Meta-analysis of non-linear effect modification using individual participant data using generalised additive mixed effects models: A new approach

Michail Belias

03 July, 2019

Background

Individual participant data(IPD) meta-analysis(MA) is considered the gold standard since a variety of opportunities are offered. The investigation of treatment-effect modification is one of them, Nevertheless, effect modification over a continuous co-variables may be challenging, as non-linear interactions may be present. Most methods either ignore non-linear effect modification, or use a forward technique to model non-linearities that relies on statistical tests with arbitrary significance levels.

Objective

We propose a new approach to model and investigate treatment-effect modification, while modelling non-linear associations using smoothing splines.

Methods

Results

Conclusion

1. Introduction

The effect of a treatment may differ depending on patient characteristics. One of the main goals of an individual participant data (IPD) meta-analysis (MA) is to investigate whether treatment effect differences are present, and how they are associated with these characteristics [1]. Patient characteristics can be represented by either categorical or continuous variables. Especially the latter, treatment effect modification over a continuous variable, may be challenging to investigate because the association between the outcome and the effect modifier and/or the interaction may not be known.

Often, regression based approaches are used in combination with several assumptions and strategies to deal with continuous effect modifiers. A naive approach is to ignore possible non-linearity either through categorization of the continuous variable, or using the continuous variable as it is in a linear regression model. Categorization involves splitting the continuous co-variable into subgroups based on clinical reasoning. For instance, the risk of developing ovarian cancer may be associated with menopause. Therefore, if age is investigated as a potential effect modifier it is reasonable to categorize the age to younger (≤ 50 years older) and older (> 50 years old) participants. Due to loss of information, categorization has been criticized for misspecification, reduced power, inflation of the type I error rates and biased results [2–6]. When clinical knowledge to define the subgroups is not available, tree-based approaches have been proposed to estimate these subgroups [7–12]. However, tree-based approaches rely on statistical tests with arbitrary levels of significance, are data sensitive and fit well only when the underlying functional form is a step function.

Another naive approach is to include the effect modifier as it is and fit a linear model, without adjusting for non-linearity. Nevertheless, if the underlying association in not linear the results may be biased. For instance, if we assume linearity between BMI and mortality while the underlying shape is quadratic (U-shape) we may draw the conclusion that they are not associated at all. Additionally, any predictions made by a misspecified linear model will be biased and increasing the sample size will only make things worse.

Often the association between potential effect modifiers and the outcome has already been investigated and a non-linear baseline functional form may have been already known. For instance, in an IPD-MA of 44 cohort studies Liu et al. showed a J-shaped association between BMI and risk of stroke [13]. Researchers investigating thrombolysis treatment and potential effect modification by BMI may include this finding as an a-priori knowledge and introduce exponential or quadratic terms in their regression model. However, the association between BMI and risk of stroke may have a different shape for the treated group than for the control group. Misspecifying this functional shape may lead to the same problems as in the naive linear approach mentioned above.

The aforementioned approaches were either ignoring or assuming the functional shape of the associations known. Another approach is to estimate the functional shape from the data and investigate effect modification. One commonly applied strategy is trail and error. Thereto, researchers fit regression models including various transformations of the effect modifier, such as polynomial, trigonometric (sine, cosine), exponential and logarithmic. Then, these models are compared with each other using statistical tests such as Wald and likelihood ratio tests or criteria such as AIC and BIC. The trial and error procedure may be time-consuming and automated approaches are also available. Royston and Altman [14] proposed multi-variable fractional polynomial procedure (MFP) in order to detect the best fitting fractional polynomial. Nevertheless, their approach was initially limited to single studies and didn't include interactions. Therefore, in a subsequent article Royston and Sauerbrei [15] extended their algorithm to include also interactions between binary variables such as treatments and continuous co-variables. Their proposal included estimating treatment effect functions and using treatment effect plots for illustration purposes. Finally, Sauerbrei and Royston [16] extended fractional polynomials to IPD-MA, proposing a two-stage approach. On a first stage an appropriate statistical model is fitted per trial and either the estimated coefficients with their standard errors or the treatment effect functions with their 95% confidence intervals are extracted. On a second stage, these estimates are pooled using either multivariate or point-wise meta-analysis respectively. Nevertheless,their two-stage approach may be prone to power loss when the sample size is small and the outcome is binary [17]

or the per trial mean of the continuous effect modifier is highly heterogeneous [18]. Furthermore, both trial and error and fractional polynomial approaches rely on tests with arbitrary α significance levels and estimate global functions which may not fit well on the boundaries of a continuous effect modifier. For instance, using the previous example, BMI and mortality may on average fit well using regression models with quadratic terms, but the fit may not be adequate for extreme cases such underweight (BMI <18.5) and severely obese (BMI>40) participants. Splitting the BMI into 3 intervals of BMI [<18.5 , 18.5-40 , 40+] and fitting a model within each interval would provide better fit. Nevertheless, the resulting piece-wise functional shapes, will probably be discontinuous to the knots (18.5, 40). Therefore, piece-wise approaches are usually combined with smoothing techniques in order to show continuous functional shapes.

Wang et al. [12,19] proposed a two-stage IPD-MA moving average (sliding window) approach for binary and time-to-event outcomes, called meta-STEPP. On the first stage the continuous effect modifier is split into intervals with the same number of events and within-interval and per study treatment effect are estimated using an effect size of choice. On a second stage, these within-interval treatment effects are pooled using either fixed or random effects meta-analysis. Finally, the pooled effect sizes are tested for heterogeneity using Cochran's Q χ^2 test. This algorithm is repeated multiple times with overlapping sliding windows for smoother results. Cochran's Q is the weighted average of the distances of the within-interval estimates to the average. Therefore, Cochran's Q lacks power to detect non-linear effect modification, compared to smoother approaches. Furthermore, the size of the window and the moving step may influence our results.

A more flexible approach is to use generalized additive models (GAMs). GAMs in their simplest form also split the effect modifier into knots and fit a polynomial regression within each interval. Nevertheless, these regression models are restricted to be continuous over the knots. The most often used polynomials are cubic splines while quadratic and linear are also available. Natural splines have one further restriction to be linear beyond the extreme knots. If we don't have prior knowledge for the knots these can be estimated using a cross-validation technique. Another piece-wise approach is locally (weighted) scatterplot smoothing (loess). Loess is a non-parametric regression that uses a sliding window techique for smoothing. Specifically, a weighted linear or quadratic model is fitted per data point using as a dataset its nearest observations. The number of these observations determines the size of the sliding window controls the smoothness of the functional shape. A combination of splines and loess is smoothing splines. Hereby, the sum of the maximum likelihood and a penalty factor for wiggliness (λ) is minimised. The functional shapes are adjusted by either manually tuning the values of λ or through an automated cross-validation procedure.

Applying GAMs may be beneficial for research as they make limited assumptions and can detect complex shapes. Specifically, with GAMs we may begin with smooth functions rather than naive linearity and let the data decide the underlying functional shape. This procedure involves less assumptions, no use of multiple statistical tests and may lead to better fitted regression lines. Nevertheless, it is still unknown how should we apply GAMs, when information from multiple studies is available. Our goal is to advocate the use of flexible over naive and resticted methods, inform for the avaible approaches and provide a tutorial how to fit them in empirical examples of IPD-MA.

2. Methods

In our study we advocate the use of generalized additive models (GAMMs) to model and investigate treatment effect modification by a continuous variable, whilst accounting for non-linear functional shapes and within study clustering of the participants. We will present two approaches to model and two approaches to pool the results of a GAMs.

2.1 Statistical approaches

As described in section 2, both empirical example datasets are composed of multiple randomized trials. We will adopt the following notation throughout our manuscript:

- The trials as j = 1,2, ...,N_j,
- Trial participants as i = 1,2, ...,n_i,
- Standardised $Age_{ij} = \frac{Age_{ij}-30}{60}$
- The per trial mean of age as $\overline{Age_j} = \frac{\overline{Age_{ij}-30}}{60}$ $Age_{ij} - \overline{Age_j}$

2.1.1 Natural cubic splines

Generalized additive models are a form of penalized generalised linear models. One type of penalization is identical to the random effects' technique introduced by McCullagh et al. [20] in the generalised mixed effects models. There we can estimate a separate functional form per trial, given the fixed effects parameters. Therefore, per trial and treatment regime a different shape will be estimate.

The statistical model will be:

$g(Y_{ij}) = s(Age)_{kj}$, where k the treatment regime.

2.1 Data-set

2.1.1 Simulated data-set

We simulated 1 IPD-set which we considered an extreme case with 5 RCTs with 100 participants each, including 4 variables: 1) a binary variable **T** which corresponds to treatment 2) a uniformly distributed co-variable **Age** ranging from 30 to 60 years old and 3) a variable to indicate trial membership 4) a continuous outcome **Y**. We assumed equal treatment allocation for all trials, but different age groups randomized per trial, see table 1. **Age** was linearly associated with **Y** in the control group and quadratically in the treated group.

We considered the following example. Patients with liver disease are randomly allocated to *placebo* and *drug*. After a period of time the difference in liver volume is measured. We assumed that patients that received *placebo* had an 2.5% decrease in all age categories. The treatment had no-effect on young patients (around 30), and as Age increased the treatment effect increased with a quadratic shape we assumed t.

The linear predictor for the whole IPD-set was: $Y_i = \beta_0 + \beta_{Age} \times \frac{Age_i - 30}{60} + \beta_{Treat} \times Treatment_i + \beta_x \times (\frac{Age_i - 30}{60})^2 \times Treatment_i + \epsilon_i$

Notice that we don't vary the βs per trial. β_0 = -2.5, β_{Treat} = 0 , β_{Age} = 0, β_X = 20

2.1.2 Empirical data-sets

We use 2 IPD-sets to illustrate the aforementioned method. The first IPD-MA investigates the effect of antibiotics in children with acute otitis media [21]. Rovers et al. collected IPD from 6 randomized 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 IPD-set [22] considers an IPD-MA to investigate the effect of somatostatin on liver volume reduction. Gevers et al. collected IPD from 3 randomized 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.

Results

Discussion

In our paper, we described and illustrated a new approach to model and investigate effect modification when the potential effect modifier is not linearly associated with the outcome. Furthermore, we applied our method on two empirical examples and one extreme-case simulated IPD-set. Clinical decision making may be based on either relative or absolute treatment effects. Our results show that it may be important to account for the outcome-variable functional shape. Two-stage methods suffered from ecological bias in our simulated example. Finally, we showed that effect modification may not be linear and therefore not possible to be encapsulated in a single interaction term. Thus treatment effect functions along with illustrative methods such treatment effect plots may be better options for decision making.

5.1 Comparison with literature

5.2 Strengths and limitations

The major strength of our paper is that we propose a novel approach for IPD-MA of RCTs. Particularly, we considered generalised additive mixed effects models with smoothing splines. We showed that smoothing splines make minimal shape assumptions as they minimize the sum of maximum likelihood and a penalty term for the wiggliness of the line.

5.3 Implications for practice

We believe that our approach may change the point of view of IPD-MA conducted. Specifically, the naive idea that. . . .

5.4 Conclusions

We propose the use of one-stage generalised additive model with smoothing spline. This approach makes no assumptions over the functional shape and the cut-point where it changes. Finally, when combined with the treatment effect plot researchers .

References

[1] Simmonds M, Stewart G, Stewart L. A decade of individual participant data meta-analyses: A review of current practice. Contemporary Clinical Trials 2015;45:76–83. doi:10.1016/j.cct.2015.06.012.

[2] 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.

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

[4] 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.

[5] 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.

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

[7] Zeileis A, Hothorn T, Hornik K. Model-based recursive partitioning. Journal of Computational and Graphical Statistics 2008;17:492–514. doi:10.1198/106186008x319331.

[8] Seibold H, Hothorn T, Zeileis A. Generalised linear model trees with global additive effects 2016.

[9] Su X, Tsai C-L, Wang H, Nickerson DM, Li B. Subgroup analysis via recursive partitioning. SSRN Electronic Journal 2009. doi:10.2139/ssrn.1341380.

[10] Mistry D, Stallard N, Underwood M. A recursive partitioning approach for subgroup identification in individual patient data meta-analysis. Statistics in Medicine 2018. doi:10.1002/sim.7609.

[11] Fokkema M, Smits N, Zeileis A, Hothorn T, Kelderman H. Detecting treatment-subgroup interactions in clustered data with generalized linear mixed-effects model trees. Behavior Research Methods 2017. doi:10.3758/s13428-017-0971-x.

[12] 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.

[13] Liu X, Zhang D, Liu Y, Sun X, Hou Y, Wang B, et al. A j-shaped relation of BMI and stroke: Systematic review and doseresponse meta-analysis of 4.43 million participants. Nutrition, Metabolism and Cardiovascular Diseases 2018;28:1092–9. doi:10.1016/j.numecd.2018.07.004.

[14] Royston P, Altman DG. Regression using fractional polynomials of continuous covariates: Parsimonious parametric modelling. Applied Statistics 1994;43:429. doi:10.2307/2986270.

[15] Royston P, Sauerbrei W. A new approach to modelling interactions between treatment and continuous covariates in clinical trials by using fractional polynomials. Statistics in Medicine 2004;23:2509–25. doi:10.1002/sim.1815.

[16] 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.

[17] 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.

[18] Simmonds MC, Higgins JPT. Covariate heterogeneity in meta-analysis: Criteria for deciding between meta-regression and individual patient data. Statistics in Medicine 2007;26:2982–99. doi:10.1002/sim.2768.

[19] Wang XV, Cole B, Bonetti M, Gelber RD. Meta-STEPP with random effects. Research Synthesis Methods 2018;9:312–7. doi:10.1002/jrsm.1288.

[20] McCullagh P, Nelder JA. An outline of generalized linear models. In:. Generalized linear models, Springer US; 1983, pp. 15–34. doi:10.1007/978-1-4899-3244-0_2.

[21] 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.

[22] 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.