## Workshop Agenda

<table>
<thead>
<tr>
<th>Order</th>
<th>Presenter</th>
<th>Opportunities and Considerations of Network Meta-Analysis</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Keith Abrams</td>
<td>An Introduction to Network Meta-Analysis (NMA)</td>
<td>10 mins</td>
</tr>
<tr>
<td>2</td>
<td>Bob Cuddihy</td>
<td>A Clinician’s Perspective</td>
<td>10 mins</td>
</tr>
<tr>
<td>3</td>
<td>Maud Pacou</td>
<td>An Example of Bayesian NMA in T2d</td>
<td>10 mins</td>
</tr>
<tr>
<td>4</td>
<td>Mel Schroeder</td>
<td>HTA Experience</td>
<td>10 mins</td>
</tr>
<tr>
<td>5</td>
<td>All (+audience participation)</td>
<td>Q&amp;A</td>
<td>20 mins</td>
</tr>
</tbody>
</table>
NMA – a statistician’s perspective
Keith R Abrams, PhD CStat

Disclaimer

- I have received financial help to attend this conference from J&J/Janssen.

- I have acted as a paid consultant to a number of pharmaceutical companies, including; AZ, GSK, J&J/Janssen, Merck, Novartis & Roche.
Indirect comparisons

- Interest in treatment comparison A v B
- No direct evidence comparing A v B
- Treatment C is common comparator of A & B
- We can indirectly estimate the effect of A v B

Network meta-analysis

- Utilises both direct and indirect information
- Loops of evidence e.g. AB, AC, BC
- Assess inconsistency
- Any connected network
NMA – the key assumption

- IC and NMA assume that the “Direct” and “Indirect” evidence estimate the same parameter (effect)

- Treatment effect $d_{BC}$ estimated by the BC trials, would be the same as the treatment effect estimated by the AC and AB trials (if they had included B and C arms, respectively), i.e. they were 3 arm trials

- Nearly all the doubts about IC and NMA are doubts about this assumption

NMA – models

- Fixed effect model
  - common treatment effect

- Random effects models
  - Studies have an underlying effect drawn from a population of effects
  - Allow for between-study heterogeneity
  - Typically will produce greater uncertainty in overall (population) effects

- Comparison of fixed and random effect models
  - Use common relative model comparison statistics, e.g. DIC
  - Assess between-study heterogeneity, e.g. SD
  - RE models require more data!
  - Absolute model fit is important too! e.g. residual deviance
Heterogeneity & Inconsistency – 1

- In NMA also need to explore potential sources of variability:

  **Heterogeneity** - variation in treatment effects *between trials within pairwise contrasts*,

  &

  **Inconsistency** - variation in treatment effects *between pairwise contrasts*

- **Random effects** – allows for heterogeneity but does NOT explain it nor ensure inconsistency is addressed

- **Inconsistency models** – allows for inconsistency but does NOT explain causes (Dias et al, 2013)

Heterogeneity & Inconsistency – 2

- **Incorporation of study-level covariates** (including baseline risk) can reduce both heterogeneity and inconsistency by allowing systematic variability between trials to be explained (Achana et al, 2013; Cooper et al, 2009)

  - **BUT** there has to be sufficient data for this – “rule of thumb”: 10 data points per additional parameter in model => 10 trials

  - **AND** we need to make assumptions in NMA regarding how covariates act on each treatment – common effect, independent effects or exchangeable effects (Cooper et al, 2009)

- Baseline risk can pose problems because very often different trials (with different interventions) have used different comparators => also need to develop a baseline risk model to predict effects of a common comparator (Achana et al, 2013)
Bayesian vs Frequentist methods

- Bayesian methods can be more flexible than Frequentist approaches, and allow appropriate probability statements to be made.
  - **BUT** these need to be understood by decision makers
  - **BUT** require specification of prior distributions for model parameters
- Crucial to explore the impact that prior distributions have AND
  - Assess evidence of non-convergence of simulation methods, e.g. Markov Chain Monte Carlo (MCMC)
- Common mistake is to use excessively vague prior distributions (especially with sparse data/networks)
  - Use plausibly vague prior distributions (Spiegelhalter et al, 2003)
  - Example: log OR scale – between-study SD of 0.5 => ratio of ORs of any random pair of studies is 1.7 and of range (max:min) is 7.0

Further developments

- **Network geometry:**
  - Size of network – evidence-space vs decision-space (Dequen et al, 2014)
- **Interventions:**
  - Class effects/hierarchical NMA (Warren et al, 2014; Owen et al, 2015)
  - Complex interventions (Welton et al, 2009)
- **Outcomes:**
  - Multivariate NMA (Bujkiewicz et al, 2013)
  - Multiple time-points (Jansen et al, 2015)
- **Study design & data availability:**
  - Inclusion of RWE (Jenkins et al, 2014)
  - Summary data and IPD (Saramago et al, 2012)
- **Reporting:**
  - Predictive vs average effects (Ades et al, 2005)
References


Opportunities and Considerations of Network Meta-Analysis: A Clinician’s Perspective

Robert Cuddihy, MD

Janssen Global Services, LLC

ISPOR 21st Annual Meeting

The Art and Science of Performing Network Meta-analysis to Aid Decision Making for Type 2 Diabetes Treatment: How Does It Work in Practice?

May 23, 2016
Recommendations for Hyperglycemia Management in Type 2 Diabetes Mellitus (T2DM)

- The American Diabetes Association (ADA) and European Association for the Study of Diabetes (EASD) issued an updated position statement on hyperglycemia management in T2DM in 2015.

- Key recommendations:
  - Individualize treatment targets and strategies, with emphasis on patient-centered care and shared decision making
  - Balance benefits of treatment with risks (eg, age, side effects)
  - Consider the impact of comorbidities
  - Provide nutritional counseling and diabetes self-management education

ADA/EASD Treatment Algorithm

- 7 different classes are included in the treatment algorithm (~40 unique agents!)
- Impractical to conduct head-to-head trials of all drugs
Increased Use of Indirect Comparisons/Network Meta-analysis (NMA) in T2DM

- Growing number of indirect comparison/NMA manuscripts in diabetes
- PubMed search for ‘diabetes’ and ‘indirect comparison,’ ‘network meta-analysis,’ or ‘Bayesian network meta-analysis’
- Included all manuscripts published through 2015

![Number of publications over time (N = 39)](image)

- Many indirect comparisons/NMAs are presented at congresses
  - 25 ‘indirect comparison’ and 37 ‘network meta-analysis’ presentations at global ISPOR meetings from 1998 to 2015
    - Few before 2011; ~30-40% in 2015

Need for Clinical Interpretation of NMAs:

National Institute for Health and Care Excellence (NICE) Draft Guidance on T2DM as an Example

- In the UK, NICE issued an updated draft guidance for T2DM management to various stakeholders for consultation in early 2015

- Most recommendations were considered to be sensible, including those related to lifestyle modifications and patient education

- Recommendations for blood glucose management based on NMA results were controversial

A Controversial Recommendation

"1.6.19 If standard release metformin is contraindicated or not tolerated consider repaglinide as initial drug therapy. Advise the person that if treatment with repaglinide does not control HbA1c, then the person would need to change to pioglitazone, a sulphonylurea, or a DPP-4 inhibitor before adding another treatment."—NICE

- Recommendations were informed by NMAs, but appeared not to give proper consideration to certain clinical issues and ignore some important features of these drugs.
- Recommendations also seemed to be inconsistent and relied heavily on the actual point estimates of cost-effectiveness, even though there was acknowledgment of important issues that could not be fully captured in simulation exercises.

Importance of Study Selection and Clinical Interpretation

- Of the 7 monotherapy studies with repaglinide included in the NMA, 4 of the studies were found to be:
  - Underpowered
  - Open to treatment bias in their design
  - Not relevant to the UK population
- Inclusion of these studies produced relative efficacy/safety NMA findings that led to the controversial recommendation
- Guidance was revised based on stakeholder feedback and published in December 2015.
Clinical Considerations for NMAs

- Trial design
- Trial duration
- Outcomes of interest
- Inclusion/exclusion criteria
- Population characteristics
- Regional differences
- Comparators
- Titration schemes

Summary

- NMAs can provide a robust and efficient method for comparing data from a large number of T2DM treatments
- Optimal T2DM management needs to be patient-centered
- Clinicians can provide valuable perspective to ensure studies are appropriately selected for NMAs, so findings can be appropriately contextualized

“..Complex network meta-analyses used are only fit for purpose when sensible clinical judgement provides the context for their use…”—O’Hare and colleagues

AN EXAMPLE OF BAYESIAN NETWORK META-ANALYSIS

Canagliflozin in type 2 diabetes mellitus

Amaris, UK

Maud Pacou, MSc

INTRODUCTION & CONTEXT

- A Bayesian NMA was conducted to estimate the relative efficacy and safety of canagliflozin in type 2 diabetes mellitus (T2DM)
- Results were used as part of HTA submissions to synthesize clinical evidence and inform cost-effectiveness models
- Canagliflozin has been assessed in clinical trials with different background therapies
  - Monotherapy
  - Dual therapies: metformin (MET), sulfonylurea (SU)
  - Triple therapies: MET+SU, MET+pioglitazone (PIO)
  - With insulin +/- oral anti-diabetic agent
- Objectives of this presentation
  - Present the assumptions required to conduct the NMA
  - Discuss 3 challenges that we have experienced in this project
  - Discuss the interpretation of results from a Bayesian NMA
METHODS

- Bayesian NMA was conducted in line with NICE and ISPOR guidelines.

- Prior distributions for the analysis of the HbA1c change from baseline:
  - Trial baseline: Normal (mean=0, precision=10^{-4} or 10^{-3} depending on the paucity of the network)
  - Relative treatment effects: Normal (mean=0, precision=10^{-4} or 10^{-3} depending on the paucity of the network)
  - Between-study standard deviation: Uniform (0, 5)

- Networks were based on:
  - Background therapy
    - One network of evidence was built by background therapy to match the study population of each canagliflozin trial as well as the treatment pathway of T2DM patients.
  - Timepoint: 26, 52 and 104 weeks +/- 4 weeks
  - Outcomes
    - Glycaemic control: HbA1c change from baseline, % of patients reaching HbA1c <7%, fasting plasma glucose (FPG) change from baseline
    - Weight and BMI change from baseline
    - Systolic blood pressure (SBP) change from baseline
    - Proportion of patients who experienced hypoglycaemic events

NETWORKS

- MET background – 26 weeks
- SU background – 26 weeks
- INS background – 26 weeks
- MET background – 52 weeks
- MET+SU background – 26 weeks
- MET+PIO background – 26 weeks
- Monotherapies – 26 weeks
SPECIFIC CHALLENGES (1)

- **Paucity of evidence**, which may result in a lack of power of the analyses
  - Important issue for the 104-week analysis since long-term data were missing or not published for most comparators

SPECIFIC CHALLENGES (2)

- **Heterogeneity** across trials with regards to baseline characteristics, in particular HbA1c level and weight
  - Meta-regression was not feasible in the 104-wk analysis (due to the paucity of evidence)
  - Meta-regression was only feasible in the 26-week analysis adjusting for 1 single study-level covariate based on placebo-controlled studies only (model assuming the same interaction effect for all treatments)

Full network
MET background – 26 weeks

Only placebo-controlled studies
MET background – 26 weeks
SPECIFIC CHALLENGES (3)

- Challenges associated with **secondary efficacy and safety endpoints**
  - High level of missing data on secondary efficacy endpoints such as BMI change from baseline
  - Analyses conducted on restricted networks

HbA1c change from baseline  
MET background – 26 weeks

BMI change from baseline  
MET background – 26 weeks

**HBA1C MEAN CHANGE FROM BASELINE AT 104 WEEKS**

- Interpretation of Bayesian results
  - **95% credibility interval** (CrI): there is 95% probability that the parameter lies in the credibility interval
  - **Bayesian pairwise probability** (P): probability for canagliflozin to perform better than each comparator
CONCLUSION

- Difficulties associated with the analyses of multiple endpoints across time points, adjusting for heterogeneity across trials

- Major assumptions required for the NMA to be feasible
  - Validation of assumptions based on
    - Clinical opinion
  - Analyses of patient data from the canagliflozin trial(s), e.g. LOCF vs. MMRM analyses at 104 weeks
  - Sensitivity analyses to ensure the robustness of the results

Opportunities and Considerations of Network Meta-Analysis: HTA experience

Melanie Schroeder, MSc

Janssen-Cilag Ltd, UK

ISPOR 21st Annual Meeting
The Art and Science of Performing Network Meta-analysis to Aid Decision Making for Type 2 Diabetes Treatment: How Does It Work in Practice?
May 23, 2016
Overview

- Making an NMA reimbursement ready
  - Which networks of evidence are used, how comparators were deemed relevant, etc.
- Examples of reimbursement group challenges
  - Experience & Priorities of Scottish Medicines Consortium (SMC), and Consistency and Subsequent Confidence by the National Institute for Health and Care Excellence (NICE).

Considerations when Targeting an NMA to a Particular Healthcare Environment

1. Clinical Understanding and Value
   - Complexity of the network
   - Relevant outcomes
2. Market Access Environment
   - Reimbursed medicines
   - Prescribed medicines
3. Reflective Input Data
   - Patient population
   - Trial design
4. Different HTA Bodies have different Experience & Priorities
Using NMA to Inform Technology Appraisals in the UK

There have been two rounds of assessments for the reimbursement of SGLT2 inhibitors in England, Wales, and Scotland

1. Individual technology appraisals for combination therapy by the NICE and the SMC for the treatment of adults with T2d

2. A multiple technology appraisal (MTA) for the monotherapy use of SGLT2 inhibitors by NICE for the treatment of adults with T2d

During this time there have been a minimum of 8 NMA sets submitted to reimbursement committees for the review of SGLT2 inhibitors for the treatment of T2DM, none of which are the same and all informed by different data sets.

During this time there was also a therapy area treatment guideline update, which generated even more NMAs.

Varying Degrees of Familiarity with Bayesian NMA: an Example from the SMC Review

1. Uncertainty around Bayesian statistic methods used to inform the NMA

- Further justification for the statistical approach taken in the original submission, explored the strengths and limitations of using Bayesian NMA to indirectly compare data across clinical trials.
- Clinical opinion, confirming acceptance of the approach taken.

2. Manipulation of NMA inputs in the economic model that causes an overestimation of treatment effect

- The methodology for diabetes modelling is complex and further clarification confirmed the robustness of the approach.
- ECHO-T2DM uses a micro-simulation approach and captures second order (i.e. parameter) uncertainty. A probabilistic approach consistent with the results of the Bayesian NMA.
Varying Degrees of Familiarity with Bayesian NMA Leads to Challenging Questions: e.g. “Statistical Significance”

- INVOKANA® 300mg has greater HbA1c reductions compared to sitagliptin with changes ranging from –0.20 and –0.39 (all P=100%) and INVOKANA® 100mg had a similar HbA1c reduction compared to sitagliptin (Δ=0.05; P=33%)

- Alternative approaches to demonstrate the confidence in these results:
  - After the 1st time of asking:
    - Simulations results were presented only for the data where head-to-head information was available.
  - After the 2nd time of asking:
    - An analysis was run whereby any differences in treatment effect that are small relative to their variability (and thus more uncertain) were excluded.

Bayesian NMA result in a posterior distribution of the treatment effect; this needs to be interpreted as a probability distribution.

It is not correct, to conclude that a credible interval containing zero means a non-significant result.

Replicability is Key to Understanding: An Example from the NICE MTA Submission

Replicated analysis by Janssen:

The Original Network by the Assessment Group:

The Original Network by Janssen:
Different Approaches Lead to Differences in Results

**Comparison of NMA Results: HbA1c (%) change versus placebo at 26 weeks ±4 weeks.**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Janssen</th>
<th>Astra Zenica</th>
<th>Assessment Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>CANAGLIFLOZIN 100 mg</td>
<td>-0.97</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>CANAGLIFLOZIN 300 mg</td>
<td>-1.20</td>
<td>-0.74</td>
<td>-1.153</td>
</tr>
<tr>
<td>DAPAGLIFLOZIN 10 mg</td>
<td>-0.64</td>
<td></td>
<td>-0.704</td>
</tr>
<tr>
<td>EMPAGLIFLOZIN 10 mg</td>
<td>-0.73</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>EMPAGLIFLOZIN 25 mg</td>
<td>-0.85</td>
<td></td>
<td>-0.870</td>
</tr>
<tr>
<td>SITAGLIPTIN 100 mg</td>
<td>-0.72</td>
<td>-0.64*</td>
<td>-0.723</td>
</tr>
<tr>
<td>PIOGLITAZONE†</td>
<td>-0.76</td>
<td>-0.90</td>
<td>-1.200</td>
</tr>
<tr>
<td>SULPHONYLUREA†</td>
<td>-1.04</td>
<td>-0.95</td>
<td>-1.301</td>
</tr>
<tr>
<td>REPAGLINIDE</td>
<td>-1.28</td>
<td>-</td>
<td>-1.200*</td>
</tr>
</tbody>
</table>

* *pooled* variable dose; † Assumed as no estimate within NMA; *pooled gliptins.


**Conclusions**

- NMAs need to be **tailored** to the reimbursement Committee
- In a dynamic therapy area, be **prepared** to re-run analyses
- This new technique needs **explanation** and accept there will be **alternative interpretations** of the results
- **Leverage** your Medical Affairs colleagues and Clinical Experts to build medically sensible indirect comparison networks
THE SCIENCE AND ART OF PERFORMING NETWORK META-ANALYSES (NMA) TO AID DECISION MAKING FOR TYPE 2 DIABETES

Q&A