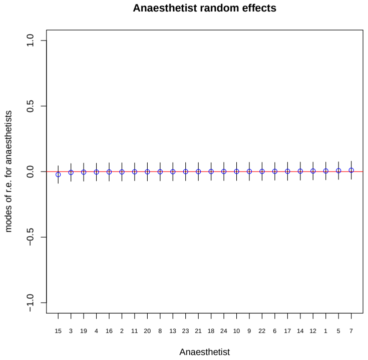


(b) Predicted probability of event per surgeon for different risk scores

Results



(c) Surgeon random effects



(d) Anaesthetist random effects

Results - MMMC Model

↪ 550000 iterations with burnin=100000.

↪ Thinning at 100 observations.

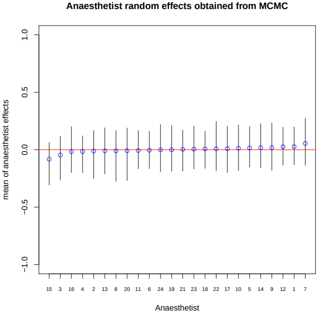
Fixed effects	Estimate	Credible interval
Intercept $\hat{\alpha}$	-4.48	(-4.68,-4.3)
logit(EuroSCORE) $\hat{\beta}$	0.99	(0.92, 1.06)
Random effects	$\hat{\sigma}_u^2$	Credible interval
Surgeon	20.1	(0.913,46.7)
Anaesthetist	0.0134	(0,0.0472)
Diagnostics	*DIC*	4241

↪ Convergence assessed by Gelman plots.

↪ Using both expanded uniform priors on σ_u and σ_u^2 did not affect the estimates.

Results

- Number of higher level units close to that of lower level ⇒ difficulty in obtaining robust results.



- Correlation between anaesthetists similar irrespective of their collaborations' profile e.g. $ICC(A_7, A_1) = 0.36$ and $ICC(A_7, A_{15}) = 0.33$

Choice of Exploration Stage design

Design Effect

$$\textbf{Deff} = \begin{cases} < 1 & \Longrightarrow \text{gain in power} \\ > 1 & \Longrightarrow \text{loss in power} \end{cases}$$

$\hookrightarrow$ The greater the degree of clustering, the more carefully we need to consider our choice of trial design.

Choice of Exploration Stage design

Design Effect

$$\textbf{Deff} = \begin{cases} < 1 & \implies \text{gain in power} \\ > 1 & \implies \text{loss in power} \end{cases}$$

$\hookrightarrow$ The greater the degree of clustering, the more carefully we need to consider our choice of trial design.

1) Stratified individually randomised trial

$Deff = 1 - ICC_{surgeon} = 0.972$ (3 $s.f.$).

Choice of Exploration Stage designs

↪ An individually randomised trial would be infeasible if:

- all participating surgeons are not equally expert in the procedures under study.
- surgeon equipoise does not hold.

Choice of Exploration Stage designs

↪ An individually randomised trial would be infeasible if:

- all participating surgeons are not equally expert in the procedures under study.
- surgeon equipoise does not hold.

Choice of Exploration Stage designs

$\hookrightarrow$ An individually randomised trial would be infeasible if:

- all participating surgeons are not equally expert in the procedures under study.
- surgeon equipoise does not hold.

2) Expertise-based trial

$Deff = 1 + (m-1)ICC_{surgeon}$ where m=average cluster size.

$\hookrightarrow$ In a trial setting with surgeons operating, on average, 10 patients:

$Deff = 1.25\ (3\ s.f.) \Longrightarrow$ Sample size must be increased by 25%.

Conclusions

- In-hospital mortality after cardiac surgery is mainly attributed to the patient risk-profile.

Conclusions

- In-hospital mortality after cardiac surgery is mainly attributed to the patient risk-profile.
- Surgeon expertise has a small, yet significant influence.

Conclusions

- In-hospital mortality after cardiac surgery is mainly attributed to the patient risk-profile.
- Surgeon expertise has a small, yet significant influence.
- In this setting, anaesthetist effect was negligible.

Conclusions

- In-hospital mortality after cardiac surgery is mainly attributed to the patient risk-profile.
- Surgeon expertise has a small, yet significant influence.
- In this setting, anaesthetist effect was negligible.
- No systematic pairings to induce stronger correlations $\Rightarrow$ almost half of the anaesthetists worked with all surgeons.

IDEAL framework for surgical interventions

↪ Developed to identify and describe the stages of surgical innovation.

↪ Served as a basis for developing and recommending the discussed methodology.

Proposed Methodology

- can be used to detect under-performance thus offering the potential to correct it.
- informs the choice of stage designs and consequently sample size calculations.
- enables the investigation of non-hierarchical data structures.

References I

1. Papachristofi O, Mackay J, Powel S, Nashef S and Sharples L: Impact of anesthetist and surgeon on cardiac surgical outcomes. JCVA 28(1): 103-109, 2014.
2. Barkun JS, Aronson JK, Feldman LS, et al: Evaluation and stages of surgical innovations. Lancet 374: 1089-1096, 2009.
3. Ergina PL, Cook JA, Blazeby JM, et al: Challenges in evaluating surgical innovation. Lancet 374: 1097-1104, 2009.
4. McCulloch P, Altman DG, Campbell BW, et al: No surgical innovation without evaluation: the IDEAL recommendations. Lancet 374: 1105-1112, 2009.
5. Rodriguez G, Elo I: Intra-class correlation in random-effects models for binary data. STATA Journal 3: 32-46, 2003.
6. Nashef SAM, Roques F, Michel P, et al: European system for cardiac operative risk evaluation(EuroSCORE). Eur J Cardiothorac Surg 16: 9-13, 1999.
7. Lee KJ, Thompson SG: The use of random effects models to allow for clustering in individually randomized trials. Clin Trials 2(2): 163-173, 2005.

References II

8 Browne WJ, Goldstein H and Rasbach J: Multiple membership multiple classification (MMMC) models . Stat Modelling 1: 103-124, 2001.

9 Browne WJ: Applying MCMC Methods to Multi-level Models. PhD thesis, 1998.

10 Hadfield J: MCMC Generalised Linear Mixed Models. CRAN repository.2012.

11 Gelman A: Prior ditributions for variance parameters in hierarchical models. Bayesian Analysis 1(3): 515-533, 2006.

12 Cook JA, Bruckner T, MacLennan GS, et al: Clustering in surgical trials database of intracluster correlations. Trials 13, 2012.

13 Vierron E, Giraudeau B: Design effect in multicenter studies: gain or loss of power? BMC Med Res Methodol 9, 2009.

Complex Interventions

Complex Interventions

↪Interventions composed by several interacting components.

Complex Interventions

↪Interventions composed by several interacting components.

Dimensions of Complexity

1. Number of and interactions between constituent components within the intervention.
2. Number and difficulty of behaviours required by those delivering or, receiving the intervention.
3. Number and variability of outcomes.
4. Number of groups or organisational levels targeted.
5. Degree of flexibility or tailoring permitted.

Random effects models

Advantages:

- clusters treated as a random sample from a general population allowing the generalisation of the model-fitting results to the whole population.
- in traditional regression methods, the standard errors of the regression coefficients are wrongly estimated.
- the separate estimation of cluster effects and the effects of covariates at the cluster level is achieved.
- measure of the degree of correlation amongst observations within a cluster.
  $\hookrightarrow$ *Intra-Class Correlation coefficient* (ICC).

Partition of variation

$$VPC = \frac{\sigma_u^2 \sum_{j \in Sur_{(i)}} (w_{i,j})^2}{\sigma_u^2 \sum_{j \in Sur_{(i)}} (w_{i,j})^2 + \sigma_e^2}$$

$$ICC = \frac{\sigma_u^2 \sum_{j \in Sur_{(i)} \cup Sur_{(k)}} w_{i,j} w_{k,j}}{\sqrt{\sigma_u^2 \sum_{j \in Sur_{(i)}} (w_{i,j})^2 + \sigma_e^2} \sqrt{\sigma_u^2 \sum_{j \in Sur_{(k)}} (w_{k,j})^2 + \sigma_e^2}}$$

Correlation Coefficients

Partition of variation

$$VPC = \frac{\sigma_u^2 \sum_{j \in Sur_{(i)}} (w_{i,j})^2}{\sigma_u^2 \sum_{j \in Sur_{(i)}} (w_{i,j})^2 + \sigma_e^2}$$

$$ICC = \frac{\sigma_u^2 \sum_{j \in Sur_{(i)} \cup Sur_{(k)}} w_{i,j} w_{k,j}}{\sqrt{\sigma_u^2 \sum_{j \in Sur_{(i)}} (w_{i,j})^2 + \sigma_e^2} \sqrt{\sigma_u^2 \sum_{j \in Sur_{(k)}} (w_{k,j})^2 + \sigma_e^2}}$$

Two-level model

$$VPC = ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

Parameter Expansion

- Used in single-response models when a variance component is small and chain gets stuck at values close to zero.
- Originally applied to Gibbs sampling to speed up convergence and mixing properties of the chain.
- Achieved by introducing parameters α not identified in the likelihood, for which all information comes from the prior. Placing priors on these, induces different prior distributions for the variance components.
- All priors from the non-central scaled F distribution $\rightarrow$ prior for the standard deviation is a non-central folded scaled t-distribution (Gelman, 2006).
- Essential to specify the prior means $\alpha.mu$ and prior covariance $\alpha.V$ in the prior.

Slice Sampling

- Can be used when the distribution can be factored such that one factor is a distribution from which truncated random variables can be drawn.
- The latent variables in univariate binary models can be updated in this way.
- In these models, slice sampling is only marginally more efficient than adaptive Metropolis-Hastings updates when the residual variance is fixed.
- For parameter expanded binary models where the residual variance is not fixed, the slice sampler can be much more efficient.