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) is considered the gold standard for evidence based inference. It is well established that IPD-MA offers great advantages over both aggregate MA and single RCTs. Nevertheless, it is unclear which advantages are mostly addressed when IPD-MAs are conducted **(for instance subgroup analysis, overall treatment effect pooling, modeling etc)**. Furthermore, it is unclear which statistical approaches are preferred, how they are the results presented, which medical fields are involved and to what extent guidelines are flowed….

## Objective

Our objective is to conduct a scoping review of existing IPD-MA methods, and summarise their properties. Furthermore, we aim to inform when and how IPD-MA are performed, whether state-of the art methods are used and if they are clearly described.

*We can propose for a meta-data ID*

## Methods

We searched MEDLINE and the Cochrane Library for studies since 01/01/2010 to 01/04/2019 using the “individual participant data meta-analysis” OR “IPD” OR “individual participants data meta-analysis” query **(we may put more databases)**.

## Results

Our search resulted in 4137 articles. We excluded 2407 IPD-MAs of non-RCTs. We included only IPD-MAs with at least one treatment comparison. A short decline in the published IPD-MAs in 2018 has been showed. The two most predominant medical fields were Cancer *(16%)*, Cardiovascular diseases *(16%)* and Mental health *(10%)*. Most of the IPD-MAs had as a primary goal to pool an overall treatment effect and only few to investigate for subgroups. An increasing trend in one-stage multi-level mixed effects frequency by year has been showed. Nevertheless, more information should be provided in both the abstract and the article over the statistical approaches followed. *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

Description of the goal and statistical approach performed is still unclear. Mixed effects approaches are increasing by year. Subgroups analysis is not the primary goal of IPD-MAs .

# Figures and tables

library(readxl)  
library(ggpubr)

## Loading required package: ggplot2

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

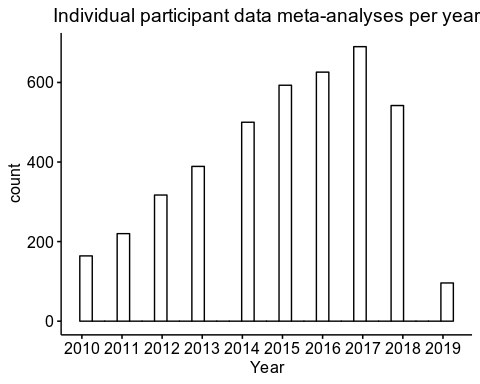
## Loading required package: magrittr

IPD\_MA = read\_xlsx(path = "IPD-MA Cochrane papers/IPD-MAs in General.xlsx", sheet = "Pub med IPD-MA articles")

## New names:  
## \* `` -> ...26  
## \* `` -> ...28  
## \* `` -> ...29

gghistogram(data = IPD\_MA, y = "..count..",x = "Year", breaks = 10, title = "Individual participant data meta-analyses per year") + scale\_x\_continuous(breaks = seq(2010, 2019,by = 1)) + theme(plot.title = element\_text(hjust = 0.5))

## Warning: Using `bins = 30` by default. Pick better value with the argument  
## `bins`.



# Introduction

Meta-analysis (MA) is a statistical method that involves combining information from multiple studies. While initially, meta-analyses were limited in aggregated data (AD) in the early 1990s 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 opportunities (Walraven 2010). Besides when pooling an overall treatment effect where AD-MA and IPD-MA are mathematically equal, IPD-MA offers (1) the possibility to standardize subgroup definitions and outcomes across studies, (2) higher validity and credibility of subgroup findings, (3) increased flexibility to search for subgroups based on combinations of patient and/or disease characteristics (4) the possibility to avoid ecological BIAS and (5) model more complicated than linear functional forms.

*Subgroup Analysis* For the opportunities mentioned above guidance on how to perform an IPD-MA exist. For instance, when the research goal is to detect effect modification several statistical approaches are used. Simmonds et al. (Simmonds, Stewart, and Stewart 2015) and Fisher et al. (Fisher et al. 2011) classified the available statistical approaches into one and two stages. One-stage IPD-MA involves a mixed-effects model, where all available data are analysed simultaneously, whilst taking into account the within trial clustering of the patients. On the other hand two-stage approaches first estimate either a main treatment effect, or the different effects observed per subgroup, or the treatment-covariate interaction effect. Subsequently, these effects are either modelled over the levels of the potential effect modifier (meta-regression) or pooled using standard meta-analysis leadind to per-subgroup meta-analysis (PS-MA) and meta-analysis of interaction terms (MA-IT). Simmnond and Higgins (Simmonds and Higgins 2007) mathematically proved that one-stage IPD-MA are always more powerful than meta-analysis of interaction terms and meta-regression. Nevertheless, the assumptions made were unrealistic. Fisher et al. (Fisher et al. 2011) critically reviewed one-stage IPD-MA of them on how to detect patient level effect modification in IPD-MA. They concluded that one-stage or mixed-effects modelling allows for more complex analysis, but is more difficult to perform than pooling within-trial interaction terms. Subsequently, Hua et al. (Hua et al. 2016) suggested that these mixed effects models should also centre the effect modifiers to their mean, in order to separate across and within trial information. Finally, Legha et al. (Legha et al., n.d.)

*Modeling functional forms* On the other hand when the goal of our research is to model a non-linear functional relationship between a continuous covariate and the outcome Sauerbrei and Royston (Sauerbrei and Royston 2011) suggested to use a two stage approach involving fitting per trial a fractional polynomial and pooling their estimates through a point-wise weighted meta-analytical process. Subsequently they extended these non-linear associations to include interactions (Royston and Sauerbrei 2013).

Finally, when reporting results (Stewart et al. 2015)

Wallach et al. (Wallach et al. 2017) performed a

Nevertheless, it is still unclear which are the goals for IPD-MA, how are they described and how are the statistical analyses performed.

# Methods

library(readxl)  
library(knitr)  
library(ggpubr)  
IPD\_MA\_Cochrane1 <- read\_excel("IPD-MA Cochrane papers/IPD-MA\_Cochrane.xlsx", sheet = "Meta-analysis size")  
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")

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 |
| 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")

Figure 2

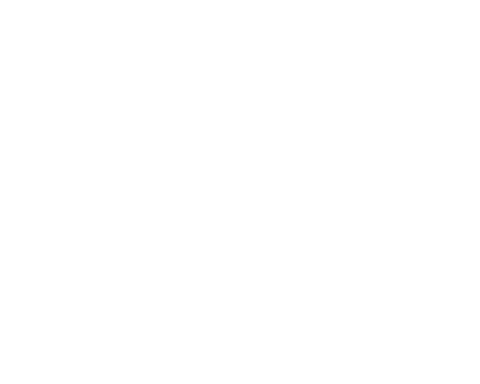
|  |  |
| --- | --- |
| Cancer related studies | Number of Cancer Studies |
| Bladder | 8 |
| Breast | 6 |
| Childhood cancers | 5 |
| Colorectal | 10 |
| Generic cancer care | 3 |
| Gynaecological | 8 |
| Haematological malignancies | 13 |
| Head & neck | 11 |
| Lung | 23 |
| Neurological | 2 |
| Oesophagus | 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")

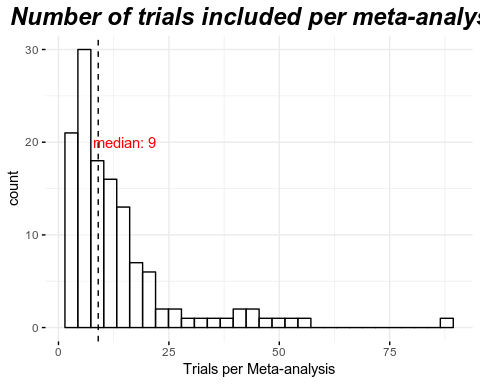
## Loading required package: DiagrammeR

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'  
 ")

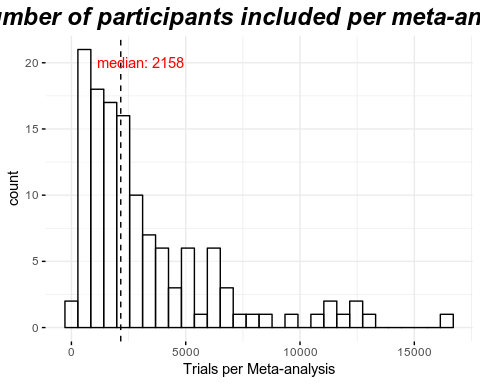
## TypeError: Attempting to change the setter of an unconfigurable property.  
## TypeError: Attempting to change the setter of an unconfigurable property.



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(), title = "Number of trials included per meta-analysis", add = "median", xlab = "Trials per Meta-analysis" ) + annotate(geom="text", x=15, y=20, label= paste("median:", median(IPD\_MA\_Cochrane1$Studies)), color="red") + theme(plot.title = element\_text(hjust = 0.5,size = 18, face = "bold.italic"))



gghistogram(data = IPD\_MA\_Cochrane1, x = "Participants", y = "..count..", bins = 30, ggtheme = theme\_minimal(), title = "Number of participants included per meta-analysis", add = "median", xlab = "Trials per Meta-analysis" ) + annotate(geom="text", x=3000, y=20, label= paste("median:", median(IPD\_MA\_Cochrane1$Participants, na.rm = T)), color="red") + theme(plot.title = element\_text(hjust = 0.5,size = 18, face = "bold.italic"))



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>.

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>.

Hua, Hairui, Danielle L. Burke, Michael J. Crowther, Joie Ensor, Catrin Tudur Smith, and Richard D. Riley. 2016. “One-Stage Individual Participant Data Meta-Analysis Models: Estimation of Treatment-Covariate Interactions Must Avoid Ecological Bias by Separating Out Within-Trial and Across-Trial Information.” *Statistics in Medicine* 36 (5): 772–89. <https://doi.org/10.1002/sim.7171>.

Legha, Amardeep, Richard D. Riley, Joie Ensor, Kym I. E. Snell, Tim P. Morris, and Danielle L. Burke. n.d. “Individual Participant Data Meta-Analysis of Continuous Outcomes: A Comparison of Approaches for Specifying and Estimating One-Stage Models.” *Statistics in Medicine*. <https://doi.org/10.1002/sim.7930>.

Royston, Patrick, and Willi Sauerbrei. 2013. “Interaction of Treatment with a Continuous Variable: Simulation Study of Significance Level for Several Methods of Analysis.” *Statistics in Medicine* 32 (22): 3788–3803. <https://doi.org/10.1002/sim.5813>.

Sauerbrei, Willi, and Patrick Royston. 2011. “A New Strategy for Meta-Analysis of Continuous Covariates in Observational Studies.” *Statistics in Medicine* 30 (28): 3341–60. <https://doi.org/10.1002/sim.4333>.

Simmonds, Mark, Gavin Stewart, and Lesley Stewart. 2015. “A Decade of Individual Participant Data Meta-Analyses: A Review of Current Practice.” *Contemporary Clinical Trials* 45 (November): 76–83. <https://doi.org/10.1016/j.cct.2015.06.012>.

Simmonds, M. C., and J. P. T. Higgins. 2007. “Covariate Heterogeneity in Meta-Analysis: Criteria for Deciding Between Meta-Regression and Individual Patient Data.” *Statistics in Medicine* 26 (15): 2982–99. <https://doi.org/10.1002/sim.2768>.

Stewart, Lesley A., Mike Clarke, Maroeska Rovers, Richard D. Riley, Mark Simmonds, Gavin Stewart, and Jayne F. Tierney. 2015. “Preferred Reporting Items for a Systematic Review and Meta-Analysis of Individual Participant Data.” *JAMA* 313 (16): 1657. <https://doi.org/10.1001/jama.2015.3656>.

Wallach, Joshua D., Patrick G. Sullivan, John F. Trepanowski, Kristin L. Sainani, Ewout W. Steyerberg, and John P. A. Ioannidis. 2017. “Evaluation of Evidence of Statistical Support and Corroboration of Subgroup Claims in Randomized Clinical Trials.” *JAMA Internal Medicine* 177 (4): 554. <https://doi.org/10.1001/jamainternmed.2016.9125>.

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>.