A new approach to investigate treatment effect modification over continuous co-variables in IPD-MA by using smoothing splines

Michail Belias

12 June, 2019

Abstract Statistics in Medicine)

Background

Individual participant data(IPD) meta-analysis(MA) of randomised clinical trials are considered the golden standard to investigate effect modification. Nevertheless, detecting and investigating treatment-effect modification can be lead to evidence-based personalised treatment. Treatment modified by continuous variables may be challenging to investigate when non-linear associations are present.

Objective

We propose a new approach to detect treatment-effect modification, when non-linear association are present.

Methods

We apply mixed effects models with smoothing splines.

Results

Conclusion

1. Introduction

Individual participant data meta-analysis (IPD-MA) is a type of statistical analysis, where data are gathered from multiple studies are combined and analysed. The possibility to standardise subgroup definitions and outcomes across studies, the increased power to investigate other than linear associations, the increased validity and reliability of the subgroups and the flexibility to search for subgroups based on combinations of patient and/or disease characteristics are some of the benefits of using IPD of multiple trials rather than traditional (aggregate) meta-analysis [1–3]. A vivid field of research towards personalised healthcare is the investigation of effect modification. For this, IPD-MA is considered the gold standard as single trials rarely have sufficient power to properly detect effect modification.

Effect modification may be present in both categorical and/or continuous covariates. For instance, differences in the treatment effect may be present between smokers and non-smokers, or across the age of the patients. If subgroups have been predefined, hypothesis testing may be performed using traditional regression-based methods. For single trials, this typically involves the estimation of (linear) interaction terms between treatment and the modifier of interest. In an IPD-MA, these interaction effects may be estimated either separately within studies and subsequently be pooled across studies (two-stage IPDMA), or directly across all studies (one-stage IPDMA. However, effect modification across a continuous variable is more challenging, since we may need to evaluate effect modification and at the same time want to define a threshold from which point the treatment effect is relevantly different. This way the continuous variable is categorised into subgroups, where the clinician can make decision whether to treat or not treat. Reversely, the same notion is used to analyse data. Categorization is a common technique to investigate effect modification, by splitting the continuous covariate into subgroups. Nevertheless, these subgroups should always be created based on good prior knowledge from literature. If so, this approach can be meaningful. In all other cases, categorization has been criticised for misspecification, loss of information and power, inflation of the type I error rate and even biased results [4–8]. Another common practice is using the continuous variable as it is, and assume linearity on the linear predictor scale . This approach may also lead to deterioration of power, misspecification, and even spurious results if the true relationship is not linear [9]. Both categorisation and false functional form assumptions are prone to significant ecological bias. For instance, if the functional form of mortality and age is exponential and some trials have old participants while other young, both approaches can lead to biased pooled results. Ideally, when continuous covariates are included, we would like to account for their functional form, while simultaneously making inferences over the presence of the effect modification and avoiding ecological. Furthermore, although the association between the outcome and the continuous effect modifier is highly informative, clinical decisions are based on subgroups of participants the differ in treatment response. Finally, subgroups generated from continuous variables are defined by the cut-points were the treatment effect is considered to change. These cut-points may be based on the treatment effect function [10], i.e. the difference between the two treatments over the range of the covariable or the treatment-effect modifier interaction terms [11]. For this, various approaches to account for non-linear associations have been developed, such as splines and fractional polynomials (FP) [12].

For IPD-MA, regression-based approaches such as linear models, piecewise polynomials, FPs and smoothing splines may be performed either in one or two stages. In a two-stage approach, each trial is first modelled separately using an appropriate statistical model. Subsequently, we pool either the extracted coefficients if shared across the trials or their fitted functions, using standard meta-analytical tools. In contrast, in one-stage IPD-MA the IPD from all trials are analysed simultaneously whilst accounting for the clustering of participants within studies . Hereto, we model interactions between treatment and patient-level variables while accounting also for the shape of the associations with the outcome. Recent recommendations suggest mean-centring the potential effect modifiers per trial in order to account for potential ecological bias due to unadjusted confounding. In such a one-stage model, within-trial clustering can be accounted for using either a fixed effect (common intercept/slope), fixed effects (stratified intercept/slope), or random effects [13]. Other methods to explore effect modification are plot- and tree-based methods such as the generalised

linear mixed-effects model tree (GLMM-tree) method [14] or meta-stepp, a moving average (sliding window) method.

Although there is a large variety of methods to explore effect modification for continuous covariates, little guidance exists on their use. We aim to describe and illustrate the aforementioned methods by applying them on two empirical examples, while discussing their (potential) advantages and limitations.

2. Empirical examples

WWe use 2 IPD-sets to illustrate aforementioned methods. The first empirical example [15] considers an IPD-MA where the effect of antibiotics in acute otitis media was investigated in children. Rovers et al. collected IPD from 6 randomised clinical trials with a total of 1643 children, aged from 0-12 years old. The primary outcome was fever and/or ear-pain after 3-7 days (yes/no). They concluded that antibiotics were more beneficial in younger children (less than 2 years old) with bilateral acute otitis media. Bilateral acute otitis media (yes/no), age, otorrhea were investigated also separately for potential effect modification and only bilateral acute otitis media showed a significant result. The second empirical example [16] considers an IPD-MA to investigate the effect of Somatostatin on liver volume reduction. Gevers et al. collected IPD from 3 randomised placebo-controlled trials with a total of 107 participants. In this example, the outcome was continuous (liver volume reduction), and age, sex, baseline liver volume, and diagnosis of either autosomal dominant polycystic liver or kidney disease were investigated for effect modification . They concluded that use of Somatostatin was more beneficial for younger (<47) female patients. One of the 3 trials had a cross-over design, therefore participants were treated both with the active and the control treatment in different time periods. In order to use these data for our illustrative purposes, we removed the cross-over design and used all patients only once, by selecting half of the patients from the active period and the other half (sex and age-matched) from the control period. Therefore, differences between our results and those reported in the original article may occur.

Table 1: Baseline Characteristics (Gevers et al.)

	van Keimpema et al		**Hogan et al**		**Caroli et al**	
	Placebo	**Somatostatin**	**Placebo**	**Somatostatin**	**Placebo**	**Somatostatin**
Number of participants	27	27	14	28	6	6
Age, median [range]	48 [36,68]	49 [32,65]	49 [38,65]	47 [34,69]	39 [29,53]	40 [30,59]
Female	23 (85%)	24 (89%)	13 (93%)	23 (82%)	2 (33%)	1 (17%)
Male	4 (15%)	3 (11%)	1 (7%)	5 (18%)	4 (67%)	5 (83%)
Log-scaled Liver volume difference, median [range]	0.011 [-0.038,0.079]	-0.018 [-0.19,0.043]	-0.006 [-0.109,0.166]	-0.034 [-0.172,0.092]	-0.006 [-0.096,0.158]	-0.05 [-0.09,0.031]

Table 2: Baseline Characteristics (Rovers et al.)

	Damoiseaux		**Burke**		**Appelman**		**Little**		**Saux**		**McCormick**	
	Placebo	**Antibiotics**	**Placebo**	**Antibiotics**	**Placebo**	**Antibiotics**	**Placebo**	**Antibiotics**	**Placebo**	**Antibiotics**	**Placebo**	**Antibiotics**
Number of participants	123	117	118	114	54	67	164	151	254	258	111	112
Age, median [range]	1 [1,2]	1 [0,2]	5 [3,9]	5 [3,9]	4.16 [1.07,10.22]	3.83 [0.65,10.19]	4.74 [0,10.87]	5.03 [0.48,11.1]	2.71 [0.55,5.97]	2.86 [0.5,6.01]	1.82 [0.51,12.66]	1.46 [0.5,12.65]
Males (%)	66 (54%)	64 (55%)	50 (42%)	59 (52%)	25 (46%)	31 (46%)	81 (49%)	74 (49%)	131 (52%)	129 (50%)	58 (52%)	54 (48%)
Bilateral AOM (%)	76 (62%)	75 (64%)	16 (14%)	16 (14%)	6 (11%)	14 (21%)	NA (NA%)	NA (NA%)	74 (29%)	79 (31%)	48 (43%)	52 (46%)
Otorrhoea (%)	19 (15%)	16 (14%)	NA (NA%)	NA (NA%)	NA (NA%)	NA (NA%)	46 (28%)	35 (23%)	NA (NA%)	NA (NA%)	NA (NA%)	NA (NA%)

[1] Debray TPA, Moons KGM, Valkenhoef G van, Efthimiou O, Hummel N, Groenwold RHH, et al. Get real in individual participant data (IPD) meta-analysis: A review of the methodology. Research Synthesis Methods 2015;6:293–309. doi:10.1002/jrsm.1160.

[2] Rovers M, Reitsma JB. [The meta-analysis of data from individual patients]. Nederlands Tijdschrift Voor Geneeskunde 2012;156:A4743.

[3] Tierney JF, Vale C, Riley R, Smith CT, Stewart L, Clarke M, et al. Individual participant data (IPD) meta-analyses of randomised controlled trials: Guidance on their use. PLOS Medicine 2015;12:e1001855. doi:10.1371/journal.pmed.1001855.

[4] Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: A bad idea. Statistics in Medicine 2005;25:127–41. doi:10.1002/sim.2331.

[5] Altman DG. The cost of dichotomising continuous variables. BMJ 2006;332:1080–0. doi:10.1136/bmj.332.7549.1080.

[6] Austin PC, Brunner LJ. Inflation of the type i error rate when a continuous confounding variable is categorized in logistic regression analyses. Statistics in Medicine 2004;23:1159–78. doi:10.1002/sim.1687.

[7] Maxwell SE, Delaney HD. Bivariate median splits and spurious statistical significance. Psychological Bulletin 1993;113:181–90. doi:10.1037/0033-2909.113.1.181.

[8] Weinberg C. How bad is categorization? Epidemiology 1995;6:345–6. doi:10.1097/00001648-199507000-00002.

[9] Jørgensen TSH, Osler M, Ängquist LH, Zimmermann E, Christensen GT, Sørensen TIA. The u-shaped association of body mass index with mortality: Influence of the traits height, intelligence, and education. Obesity 2016;24:2240–7. doi:10.1002/oby.21615.

[10] Royston P, Sauerbrei W. Interaction of treatment with a continuous variable: Simulation study of significance level for several methods of analysis. Statistics in Medicine 2013;32:3788–803. doi:10.1002/sim.5813.

[11] Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. BMJ 2010;340:c117–7. doi:10.1136/bmj.c117.

[12] Sauerbrei W, Royston P. A new strategy for meta-analysis of continuous covariates in observational studies. Statistics in Medicine 2011;30:3341–60. doi:10.1002/sim.4333.

[13] Legha A, Riley RD, Ensor J, Snell KIE, Morris TP, Burke DL. Individual participant data meta-analysis of continuous outcomes: A comparison of approaches for specifying and estimating one-stage models. Statistics in Medicine n.d. doi:10.1002/sim.7930.

[14] Wang XV, Cole B, Bonetti M, Gelber RD. Meta-STEPP: Subpopulation treatment effect pattern plot for individual patient data meta-analysis. Statistics in Medicine 2016;35:3704–16. doi:10.1002/sim.6958.

[15] Rovers MM, Glasziou P, Appelman CL, Burke P, McCormick DP, Damoiseaux RA, et al. Antibiotics for acute otitis media: A meta-analysis with individual patient data. The Lancet 2006;368:1429–35. doi:10.1016/s0140-6736(06)69606-2.

[16] Gevers TJG, Inthout J, Caroli A, Ruggenenti P, Hogan MC, Torres VE, et al. Young women with polycystic liver disease respond best to somatostatin analogues: A pooled analysis of individual patient data. Gastroenterology 2013;145:357–365.e2. doi:10.1053/j.gastro.2013.04.055.