Individual participant data meta-analysis. When? Why? How? A scoping review

Michail Belias

May 13, 2019

I used the BMC Medical Research Methodology format

Abstract (200 words)

Background

Individual participant data(IPD) meta-analysis(MA) of randomised clinical trials (RCTs) offers advantages over aggregate MA and single RCTs. IPD-MA is considered the gold standard for evidence based inference especially when personalised treatment decision making is involved. Nevertheless, IPD-MA needs careful modelling and statistical expertise. Although literature over the advantages and disadvantages of the available statistical methods is available, it is unclear to what extent researchers have adopted this guidance. Our objective is to conduct a scoping review of existing IPD-MA methods, and summarise their properties. Furthemore, we aim to inform when and how IPD-MA are performed and whether state-of the art methods have been used.

Methods

We searched MEDLINE and the Cochrane Library **(we may put more databases)** . We included IPD-MA of RCTs published from 01/01/2010 to February 2019.

Results

Out of 4137 results we identified 720 papers **until now** including RCTs only with at least one treament comparison. A short decline in the published IPD-MAs in 2018 has been showed. The two most prodominant medical fields were Cancer *(16%)*, Cardiovascular diseases *(16%)* and Mental health *(10%)*. Effect modification was investigate in most of the articles (85%) though only in a few was that described in their goals. One-stage methods with mixed-effect models are increasing in numbers per year, while subgroup analysis is decreasing. Two-stage methods are almost constant. Although, modelling techinque used is often reported, it is unclear what type specifically was used. *For instance, some report a one-stage mixed effects model has been used to account for within trial clustering, but they don't provide more info such as (stratified or random effects intercept is the interaction term random effects or fixed? etc)*

Conclusions

Introduction

Meta-analysis (MA) is a statistical method that involves combining information from multiple studies. Initially, meta-analyses were limited in aggregated data (AD), until the early 1990s when individual participant data meta-analysis (IPD-MA) was introduced (CHALMERS 1993). In IPD-MA the participant level information is available and evidence from multiple studies can be analysed centrally. IPD-MA is considered the gold standard in evidence synthesis and offers great rewards (Walraven 2010). Besides when pooling an overall treatment effect where AD-MA and IPD-MA are mathematically equal, IPD-MA offers increased power to perform more complicated analyses. This is of great value, especially when the studies included are randomised clinical trials (RCTs). Thereto, RCTs are designed to barely have enough power to detect an overall treatment effect. Nevertheless, the one treatment fits all may not be true. Different patients characteristics may alter the effect of a treatment. These characteristics are often called effect modifiers and their investigation can lead to better clinical decision-making, whether to treat or not a patient. Effect modifier can be either categorical (Smoking (Yes/No), Age group (<30,30-60,60+), ethnicity (Caucasian, Black, Asian)) or continuous (Age, blood pressure, tumour diameter). In the first case the potential effect modifier may be detected either by measuring the treatment effect across its levels (Altman 2003) or by introducing interaction terms in a (generalised) linear model. Apparently, the first choice may be performed only with categorical effect modifiers where both approaches coincide. Nevertheless, that is not true in IPD-MA, where pooling per-subgroup the across trials estimates and then comparing them, we will produce different results than pooling the across trials interaction terms (Fisher et al. 2011). On the other hand, continuous effect modifiers are not without challenges, as modelling the functional shape of their association with the outcome may be needed. One approach is to neglect that and either categorise the continuous effect modifier or make linearity assumptions. Nevertheless, both approaches have been criticized.

A straight forward approach is to include more than one RCTs in

Given that data originate from different sources heterogeneity should be investigated and adjusted for, in order to end up with unbiased results. Therefore, it is essential to account for within study clustering of the participants (Riley, Lambert, and Abo-Zaid 2010). This can be accomplished either by performing a two-stage IPD-MA or by fitting a multi-level (mixed effects) model accounting for the within study clustering of patients through random intercept and/or slope.

In two-stage IPD-MA, a first stage each trial is analysed separately using an appropriate statistical model and thus estimating an effect of interest. Subsequently these effects are pooled into a summary estimate in the second stage of the meta-analysis. Another approach is to perform a one-stage IPD-MA, whilst accounting for the clustering of participants within studies with a combination of random intercepts and/or slopes.

Although guidelines exist on how to investigate effect modification in both categorical and continuous variables, it is unclear to what extent these are followed. Our goal is perform a scoping review on IPD-MAs and report whether 1) effect modification has been investigated 2) which method was chosen (one or two-stage IPD-MA) 3) whether the effect modifier was categorical or continuous and 4) which modelling assumptions were made (Centring per trial, categorisation, linearity assumptions etc). Since IPD-MA is a challenging task we searched for IPD-MAs in the Cochrane IPD-MA methods, as all their IPD-MAs involve at least one statistician member of their IPD. and report if they investigated for effect modification over patient level characteristics. We chose the Cochrane library as Cochrane reviews and meta-analyses are considered on average of high quality. Finally, IPD-MA may be conducted either in one or two stages (Debray et al. 2015). In two-stage IPD-MA, each study is first analysed separately, using an appropriate statistical model. Subsequently, the results extracted in the first stage are pooled into a summary estimate in the second stage of the meta-analysis. In the other hand, one-stage IPD-MA can be conducted with mixed effects model adjusting for within trial clustering.

Methods

We investigated whether effect modification has been investigated and how in meta-analyses with individual participant data available. Since numerous IPD-MAs are conducted worldwide we narrowed our search

into cochrane methods IPD meta-analysis group. Therefore, our sample may not be representative for all IPD-MAs worldwide, but for high quality IPD-MAs.

We searched in the Cochrane library for IPD meta-analyses (https://methods.cochrane.org/ipdma/ipd-meta-analyses). The cochrane library for IPD-MA is divided into 18 medical fields (Cancer, Cardiovascular disease, Child health, Ear nose and throat, Endocrine and metabolic, Eyes and vision, Gastroenterology, Gynaecology, Infectious disease, Lungs and airways, Mental health, Multiple clinical areas, Neonatal care, Neurology, Pregnancy and childbirth, Renal disease, Rheumatology, Wounds). In total 202 studies are reported from 1991 to 2018. Almost half of them (96) were investigating some type of cancer(47.25%). 34 studies were over cardiovascular diseases, 16 over Neurology and 10 over Pregnancy and childbirth. All other medical fields had less than 10. Particularly, 1 was over Child health, 5 Ear nose and throat, 1 Endocrine and metabolic, 2 Eyes and vision,5 Gastroenterology, 5 Gynaecology, 3 Infectious disease, 1 Lungs and airways, 9 Mental health, 1 Multiple clinical areas, 3 Neonatal care, 10 Pregnancy and childbirth, 6 Renal disease, 2 Rheumatology, 2 Wounds. Out of the 201 studies we excluded 11 as they were ongoing and no-results were showed and 26 studies from the Cardiovascular category, as 10 were investigating risk factors and 16 prevention methods.

Cancer related studies were further divided into 14 categories depending on the infected organ, see Figure 1.

```
library(readxl)
library(knitr)
library(ggpubr)
```

```
## Loading required package: ggplot2
```

```
## Registered S3 methods overwritten by 'ggplot2':
##   method         from
##   [.quosures     rlang
##   c.quosures     rlang
##   print.quosures rlang
```

```
## Loading required package: magrittr
```

```
IPD_MA_Cochrane1 <- read_excel("IPD-MA Cochrane papers/IPD-MA_Cochrane.xlsx",  sheet = "Meta-analysis s
IPD_MA_Cochrane1 <- as.data.frame(IPD_MA_Cochrane1)
IPD_MA_Cochrane1 <- as.data.frame(IPD_MA_Cochrane1[IPD_MA_Cochrane1$Remarks == "RCTs",])
IPD_MA_Cochrane2 <- read_excel("IPD-MA Cochrane papers/IPD-MA_Cochrane.xlsx",  sheet = "Medical fields"
kable(IPD_MA_Cochrane2[,3:4], caption = "Figure 1")
```

Table 1: Figure 1

All medical fields	Number of all medical fields studies
Cancer	96
Cardiovascular disease	34
Child health	1
Ear, nose and throat	5
Endocrine and metabolic	1
Eyes and vision	2
Gastroenterology	5
Gynaecology	5
Infectious disease	3
Lungs and airways	1

All medical fields	Number of all medical fields studies
Mental health	9
Multiple clinical areas	1
Neonatal care	3
Neurology	16
Pregnancy and child birth	10
Renal disease	6
Rheumatology	2
Wounds	2
Total	202

```
kable(IPD_MA_Cochrane2[,1:2], caption = "Figure 2")
```

Table 2: Figure 2

Cancer related studies	Number of Cancer Studies
Bladder	8
Breast	6
Childhood cancers	5
Colorectal	10
Generic cancer care	3
Gynaecological	8
Haematological malignancies	13
Head & neck	11
Lung	23
Neurological	2
Oesophagus	2
Prostate	2
Soft tissue sarcoma	2
Stomach cancer	1
Total	96
NA	NA
NA	NA
NA	NA
NA	NA

```
if(!require("DiagrammeR")) install.packages("DiagrammeR")
```

```
## Loading required package: DiagrammeR
```

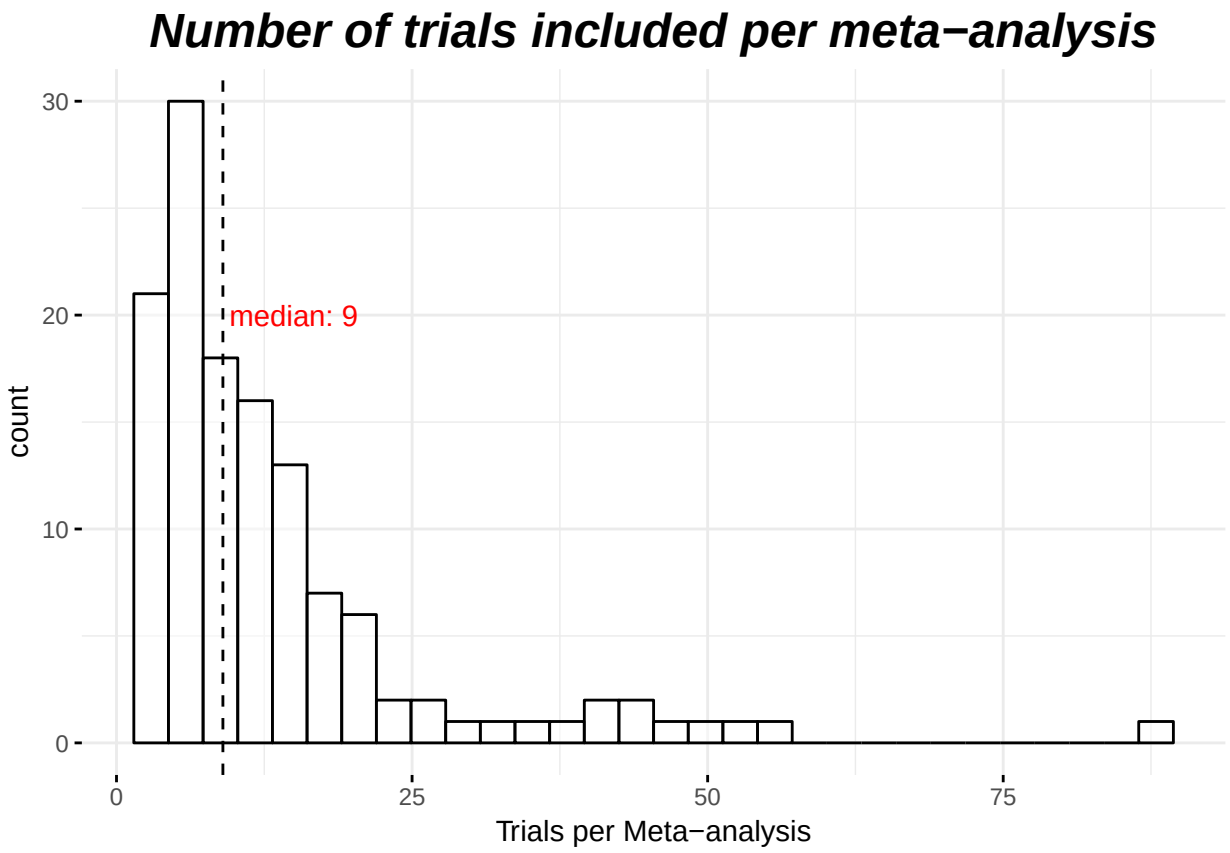
```
grViz("digraph flowchart {
      # node definitions with substituted label text
      node [fontname = Helvetica, shape = rectangle]
      tab1 [label = '@@1']
      tab2 [label = '@@2']
      tab3 [label = '@@3']
      tab4 [label = '@@4']
      tab5 [label = '@@5']
      tab6 [label = '@@6']
      tab7 [label = '@@7']
```

```
      tab8 [label = '@@8']

      # edge definitions with the node IDs
      tab1 -> tab2
      tab1 -> tab3
      tab3 -> tab4
      tab3-> tab5
      tab3-> tab6
      tab3-> tab7
      tab3-> tab8 ;
      }

      [1]: 'Studies identified through searching in cochrane IPD-MA methods group [202]'
      [2]: '37 Excluded studies'
      [3]: 'Studies included [165]'
      [4]: '95 Cancer'
      [5]: '16 Neurological'
      [6]: '7 Cardiovascular'
      [7]: '6 Cardiovascular'
      [8]: '41 Others'
      ")
```

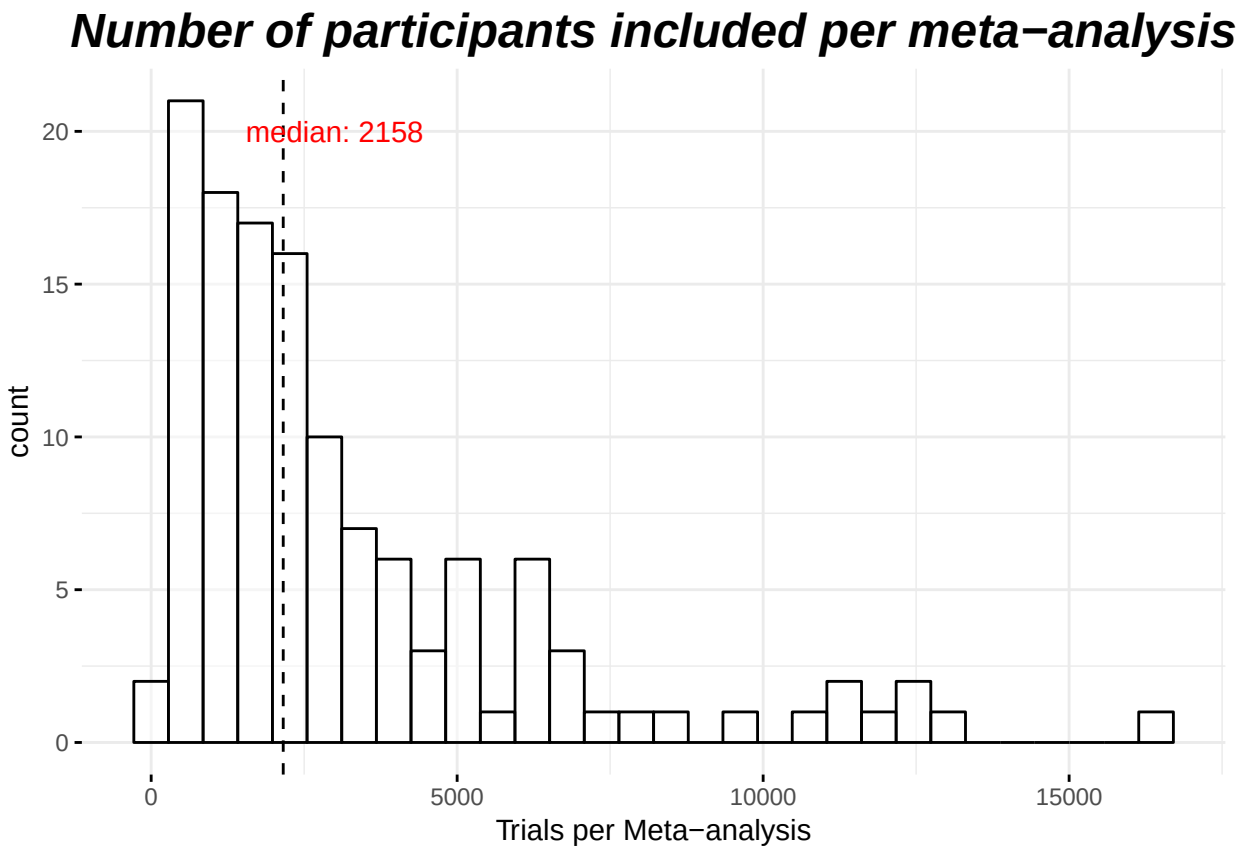
```
## TypeError: Attempting to change the setter of an unconfigurable property.
## TypeError: Attempting to change the setter of an unconfigurable property.
```

```
IPD_MA_Cochrane1[,1:3] =  apply(IPD_MA_Cochrane1[,1:3], 2, as.numeric)
```

```
gghistogram(data = IPD_MA_Cochrane1, x =  "Studies", y = "..count..", bins = 30, ggtheme = theme_minimal
```




```
gghistogram(data = IPD_MA_Cochrane1, x =  "Participants",  y = "..count..", bins = 30, ggtheme = theme_m
```



Altman, D. G. 2003. "Statistics Notes: Interaction Revisited: The Difference Between Two Estimates." *BMJ* 326 (7382): 219–19. https://doi.org/10.1136/bmj.326.7382.219.

CHALMERS, IAIN. 1993. "The Cochrane Collaboration: Preparing, Maintaining, and Disseminating Systematic Reviews of the Effects of Health Care." *Annals of the New York Academy of Sciences* 703 (1 Doing More Go): 156–65. https://doi.org/10.1111/j.1749-6632.1993.tb26345.x.

Debray, Thomas P. A., Karel G. M. Moons, Gert van Valkenhoef, Orestis Efthimiou, Noemi Hummel, Rolf H. H. Groenwold, and Johannes B. Reitsma and. 2015. "Get Real in Individual Participant Data (IPD) Meta-Analysis: A Review of the Methodology." *Research Synthesis Methods* 6 (4): 293–309. https://doi.org/10.1002/jrsm.1160.

Fisher, D. J., A. J. Copas, J. F. Tierney, and M. K.B. Parmar. 2011. "A Critical Review of Methods for the Assessment of Patient-Level Interactions in Individual Participant Data Meta-Analysis of Randomized Trials, and Guidance for Practitioners." *Journal of Clinical Epidemiology* 64 (9): 949–67. https://doi.org/10.1016/j.jclinepi.2010.11.016.

Riley, R. D., P. C. Lambert, and G. Abo-Zaid. 2010. "Meta-Analysis of Individual Participant Data: Rationale, Conduct, and Reporting." *BMJ* 340 (feb05 1): c221–c221. https://doi.org/10.1136/bmj.c221.

Walraven, Carl van. 2010. "Individual Patient Meta-Analysisrewards and Challenges." *Journal of Clinical Epidemiology* 63 (3): 235–37. https://doi.org/10.1016/j.jclinepi.2009.04.001.