Retrospective study of Multivariate Meta-analysis and A Case analysis of Nintedanib

Ziyue Liu

School of Mathematical Sciences, Nankai University, Tianjin, China

Keywords: Multivariate Meta-analysis, Robust Variance Estimator, Restricted Maximum Likelihood, Nintedanib, IPF

Abstract: In this paper, a detailed and retrospective study on multivariate meta-analysis theory is carried out, including the proposed improvement of model hypothesis, various methods of parameters estimation, and a case analysis of multivariate meta-analysis using robust variance estimator.

1. Introduction

The first part of the paper is based on model hypothesis and theoretical derivation. It requires readers having a mathematical background and a certain understanding of meta-analysis. We discuss two major multivariate models and the comparison is made, after briefly introducing the univariate model. During the section of parameters estimation, multiple methods are introduced, including generalized-least-squares multiple-regression, maximum likelihood, and restricted maximum likelihood, to make a good estimate of the quintuple parameters of the model when the IPD (Individual Patient Data) and the within-study correlation, both necessary for standard inference procedures, are unavailable.

The second part of the article uses the theoretical basis of the first part to conduct a practical analysis which is aimed at Nintedanib, one of the mainstream drugs for treating IPF (Idiopathic Pulmonary Fibrosis). With strictly following the complete procedures of meta-analysis, we perform multiple sets of bivariate meta-analysis using annual rate of decline in FVC, change from baseline in SGRQ, etc. as outcomes of studies, and improved multivariate model, restricted maximum likelihood and robust variance estimator as methods. In comparing results with corresponding univariate meta-analysis from another paper [7], the multivariate meta-analysis shows better statistical properties. The largest improvement of accuracy of annual rate of decline in FVC, change from baseline in FVC (% predicted) and change from baseline in SGRQ is, respectively, 21%, 41% and 12%.

2. Review of Multivariate Meta-analysis

2.1 Introduction of Meta-analysis

Meta-analysis is a statistical method which can be defined as the quantitative synthesis of results from multiple studies with certain conditions. Since the “combined p-value” proposed by Fisher in 1920, meta-analysis has been widely used in medicine and biology research, which can be largely contributed to its integrated and rigorous procedures which ensure the validity and effectiveness of meta-analysis. The practical steps are as follows:

2.1.1 Make the main question clear

2.1.2 Extensively collect relevant research literatures, including articles, books, experiments and even unpublished literatures

2.1.3 Strictly screen literatures with given requirements

2.1.4Extract information from the literatures which are selected by previous step (data, charts,

  DOI: 10.25236/medsbe.2019.066

etc.)

2.1.5 Choose appropriate summary statistics given study outcomes

2.1.6 Parameters estimation

2.1.7 Diagnose Publication bias when the number of included literatures is relatively large

Step five and six will be discussed later in detail.

2.2 Model of Meta-analysis

2.2.1 Univariate Model

Under univariate condition, the model is relatively simple and can be divided into two major cases, which is based on the results of homogeneity test. [1]

The first one is Fixed-Effects Model:

$$Y_i \sim N(\theta, s_i^2), \quad for\ i = 1, 2, ...k \quad (1.1)$$

Assume that $E(Y_i) = \theta,$ and let $s_i^2 = Var(Y_i)$ be the variance of the summary statistic in the i^{th} study.

And the Random-Effects Model:

$$\begin{aligned} Y_i \mid \theta_i,\ s_i^2 &\sim N(\theta_i, s_i^2) \\ \theta_i \mid \theta,\ \tau^2 &\sim N(\theta, \tau^2) \end{aligned} \quad (1.2)$$

Here assume that Y_i is a draw from $N(\theta_i, s_i^2)$, which are depended on specific study. Moreover, each θ_i is a draw from $N(\theta, \tau^2)$, thus θ and τ^2 present, respectively, the average treatment and between-study variance.

2.2.2 Multivariate Model

Under the condition of multiple summary statistics of outcomes from each study, random-effects model is more likely used because the only possible case for fixed-effects model is that every endpoint of the vector of summary statistics is proved to be homogeneous, which hardly happens in practice, thus we just introduce the random-effects model. Using bivariate model as example [2]

$$\begin{aligned} \begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} &\sim N\left(\begin{pmatrix} \theta_{i1} \\ \theta_{i2} \end{pmatrix}, \mathbf{S}_i \right), \quad \mathbf{S}_i = \begin{pmatrix} s_{i1}^2 & s_{i1} s_{i2} \rho_{\mathrm{W}i} \\ s_{i1} s_{i2} \rho_{\mathrm{W}i} & s_{i2}^2 \end{pmatrix} \\ \begin{pmatrix} \theta_{i1} \\ \theta_{i2} \end{pmatrix} &\sim N\left(\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \Omega \right), \quad \mathbf{\Omega} = \begin{pmatrix} \tau_1^2 & \tau_1 \tau_2 \rho_{\mathrm{B}} \\ \tau_1 \tau_2 \rho_{\mathrm{B}} & \tau_2^2 \end{pmatrix} \end{aligned} \quad (1.3)$$

Since the parameters θ_{ij} are of little interest, there is a marginal model written as

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \sim N\left(\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \mathbf{V}_i \right) \quad \mathbf{V}_i = \mathbf{S}_i + \mathbf{\Omega} = \begin{pmatrix} s_{i1}^2 + \tau_1^2 & s_{i1} s_{i2} \rho_{\mathrm{W}i} + \tau_1 \tau_2 \rho_{\mathrm{B}} \\ s_{i1} s_{i2} \rho_{\mathrm{W}i} + \tau_1 \tau_2 \rho_{\mathrm{B}} & s_{i1}^2 + \tau_1^2 \end{pmatrix} \quad (1.4)$$

The within-study standard errors s_{ij} are assumed known, which are actually estimated or reported by the individual studies in practice. The within-study correlations ρ_{Wi} are assumed known to fit the model but are rarely available in practice.

2.2.3 An Alternative Model

To avoid using the within-study correlations, another marginal model was proposed in [2] as

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \sim N\left(\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \Phi_i \right), \quad \Phi_i = \begin{pmatrix} \psi_1^2 + s_{i1}^2 & \rho \sqrt{\left(\psi_1^2 + s_{i1}^2\right)\left(\psi_2^2 + s_{i2}^2\right)} \\ \rho \sqrt{\left(\psi_1^2 + s_{i1}^2\right)\left(\psi_2^2 + s_{i2}^2\right)} & \psi_2^2 + s_{i2}^2 \end{pmatrix} \quad (1.5)$$

In this model, s_{ij}^2 can also be assumed known, ψ_i^2 are additional variation beyond within-study variances and ρ accounts for an overall correlation which is a combination of

within-study and between-study correlation. Comparing with previous models, the within-study correlations ρ_{Wi} are no long needed, which means the model only, requires as known Y_i and s_{ij}^2 which are the same information to fit separate univariate model. With certain relative difference between ψ_i^2 and s_{ij}^2 in magnitude, the alternative model can closely approximate the general model. [2]

2.3 Parameters Estimation

2.3.1 Why?

We mentioned that meta-analysis model can be divided into two different models depended on the results of homogeneity test. Under the condition of fixed-effects model, the only thing to do is weighing each study with its reciprocal of within-study variance which is estimated by sample variance in practice. Using univariate model in [1] as example

$$\hat{\theta}_{\text{MLE}} = \frac{\sum_{i=1}^{k} W_i Y_i}{\sum_{i=1}^{k} W_i}; \quad W_i = \frac{1}{s_i^2} \qquad (1.6)$$

There is no actual parameter to be estimated by observing the equation above. When considering random-effects model, we need to estimate the between-study variance. It can be naturally generalized to multivariate model with assuming all studies provide all outcomes. The pooled estimate $\hat{\boldsymbol{\beta}}$ is given in [3] as

$$\hat{\boldsymbol{\beta}} = \left(\sum_{i=1}^{n} \left(\mathbf{S}_i + \widehat{\boldsymbol{\Omega}} \right)^{-1} \right)^{-1} \left(\sum_{i=1}^{n} \left(\mathbf{S}_i + \widehat{\boldsymbol{\Omega}} \right)^{-1} \mathbf{Y}_i \right) \quad (1.7)$$

Here we can notice two difficulties. First the off-diagonal components of $\mathbf{S}_i$ are always unavailable due to the unknown within-study correlations. Second as well as the univariate model, between-study covariance $\boldsymbol{\Omega}$ need to be estimated, which can be quite complicated.

2.3.2 GLS Multiple Regression

In this part of discussion, we assume that the within-study correlation is known thus using the model of (1.4), and the model of random-effects multiple-regression is given in [4] as

$$\mathbf{Y}_i = X_i \boldsymbol{\beta} + \boldsymbol{\delta}_i + \mathbf{e}_i \quad (1.8)$$

$\mathbf{Y}_i$ is a vector of p outcomes reported by study i, X_i is the corresponding design matrix, $\boldsymbol{\beta}$ is the vector of regression coefficients, $\boldsymbol{\delta}_i$ which is a draw from $N(0, \boldsymbol{\Omega})$, is a vector of p random effects associated with study i, and $\mathbf{e}_i$, which is approximately $N(0, \mathbf{S}_i)$ when the number of sample n_i is sufficiently large, is a vector of random sampling errors within study i.

Under the assumption of fixed-effects, $\mathbf{Y}_i \sim N(X_i \boldsymbol{\beta}, \mathbf{S}_i)$ and the model can be simply written as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \quad (1.9)$$

Considering k included studies, $\mathbf{Y}$ and $\mathbf{X}$ are stacked by $\mathbf{Y}_i$ and $\mathbf{X}_i$ respectively, the $kp \times kp$ block-diagonal matrix $\mathbf{S}$ is the estimate of the covariance of $\mathbf{Y}$. Therefore the estimate of $\boldsymbol{\beta}$ is written as

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^{\mathrm{T}} \mathbf{S}^{-1} \mathbf{X} \right)^{-1} \mathbf{X}^{\mathrm{T}} \mathbf{S}^{-1} \mathbf{Y} \quad (1.10)$$

When considering the assumption of random effects, we need to substitute $\mathbf{S}_i + \boldsymbol{\Omega}$ for $\mathbf{S}_i$ and to estimate $\boldsymbol{\Omega}$ and $\boldsymbol{\beta}$ by iteration. Here we just give $\widehat{\boldsymbol{\Omega}}$ as

$$\widehat{\boldsymbol{\Omega}} = (k-2)^{-1} (\mathbf{Y} - X\hat{\boldsymbol{\beta}})^{\mathrm{T}}_{p \times k} (\mathbf{Y} - X\hat{\boldsymbol{\beta}})_{k \times p} - k^{-1} \sum_{i=1}^{k} S_i \quad (1.11)$$

while details of iteration are discussed in [4].

2.3.3 Maximum Likelihood Estimate (MLE)

As a method of statistical estimation, MLE, which estimate parameters by resolving the maximized likelihood function, is quite important and widely used. The restricted maximum likelihood estimate (REML), which is better for estimating variance and covariance parameters, is a particular form of MLE thus more likely used in meta-analysis.

In this part of discussion, within-study correlations ρ_{Wi} are no longer available thus the model of (1.5) is fitted. Considering writing convenience, we assume that each study provides two outcomes for making design matrices X_i be identity matrices thus simplifying the expression. The restricted log-likelihood of residuals, $\log \lambda_{REML}$, is given in [5] as

$$\log \lambda_{\mathrm{REML}} = \text{ Const. } -\frac{1}{2}\sum_{i=1}^{m}\log\left|\Phi_i\right| - \frac{1}{2}\log\left|\sum_{i=1}^{m}\Phi_i^{-1}\right| - \frac{1}{2}\sum_{i=1}^{m}(Y_i-\beta)^T\Phi_i^{-1}(Y_i-\beta) \quad (1.12)$$

2.3.4 Robust Variance Estimator

Let $\boldsymbol{\eta}$ denote $(\beta_1, \tau_1^2, \beta_2, \tau_2^2, \rho)$, $\nabla \log \lambda_{REML}$ and $\nabla^2 \log \lambda_{REML}$ denote, respectively, the first and second partial derivatives of $\log \lambda_{REML}$, $\tilde{\boldsymbol{\eta}}$ denote the solution of equation $\nabla \log \lambda_{REML} = 0$, $H(\boldsymbol{\eta})$ and $J(\boldsymbol{\eta})$ denote $E\left\{-\nabla^2 \log \lambda_{REML}\right\}$ and $\operatorname{cov}\left\{\nabla \log \lambda_{REML}\right\}$. By Taylor expansion around $\boldsymbol{\eta}$, the approximate distribution is given in [5] as

$$\sqrt{m}(\tilde{\boldsymbol{\eta}} - \boldsymbol{\eta}) \xrightarrow{D} N\left(0, H(\boldsymbol{\eta})^{-1} J(\boldsymbol{\eta}) H(\boldsymbol{\eta})^{-1}\right) \quad (1.13)$$

In practice, we use empirical estimates $\hat{H}(\boldsymbol{\eta})$ and $\hat{J}(\boldsymbol{\eta})$ substituting for $H(\boldsymbol{\eta})$ and $J(\boldsymbol{\eta})$, thus the robust estimate for $\tilde{\boldsymbol{\eta}}$ is written as

$$\hat{H}(\tilde{\boldsymbol{\eta}})^{-1} \hat{J}(\tilde{\boldsymbol{\eta}}) \hat{H}(\tilde{\boldsymbol{\eta}})^{-1} \quad (1.14)$$

3. Case Analysis for Nintedanib

3.1 Information and Method

In this section, we will perform a complete multivariate meta-analysis for IPF drug Nintedanib step by step as described previously and compare the results with other literatures.

3.1.1 Criteria for Collecting Literature

3.1.1.1 Published randomized controlled trials, RCTs, in English

3.1.1.2 Patients are all recruited according to international diagnostic criteria for IPF

3.1.1.3 Experimental: Nintedanib 150mg bid, Controlled: Placebo

3.1.1.4 Outcomes: Annual rate of decline in Forced Vital Capacity (FVC) over 52 weeks, change from baseline in FVC (% predicted) over 52 weeks, change from baseline in Saint-George's Respiratory Questionnaire (SGRQ) total score at 52 weeks, change from baseline in SpO2 (Oxygen Saturation, Expressed in Percent) at rest up over 52 weeks, change from baseline in Carbon Monoxide Diffusion Capacity (DLCO) at rest over 52 weeks.

3.1.1.5 Preclude literatures published more than one time, retrospective studies and literatures unavailable

3.1.2 Literature Collecting

For collecting literatures as extensive as possible, we search on several major databases like PubMed, EMbase, The Cochrane Library and Clinicaltrials.gov, including keywords like Nintedanib, BIBF1120, OFEV, IPF, Idiopathic Pulmonary Fibrosis. In this process, we firstly collect literatures according to the titles and abstracts, then screen them by reading the full text. The

information extracted includes basic information of studies, baselines, sample sizes, average ages, sexes, races, smoking conditions, weights, time since diagnosis, etc.

3.1.3 Results of Collecting

After the process of collecting described previously, we have five highly relative literatures. Three of them, which are INPULSIS-1 [8], INPULSIS-2 [8] and TOMORROW [9], are suitable for meta-analysis because the types of data they provided are highly homogeneous, while others tended to provide different types or time frames of data. The trial NCT01170065 [10] likely provides the time frame of data for three years, which are too long to be considered comparing with other trials. Trials NCT01136174 [11] tends to provide data of two months, which is shorter than others.

3.2 Result analysis

The summary statistics of studies are presented by Mean Differences (MD), which are calculated by subtracting mean of controlled group from mean of experimental group. As for indicators positively related to the health condition, treatment is reflected effective by positive value of MD.

We perform multiple sets of bivariate meta-analysis using alternative model (1.5), restricted maximum likelihood estimation (1.12) and robust variance estimator (1.14), then compare the results with other literatures. The estimations of $\boldsymbol{\beta}$, which is defined as MD, are quite close with other papers thus assure the efficacy of Nintedanib based on multiple indicators for another time. Meanwhile, the estimated variances of $\hat{\boldsymbol{\beta}}$ are relatively smaller, which suggest that our estimations of $\boldsymbol{\beta}$ are more accurate and that multivariate meta-analysis has better statistical properties.

3.2.1 Annual rate of decline in FVC & Change from baseline in FVC (%Predicted)

FVC, refers to forced vital capacity, is an important indicator for measuring lung function which is better with higher FVC value. As presented in our results, Nintedanib is effective for slowing the decline of FVC. The $\boldsymbol{\beta}$ estimated in our multivariate meta-analysis is (111.68, 3.31), estimated ρ is 0.65, 95% confidence intervals (CI) are (91.59, 131.76) and (2.93, 3.69), the standard deviations are 10.25 and 0.19. With exactly the same data, the $\boldsymbol{\beta}$ estimated in univariate meta-analysis is (112.0, 3.3), 95% CIs are (79.6, 144.3) and (2.5, 4.7), the standard deviations are 13.04 and 0.32. Thus the standard deviation of estimation decreases 21% and 41% respectively, which means the accuracy of estimation is improved. Details are shown in Table 1 and Table 2, forest plots are shown in Figure 1, the comparisons of estimated distributions between multivariate meta-analysis and univariate meta-analysis are shown in Figure 2.

Table 1: Annual Rate of Decline in FVC

	Nintedanib			Placebo			
Trial[1]	Mean	SE[2]	Total	Mean	SE	Total	MD[95%CI]
INPULSIS-1	-114.65	16.329	309	-219.91	18.709	204	125.30[77.81, 172.70]
INPULSIS-2	-113.59	15.726	329	-207.32	19.309	219	93.73[90.66, 96.80]
TOMORROW	-60.0	39	84	-190.0	36	83	130.80[27.0, 234.6]

Meta	Total	Total	Estimated[95%CI]
Univariate	772	506	112.0[79.6, 144.3]
Multivariate	772	506	111.68[91.59, 131.76]

Table 2: Change from Baseline in FVC (%Predicted)

	Nintedanib			Placebo			
Trial	Mean	SE	Total	Mean	SE	Total	MD[95%CI]
INPULSIS-1	-2.76	0.408	307	-5.98	0.474	204	3.06[1.87, 4.25]
INPULSIS-2	-3.09	0.433	327	-6.15	0.505	217	3.22[2.11, 4.33]
TOMORROW	-1.04	0.990	84	-6.00	1.019	84	4.97[2.37, 7.56]

Meta	Total	Total	Estimated[95%CI]
Univariate	818	505	3.3[2.5, 4.1]
Multivariate	818	505	3.31[2.93, 3.69]

3.2.2Change from baseline in SGRQ & Change from baseline in FVC (%Predicted)

SGRQ is a questionnaire designed for measuring the condition of a patient with respiratory disease, and the condition is suggested more serious with higher change from baseline in SGRQ. In our results, it can be clearly seen that the average change from baseline in SGRQ in patients using Nintedanib is reduced relative to patients using placebo, indicating that Nintedanib can improve the function of overall respiratory system.

The $\boldsymbol{\beta}$ estimated in our multivariate meta-analysis is (-2.38, 3.37), overall correlation ρ is estimated as -0.83, 95% CIs are (-4.40, -0.36) and (2.94, 3.80), standard deviations are 1.03 and 0.22. The corresponding $\boldsymbol{\beta}$ estimated in univariate meta-analysis is (-2.50, 3.3), standard deviations are 1.17 and 0.32, 95% CIs are (-5.41, 0.41) and (2.5, 4.1). Thus the standard deviation decreases 12% and 31% respectively. Details are shown in Table 3 and Table 4, forest plots are shown in Figure 3, the comparisons of estimated distributions between multivariate meta-analysis and univariate meta-analysis are shown in Figure 4.

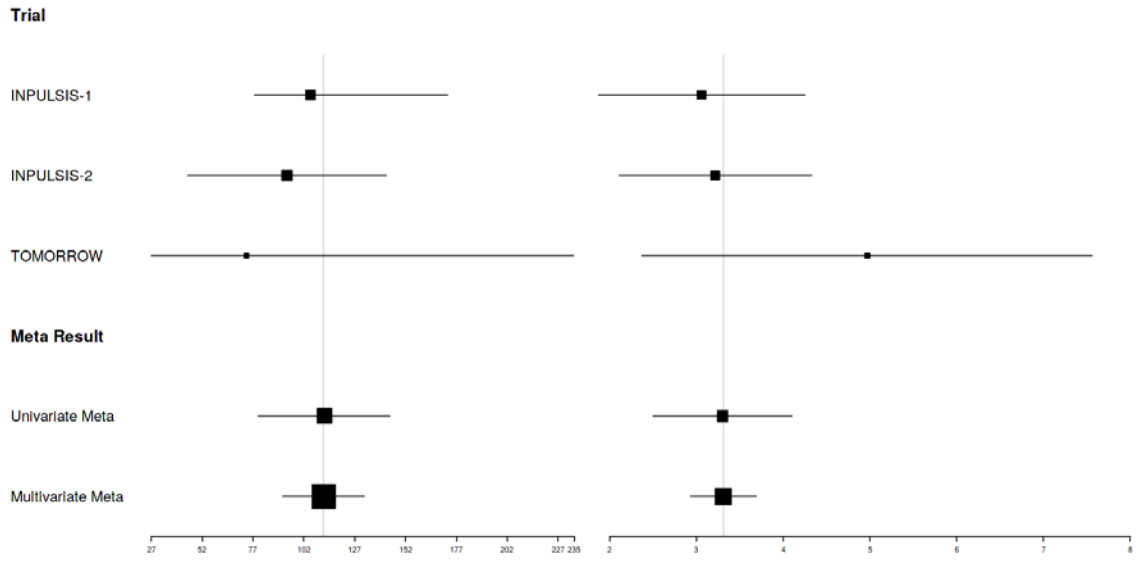

Figure 1: Forest Plot of Annual Rate of Decline in FVC (Left) & Change from Baseline in FVC (%Predicted) (Right)

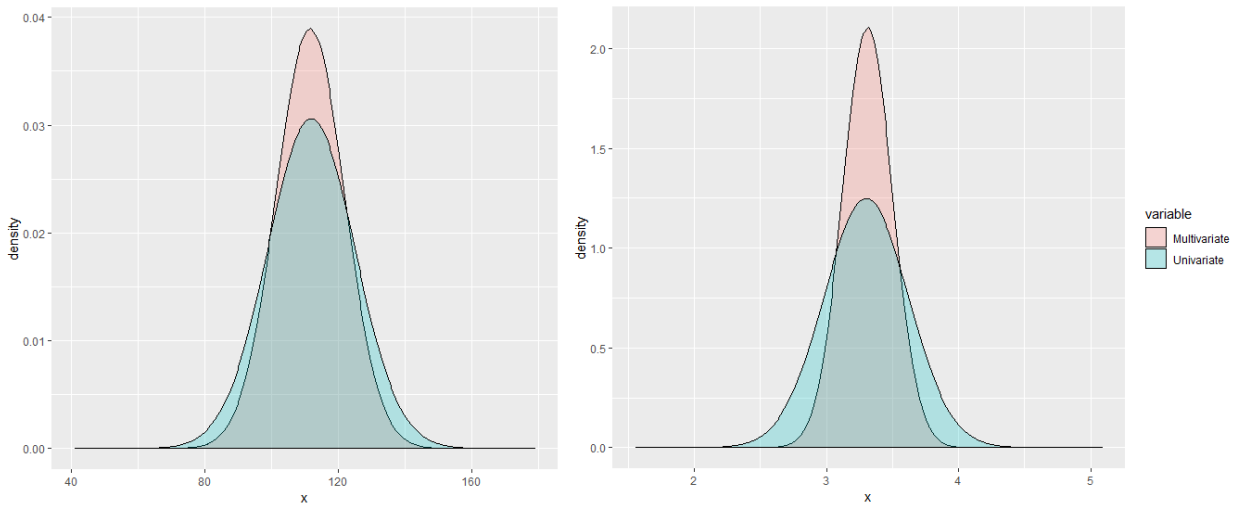

Figure 2: Comparison of Estimated Distributions of Mean Difference of Annual Rate Decline in FVC (Left) & Change from Baseline in FVC (%Predicted) (Right) between Multivariate and Univariate Meta-analysis

Table 3: Change from Baseline in SGRQ

	Nintedanib			Placebo			
Trial	Mean	SE	Total	Mean	SE	Total	MD[95%CI]
INPULSIS-1	4.34	0.799	289	4.39	0.960	200	-2.69[-4.95, -0.43]
INPULSIS-2	2.80	0.730	320	5.48	0.891	213	-0.05[-2.50, 2.40]
TOMORROW	-0.66	1.712	75	5.46	1.731	79	-6.12[-10.55, 1.69]

Meta	Total	Total	Estimated[95%CI]
Univariate	684	492	-2.50[-5.41, 0.41]
Multivariate	684	492	-2.38[-4.40, 0.36]

Table 4: Change from Baseline in FVC (%Predicted)

	Nintedanib			Placebo			
Trial	Mean	SE	Total	Mean	SE	Total	MD[95%CI]
INPULSIS-1	-2.76	0.408	307	-5.98	0.474	204	3.06[1.87, 4.25]
INPULSIS-2	-3.09	0.433	327	-6.15	0.505	217	3.22[2.11, 4.33]
TOMORROW	-1.04	0.990	84	-6.00	1.019	84	4.97[2.37, 7.56]

Meta	Total	Total	Estimated[95%CI]
Univariate	818	505	3.3[2.5, 4.1]
Multivariate	818	505	3.37[2.94, 3.80]

3.2.3Change from Baseline in SpO2 & Change from Baseline in DLCO

SpO2, refers to Oxygen Saturation, and DLCO, refers to Carbon Monoxide Diffusion Capacity, are indicators for the degree to which IPF have exacerbated. The lung function of patient, as well as the condition of IPF, is suggested worse or more serious with lower value of SpO2 and DLCO. In our result-analysis, it shows that Nintedanib is able to reduce the degree to which SpO2 and DLCO decrease thus slowing the exacerbation of IPF.

The $\boldsymbol{\beta}$ estimated in our meta-analysis is (0.34, 0.03), the overall correlation ρ is estimated as -0.75, 95% CIs are (0.21, 0.46) and (-0.05, 0.10), standard deviations are 0.06 and 0.04 which are small. Comparison is not made because these two indicators were not included in the paper we

compare. Details are shown in Table 5 and Table 6, forest plots are shown in Figure 5.

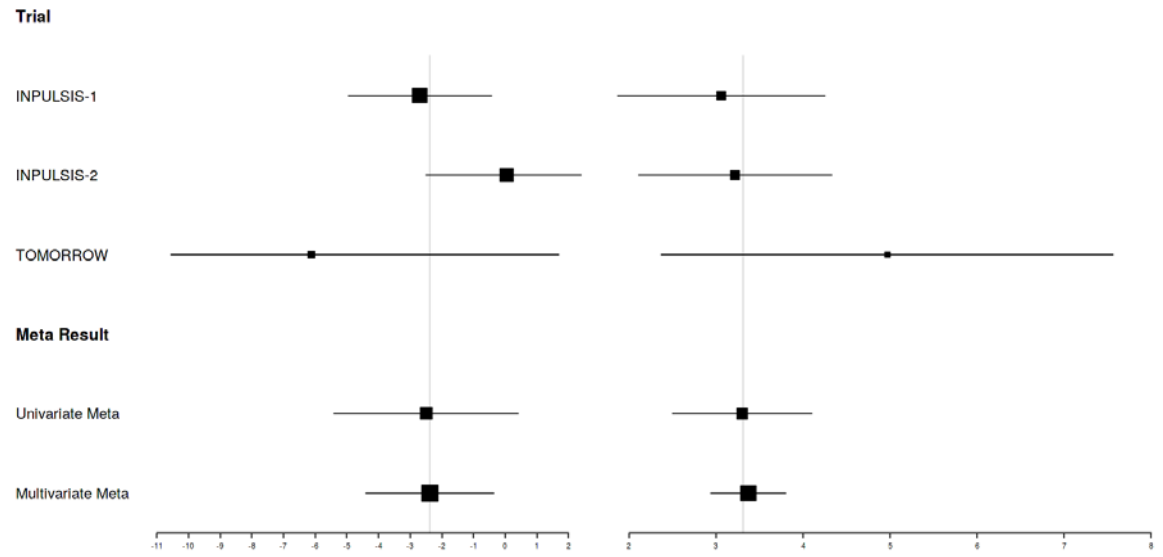

Figure 3: Forest Plot of Change from Baseline in SGRQ (Left) & Change from Baseline in FVC (%Predicted) (Right)

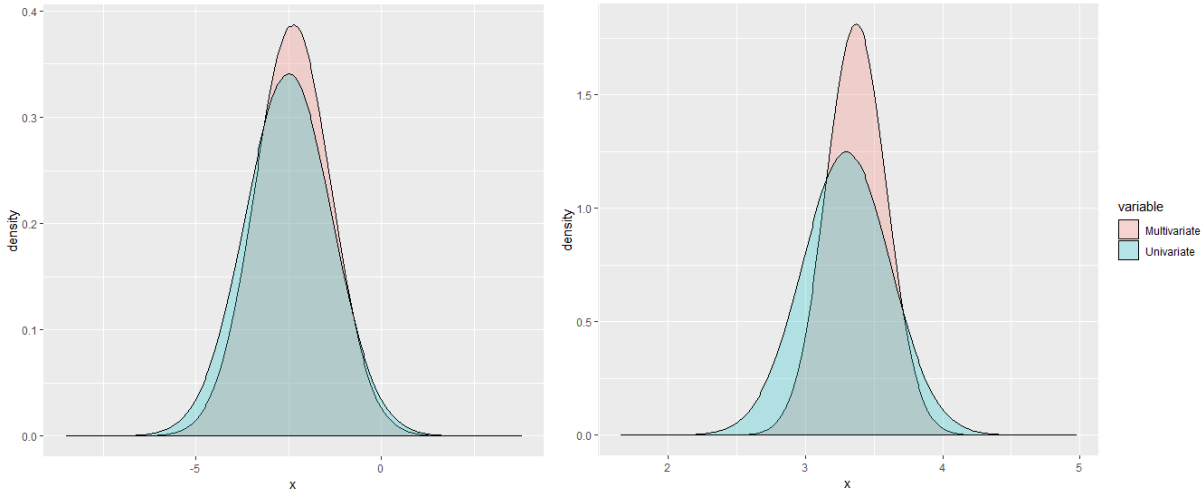

Figure 4: Comparison of Estimated Distributions of Mean Difference of Change from Baseline in FVC (%Predicted) (Left) & Change from Baseline in SGRQ (Right) between Multivariate and Univariate Meta-analysis

Table 5: Change from Baseline in SpO2

	Nintedanib			Placebo			
Trial	Mean	SE	Total	Mean	SE	Total	MD[95%CI]
INPULSIS-1	-0.24	0.129	299	-0.53	0.150	199	0.29[0.26, 0.32]
INPULSIS-2	-0.39	0.149	320	-0.66	0.174	212	0.27[0.24, 0.30]
TOMORROW	-0.18	0.360	83	-1.29	0.373	82	1.12[0.17, 2.07]

Meta	Total	Total	Estimated[95%CI]
Univariate	702	493	0.34[0.21, 0.46]

Table 6: Change from Baseline in DLCO

	Nintedanib			Placebo			
Trial	Mean	SE	Total	Mean	SE	Total	MD[95%CI]
INPULSIS-1	-0.38	0.0644	286	-0.365	0.0750	195	0.113[-0.084, 0.310]
INPULSIS-2	-0.286	0.729	320	-0.4	0.843	213	-0.015[-0.191, 0.161]
TOMORROW	-0.576	0.1111	69	-0.455	0.1098	75	-0.121[-0.406, 0.164]

Meta	Total	Total	Estimated[95%CI]
Univariate	675	483	0.03[-0.05, 0.10]

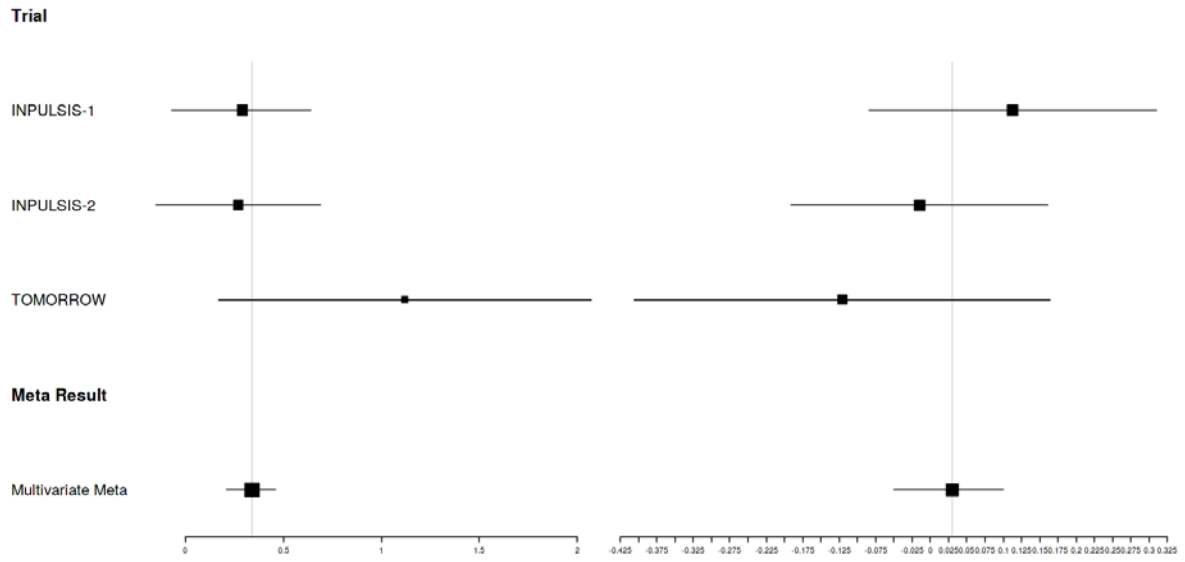

Figure 5: Forest Plot of Change from Baseline in SpO2 (Left) & Change from Baseline in DLCO (Right)

3.3 Summary

With all the results showing above, we can conclude that Nintedanib is effective for treating IPF considering multiple indicators related to lung functions. The number of clinical trials, however, are still relatively small, which is insufficient for evaluating more aspects like long-term safety and recessive side-effects meanwhile limits the reliability of meta-analysis results. Moreover, nearly all the trials existed are conducted by the company Boehringer Ingelheim, producer of Nintedanib, so we have reason to be skeptical about the absolute objectiveness thus the practical usage should still be cautious.

4. Conclusion

As for the method we use, multivariate meta-analysis obviously has better statistical properties with giving similar estimated values like univariate method. Problem is that multivariate method is relatively complicated and the complexity is escalated with more than two summary statistics included. For medical researchers, using specific software designed for univariate meta-analysis is much more convenient and feasible. Furthermore, the improvement of statistical properties is highly depended on the number of studies included and also decided by other factors which are discussed in other papers thus the specific method should be decided by actual conditions. At last but not least, meta-analysis is mainly used under the assumption of normal distribution, which can be somewhat dubious because the actual distributions of summary statistics can be quite different and sometimes hardly be tested due to the few studies included. The more reasonable assumption of meta-analysis, therefore, should be improved to consider other possible distributions, even involved with nonparametric methods.

Annotation

[1] Trial refers to the three studies included.

[2] SE refers to Standard Error reported from studies.

References

[1] Normand, S. L. T. (1999). Meta-analysis: formulating, evaluating, combining, and reporting. Statistics in medicine, 18(3), 321-359.

[2] Riley, R. D., Thompson, J. R., & Abrams, K. R. (2007). An alternative model for bivariate random-effects meta-analysis when the within-study correlations are unknown. Biostatistics, 9(1), 172-186.

[3] Jackson, D., Riley, R., & White, I. R. (2011). Multivariate meta-analysis: potential and promise. Statistics in medicine, 30(20), 2481-2498.

[4] Berkey, C. S., Hoaglin, D. C., Antczak-Bouckoms, A., Mosteller, F., & Colditz, G. A. (1998). Meta-analysis of multiple outcomes by regression with random effects. Statistics in medicine, 17(22), 2537-2550.

[5] Hong, C., D. Riley, R., & Chen, Y. (2018). An improved method for bivariate meta-analysis when within-study correlations are unknown. Research synthesis methods, 9(1), 73-88.

[6] Jennrich, R. I., & Schluchter, M. D. (1986). Unbalanced repeated-measures models with structured covariance matrices. Biometrics, 805-820.

[7] Richeldi, L., Cottin, V., du Bois, R. M., Selman, M., Kimura, T., Bailes, Z., ... & Brown, K. K. (2016). Nintedanib in patients with idiopathic pulmonary fibrosis: combined evidence from the TOMORROW and INPULSIS® trials. Respiratory medicine, 113, 74-79.

[8] Richeldi, L., Cottin, V., Flaherty, K. R., Kolb, M., Inoue, Y., Raghu, G., ... & Stowasser, S. (2014). Design of the INPULSIS™ trials: two phase 3 trials of nintedanib in patients with idiopathic pulmonary fibrosis. Respiratory medicine, 108(7), 1023-1030.

[9] Richeldi, L., Costabel, U., Selman, M., Kim, D. S., Hansell, D. M., Nicholson, A. G., ... & Brun, M. (2011). Efficacy of a tyrosine kinase inhibitor in idiopathic pulmonary fibrosis. New England Journal of Medicine, 365(12), 1079-1087.

[10] Richeldi, L., Kreuter, M., Selman, M., Crestani, B., Kirsten, A. M., Wuyts, W. A., ... & Costabel, U. (2018). Long-term treatment of patients with idiopathic pulmonary fibrosis with nintedanib: results from the TOMORROW trial and its open-label extension. Thorax, 73(6), 581-583.

[11] Ogura, T., Taniguchi, H., Azuma, A., Inoue, Y., Kondoh, Y., Hasegawa, Y., ... & Klüglich, M. (2015). Safety and pharmacokinetics of nintedanib and pirfenidone in idiopathic pulmonary fibrosis. European Respiratory Journal, 45(5), 1382-1392.