Disclosure

• Maria Costa is an employee and shareholder of GSK

• Data presented is based on human research studies funded and sponsored by GSK
Outline

- Motivation
- Bayesian Joint Modelling of Mixed Outcomes
- Simulation Study
- A Case Study in Type 1 Diabetes
- Summary
Motivation

Quantitative Benefit-Risk as a Strategy for Risk Mitigation

- Goal of NDA Safety Review: To determine the significance of the adverse events and their impact on the approvability of the drug
  - “To show whether or not such drug is safe for use under the conditions prescribed, recommended, or suggested in the proposed labelling” (Food, Drug, and Cosmetic Act (Section 505))

- How do we know that the Benefit-Risk balance is “positive”? If positive, under which context (which population, etc)?

- How to enhance the transparency, reproducibility and communication of the Benefit-Risk balance of medicines?

- How to assess the impact of uncertainty in Benefit-Risk assessments?

Quantitative Benefit-Risk assessments can support decision makers with these questions...
Motivation
Multivariate Modelling, Bayesian Inference & Quantitative Benefit-Risk

Multivariate Modelling
- Potential for efficacy and safety signals to be linked via exposure to active drug
- Joint modelling of efficacy and safety endpoints enables efficient data driven BR analyses

Bayesian Inference
- Provides direct framework to build relevant and intuitive probability statements in the context of BR that can be used to quantify uncertainty and risk
- Bayesian updating mechanism naturally supports “Learn & Confirm” drug development paradigm – crucial when assessing BR

Quantitative Benefit-Risk Assessment
- Can help team gain insight into specific BR questions about key endpoints of interest
- Important to communicate BR to stakeholders in a way that supports decision-making
- Important to quantify uncertainty in BR profile – particularly if aim is to discharge risk
Bayesian Joint Modelling of Mixed Outcomes

Motivation

- Strength of efficacy and safety signals likely to be linked at subject level via exposure to active drug:

- Approach that accounts for observed correlation at subject level between efficacy and safety signals is desirable – more efficient and realistic assessment

- Often efficacy and safety endpoints modelled using different distributions

- Should focus on key endpoints – primary efficacy and key safety finding(s) identified by safety team?
Bayesian Joint Modelling of Mixed Outcomes

Option 1: Use generalised linear mixed models (GLMM)
- Assume $J$ different observations on same subject (each following some distribution)
- For subject $i$ with mean response $\mu_i$, $g(\mu_i) = X_i b + Z_i u_i$, $u_i \sim N(0, G(X_i))$

- Random effect $u_i$ is shared across all $J$ observations for subject $i$ thus modelling potential correlation

- When $g_j(.)$ is not identity function the fixed effects $b$ are conditional on random effects $u_i$
  - Monte Carlo integration can be used to obtain marginal population effects – important when making inferences at the population level

- Constraints may be necessary to ensure identifiability for certain distributions
Bayesian Joint Modelling of Mixed Outcomes

Approaches to Linking Mixed Outcomes: Copulas

- Option 2: Use copulas
  - Copulas - distribution functions used to form new multivariate distributions given set of marginal distributions of interest (which are preserved)
  - E.g., \( H(y_1, y_2) = C(F(y_1), G(y_2) | \theta) \), with \( F(.) \) and \( G(.) \) the CDF of the marginal distributions of \( y_1 \) and \( y_2 \)
  - \( C(. , . | \theta) \) is the copula function (e.g., Gaussian CDF)
  - \( \theta \) measures association between \( y_1 \) and \( y_2 \)

- Directly obtain marginal population effects for parameters of interest

- Choice of copula \( C(.) \) may impact results through different dependency assumptions

- Difficult to interpret beyond 3 dimensions (non-unique model definition)
Simulation Study

Set Up

- Two treatment arms: new drug (treatment 2) vs comparator (treatment 1)

- Endpoints and parameters:

<table>
<thead>
<tr>
<th>BR Endpoints</th>
<th>Endpoint Type</th>
<th>Parameter Values</th>
<th>Correlation between endpoints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary efficacy endpoint</td>
<td>Continuous, N (μ, σ²)</td>
<td>μ₁ = -150</td>
<td>ρ₁ = 0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>μ₂ = -50</td>
<td>ρ₂ = 0.6</td>
</tr>
<tr>
<td>Key AE endpoint - AESI</td>
<td>Binary, Bernoulli (p)</td>
<td>p₁ = 0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>p₂ = 0.4</td>
<td></td>
</tr>
</tbody>
</table>

- Comparisons of interest as follows: μ₂ - μ₁ and ρ₂ - p₁

- Non-informative priors assumed for all model parameters

- 100 simulated datasets generated

- Bayesian inference performed using MCMC
Aim is to assess the level of evidence (i.e., posterior probability) associated with BR profiles of interest and to understand trade-off between efficacy & safety.

Different BR profiles can be set up using range of clinically meaningful efficacy and safety thresholds:
- $\Delta_e$ represents minimum improvements in efficacy with the new drug relative to comparator
- $\Delta_s$ represents maximum increases in risk with the new drug relative to comparator

$\Delta_e$ and $\Delta_s$ are independent and set by the project team - can be viewed as clinical Go/No-go boundaries.

Trade-off between efficacy and safety represented by following probability statement:
- Prob ($\mu_2 - \mu_1 > \Delta_e$ and $p_2 - p_1 < \Delta_s$ | Data, prior)
Simulation Study

Results for a typical simulated dataset

Posterior Distribution for $\mu_2 - \mu_1$ and $p_2 - p_1$
(Joint and Marginal)

Elliptical shape of joint posterior reflects correlation between $\mu_2$ and $p_2$

BR Contour
Prob $(\mu_2 - \mu_1 > \Delta_e$ and $p_2 - p_1 < \Delta_s |$ Data)

Example: 84% posterior probability that difference active vs comparator in risk of an AE is at most 0.35 ($\Delta_s = 0.35$) and in efficacy at least 80 units ($\Delta_e = 80$)
Simulation Study

Impact of correlation

- To assess impact of correlation on $\text{Prob}\left(\mu_2 - \mu_1 > \Delta_e \text{ and } p_2 - p_1 < \Delta_s \mid \text{Data, prior}\right)$ simulations were run for different values of $\rho_2$

- We fix $\Delta_e = 100$ and $\Delta_s = 0.3$

<table>
<thead>
<tr>
<th>$\rho_2$</th>
<th>GLMM Model</th>
<th>Gaussian Copula Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>24.95%</td>
<td>26.06%</td>
</tr>
<tr>
<td>0.2</td>
<td>23.06%</td>
<td>23.93%</td>
</tr>
<tr>
<td>0.4</td>
<td>20.72%</td>
<td>21.68%</td>
</tr>
<tr>
<td>0.6</td>
<td>18.42%</td>
<td>19.77%</td>
</tr>
<tr>
<td>0.75</td>
<td>17.89%</td>
<td>19.57%</td>
</tr>
</tbody>
</table>

- Increasing values of $\rho_2$ leads to lower posterior probability values for the BR profile defined by $\Delta_e = 100$ and $\Delta_s = 0.3$
  - Accounting for correlation results in more realistic BR assessments
Choosing dose with optimal BR profile is major hurdle in drug development
- Too high dose may result in an unacceptable risk profile
- Too low dose may decrease the chances of achieving the desired level of efficacy in a phase 3 trial
- 16% of NME applications fail due to uncertainty related to dose selection (Sacks et al, 2014)

Previous simulation study expanded to dose-response setting:
- Same 2 endpoints for efficacy and safety
- 5 active doses (from d2 = 0.3 to d6 = 6 units) and comparator (d1 = 0)
- Correlation $\rho_d$ between efficacy and safety at subject-level increases with dose such that $\rho_{d1} = 0$, and $\rho_{d1} \sim 0.6$
- Emax model (3 parameter) used to generate data for efficacy
  - $E_0 = -150$, $E_{\text{max}} = 150$, $ED_{50} = 0.5$
- Linear regression model on probit scale used to generate safety data
  - $\text{Prob (AE in dose } d\text{)} = p_d = \Phi (-1.28 + 0.26 \times d)$
- Bayesian inference via MCMC
Simulation Study: Dose-Response
Minimum Effective Dose vs Critical Effective Dose

- We define the following quantities:
  - Minimum Effective Dose (MED) = the smallest dose \( d \) that produces an improvement of size \( \Delta_e \) or larger compared to placebo with posterior probability > \( p\% \)
  - Critical Effective Dose (CED) = the largest dose \( d \) that produces an increase in toxicity no greater than \( \Delta_s \) compared to placebo with posterior probability > \( p\% \)

How to select the dose with optimal BR profile?
(given \( \Delta_e, \Delta_s, \) and \( p \))

- If MED < CED
  - Any dose within [MED, CED] will satisfy the desired BR profile
- If MED = CED
  - This corresponds to the single optimal dose
- If MED > CED
  - Will need to either modify \( \Delta_e, \Delta_s, \) or \( p \), or assess whether clinical program is viable
Simulation Study: Dose-Response
Choosing the dose with optimal BR profile – results for a typical simulated dataset

- In general, as $\Delta_e$ increases and $\Delta_s$ decreases, MED $\leq$ CED only by lowering the posterior probability $p\%$ - so the team will need to accept more uncertainty going to phase 3

- If existing correlation is not accounted for, data may erroneously suggest that MED $\leq$ CED with high probability $p$, when in fact this is not the case (simulation results not shown)

If $\Delta_e=80$ $\Delta_s=0.3$ and $p = 70\%$ then MED $= 2.5$ and CED $= 4.0$

Any dose in the range $[2.5, 4.0]$ can be considered “optimal”

If $\Delta_s= 0.3$ is considered too high increase in risk of AE and team sets $\Delta_s= 0.05$, then MED $\leq$ CED only if $p = 30\%$

This means there will be considerably more uncertainty with this more stringent BR profile
Treatment X was a monoclonal antibody targeting CD3 receptors that was being developed as a potential treatment for new-onset (<3 months) Type 1 diabetes mellitus.

- Significant associated morbidity and mortality (neuropathy, ischemic heart disease, among others)

A clinical trial (PoC) was designed to assess efficacy and safety of X over an 18 month period in patients with new-onset type 1 diabetes.

- **Primary efficacy** endpoint was the decline of C-peptide levels at 6 months (measurement of beta-cell function) – treated as continuous outcome

- **Key safety** events of interest included infection and Cytokine Release Syndrome (CRS) – treated as binary outcomes

- A total of 73 subjects had C-peptide levels recorded at 6 months (39 received X, 34 placebo)
Case Study – Treatment X for New Onset Type 1 Diabetes

1 Efficacy & 1 Safety Endpoint – Bayesian Inference of Multivariate Model

- For safety, focus initially on risk of infection, modelled as binary outcome

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Posterior Median</th>
<th>95% Credible Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Efficacy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFB C-Peptide (X - Placebo)</td>
<td>0.63</td>
<td>(0.27, 0.99)</td>
</tr>
<tr>
<td><strong>Safety</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prob (Infection) (X - Placebo)</td>
<td>0.24</td>
<td>(0.07, 0.42)</td>
</tr>
</tbody>
</table>

CFB = Change from baseline at 6 months

- GLMM and Bayesian inference used to obtain parameter estimates of interest

Observed correlation suggests that for patients receiving X, more stable C-peptide levels tend to be associated with the occurrence of at least one infection event.

Patients receiving X have more stable levels of C-Peptide

Patients receiving X have higher risk of serious infection
Case Study – Treatment X for New Onset Type 1 Diabetes

1 Efficacy & 1 Safety Endpoint – Benefit-Risk Assessment

BR Contour Plot

BR profiles with “high” posterior probability correspond to scenarios with a substantial increase in risk of infection.

The data does not support BR profiles for which $\Delta_e > 0.8$ and $\Delta_s < 0.1$.
“Given a patient’s baseline C-peptide level, what is his/her likely BR profile with drug X compared to placebo?”

The BR profile of X is robust to a patient’s baseline C-Peptide level.

In placebo group, subjects with lower baseline C-Peptide levels have a more favourable BR profile.
Case Study – Treatment X for New Onset Type 1 Diabetes

Benefit-Risk Assessment

- Does the BR assessment of drug X PoC study support further development?
  - BR analysis presented here suggests that high efficacy levels with low increases in risk are unlikely (< 10% probability)

- GSK run phase 3 program with lower dose of drug X – studies failed to achieve their primary endpoints, although risk profile improved
  - This is coherent with BR analysis conducted on PoC data – could the expensive and time consuming phase 3 program been avoided by looking quantitatively at chances of positive benefit-risk profile?
Bayesian inference based on joint models of mixed outcomes is a powerful tool for Benefit-Risk assessment

- Explore dependency between benefit and risk thresholds for decision-making
- Joint (and conditional) probabilistic statements that help quantify risk in development program
- Predicting responses for a new subject conditional on what was learned from study data

Benefit-Risk profile is a combination of two different quantities:

- Set of thresholds for efficacy and safety – define Benefit-Risk profile of interest (qualitative)
- Level of evidence (posterior probability) to support Benefit-Risk profile – quantify risk (quantitative)

Methods have been successfully applied to 3-dim setting as well (mixture of continuous, binary and count endpoints)

- Beyond 3 dimensions it is difficult to interpret and visualise quantitative BR assessments
References

- Costa & Drury (2017), Bayesian Joint Modelling of Benefit and Risk in Drug Development (submitted)


Acknowledgements

- Thomas Drury
- Nigel Dallow
- Graeme Archer
- Nicky Best
- James Roger
Thank you