Living network meta-analysis

Living Evidence Network “state of the science” webinar

21 Mar 2019

Dr Anna Chaimani
Senior Research Fellow
Research Centre of Epidemiology and Statistics,
Paris Descartes University

Dr Sam Whittle
Senior Consultant Rheumatologist, Queen Elizabeth Hospital
Senior Lecturer, University of Adelaide

Trusted evidence.
Informed decisions.
Better health.
LIVING NETWORK META-ANALYSIS
the next step in evidence synthesis

Anna Chaimani
Research Center of Epidemiology & Statistics, Sorbonne Paris Cité (CRESS-UMR1153), Paris Descartes University

Webinar on living network meta-analyses
21 March 2019

Acknowledgements: Georgia Salanti, Dimitris Mavridis, Philippe Ravaud, Perrine Crequit, Andrea Cipriani
Outline

• Introduction to network meta-analysis (NMA)

• Example of a living NMA (LNMA)

• Challenges in the process of LNMA

• Future work on LNMA
Introduction to network meta-analysis
Why we need network meta-analysis?

“Although Mirtazapine is likely to have a faster onset of action than Sertraline and Paroxetine no significant differences were observed…”

“…meta-analysis highlighted a trend in favour of Sertraline over other Fluoxetine”

“…statistically significant differences in terms of efficacy …. between Fluoxetine and Venlafaxine, but the clinical meaning of these differences is uncertain…”

“Venlafaxine tends to have a favorable trend in response rates compared with duloxetine”
Indirect and mixed comparisons

- Indirect effect
- Direct effect
- Mixed effect
Indirect and mixed comparisons
Full network of interventions

which are the most appropriate treatments to recommend, for which population and under which setting?
Full network of interventions

which are the most appropriate treatments to recommend, for which population and under which setting?
Full network of interventions

which are the most appropriate treatments to recommend, for which population and under which setting?
Full network of interventions

which are the most appropriate treatments to recommend, for which population and under which setting?
Full network of interventions

which are the most appropriate treatments to recommend, for which population and under which setting?
Assumptions underlying network meta-analysis

Single Assumption
underlying indirect and mixed comparison

Conceptual definition

Transitivity

Manifestation in the data

Coherence
Assumptions underlying network meta-analysis

In the outset

The treatments we compare are *in principle* jointly randomizable

They have the same indication, I can imagine a mega-trial with all treatments being compared etc

When you find the studies

The groups of studies that compare them do not differ with respect to the distribution of effect modifiers

You can test this assumption if you have enough studies per comparison

When you extract the outcomes

Direct and indirect treatment effects are in statistical agreement

Various statistical tests
Example of a living network meta-analysis
2nd line treatments of advanced non-small cell lung cancer

Treatments (n=58)

Trials (n =92)

Patients ( n=32 434)

29 Systematic reviews
2nd line treatments of advanced non-small cell lung cancer

• Each year, between 2009 and 2015, the evidence covered by all existing systematic reviews was consistently incomplete
  - 40% to 66% of treatments missing
  - 45% to 70% of trials missing.
  - 30% to 58% of patients missing
Challenges in the process of living network meta-analyses
The concept of living network meta-analysis
Work-load for the NSCLC example

Investing a massive amount of resources to produce a NMA and not maintaining it afterwards does not make sense.
Issues with updating network meta-analyses

• The initial research question might become outdated over time

• Adding new treatments might threaten the validity of the NMA assumptions
  ▪ treatments evaluated only in one trial and connected weakly to the network often introduce heterogeneity and incoherence

• How to exploit any type of information within the NMA framework

• How useful are in the network old treatments evaluated in possibly low-quality studies?
Building a research community

Any individuals interested in a given condition

Identification of new treatments and trials

Group of experts in a given condition (clinicians, trialists and members of cooperative groups)

+ Validation of reported treatments and trials
+ Definition of nodes in the network of trials
+ Screening and selection of records

Group of trained reviewers

+ Manual search of additional sources
+ Contact trialists
+ Identifying multiple reports from the same trial
+ Data extraction
+ Assessment of risk of bias
Building a research community

• Developing a living community for one condition

• A community including systematic reviewers but also clinicians, patients, trialists, methodologists, statisticians and guidelines experts

• Leveraging this community to improve beyond evidence synthesis the whole production of evidence
Future work on LNMAss
Treatments for chronic plaque psoriasis
Methodological work

• Development of formal statistical and non-statistical criteria upon the inclusion and exclusion of treatments at each iteration
  ▪ trade-off between increasing the amount of data and threatening the assumptions of the analysis

• Development of methods allowing to share information across networks and to incorporate information from external evidence
  ▪ networks with sparse data often fail to provide useful and meaningful results

• How new evidence can affect the results
  ▪ can results change?
We need a comprehensive, up-to-date synthesis of evidence for all treatments available for a given disease

For many conditions, multiple competing treatments are available, many of which have been assessed in randomized trials. Clinicians and patients who are making medical decisions need to know which treatments work best among all treatments available for the condition of interest. They increasingly use meta-analyses that synthesize the results of randomized trials to inform the relative efficacy and safety of the different treatments.

But conventional meta-analyses do not provide an exhaustive up-to-date synthesis of all available treatments, and thus prevent from answering easily to the real questions of interest.

We propose to switch:

- from a series of conventional meta-analyses focusing on specific treatments (many treatments being not considered), performed at a given time and frequently out-of-date
- to a single systematic review and evidence synthesis (with meta-analyses and network meta-analyses) covering all treatments and systematically updated when new trial results become available

We call this approach ‘live cumulative network meta-analysis’.
References

- Créquit P, Trinquart L, Yavchitz A, Ravaud P. Wasted research when systematic reviews fail to provide a complete and up-to-date evidence synthesis: the example from lung cancer BMC Med. 2016; 14:8.
- Chaimani A, Caldwell DM, Li T, Higgins JPT, Salanti G. Additional considerations are required when preparing a protocol for a systematic review with multiple interventions. Journal of Clinical Epidemiology 2017; 83:65-74
Planning for a Living Network
Meta-Analysis of Treatments for Rheumatoid Arthritis

Dr. Glen Hazlewood & Dr Samuel Whittle
March 2019
Rationale

- RA is common, and causes pain, disability and reduced quality of life

- Disease-modifying drugs (DMARDs) are effective in reducing symptoms and signs of RA and improve long-term outcomes

- In recent years, novel DMARDs (bDMARDs & tsDMARDs) have joined older drugs (csDMARDs) as potential therapies for individuals with RA

- There are now multiple therapeutic options, with varying effects on important outcomes, and different adverse effect profiles

- New drugs continue to emerge
Rationale

- The array of treatment options makes NMA an attractive option

- A living approach to evidence synthesis is important in RA:
  - High burden of disease
  - Rapidly-evolving evidence base
  - Up-to-date synthesis of all available therapies is important for shared decision-making in this chronic disease

- Opportunity to develop living treatment guidelines
Existing NMA
Living Guidelines Project in Rheumatoid Arthritis

- Collaborative approach to ongoing evidence reviews: CRA, Cochrane, Australia & NZ (ANZMUSC), ACR
- Each group makes their own recommendations
- Opportunity to add other collaborators as project evolves
Globalizing the evidence, localizing the decision

- Global living summary of the evidence
- Country/region-specific contextual factors
  - Patient preferences
  - Equity
  - Health economics
- Treatment recommendations
Planned methods

- RCTs, adults with RA, any DMARD, outcomes under discussion
- **PICO sorting** – 3 groups (MTX-naive, MTX-IR, bDMARD-IR)
- Covidence for PICO annotation, data extraction, RoB assessment
- Data synthesis: Random-effect Bayesian network meta-analysis
- Node-splitting analysis for consistency; meta-regression for heterogeneity
Unscreened records

Machine Learning and Cochrane Crowd

Probable RCTs

Crowdsourcing (controlled)

PICO 1

PICO 2

PICO 3

Data analysis - automated meta-analyses and network meta-analyses
Question 1: It is an RCT?

Question 2: Is the population adults with rheumatoid arthritis?

Question 3: Is the intervention a DMARD?

Question 4: Are corticosteroids/glucocorticoids being used as part of the intervention?

Question 5: Is reduction of treatment the randomized intervention?

Question 6: What type of population is included?
1) DMARD naïve/minimally exposed;
2) DMARD inadequate response;
3) Biologics inadequate response;
4) Mixed populations;
5) Unclear from abstract

RCTs of non-DMARDs (future classification)

PICO sorting
Budesonide inhaled via Turbuhaler: a more effective treatment for asthma than beclomethasone dipropionate via Rotahaler [1995261770]

BACKGROUND: Chlorofluorocarbon-propelled metered dose inhalers are facing a worldwide ban. Dry powder inhalers have been developed for the agents used in treatment of asthma. OBJECTIVE: Our objective was to compare the effects of two inhaled glucocorticosteroids in dry power inhalers: budesonide (delivered via Turbuhaler) and beclomethasone dipropionate (delivered via Rotahaler).

METHODS: A randomized, crossover study with two steroid-treatment periods of 8 weeks. At the end of the study, the treatment with the inhaled steroid was stopped for 4 weeks. Sixteen adult patients with moderately severe asthma participated. Before the study all patients were treated with an inhaled steroid in a median dose of 0.60 mg/day (range 0.15-0.80); during the study they received 0.20 mg twice daily. Peak expiratory flow rate was measured twice daily at home throughout the study, lung function was assessed every fourth week and airway responsiveness was measured before and after each period. Preference concerning efficacy and inhaler type was assessed at the end of the study.

RESULTS: Twelve patients completed the study. Lung function, airway responsiveness, and symptoms deteriorated significantly in the steroid-free washout period; this period had to be shortened in 5/12 patients. Mean morning peak expiratory flow was significantly higher during budesonide treatment than during beclomethasone dipropionate treatment, the difference being 17 L/min (95% C.I.: 2-32 L/min, P = .026). Airway responsiveness improved 1.1 doubling concentrations after budesonide treatment, but decreased 0.3 doubling concentrations after beclomethasone dipropionate treatment. The difference between the values after budesonide and beclomethasone dipropionate treatment was 1.4 doubling concentrations (95% C.I.: 0.4-2.4 doubling concentrations, P = .033). Forced expiratory flow in one second improved slightly more during budesonide than during beclomethasone treatment. The difference was 4.3% predicted (95% C.I.: -0.7-9.3%). Most patients reported budesonide Turbuhaler to be more effective (10 versus 0) and easier to use (11 versus 1) than beclomethasone dipropionate Rotahaler.

CONCLUSIONS: As a consequence of the difference in local potency of the steroids and the fact that Turbuhaler deposits more drug particles in the lung than
Initial Search Results
### Baseline Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Arm 1</th>
<th>Arm 2</th>
<th>Arm 3</th>
<th>Arm 4</th>
<th>Arm 5</th>
<th>Arm 6</th>
<th>Arm 7</th>
<th>Arm 8</th>
<th>Arm 9</th>
<th>Arm 10</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>50.9</td>
<td>52.5</td>
<td>48.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50.0</td>
</tr>
<tr>
<td>Female</td>
<td>0.78</td>
<td>0.84</td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.79</td>
</tr>
<tr>
<td>Disease duration, years (mean/median)</td>
<td>7.9</td>
<td>5.8</td>
<td>6.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.9</td>
</tr>
<tr>
<td>Proportion RF</td>
<td>0.68</td>
<td>0.60</td>
<td>0.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.68</td>
</tr>
<tr>
<td>Proportion prior MTX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion prior DMARD including MTX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of prior DMARDs including MTX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion prior TNF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion prior non-TNF biological</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion taking MTX during study</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dose MTX target</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dose MTX min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dose MTX max</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean MTX dose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Training

● Canadian and Australian early career rheumatologists and trainees
● Initial e-mail expression of interest
● Incentives:
  ○ Learning
  ○ Listed as collaborator/author
  ○ Individual thank-you letter of contribution?
Cochrane classmate

Classmate enables you to use the Cochrane Crowd tasks as learning activities
Automating the analyses

- Planned project for PhD student (Kamso Mujaab)
- Goal: Take the data in covidence and automate the NMA analyses in R, including:
  - Main results
  - All data necessary to inform the quality of evidence ratings (GRADE), including inconsistency testing
  - Formatting into a summary of findings table
Auto-generation of SoF tables

Extracted data

GRADE Summary of findings table

NMA-SoF table example 1

Estimates of effects, credible intervals, and certainty of the evidence for comparisons fluid resuscitation in patients with sepsis
Auto-generation of analyses necessary to inform GRADE quality appraisal

- Rating the quality of evidence requires human judgements, but these can be facilitated by automated data summaries for the relevant GRADE domains, and even algorithms to provide ‘expected’ judgements based on set criteria

**Appendix D – GRADE Quality Appraisal**

**Table D1. Treatment effects for direct, indirect and network meta-analysis evidence and GRADE quality appraisal**

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Comparator</th>
<th>Number of trials for direct comparison</th>
<th>Treatment effect (95% CI)</th>
<th>Quality of Evidence</th>
<th>Treatment effect (95% CI)</th>
<th>P (indirect)</th>
<th>Quality of evidence</th>
<th>Treatment effect (95% CI)</th>
<th>Quality of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTX-native AGRAD response</td>
<td>OR</td>
<td>1</td>
<td>1.33 (1.23 to 2.68)</td>
<td>High</td>
<td>1.34 (1.08 to 1.55)</td>
<td>High</td>
<td>Moderate (study limitation)</td>
<td>1.49 (0.85 to 2.64)</td>
<td>Moderate (study limitation)</td>
</tr>
<tr>
<td>MTX-XRA81 (4)</td>
<td>MTX</td>
<td>1</td>
<td>1.34 (1.23 to 2.68)</td>
<td>High</td>
<td>1.34 (1.08 to 1.55)</td>
<td>High</td>
<td>Moderate (study limitation)</td>
<td>1.49 (0.85 to 2.64)</td>
<td>Moderate (study limitation)</td>
</tr>
<tr>
<td>MTX-NB (3c)</td>
<td>MTX</td>
<td>1</td>
<td>2.13 (1.18 to 2.53)</td>
<td>High</td>
<td>2.32 (1.50 to 3.56)</td>
<td>High</td>
<td>Moderate (study limitation)</td>
<td>3.67 (0.82 to 17.57)</td>
<td>Low (poor precision)</td>
</tr>
<tr>
<td>MTX4-QAR</td>
<td>MTX</td>
<td>0</td>
<td>NA</td>
<td>High</td>
<td>2.32 (1.50 to 3.56)</td>
<td>High</td>
<td>Moderate (study limitation)</td>
<td>3.67 (0.82 to 17.57)</td>
<td>Low (poor precision)</td>
</tr>
</tbody>
</table>
Challenges & Considerations

- How do we create methods for surveillance for new drugs (eg experimental or novel therapies with names that do not appear in our original search strategy)?
- How do we handle the situation in which a new trial might mean that a trial that was previously excluded due to lack of an indirect evidence link, is now eligible due to a common comparator? Is this a matter of classifying exclusions in a particular way so that they can be easily searched again when the network changes, or some other way?
- What is the best method for balancing machine and human tasks once we enter living mode: ie can we use the data from our Crowd tasks to train machine automation tools for ongoing tasks in the future? How would this look in practice?
- What is the best method for screening new citations when in living mode: manual vs automation (if plans to use trained Crowd, how to ensure durability of participation and acquired skills)? Are there certain tasks that are likely to be better suited to humans vs machines?
Challenges & Considerations

- How to ensure that we maintain a durable workforce with sufficient expertise to continue the review in living mode, assuming that it may remain in this mode for years?
- What areas specific to LNMA are amenable to automation?
- Need for *a priori* decisions on when to incorporate new evidence, when to re-publish (eg if change in major outcome or a threshold for new interventions), search frequency & when to take out of living mode; also, the frequency of updating of the review scope and methods (eg search terms, eligibility criteria)
- Where to host our data to ‘future proof’ the current work and prevent the need for duplication of work in the future (eg is Covidence the best data repository?); including clear rules for ‘ownership’ of the data and access as the work evolves
The team

- Cochrane: Jordi Pardo, Peter Tugwell
- Australia/NZ (ANZMUSC): Rachelle Buchbinder, Samuel Whittle
- ACR: Liana Fraenkel, Amy Turner, Elie Akl
- Canada:
  - Pauline Hull, Megan Thomas, Kamso Mujaab
This webinar was presented on behalf of the Living Evidence Network.

To join the Living Evidence Network email: lsr@cochrane.org