An introduction to individual patient data (IPD) meta-analysis

Dr Sarah Nevitt

Department of Biostatistics
University of Liverpool

Email: sjn16@liverpool.ac.uk  
Twitter: @sjn_16
Outline of the webinar

- **IPD meta-analysis – What? When? Why?**
- **Practical aspects of conducting an IPD meta-analysis**
- **Data retrieval for IPD meta-analysis in an era of ‘data transparency.’**
- **What if a subset of IPD cannot be retrieved?**
- **Illustrative examples from Cochrane Epilepsy IPD reviews**
Introduction to Meta-analysis

Statistical technique for combining sources of quantitative evidence

Pairwise meta-analysis

• Evaluation of two interventions (or intervention and control) which have been compared head-to-head in two or more clinical studies.

Network meta-analysis

• Evaluation of more than two interventions (multiple intervention and/or controls) which have been compared head-to-head in two or more clinical studies.
• Also allows indirect evaluation of interventions which have not been compared head-to-head within any clinical studies.
• Network meta-analysis can provide an estimate of the relative effectiveness of all interventions in the network – helpful for medical decision making

Introduction to Meta-analysis

- Meta-analyses have traditionally been performed with aggregate data (AD)

  - Summary statistics (mean differences, event counts, odds ratios, hazard ratios etc.) extracted from published journal articles, conference abstracts, trial registries (e.g. clinicaltrials.gov).
  
  - Unpublished documentation such as protocols, statistical analysis plans, clinical study reports


Table 2: Time to treatment withdrawal and time to first seizure by VPA and CBZ strata for subgroups of patients with focal or generalised seizures only (intent-to-treat population excluding unclassified/unknown seizure types)

<table>
<thead>
<tr>
<th></th>
<th>LEV</th>
<th>Standard AEDs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time to treatment withdrawal</strong></td>
<td>N</td>
<td>Event</td>
</tr>
<tr>
<td>VPA stratum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Focal seizures only</td>
<td>99</td>
<td>15</td>
</tr>
<tr>
<td>Generalised seizures only</td>
<td>226</td>
<td>54</td>
</tr>
<tr>
<td>CBZ stratum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Focal seizures only</td>
<td>418</td>
<td>109</td>
</tr>
<tr>
<td>Generalised seizures only</td>
<td>46</td>
<td>5</td>
</tr>
<tr>
<td><strong>Time to first seizure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VPA stratum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Focal seizures only</td>
<td>99</td>
<td>44</td>
</tr>
<tr>
<td>Generalised seizures only</td>
<td>226</td>
<td>78</td>
</tr>
<tr>
<td>CBZ stratum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Focal seizures only</td>
<td>418</td>
<td>198</td>
</tr>
<tr>
<td>Generalised seizures only</td>
<td>46</td>
<td>15</td>
</tr>
</tbody>
</table>

*HR of <1 favours LEV.

AED, antiepileptic drug; CBZ, carbamazepine; LEV, levetiracetam; VPA, sodium valproate; N, number of patients; Event, number of patients who had the event (treatment withdrawal or seizure); HR, HR for time to event (treatment withdrawal or first seizure).

Introduction to Meta-analysis

• Alternative: **Individual patient / participant data (IPD)** approach

<table>
<thead>
<tr>
<th>Patient Number</th>
<th>Treatment</th>
<th>Withdrawal Time (days)</th>
<th>Status</th>
<th>Age</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>44</td>
<td>Withdrew</td>
<td>67</td>
<td>M</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>54</td>
<td>Withdrew</td>
<td>64</td>
<td>M</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>67</td>
<td>Completed</td>
<td>55</td>
<td>F</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>43</td>
<td>Completed</td>
<td>79</td>
<td>M</td>
</tr>
<tr>
<td>5</td>
<td>A</td>
<td>70</td>
<td>Completed</td>
<td>62</td>
<td>M</td>
</tr>
<tr>
<td>6</td>
<td>B</td>
<td>88</td>
<td>Withdrew</td>
<td>60</td>
<td>F</td>
</tr>
<tr>
<td>7</td>
<td>A</td>
<td>99</td>
<td>Completed</td>
<td>57</td>
<td>F</td>
</tr>
<tr>
<td>8</td>
<td>B</td>
<td>45</td>
<td>Withdrew</td>
<td>66</td>
<td>F</td>
</tr>
<tr>
<td>9</td>
<td>B</td>
<td>90</td>
<td>Completed</td>
<td>59</td>
<td>M</td>
</tr>
<tr>
<td>10</td>
<td>A</td>
<td>23</td>
<td>Withdrew</td>
<td>53</td>
<td>F</td>
</tr>
</tbody>
</table>

- Original participant level data is requested and re-analysed
- Widely regarded as the ‘gold-standard’ approach to meta-analysis
- Many advantages to an IPD approach
IPD approach to meta-analysis

• Advantages: allows a more flexible and complex analysis approach
  o Standardising outcomes or re-defining outcomes
  o Reinstate participants who may have been excluded
  o Reduced publication, reporting and ecological biases
  o Allows detailed checks of any analysis assumptions (e.g. normality or proportional hazards).
  o Modelling heterogeneity (within and between studies)
  o Consideration of covariates and treatment-covariate interactions
  o Modelling of prognostic and diagnostic data in synthesis

IPD approach to meta-analysis

• May be the only option where **summary data is not available**
  
  o Summary statistics needed for meta-analysis are not reported

  o Previous work in oncology has shown that **time-to-event (or survival)** statistics are **poorly and inconsistently reported** in published literature


  o Hazard ratios (HRs) and measures of precision are rarely reported

  o IPD approach particularly **‘gold standard’** for time-to-event outcomes.
IPD approach to meta-analysis

Number of IPD meta-analyses published by year (up to August 2015)

An average of 105 IPD meta-analyses published per year from 2009 to (August) 2015

➢ see Nevitt et al. BMJ 2017;357:j1390
IPD approach to meta-analysis

• IPD approach increasing, but still used in only ~ 5% of published meta-analyses
  • But why? Gold standard approach, so many advantages

• IPD meta-analyses are hard work!
  • Multi-disciplinary teams with statistical expertise required
  • Resource intensive and time consuming
  • Open to ‘availability bias’ where a subset of IPD is not available

• If an AD approach can be shown to be mathematically equivalent* to an IPD approach (and adequate AD is available)
• The resource & time savings would make using AD the approach of choice

Considerations for analysis approach – IPD or AD?

• What is the clinical question or objective of the review?
  ➢ Relative or absolute treatment effect only? ➔ AD approach
  ➢ Interest in patient subgroups or covariates?
  ➢ Variability or heterogeneity likely that should be taken account of?
  ➢ Complex data types (e.g. prognostic, diagnostic?) of interest?
  ➢ Complex statistical modelling approaches needed?

• What are the outcomes of interest?
  ➢ Is required outcome data likely to be reported? (e.g. time to event statistics)?
  ➢ Are outcomes defined consistently?
  ➢ If the answer to both questions is YES ➔ AD approach?
  ➢ If the answer to either question is NO ➔ IPD approach
Illustrative examples: Cochrane Epilepsy Group

Individual participant data network meta-analysis (IPD-NMA) in Epilepsy

Ten Antiepileptic Drugs (AEDs): Monotherapy treatment

Carbamazepine (CBZ), Phenytoin (PHT), Phenobarbitone (PB), Valproate (VPA), Oxcarbazepine (OXC), Lamotrigine (LTG), Gabapentin (GBP), Topiramate (TPM), Levetiracetam (LEV), Zonisamide (ZNS)

45 pairwise comparisons (direct and indirect evidence)
Illustrative examples: Cochrane Epilepsy Group

Individual participant data network meta-analysis (IPD-NMA) in Epilepsy

Pairwise Cochrane IPD reviews of AED monotherapy

Lamotrigine vs Carbamazepine for epilepsy: an individual participant data review

(doi: 10.1002/14651858.CD001031.pub4)
Illustrative examples: Cochrane Epilepsy Group

Individual participant data network meta-analysis (IPD-NMA) in Epilepsy

Pairwise Cochrane IPD reviews of AED monotherapy

Topiramate vs Carbamazepine for epilepsy: an individual participant data review

(doi: 10.1002/14651858.CD012065.pub2)
Individual participant data network meta-analysis (IPD-NMA) in Epilepsy

Pairwise Cochrane IPD reviews of AED monotherapy

Carbamazepine vs Sodium Valproate for epilepsy: an individual participant data review

(doi: 10.1002/14651858.CD001030)
Illustrative examples: Cochrane Epilepsy Group

Individual participant data network meta-analysis (IPD-NMA) in Epilepsy

Pairwise Cochrane IPD reviews of AED monotherapy

Carbamazepine vs Phenobarbitone for epilepsy: an individual participant data review

(doi: 10.1002/14651858.CD001904.pub4)
Illustrative examples: Cochrane Epilepsy Group

Individual participant data network meta-analysis (IPD-NMA) in Epilepsy

Pairwise Cochrane IPD reviews of AED monotherapy

Oxcarbazepine vs Carbamazepine for epilepsy: an individual participant data review

(in progress)
Illustrative examples: Cochrane Epilepsy Group

Individual participant data network meta-analysis (IPD-NMA) in Epilepsy

Pairwise Cochrane IPD reviews of AED monotherapy

*Carbamazepine vs Phenytoin for epilepsy: an individual participant data review*

(doi: [10.1002/14651858.CD001911.pub3](https://doi.org/10.1002/14651858.CD001911.pub3))
Individual participant data network meta-analysis (IPD-NMA) in Epilepsy

Pairwise Cochrane IPD reviews of AED monotherapy

Phenobarbitone vs phenytoin for epilepsy: an individual participant data review

(doi: 10.1002/14651858.CD002217.pub2)
Illustrative examples: Cochrane Epilepsy Group

Individual participant data network meta-analysis (IPD-NMA) in Epilepsy

Pairwise Cochrane IPD reviews of AED monotherapy

Oxcarbazepine vs phenytoin for epilepsy: an individual participant data review

(doi: 10.1002/14651858.CD003615.pub3)
Illustrative examples: Cochrane Epilepsy Group

Individual participant data network meta-analysis (IPD-NMA) in Epilepsy

Pairwise Cochrane IPD reviews of AED monotherapy

Sodium valproate vs phenytoin for epilepsy: an individual participant data review

(doi: 10.1002/14651858.CD001769.pub4)
Cochrane Epilepsy Reviews: Rationale for IPD approach

• What are the outcomes of interest?

• Primary outcome: Time to withdrawal of allocated treatment
  - Secondary outcome: Time to first seizure recurrence after randomisation
  - Secondary outcomes: Time to 6 and 12 month remission of seizures

Would the summary data for an AD approach be reported for these time-to-event outcomes?

• Primary outcome: Time to withdrawal of allocated treatment – definition?
  - Reason for withdrawal: lack of efficacy? Adverse events? Other?
  - Standardised reason for withdrawal: treatment related or not?

If summary data is reported, should it be combined when outcome definitions are inconsistent?
Cochrane Epilepsy Reviews: Rationale for IPD approach

• What is the clinical question or objective of the review?

• Relationship between treatment effect and a covariate (epilepsy type).

  o NICE (UK) recommends different first-line AEDs in monotherapy for:
    ▪ focal seizures (~60% of people with epilepsy) : CBZ and LTG
    ▪ generalised seizures (~40% of people with epilepsy) : VPA

  o Pairwise meta-analyses stratify by epilepsy type
  o NMA to incorporate a treatment-covariate interaction
Cochrane Epilepsy Reviews: Rationale for IPD approach

• What is the clinical question or objective of the review?
  • Interest in treatment-covariate interactions (i.e. the relative treatment effect of AEDs for people with focal and people with generalised seizures).

• What are the outcomes of interest?
  • Complex time to event outcomes
  • Time to event statistics often not reported
  • Definitions are not consistently reported

Conclusion: IPD approach is most feasible for this example
Data requesting for IPD meta-analysis

First IPD meta-analysis (IPD-MA) published in 1987

Data requesting: Early years of IPD meta-analysis (pre-2000)
Data requesting for IPD meta-analysis

First IPD meta-analysis (IPD-MA) published in 1987

Data requesting: IPD meta-analysis post 2000
Data requesting for IPD meta-analysis

**Data requesting:** the present (and the future?)

CSDR: [www.clinicalstudydatarequest.com](http://www.clinicalstudydatarequest.com)

Vivli: [www.vivli.org](http://www.vivli.org)

The YODA project: [www.yoda.yale.edu](http://www.yoda.yale.edu)
Data requesting for IPD meta-analysis

• Data sharing platforms (since 2013)
• Data transparency policies (since 2015)
  o Institute of Medicine (IoM)
  o European Medicines Agency (EMA)
  o International Committee of Medical Journal Editors (ICMJE)

• Surveys of trialists and of patients since 2011 show support of clinical trial data sharing

• Era of data transparency across the research community as a whole
Data requesting: Cochrane Epilepsy IPD reviews


IPD retrieval rates for this project (RCTs comparing two of more of eight AEDs):

1995 – 2005 (approx.): 19 out of 30 trials (63%)
4703 out of 5887 participants (80%)

• Plus IPD from two trials conducted at University of Liverpool (n=2437)

➢ Plans to update as a Cochrane IPD review with newly published evidence, including two newly licenced drugs started between 2010 and 2011

• Search: 45 new trials identified from journal articles / published abstracts
Data requesting: Cochrane Epilepsy IPD reviews (2012-13)

• **45 new trials: data requests started in January 2012**

• **How did we request IPD from these 45 trials?**
  o A short invite letter and a data request form sent to corresponding authors of all journal articles by email and by physical mail
  o No response: other authors with a traceable e-mail address contacted
  o Data of any format accepted
  o No specific deadline for providing data was set.

• **What happened?**
  o Progress was very slow - Who ‘owns’ the data? Who can provide access to data?
    ▪ IPD provided for only two publically funded (academic) trials by the end of 2013
    ▪ Only 3% of total data requested provided
Data requesting: Cochrane Epilepsy IPD reviews (2014-16)

- **January 2014:** Clinical Study Data Request.com (CSDR) launched
  - Initially GlaxoSmithKline and Roche
- Throughout 2014: Other pharmaceutical companies developed ‘data sharing policies’ and ‘data transparency teams’ who could be contacted to request data
  
  - We made requests again for 16 pharmaceutical sponsored trials
    - Requests for six trials made in CSDR
    - Requests for ten trials made directly to sponsors
      - *These sponsors were not part of a data sharing platform at the time, but all of these sponsors now are part of data sharing platform*
  - Interest in data sharing across the whole research community – outstanding requests to 27 publically funded (academic) / government funded trials made again
Data requesting: Cochrane Epilepsy IPD reviews (2014-16)

- How is IPD requested via a data sharing platform?
  - Sponsors provide lists of studies (sponsor ID or NCT number) with IPD available to request
  - Study not listed? Submit an enquiry for the availability of the IPD of that study
    - Feasibility of providing requested IPD checked by trial sponsors
  - Submit a scientific research proposal and details of all research team members
    - At least one member of the team must be a qualified statistician
    - Research proposals must include a publication plan
  - Research proposals go to independent review boards for approval
    - Modifications or clarification may be requested prior to approval
  - A data sharing agreement is signed by both the sponsor and the team requesting data
  - Anonymised data is provided in a remote analysis platform
    - Statistical software available (SAS, R, Stata) within the platform for analysis
    - Results can be exported from the platform, but (mostly) not the data
Data requesting for Cochrane Epilepsy IPD reviews

• **What happened next?**
  
  • Data requesting stopped at the end of December 2016 (~5 years of requesting)
  
  • IPD retrieval rate by the end of 2016
    
    o 15 out of 45 trials *(33%)*
    
    o 7 *publically funded (academic) trials*
    
    o 8 *pharmaceutical funded trials*
    
    o 5251 out of 9637 participants *(55%)*

*1 trial (136 participants) - IPD provided in remote analysis platform (CSDR)*

➢ *Adding this new evidence to the data from the previous IPD-NMA...*
**Individual participant data network meta-analysis (IPD-NMA) in Epilepsy**

**Ten Antiepileptic Drugs (AEDs): Monotherapy treatment**

- Carbamazepine (CBZ)
- Phenytoin (PHT)
- Phenobarbitone (PB)
- Valproate (VPA)
- Oxcarbazepine (OXC)
- Lamotrigine (LTG)
- Gabapentin (GBP)
- Topiramate (TPM)
- Levetiracetam (LEV)
- Zonisamide (ZNS)

**Total IPD for the NMA: 12,391 out of 17,961 eligible participants (69% of total data) from 36 out of the 77 eligible trials (47%)**
IPD synthesis in practice: Preparing IPD for analysis

• IPD for 12,391 participants from 36 trials published over 30 years
  o Different formats, different languages, different methods, different quality
  o Standardised approach to preparing data developed
  o **Step 1**: Check IPD according to published reports and any inconsistencies or uncertainties discussed with providers of data (author or sponsor)

  o **Step 2**: Where no major inconsistencies existed, outcomes calculated:
    ▪ **Primary outcome**: Time to withdrawal of allocated treatment
    ▪ **Secondary outcomes**: Time to 6 month and 12 month remission of seizures
      Time to first seizure post randomisation

• Two statisticians independently double checked and prepared all data
IPD synthesis in practice: Preparing IPD for analysis

• IPD for 12,391 participants from 36 trials published over 30 years

• **Time to prepare each dataset not accurately measured**
  - Not a fair comparison (between statisticians or between trials)
  - More interest in data being correct than prepared quickly

• Observation: data provided from 2014 onwards, particularly industry sponsored
data required fewer checks and had very few uncertainties or inconsistencies
  - Perceived ‘**higher quality**’ of recent pharmaceutical data
  - Preparation of data for analysis was relatively quick
  - Reflective of the extra attention given to these datasets prior to sharing
IPD synthesis in practice: Preparing IPD for analysis

• IPD for 12,391 participants from 36 trials published over 30 years

  o One trial (136 participants), provided in the remote platform (CSDR)
    ▪ Had to be excluded from IPD analyses as data cannot be downloaded
    ▪ Data sharing agreements prevent us uploading other data
    ▪ 1% of overall data, unlikely to have influenced results

  o Other 7 industry sponsored trials, IPD was provided directly (via email or CD)
    ▪ Sponsors not part of any data sharing platforms at the time of data request (2014-2015)
    ▪ Today: Four of these trials would have been provided via CSDR, two via YODA project and one via Vivli – i.e. seven trials across three platforms
    ▪ 4534 participants included in these trials (37% of total data provided)
    ▪ A much larger amount to potentially exclude if data cannot be combined!
Illustrative example: Cochrane IPD-NMA

Individual participant data network meta-analysis (IPD-NMA) in Epilepsy

Results of the IPD-NMA

Results support NICE guidelines that CBZ and LTG are suitable first line treatments for people with focal seizures.

LEV may also be a good alternative for people with focal seizures.

Illustrative example: Cochrane IPD-NMA

Individual participant data network meta-analysis (IPD-NMA) in Epilepsy

Results of the IPD-NMA

Results support NICE guidelines that VPA is most effective for generalised onset seizures.

But teratogenic VPA is not recommended for women of childbearing potential.

LTG and LEV may be good alternatives to VPA for people with generalised onset seizures.

What about the IPD we did not get?

- Remember: Results are based on 69% of eligible data
- What should we do about the 31% missing IPD?
  - Ignore it and interpret results as if we got 100% IPD?
  - 25% of published IPD meta-analyses published to August 2015 do this (see Nevitt et al. BMJ 2017;357:j1390)

- But what about the risk of ‘availability bias’?
  - Available data is not reflective of the entire evidence base
    - What if the trials without IPD are different to those providing IPD?
      - Trial characteristics, trial results?
      - Is any aggregate data available for trials not providing IPD?
    - What if the reasons IPD was not available were informative?
    - What if our results and conclusions are biased?
What about the IPD we did not get?

1) What if the trials without IPD are different to those providing IPD?

• Consider the trial characteristics and results:
  o E.g. Demographics of included patients, sample size, year of publication, source of funding, number of centres, risk of bias, positive or negative results etc.
  
  o Epilepsy IPD-NMA: Trials with IPD available (n=36) vs No IPD available (n=41)
  o Provision of IPD seems to be associated with resources of the trial
  o Large, multi-centre, pharmaceutical trials providing more IPD

• Majority of epilepsy trials did not report the time-to-event outcomes of interest
  o Assessment of positive or negative results difficult
What about the IPD we did not get?

• What if the reasons IPD was not available were informative?
  ➢ Is the unavailability of IPD related to the study results?

• Reasons why IPD was not available (if reasons given):

  1995 – 2005 (approx.) Data lost (academic trials), Data not recorded,
  and 2012- 2016 Unable to contact investigator

  2012-2016 only Data could not be anonymised
  Data sharing not permitted by original ethics or
  patient consent forms
  Costs of providing data prohibitive

Consider: Are these reasons ‘informative’?
What about the IPD we did not get?

- What if the results and conclusions are biased?
  - Explore availability bias visually and statistically
  - Comparison-adjusted network funnel plot
    - *Stata: netfunnel*
    - *Stata: metafunnel*
  - Is publication bias (availability bias) present here?
What about the IPD we did not get?

- What if the results and conclusions are biased?
  - Explore availability bias visually and statistically
  - Comparison-adjusted network funnel plot
    - \textit{Stata: netfunnel}
    - \textit{Stata: metafunnel}
  - Test the gradient of the line
    - Significant p values suggest asymmetry
    - \textbf{No evidence of publication bias (availability bias) here}
What about the IPD we did not get?

• Is any aggregate (summary data) available for trials without IPD?
  • Methods for combining IPD and AD in meta-analysis exist
  • Majority of epilepsy trials did not report the time-to-event statistics for our outcomes of interest
  • Some summary level data available in 10 trials without IPD
    ▪ Concerns over definitions for ‘Time to withdrawal of allocated treatment.’
  • Data (1 trial) in remote analysis platform treated as aggregate data
  • Little aggregate data was available by epilepsy type
    ➢ Additional 2% of data for people with focal seizures
    ➢ Additional 0.7% of data for people with generalised seizures
  • Negligible impact on our primary analysis and conclusions
What about the IPD we did not get?

Is the Cochrane Epilepsy IPD-NMA at risk of availability bias?

- Are the trials without IPD are different to those providing IPD?
  - Trial characteristics, trial results?
    - Provision of IPD may be related to the resources of the trial
    - Difficult to know if the ‘results’ were different
  - Is any aggregate data available for trials not providing IPD?
    - A very limited amount, and combining AD with IPD did not change conclusions

- What if the reasons IPD was not available were informative?
  - Probably not, but provision of IPD may be related to the resources of the trial

- What if the results and conclusions are biased?
  - No evidence of ‘publication’ (availability bias) from inspection of funnel plots

- Conclusion: The 31% missing IPD has probably not impacted on the results
- But - the results should be interpreted with availability bias in mind
Summary of the webinar

• The use of IPD in meta-analysis is increasing
  • Advantages: flexible and complex analysis opportunities with IPD
  • Truly the only option for some clinical questions
• But IPD meta-analysis is really hard work
  • Requesting and preparing data is very time consuming
  • Interpretation of results depend on available data
    ➢ Always keep availability bias in mind
  • Require expertise, not automatically free from bias
    ➢ PRISMA-IPD checklist should improve conduct and reporting
• An aggregate data meta-analysis may answer your question
  ➢ To IPD or not to IPD? That is the (first) question!
References – general

References – Cochrane Epilepsy IPD reviews


Acknowledgements and disclaimer

Many contributors: Prof Catrin Tudur Smith, Prof Anthony Marson, Graham Chan, Rachael Kelly, Dr Maria Sudell, Dr Jennifer Weston, Becky Davie, Sally Reynolds, Lisa Williams, Dr Laura Sutton

Funding: National Institute for Health Research, Liverpool Reviews and Implementation Group, University of Liverpool

The views expressed within this presentation are my own and do not necessarily reflect the views of my employer, the University of Liverpool and the National Institute for Health Research who have previously and currently fund my research time.