The use of Electronic Health Records in the design of trials of complex surgical interventions: identifying the contribution of multiple providers

Olympia Papachristofii
with Prof Linda Sharples

1 London School of Hygiene of Tropical Medicine
olympia.papachristofii@lshtm.ac.uk

March 2017, London
Complex Interventions: composed of several interacting components acting independently as well as interdependently.
Building up complex surgical evaluations:

- Rank in order of importance or influence
- Identify the main intervention components
- Construct methods considering further components in sequence

IDEAL framework

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

* IDEAL stages

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

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.

Establish the effect of the Centre and the implications of multiple clustering sources for trial design.
Explicitly account for grouped data structure i.e. patients grouped within anaesthetists and surgeons, grouped within Centres.
Random effects models

- Clusters treated as a random sample from a general population ⇒ results generalisable to the whole population.

- Improved accuracy in the standard errors and p-values of the regression coefficients.

- Separate estimation of operator effects and operator associated covariates.

- *Intra-Class Correlation coefficient (ICC)* → Proportion of variation in outcome attributed to:
  - Centre
  - Surgeons
  - Anaesthetists
  - Patients

Olympia Papachristofi
March 2017, London
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.

Cross-Classified model accounting for lower level units belonging to more than one clusters which are not nested.
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.
Application

- Large case series study (>110000 patients) in 10 of 36 specialist cardiac UK Centres during 10 years (2002-2012).

- 110769 operations conducted by 127 surgeons and 190 anaesthetists.

- Binary response variable: in-hospital death up to 3 months post-operatively.

- First level covariate: the logistic-EuroSCORE of each patient.
Three-level random intercept model

\[
\logit(\pi_{ijk}) = \beta_0_{ij} + \sum_w \beta_w (x_{ijkw} - \bar{x}_w) \quad (1)
\]

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

\( \pi_{ijk} \) the probability of an in hospital death for the \( k^{th} \) patient treated by the \( j^{th} \) surgeon and \( i^{th} \) Centre.

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

\( x_{ijkw} \) the \( w^{th} \) covariate for the \( k^{th} \) patient treated by the \( j^{th} \) surgeon and \( i^{th} \) Centre.
Three-level random intercept model

$$\logit(\pi_{ijk}) = \beta_{0ij} + \sum_{w} \beta_{w} (x_{ijkw} - \bar{x}_{w})$$  \hspace{1cm} (2)

$$\beta_{0ij} = \alpha + u_{ij} + z_{i}, \ z_{i} \sim N(0, \sigma_{z}^{2}) \ \text{the} \ i^{th} \ \text{Centre random intercept.}$$

$$u_{ij} \sim N(0, \sigma_{u}^{2}) \ \text{for the} \ j^{th} \ \text{surgeon in the} \ i^{th} \ \text{Centre.}$$
Three-level cross-classified model

\[
\text{logit}(\pi_{ijkl}) = \beta_{0ijk} + \sum_{w} \beta_w(x_{ijklw} - \bar{x}_w)
\]

(3)

\[
\beta_{0ijk} = \alpha + u_{ij} + v_{ik} + z_i, \quad z_i \sim N(0, \sigma_z^2)
\]

the \(i^{th}\) Centre random intercept.

\[
u_{ij} \sim N(0, \sigma_u^2)\]

for \(j^{th}\) surgeon in \(i^{th}\) Centre.

\[
v_{ik} \sim N(0, \sigma_v^2)\]

for the \(k^{th}\) anaesthetist.
Results

Three-level model

Difference in probability of event caused by surgeon random effects

Surgeon conditional predicted means for surgeons

(a) Surgeon "forest plot" adjusted for case-mix and Centre

\[ ICC_S = 4.06\% \]

(b) Surgeon "forest plot" adjusted for case-mix and both for Anaesthetist and Centre

\[ ICC_{S/A} = 4.01\% \]
Results

Three-level model

Three-level cross-classified model

(c) Anaesthetist "forest plot" adjusted for case-mix and Centre

\[ ICC_A = 0.72\% \]

(d) Anaesthetist "forest plot" adjusted for case-mix and both for Surgeon and Centre

\[ ICC_{A/S} = 0.24\% \]
Results

Proportion of the variation in the outcome attributed to the Centre, surgeon, anaesthetist and the patient risk profile.

<table>
<thead>
<tr>
<th></th>
<th>Centre</th>
<th>Surgeon</th>
<th>Anaesthetist</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICC (%)</td>
<td>0</td>
<td>4.01</td>
<td>0.24</td>
<td>95.75</td>
</tr>
</tbody>
</table>

Table: Percentage of outcome variation due to each group
Patient vs. Process outcomes

<table>
<thead>
<tr>
<th>Mean LOS per Surgeon</th>
<th>Mean LOS per Centre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference in LOS(days) caused by surgeon random effects</td>
<td>Difference in LOS(days) caused by centre random effects</td>
</tr>
<tr>
<td>Surgeon conditional predicted means for surgeons</td>
<td>Centre conditional predicted means for centres</td>
</tr>
<tr>
<td>1 6 12 19 26 33 40 47 54 61 68 75 82 89 96 104 113 122</td>
<td>6 4 3 9 5 1 2 8 7 10 6 7 8 9 10 11 12</td>
</tr>
</tbody>
</table>

(e) Surgeon "forest plot" adjusted for case-mix and both for Centre and Anaesthetist

\[ ICC_{S/A} = 2.79\% \]

(f) Centre "forest plot" adjusted for case-mix and both for Surgeon and Anaesthetist

\[ ICC_{C/S,A} = 1.59\% \]

Olympia Papachristofi
March 2017, London
Assessment of multiple surgical components

* **Identification of relevant setting and data sources**

→ Limited information from RCTs on the effects of different intervention components at trial design.

Example of fit-for-purpose, large collaborative registries → UK Registry of Prostate Embolization (UK-ROPE) registry.
### Choice of Trial design

#### Design Effect

<table>
<thead>
<tr>
<th>Clustering Scenario</th>
<th>Design Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) Main Component - Equal cluster sizes</td>
<td>$1 + (\bar{m}_S - 1) \rho_S$</td>
</tr>
</tbody>
</table>

**Table:** Design Effect for different clustering scenarios

$\bar{m}_S, \bar{m}_A$, the average number of patients treated per surgeon and anaesthetist.

$cv_S, cv_A$, the coefficient of variation per surgeon and anaesthetist.

$\rho_S, \rho_A$ the $ICC_{surgeon}$, $ICC_{anaesthetist}$ respectively.
## Choice of Trial design

### Design Effect

<table>
<thead>
<tr>
<th>Clustering Scenario</th>
<th>Design Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) Main Component - Equal cluster sizes</td>
<td>$1+(\bar{m}_S-1)\rho_S$</td>
</tr>
<tr>
<td>B) Main component - Unequal cluster sizes</td>
<td>$1+(\bar{m}_S(1+cv_S^2)-1)\rho_S$</td>
</tr>
</tbody>
</table>

**Table:** Design Effect for different clustering scenarios

$\bar{m}_S, \bar{m}_A$, the average number of patients treated per surgeon and anaesthetist.

$cv_S, cv_A$, the coefficient of variation per surgeon and anaesthetist.

$\rho_S, \rho_A$ the $ICC_{surgeon}$, $ICC_{anaesthetist}$ respectively.
Choice of Trial design

**Design Effect**

<table>
<thead>
<tr>
<th>Clustering Scenario</th>
<th>Design Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) Main Component - Equal cluster sizes</td>
<td>$1 + (\bar{m}_S - 1) \rho_S$</td>
</tr>
<tr>
<td>B) Main component - Unequal cluster sizes</td>
<td>$1 + (\bar{m}_S (1 + cv_S^2) - 1) \rho_S$</td>
</tr>
<tr>
<td>C) Cross-classification</td>
<td>$1 + (\bar{m}_S (1 + cv_S^2) - 1) \rho_S + (\bar{m}_A (1 + cv_A^2) - 1) \rho_A$</td>
</tr>
</tbody>
</table>

**Table:** Design Effect for different clustering scenarios

$\bar{m}_S, \bar{m}_A$, the average number of patients treated per surgeon and anaesthetist.

$cv_S, cv_A$, the coefficient of variation per surgeon and anaesthetist.

$\rho_S, \rho_A$ the $ICC_{surgeon}, ICC_{anaesthetist}$ respectively.
Choice of Trial design

→ **Primary outcome**: in-hospital death.

→ $ICC_{\text{surgeon}} = 4.01\%$, $ICC_{\text{anaesthetist}} = 0.24\%$

<table>
<thead>
<tr>
<th>Comparator</th>
<th>Design Effect</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Main component</td>
<td>Main component</td>
<td>Cross-classification</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Equal CS</td>
<td>Unequal CS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgery vs. Surgery</td>
<td>1.361</td>
<td>1.417</td>
<td>1.434</td>
<td></td>
</tr>
<tr>
<td>Surgery vs. Medical</td>
<td>1.180</td>
<td>1.209</td>
<td>1.217</td>
<td></td>
</tr>
</tbody>
</table>

→ Power is underestimated if clustering is ignored.

→ In this setting $ICC_{\text{anaesthetist}}$ very small $\Rightarrow$ minimal differences in sample size inflation.
Conclusions

* **Cardiac surgery case-study**

- Surgeon expertise had a small, yet significant influence.
- In this setting, anaesthetist effect was negligible after adjusting for surgeon effects.
- Centre did not affect mortality after adjustment for surgeon and anaesthetist but did have an impact on LOS.
- Loss in power by ignoring anaesthetist-induced clustering minimal in this example.
Conclusions

* **Importance of Electronic Health Records**

- Construct databases of risk-adjusted estimates of ICCs for different intervention components.

- Use of evidence synthesis to enrich short-term RCT outcomes → long-term cost-effectiveness analyses.

- Identify the presence of clustering for various outcomes of interest → patient vs. process outcomes.
Conclusions

* **Importance of Electronic Health Records**

- Construct databases of risk-adjusted estimates of ICCs for different intervention components.
- Use of evidence synthesis to enrich short-term RCT outcomes → long-term cost-effectiveness analyses.
- Identify the presence of clustering for various outcomes of interest → patient vs. process outcomes.

**Proposed Methodology**

- Enables the investigation of non-hierarchical data structures and the construction of ICC databases.
- Informs the choice of trial design, comparator and sample size calculations.


Alternative approach: \textit{GEEs}

\textbf{Generalised Estimating Equations}

- GEE assume observations in same cluster correlated but independent to observations from distinct clusters.
- Exchangeable correlation structure $\rightarrow$ equivalent to a random intercept model.
- Standard errors robust to misspecification of the working correlation structure.
Alternative approach: *GEEs*

**Generalised Estimating Equations**

- GEE assume observations in same cluster correlated but independent to observations from distinct clusters.
- Exchangeable correlation structure → equivalent to a random intercept model.
- Standard errors robust to misspecification of the working correlation structure.

**Inappropriate in this case**

- Do not model explicitly between and within cluster variability.
- GEE correlation measure underestimated with few clusters available → sub-optimal measure.
- Cannot be used to study specific cluster performance.
Anaesthetist performance

- Examine whether specific surgeon/anaesthetist collaborations instead of randomly assigning pairs improve collective performance.

- Anaesthetists ⇒ ”members” of more than one surgeon clusters.

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

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

**Anaesthetist Multiple Membership Multiple classification model**

\[
\text{logit}(\pi_i) = \beta_{0i} + \sum_w \beta_w(x_{iw} - \bar{x}_w) \quad (4)
\]

\[
\beta_{0i} = \alpha + v_{A(i)} + \sum_{j \in \text{Sur}(i)} w_{i,j} u_j.
\]

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.
Choice of weights

Weights

\[ w_{i,j} = \frac{\text{operations done by } A(i) \text{ with } Sur_{(i),j}}{\text{total operations done by } A(i)} \]

Fitted using simulation methods in package MCMCglmm in R.
Random Interaction

- Possibility of interacting operators within medical teams.
- Marginal effect of anaesthetist differs according to the operating surgeon and vice-versa.
- Need to extend the cross-classified model term for additional random source of variation.
Two-level cross-classification model with random interaction

\[
\logit(\pi_{ijk}) = \beta_{0ij} + \sum_w \beta_w(x_{ijkw} - \overline{x_w})
\] (5)

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

\( \beta_{0ij} = \alpha + u_i + v_j + uv_{(i,j)} \), \( u_i \sim N(0, \sigma^2_u) \) \( i^{th} \) surgeon random intercept
\( v_j \sim N(0, \sigma^2_v) \) for the \( j^{th} \) anaesthetist
\( uv_{(i,j)} \sim N(0, \sigma^2_{uv}) \) respectively for pair \((i, j)\)

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

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

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

Partition of variation

\[
VPC = \frac{\sigma_u^2 \sum_{j \in Clust(i)} (w_{i,j})^2}{\sigma_u^2 \sum_{j \in Clust(i)} (w_{i,j})^2 + \sigma_e^2}
\]

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

Two-level model

\[
VPC = ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}
\]
Choice of Priors

- **Random effects variance**

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

  ** Improper uniform on $\sigma_u^2$.

**Drawbacks:**

- bad behaviour for variances close to 0.

** miscalibration towards higher $\sigma_u$ 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 $\sigma_u^2$.**

**Drawbacks:**
- Bad behaviour for variances close to 0.
- Miscalibration towards higher $\sigma_u$ values and need more than 4 clusters to get a proper posterior.

→ **Suggest**: Improper Uniform on $\sigma_u$.

→ Parameter-expanded priors and slice sampling to ensure effective mixing.
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 $\alpha$ 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 $\alpha \cdot \mu$ and prior covariance $\alpha \cdot V$ 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.