Risk prediction with a generalized linear mixed effects model for detecting adverse drug event of drug combination

Hae Reong Kim

The Graduate School
Yonsei University
Department of Biostatistics and Computing

Risk prediction with a generalized linear mixed effects model for detecting adverse drug event of drug combination

A Dissertation Submitted to
the Department of Biostatistics and Computing
and the Graduate School of Yonsei University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in Biostatistics and Computing

Hae Reong Kim

December 2022

This certifies that the dissertation of
Hae Reong Kim is approved.

Inkyung Jung: Thesis Supervisor

Chung Mo Nam: Thesis Committee Member #1

Jae Hyun Lee: Thesis Committee Member #2

Yu Rang Park: Thesis Committee Member #3

Goeun Park: Thesis Committee Member #4

The Graduate School
Yonsei University
December 2022

Contents

List of Tables

List of Figures

Abstract

Risk prediction with a generalized linear mixed effects model for detecting adverse drug event of drug combination

Adverse drug events (ADEs) induced by multiple combination drugs are a critical concern when dealing with drug-drug interactions (DDIs).

Most studies on multiple drug combination methods mainly focus on pairwise drug reactions, so existing risk prediction approaches for high-dimensional drug interaction are inadequate. To consider ADEs of a high-dimensional situation multiple-combination drugs, we propose a generalized linear mixed-effect model with two levels to predict the multiple-combination drug risk based on a single drug. As the structure of single and multiple combination is a connected structure, it should be considered as a nested structure rather than as a multiple independent structure. In this study, a mixed-effect model was used to estimate the variation between clusters by grouping single drug units. We assume the

analyses that do not reflect these nested structures can lead to increase false-positive ADE signal detection. For performance comparison, we use the mixture drug-count response model (MDRM) (Wang, Zhang et al. 2018) for high-dimensional drug interaction. In simulation, we show that our mixed-effect model provides better performance than the conventional MDRM method with power. Then, we apply these methods to the CDM database and compare the predicted risk using the proposed method based on observed risk.

Key words: Drug adverse event, Signal detection, Spontaneous reporting system, Mixed-effect model, High-dimensional estimation, Nested structure, Drug-drug interaction, CDM database

1. Introduction

An adverse drug event (ADE) is a negative side effect that may be caused by using a medication. It can be a result of the unintended and unexpected pharmacological actions of the drug (Nguyen, Nguyen et al. 2019). When multiple drugs are administered together, the risk of ADEs may increase because of drug–drug interactions (DDIs). The Centers for Disease Control and Prevention (CDC) reports that 21.8% of the US population takes three or more prescription drugs simultaneously, and 10.7% of people take five or more simultaneously. Before 2000, these percentages were only 11.0% and 3.6%, respectively.

The ADE signal detection methods primarily exploit data from SRS using conventional statistical analysis methods (Ho, Le et al. 2016). However, SRS has several limitations and difficulties, such as under-reporting and bias, when detecting ADE (Alomar, Tawfiq et al. 2020, Ibrahim, Abdo et al. 2021). The underreporting of ADE can trigger a delay in the signal detection of ADE and raise risk estimates, resulting in false positives (Ventola 2018).

In consideration of the risk of false positives, studies are being conducted to detect side effects signals in multiple drug combinations. Conventional methods for signal detection such as proportional reporting ratio (PRR), reporting odds ratio (ROR), and log-likelihood ratio test (LRT) were able to reduce false positives through efforts to get a higher threshold for signal detection.

The use of mixture distribution allowed for the control of false positives by dividing the data into two or three components, including cases where no drug was reported and situations where the effect of the drug was constant or increased. However, there is still an issue with false positives because the structure of DDIs cannot be fully captured. In the analysis of DDI structure using the prior research method, mainly through a 2×4 contingency table, it is not possible to consider and reflect the dependent characteristics of each individual drug in the part fitted with multinomial distribution. Furthermore, the contingency table does not allow for the depiction of many structures, so it is limited to modeling likewise two-drug combination structures.

Most methods for analyzing DDIs primarily focus on pairwise interactions, as expressed in contingency tables, and there is a lack of research on methods for high-dimensional DDIs (Wang, Zhang et al. 2018). This is because it becomes challenging to intuitively represent interactions involving more than two drugs, as the number of cases to be analyzed becomes too large.

To deal with high-dimension DDIs in the SRS database, the generalized linear mixed-effect model (GLMM) was used in this study. GLMM offers a flexible approach for modeling various types of data. In general, GLMM is used in repeated measurement for longitudinal dataset (Martin, Uh et al. 2019). This repeated measurement operates within the time frame repeatedly observed within the same subject. We conducted an experiment to see if a GLMM could be used to predict the side effects of taking multiple drugs together, considering the correlations between data observed within the same subject. To do so, we

treated each single drug as a cluster or group, and each cluster contained a combination of several drugs. We then analyzed the data using a two-level GLMM, which considered the structure of the clusters containing multiple drugs.

The mixed effect model mainly considers whether the random intercept or the random slope is different for each group. Random slope models can fit the data more flexibly, so it can explain the variability between groups (Harrison, Donaldson et al. 2018). As the structure of single and multiple combination is a connected structure, it should be considered as a nested structure. Accordingly, we focus on this flexibility of this method, and we apply to our nested structure to this model.

However, given that the structure of these reported drugs itself is a nested structure, it is necessary to take this into account while modeling. When our methodology calculates false positives, this prediction tends to be close to reality because the modeling considers the baseline characteristics of an actual drug and reflects the specific risks of each drug. This can result in reducing the bias. Our proposed method can reduce false positives and reflect multiple-combination drug structures properly compared to conventional methods.

In MDRM, the conventional method, the risk of DDI can be confirmed when a drug is administered alone or in combination. However, the analysis assumes that all risks of the single drug are evenly distributed. As a single risk for a drug cannot be viewed as an evenly distributed risk, this assumption is not reasonable.

Signal detection can be performed on whether it is constant risk or drug-dependent

risk. The advantage is that we can perform signal detection in diverse single risk groups which consider as cluster. Furthermore, there is no signal detection algorithm that can be employed in this situation. As there is no algorithm that can be applied when each baseline drug has a different risk, signal detection considering the characteristics of complex drugs is a problem needing specific solutions.

It is crucial to construct a database related to adverse drug reactions in Korea. As a representative database for drug surveillance status in Korea, related information can be reported and managed through the Korea Adverse Event Reporting System (KAERS). By utilizing the information on abnormal cases collected through this system, clue information search and evaluation safety information are produced. In addition, the Korea Institute of Drug Safety and Management secures medical field data based on the Common Data Model (CDM) at multiple institutions, and through this, conducts research to verify drug side effect signals at multiple institutions. There is a need to solve this safety issue in terms of Post-Marketing Surveillance (PMS), and this can be seen as the result of using real word data, which was not confirmed in single drug clinical trials. For these reasons, there is a need for signal detection in multiple combination cases.

This method was tested using CDM data. The reason for using this approach is that it has a similar structure to electronic medical record (EMR) data, and it can be used as a correction variable with the added benefit of being able to include information about indications, drug intake orders, terms, and other baseline information. There is also the potential for this model to be extended to other institutions in the future.

2. Background

2.1 Mixture drug-count response model (MDRM)

As the practice of taking multiple medications, polypharmacy, becomes more common, the risk of ADEs caused by DDIs is becoming a major concern in clinical practice (Wang, Li et al. 2020). There is a lack of method for detecting high-dimensional DDIs in SRS databases using data mining techniques, which typically only consider pairwise DDIs. To solve this gap, MDRM was introduced to identify adverse events caused by complex, multi-drug interactions (Wang, Zhang et al. 2018).

The marginal distribution of Y_{ik} is given by $P(y_{ik}) = (1-\pi)Bin(n_{ik}, y_{ik}, q_0) + \pi Bin(n_{ik}, y_{ik}, q_1)$ where $q_0 = \frac{exp(\beta_0)}{1+exp(\beta_0)}$, $q_1 = \frac{exp(\beta_0+\beta_1 \boldsymbol{i})}{1+exp(\beta_0+\beta_1 \boldsymbol{i})}$. The i denotes the number of drugs for i-way drug combinations, k is the kth i-way drug combinations, N_{ik} is the number of total patients taking kth i-way drug combinations, Y_{ik} is the number of targets who experienced adverse event among those N_{ik} patients. If Y_{ik} follows the constant model, then $Z_{ik} = 0$ and $Z_{ik} = 1$ otherwise.

The joint distribution of (Y_{ik}, Z_{ik}) is given by $P(y_{ik}, z_{ik}) = [(1-\pi) \times Bin(n_{ik}, y_{ik}, q_0)]^{1-z_{ik}} \times [\pi \times Bin(n_{ik}, y_{ik}, q_1)]^{z_{ik}}$ where $q_0 = \frac{exp(\beta_0)}{1+exp(\beta_0)}$, $q_1 = \frac{exp(\beta_0+\beta_1 \boldsymbol{i})}{1+exp(\beta_0+\beta_1 \boldsymbol{i})}$. And the log-likelihood function is given by $l(\theta) = \sum_{i=1}^{N}\sum_{k=1}^{M_i} log\pi Bin\left(n_{ik}, y_{ik}, \frac{exp(\beta_0+\beta_1 i)}{1+exp(\beta_0+\beta_1 i)}\right) + log(1-\pi)Bin(n_{ik}, y_{ik}, \frac{exp(\beta_0)}{1+exp(\beta_0)})$

where $\theta=(\pi,\ \beta_0,\ \beta_1)$.

To find the maximum-likelihood estimates (MLE) for our analysis, we employed the expectation-maximization (EM) algorithm. We utilized the nlminb function in the R software to execute this algorithm. The nlminb function is an optimization method that can handle unconstrained as well as box-constrained problems, and it uses PORT routines, which are based on a Newton-like method, to find the MLE (Nash and Varadhan 2011).

2.2 Local false discovery rate for signal detection of MDRM

The local false discovery rate (LFDR) is defined as the probability that a given result belongs to the 'null distribution', which is a theoretical distribution of values that would be expected under the assumption that there is no real relationship between the variables being studied. This probability is calculated using posterior probability, which is the probability of an event occurring given the available evidence (Efron and Tibshirani 2002).

In the MDRM method, drug combinations are classified into two categories based on their potential give rise to adverse events. The first category is the 'null distribution', which refers to drug combinations that have a constant risk of causing adverse events. The second category is the drug-count response risk, which refers to combinations that have an increased risk of causing adverse events owing to certain risk factors, such as the number of drugs being consumed or the specific combination of drugs (Wang, Zhang et al. 2018). Signal detection was performed using the local false discovery rate method based on the

values of $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\pi}$, which MLE estimated through the EM algorithm.

LFDR is a measure of the probability that the risk of an adverse event will remain constant as the number of drugs in a combination increases. The LFDR cutoff is a predetermined threshold value that is used to determine which drug combinations are considered significant. In this case, the LFDR cutoff has been set at 0.0001 (Wang, Zhang et al. 2018). The LFDR for MDRM is given by $lfdr(y_{ik}) = \frac{(1-\pi)Bin(n_{ik},y_{ik},q_0)}{(1-\pi)Bin(n_{ik},y_{ik},q_0)+\pi Bin(n_{ik},y_{ik},q_1)}$ where $q_0 = \frac{exp(\beta_0)}{1+exp(\beta_0)}$, $q_1 = \frac{exp(\beta_0+\beta_1 i)}{1+exp(\beta_0+\beta_1 i)}$

3. Proposed method

3.1 Generalized linear mixed-effect model with two-levels

It is assumed that a combination drug generally involves a hierarchical structure. For instance, if drug A is taken alone and then drug B is added to it. Regardless of the order in which they are consumed, the resulting combination is drug A and B. It is assumed that both drug A alone and the drug A and drug B combination adhere to this hierarchical structure (See Figure 3.1). Therefore, it is assumed that clusters are formed depending on which drug is considered as a baseline drug. In this way, clusters are created as much as a single drug set, and to estimate the variation between clusters, a generalized linear mixed-effect model can be applied to perform multiple drug combination analysis. Regardless of the type of drug, if only the overall effect of whether the risk increases as the drug is considered, MDRM has an advantage in that it can predict the risk of various combinations between drugs.

By evaluating the baseline risk associated with each drug, we can more accurately predict the potential value of taking the medication. This is a useful feature of our model because it allows us to assess the potential risks of each treatment more accurately.

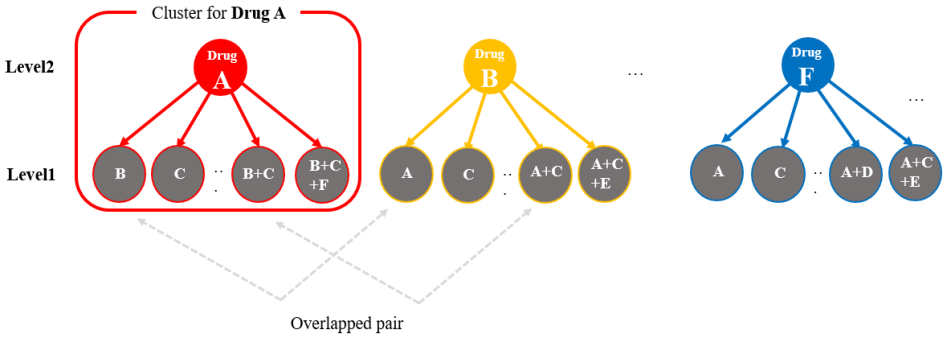


Figure 3.1 Nested structure for multiple combination drugs

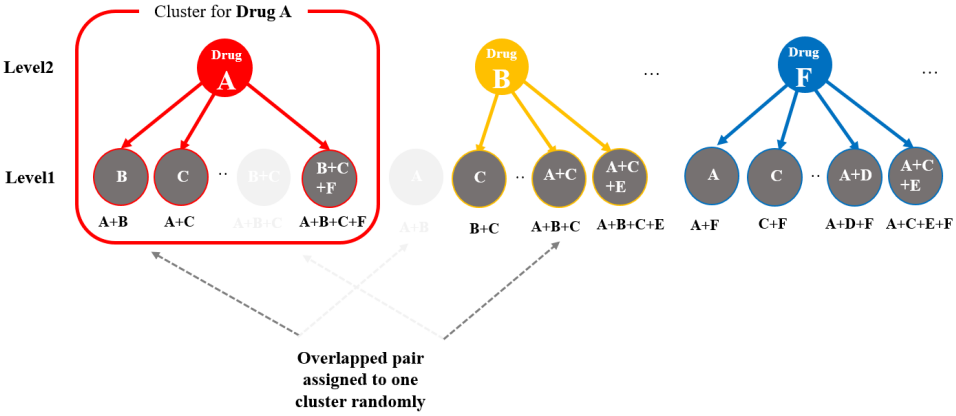


Figure 3.2 Nested structure for multiple combination drugs randomly assigned

In addition, this has the advantage that multiple estimates can be confirmed through the variance when a certain drug is included in the cluster. Not only can an overall fixed effect be seen, but unobserved hidden variance according to drug cluster can also be considered. This can be expressed as a formula as follows.

The formula is given by $\eta_{ij} = logit(\pi_{ij}) = logit(P(Y_{ij} = 1|\beta, b)) = \beta_0 + b_{0j} +$

$(\beta_1 + b_{1j})i$ where $\begin{bmatrix} b_{0j} \\ b_{1j} \end{bmatrix} = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix}\right)$ and β_0, β_1: fixed intercept and slope for drug combinations, i. The b_{0j} and b_{1j} are random intercept and slope for drug combinations i nested in cluster j. The cluster j denotes that the number of single candidate drugs ($j = 1,2,\ldots,N$) and drug combination i denotes that how many drug combination ($i = 1,2,3,4$) were considered under cluster j. The response variable Y_{ij} isassumed to be binomial distribution with probability π_{ij}

When analyzing a drug combination, such as drug A and B, as a cluster, rather than as individual drugs, it can be difficult to disentangle the overlapping issues that may affect both drugs. This can lead to issues while estimating parameters and creating incoherent constraints, as the effects of the individual drugs may be confounded with one another (Adam and Blockeel 2015). To determine the performance of our model for this overlapping pair, we evaluated the structure that is included in only one cluster using a randomly assigned set in a model without duplicated combinations.

3.2 Prediction interval for direct calculated method

We developed a formula for calculating the prediction interval for a generalized linear mixed model with a two-level structure. We focused specifically on the fixed and random slope terms for the drug combination, as this is what we were interested in examining. Based on our basic formula, $\eta_{ij} = logit(\pi_{ij}) = logit(P(Y_{ij} = 1|\beta, b)) = \beta_0 + b_{0j} + (\beta_1 + b_{1j})i$ where $\begin{bmatrix} b_{0j} \\ b_{1j} \end{bmatrix} = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix}\right)$, we expressed the prediction interval for the slope term of fixed and random effect. Based on the reference that the variance of the fixed and random terms in the existing mixed effect model is independent, we have composed the following prediction interval (compute.dtu.dk, 2022).

The 95% prediction interval for random slope is given by $(\hat{\beta}_1 + \hat{b}_{1j}) \pm z_{1-0.05/2} \times \sqrt{var(\hat{\beta}_1) + var(\hat{b}_{1j})}$ where the j denotes cluster. We aimed to detect the presence of a signal by determining whether the prediction interval, calculated based on the estimates and variances of the fixed and random effects, includes zero.

3.3 Prediction interval for bootstrap method

This is a reference based on a linear mixed-effect model in chapter 3.2, so there may be some limitations in directly applying it to a generalized linear mixed effect model. In case of calculation for the directed prediction interval, there may be an issue with the

accuracy of the estimate. Therefore, we have attempted to estimate the prediction interval through bootstrapping to overcome this limitation.

To obtain the prediction interval in the generalized mixed effect model, we proceeded with bootstrap using the replacement extraction in the dataset (middleprofessor, 2022). The prediction interval based on bootstrap below.

1. Resample n rows of the data with replacement where n is the original sample size
2. Compute estimated value $\left(\hat{\beta}_1 + \hat{b}_{1j}\right)$ for the fixed and random slope for each cluster, j
3. Repeat the step 1 and 2 for $m-1$ times where m denotes 1000 iterations
4. Given the distributions for the m estimated value and calculate the $100(1-\alpha/2)\%$ percentile. In case of 95% prediction interval, we considered 2.5% and 97.5% quantile. Considering it as a significant cluster if the prediction interval did not contain 0

4. Simulation

4.1 Data generation

We conducted a simulation study to compare the performance between two-level generalized linear mixed-effect model and conventional method, MDRM. For the simulation, specific adverse events were defined and generated from the FAERS data to fit a structure like the actual data. We constructed four-drug combinations with 30 base drugs as cluster. To proceed with such an analysis, a total of three types of data needs to be generated. Drug combination from single to four-drug combinations, total drug margin (M), event proportion (p), and count (c) according to drug margin and event proportion must be defined, and we divide the part corresponding to this value from the distribution. We extracted these values as per a specific distribution and proceeded with the simulation.

After constructing all possible combinations with the number of drugs based on GI bleeding, we checked the real-world percentage of data present among all possible combinations by considering the proportion of actual data based on the four-drug combinations. Simulation data was constructed by random sampling of all possible combinations. We first derived the results of each possible combination for the four combinations.

In the case of single drug, there are 30 possible cases, and in the case of two-drugs,

there are 435 cases $\left(=\binom{30}{2}\right)$. In the case of all possible three-drug combinations, 4060 cases $\left(=\binom{30}{3}\right)$ exist, and in the case of all possible combinations of four-drugs, 27405 cases $\left(=\binom{30}{4}\right)$ exist (Appendix table 1). Among them, we tried to extract a specific combination with the ratio from the adverse event of a specific database, and 30, 110, 30, and 7 cases were finally composed according to 100%, 25%, 0.73%, and 0.025% from single to four-drug combinations on actual data. We extracted a uniform distribution in case of drug margin, a beta distribution in case of proportion, and a binomial distribution for count variables where actual events occurred through this, and the factor used at this time was also extracted based on actual data. The drug margin was considered with reference to the actual range, and in the case of proportion, it was set based on the average value (E(p_1)=0.05) in the case of p_1 (See Appendix Table 1).

The ith drug margin, M_i is extracted from Uniform distribution, $Unif(a, b)$. And the p_i denotes the event proportion with average $\frac{\alpha}{\alpha+\beta}$ which extract from Beta distribution based on $\alpha \; and \; \beta$. And then, the ith count for adverse event, $c_{i,}$ extracted from Binomial distribution based on M_i and p_i where the *i* denotes for the number of *i*-way drug combinations.

4.2 Simulation setting

4.2.1 Simulation setting for performance

We configured the format of the data set in the specific case through the above process in Section 4.1. Based on this data configuration, we proceeded with the following data settings to compare the performances of generalized mixed-effect model with two-level as expressed in Figure 3.1 and the MDRM. Furthermore, although Figure 3.1 shows our original model, duplicated cases may be included, which can affect the performance. To compare the parts for this, we executed to compare the parts for the randomly selected cluster as a comparison group as expressed in Figure 3.2. This is not actually the model we intend to propose, but it is a method that can only be used as a suggested method for power explosion of our model.

We conducted a simulation study to compare the performance of the generalized mixed-effect model with two levels, generalized mixed-effect model with two levels randomly assigned, and MDRM. The steps were as follows,

1. Generate the drug margin (M) and the probability (p) of adverse event occurring when the drug is taken by $M_i \sim Unif(a,b)$ and $p_i \sim Beta(\alpha,\beta)$ in i-way drug combinations. If we generate M_i, p_i fixed from the distribution first, then generate count $c_i \sim Bin(M_i, p_i)$

2. Fix the drug margin with $M_1 \sim Unif$(1000,30000); $M_2 \sim Unif$(60,10000); $M_3 \sim Unif$(10,2000); $M_4 \sim Unif$(1,100) based on appendix table 1

3. We simulated by setting different event proportion rates

All data simulation is considered with $E(p_1)$=$E(p_2)$=$E(p_3)$=E(p_4), type I error, and if

at least one expected value is higher than $E(p_1)$, power, it confirms the performance. Here, $E(p_1)$ denotes the expected value for adverse events when the prescribed drug is single. Similarly, $E(p_2)$, $E(p_3)$ and $E(p_4)$ mean that the expected value when the prescribed drug is from two-, three-, and four-drug combinations, respectively.

These evaluation indicators for our simulation are defined as follows. Type I error is given by $\frac{\sum_{k=1}^{B} I(\hat{\beta}_{1k}>0, p_k<\alpha)}{B}$ where $\hat{\beta}_{1k}$ means for estimated β_1 on iteration k and p_k means for p values on iterations k, α =0.05 and $B = 1,2,\cdots,1000$ when the case is $E(p_1)$=$E(p_2)$=$E(p_3)$=E(p_4). Other setting except for the $E(p_1)$=$E(p_2)$=$E(p_3)$=E(p_4), the same index considered as power.

Based on $E(p_1) = 0.05$, the deviation of the difference from $E(p_2)\ to\ E(p_4)$, we show how the effect of our model is compared to the MDRM. All baseline risks are the same at $E(p_1) = 0.05$, and we set how they differ when the degree of difference was shown to be 0.05 to 0.12 depending on the drug combination.

Simulation design 1 has a risk of 0.05 for an event to occur in a single combination is 0.05, while the risk increases to 0.1 for two-drug combinations, 0.15 for 3-combinations, and 0.2 for 4-combinations. In simulation design 4, the risk of an event occurring in a single instance is also 0.05, but the risks for two-drug and four-drug combinations are higher at 0.17, 0.29, and 0.41, respectively. The risk of a four-drug combinations event in simulation 4 differs by more than 0.2 compared to simulation 1. From simulation 1 to 4, the difference denotes that $E(p_4) - E(p_3) = E(p_3) - E(p_2) = E(p_2) - E(p_1)$ with specific value,

such as $0.05, 0.07, 0.1$ and 1.2.

In the case, the baseline risk for all drugs is the same as 0.05, but only the degree of difference exists depending on the simulation. For $E(p_1)$ to $E(p_4)$, we show how the performance differs according to simulation design. As the risk of ADE occurring increases based on how much medicine is taken, we looked at the performance of the two-level mixed effect model and MDRM model that we propose.

The risk associated with taking a single drug may vary in real-world situations, and this is referred to as the baseline risk. To assess the effectiveness of a two-level mixed effect model, it is necessary to evaluate its performance under different baseline risks and when the risk difference between drug combinations increases significantly as the number of drugs being taken together increases.

Baseline risk was composed of a total of four types, which are considered from $E(p_1) = 0.05$ to $E(p_1) = 0.2$. In this way, as the baseline risk was different and the number of drug combinations increased, the degree of increase was divided into steady or sharply increases to examine the changes in the performance. The specific settings are listed in the Table 4.1.

Table 4.1 Simulation Design for risk based on different baselines

Simulation	**Baseline Risk**	**Drug Combinations**	**Increasing rate (Steady)**	**Increasing rate (Sharply)**
Simulation 1, 2 (Steady, Sharply)	$E(p_1)$ =0.05	1	0.05	0.05
		2	0.1	0.15

		3	0.15	0.3
		4	0.2	0.45
Simulation 3, 4 (Steady, Sharply)	$E(p_1)$ =0.1	1	0.1	0.1
		2	0.15	0.2
		3	0.2	0.35
		4	0.25	0.5
Simulation 5, 6 (Steady, Sharply)	$E(p_1)$ =0.15	1	0.15	0.15
		2	0.2	0.25
		3	0.25	0.4
		4	0.3	0.6
Simulation 7, 8 (Steady, Sharply)	$E(p_1)$ =0.2	1	0.2	0.2
		2	0.25	0.3
		3	0.3	0.45
		4	0.35	0.65

4.2.2 Simulation setting for signal detection

Simulation studies were conducted to evaluate the performance of our two-level mixed effect model and compare it with the MDRM. To evaluate the relative performance of our two-level mixed effect model and MDRM, we need to assess the ability of each model to detect signals in a realistic setting. This will involve examining the ability of each model to identify signals under various conditions.

In general, it is common to have false positives that say there is no actual signal and true positives that say there is a real signal. We have also demonstrated how our model can detect the signal where we set it up as an actual signal to calculate true positives. In this simulation, we compared the risk of true positive rate and false negative rate with that of the existing MDRM model.

$$\text{True Positive rate (TPR)} = \frac{\#\text{ of (detected Signal)}}{\#\text{ of (True Signal)}}$$

$$\text{False Positive rate (FPR)} = \frac{\#\text{ of (detected Signal)}}{\#\text{ of (Have no Signal)}}$$

We suggest indicators for two methods. The first refers to the bootstrap prediction interval obtained through the bootstrap method, while the second is the direct calculated PI obtained by the two-level mixed effect model. If the prediction interval that is set up does not include zero, we can conclude that it is a signal.

4.2.2.1 Signal Setting

We conducted signal setting to investigate how well MDRM and the two-level mixed-effect model can distinguish the combination set up as an actual signal in the actual signal. We set up the signal using the existing combination as is, and the related process is described below.

1. The total number of cases in the entire dataset ranges from single combinations to four-drug combinations, for a total of 177 cases. In the case of MDRM, 147 cases, excluding 30 single cases, are considered in total. In the case of the two-level mixed-effect model, a total of 30 clusters can be considered in total.

2. If the drug we have chosen is included in a specific cluster, the cluster corresponding to all cases with the set drug included is set as a signal in the two-level mixed effect model, and all cases with the set drug included are set as signals in MDRM. In the case of the two-level mixed effect model, the maximum number

of signals is equal to the number of clusters, which is 30, whereas in the case of MDRM, the maximum number of signals is all cases excluding single is 147. In the case of a cluster with a signal, it is set to have a signal, and for other cases, it is set to have no signal, and calculate the proportion.

3. As described in Section 4.1, random number generation was carried out with drug margin, proportion, and count.

4. The LFDR was calculated for the MDRM and bootstrap PI, and directed calculated PI was calculated for the two-level generalized mixed effect model to derive the ratio of false positive rate and true positive rate of how many sets were detected in our set signal.

4.3 Simulation result

4.3.1 Performance comparison

We consider all baseline probabilities as 0.05, and as the drug combinations increase, the part for the difference increases to 0.05, 0.07, 0.1, and 0.12, and how the performance changes compared to the existing method is confirmed. The results obtained from the simulated data set are presented in Table 4.2.

The type I errors for MDRM and the two-level mixed-effect model are 0.35 and 0.12, respectively. MDRM has a relatively high type I error compared to the other methods. Therefore, we can say that this MDRM method has a high false positive rate. Although it was not as high as the mixture model, the two-level mixed-effect model had a higher false positive rate than the general significance level ($\alpha = 0.05$). The false positives were observed in all existing clusters with pairs included. In the case of multiple combination drugs, the method that randomly included only a single cluster had a low false positive rate of 0.04.

In case of $E(\mathrm{p}_1) = \mathrm{E}(\mathrm{p}_2) = \mathrm{E}(\mathrm{p}_3) < E(p_4)$ when the difference between drug combinations is 0.05, the randomly assigned cluster value for power is still low with 0.05. As this part is included in the power, but the difference based on type I error is the only difference for $E(p_4)$, the value is likewise type I error. Similarly, the power for the two-level mixed effect model was not high at 0.14, and the power for MDRM was confirmed to be 0.383. This is because the proportion only increased in the case of four-drug

combinations, and the proportion was the same for other combinations. However, the number of cases included in the four-drug combinations was not large, so it did not result in proper power.

In case of $E(\mathrm{p}_1) = \mathrm{E}(\mathrm{p}_2) < \mathrm{E}(\mathrm{p}_3) = E(p_4)$, the power for the two-level mixed-effect model increased significantly to 0.99, compared to before. Similarly, the power for MDRM and the randomly assigned mixed-effect model also increased significantly to 0.71 and 0.85, respectively, compared to before. In the $E(\mathrm{p}_1) = \mathrm{E}(\mathrm{p}_2) < \mathrm{E}(\mathrm{p}_3) < E(p_4)$, the power of the mixed-effect model increased to 0.99, while the power of MDRM was found to be 0.71 which means that the probability of detecting the actual signal is approximately 70%.

Even when the difference in ratio was 0.07 ($\Delta E(p_i) = 0.07$), the power of MDRM was confirmed to be 0.36 in the case of $E(\mathrm{p}_1) = \mathrm{E}(\mathrm{p}_2) = \mathrm{E}(\mathrm{p}_3) < \mathrm{E}(p_4)$, and the power was confirmed to be 0.15 in the case of mixed-effect model with two-level. However, this case is possible because there is no significant difference in ratio from type I error. In this case, the situation was the same as the previous simulation with a difference of 0.05, where only the proportion of four-drug combinations was increased. It was confirmed that none of the algorithms worked properly in calculating power because of the low proportion of four-drug combinations.

It was observed that the power increased in all cases where the difference in drug combinations was extracted from three-drug combinations. In particular, the power of the two-level mixed effect model quickly converged to 1 as the degree of difference increased.

Overall, it was possible to confirm that the power increased properly in all cases except when only the proportion of four-drug combinations was increased. In addition, it was always seen that the power was set high in our proposed two-level mixed model, except for the case where only the proportion of four-drug combinations was increased in the MDRM model, which suggests that the performance of our model is better than that of MDRM.

In summary, the power of the two-level mixed effect model increased significantly compared to the MDRM model, especially when the difference between drug combinations increased. This suggests that our proposed two-level mixed effect model is more effective in detecting significant clusters compared to the MDRM model. It was also observed that the power of the two-level mixed effect model increased when there were multiple clusters containing the same drug combination, indicating that the model is robust and not simply influenced by sample size. Overall, our results show that the two-level mixed effect model is a valid and reliable approach for identifying significant clusters in the analysis of drug combination studies.

Table 4.2 Type 1 error, power obtained by these two methods (including randomly assigned cluster) according to the simulation designs from 1 to 4, which are the mean difference of proportion depend on drug combinations from 0.05 to 0.12

Simulation design	Scenario	Power*		
		MDRM	Mixed-effect model with two-level randomly assigned**	Mixed-effect model with two-level
Simulation Design1 $(\Delta E(p_i) = 0.05)$	$E(p_1)=E(p_2)=E(p_3)=E(p_4)$	0.35	0.04	0.12
	$E(p_1)=E(p_2)=E(p_3)<E(p_4)$	0.38	0.05	0.14
	$E(p_1)=E(p_2)<E(p_3)=E(p_4)$	0.71	0.85	0.99
	$E(p_1)=E(p_2)E(p_3)<E(p_4)$	0.71	0.89	0.99
	$E(p_1)<E(p_2)<E(p_3)<E(p_4)$	1	1	1
Simulation Design2 $(\Delta E(p_i) = 0.07)$	$E(p_1)=E(p_2)=E(p_3)<E(p_4)$	0.36	0.04	0.15
	$E(p_1)=E(p_2)<E(p_3)=E(p_4)$	0.82	0.87	0.99
	$E(p_1)=E(p_2)<E(p_3)<E(p_4)$	0.84	0.9	0.99
	$E(p_1)<E(p_2)<E(p_3)<E(p_4)$	1	1	1
Simulation Design3 $(\Delta E(p_i) = 0.1)$	$E(p_1)=E(p_2)=E(p_3)<E(p_4)$	0.35	0.06	0.19
	$E(p_1)=E(p_2)<E(p_3)=E(p_4)$	0.93	0.97	1
	$E(p_1)=E(p_2)<E(p_3)<E(p_4)$	0.93	0.97	1
	$E(p_1)<E(p_2)<E(p_3)<E(p_4)$	1	1	1

Simulation design	Scenario	Power*		
		MDRM	Mixed-effect model with two-level randomly assigned**	Mixed-effect model with two-level
Simulation Design4 $(\Delta E(p_i) = 0.12)$	$E(p_1)=E(p_2)=E(p_3)<E(p_4)$	0.35	0.08	0.21
	$E(p_1) = E(p_2) < E(p_3) = E(p_4)$	0.96	0.99	1
	$E(p_1)=E(p_2)<E(p_3)<E(p_4)$	0.97	0.99	1
	$E(p_1)<E(p_2)<E(p_3)<E(p_4)$	1	1	1

* This refers to type I error when the β_1 is zero (it also means $p_1=p_2=p_3=p_4$)
** In the case of a combination drug, the combination pairs are assigned within a cluster randomly

We consider all baseline probabilities as 0.05 in previous analysis. However, we also consider different baseline risks. In addition to changing the baseline risk, we also differentiated the degree of difference. We compared the power of the MDRM and mixed-effect models by separating the difference in drug combinations into two categories: steady and sharply changing. We conducted this comparison in the same manner for both cases. The results obtained from the simulated data set presented in Table 4.3.

In the sharply increasing case with an average proportion difference of 0.1, the type I errors for MDRM and the two-level mixed-effect model were 0.14 and 0.06, respectively. This is lower than the type I error when the baseline risk was higher. The MDRM method still has a relatively high type I error compared to the other methods, which means it still has a higher false positive rate. It was not as high as the mixture model, the mixed-effect model with two levels was higher than the general significance level ($\alpha = 0.05$). This indicates that the MDRM method may not be as reliable in terms of identifying significant drug combinations compared to the two-level mixed-effect model. In the simulated dataset, the mixed-effect model with two level that was randomly selected had a low false positive rate of 0.02.

In the sharply increasing case of $\mathrm{E}(p_1) = E(p_2) = \mathrm{E}(p_3) < \mathrm{E}(p_4)$ when the baseline risk of $E(p_1)$ is 0.1, the randomly assigned cluster for power is still low as 0.08. As this part is included in the power, but the difference based on type I error is the only difference for $\mathrm{E}(p_4)$, the value is likewise type I error. This is also because the number of cases included in four-drug combinations was not large, so when the proportion of four-drug

combinations was increased, it could not be properly detected owing to the characteristics of the model. This is a characteristic that can be seen in all results in general.

The mixed-effect model with two levels produced lower type I error as 0.09 compared to type I error of the MDRM method with 0.14. This may be because the signal was not particularly strong in this case, as indicated by the assumption of equal expected type I error rates $E(p_1) = E(p_2) = E(p_3) = E(p_4)$. However, it is worth noting that the mixed-effect model with two levels may have higher power owing to the presence of redundant data in the analysis.

Overall, the analysis showed that convergence occurred more quickly when the proportion of drug combinations was changed significantly, compared to when the proportion remained steady. This was true for 0.15 as well as 0.2 values. However, the overall trend was like that seen at other points, except for a larger portion when the type 1 error value was 0.2, compared to the baseline risk of 0.15. This indicates that the model is functioning effectively, regardless of the baseline risk.

Table 4.3 Type 1 error, and power obtained by two methods (including randomly assigned cluster) according to the simulation designs from 1 to 8, which are baseline risks from 0.05 to 0.2

Simulation design	**Scenario**	**Power***		
		MDRM	**Mixed-effect model with two-level randomly assigned****	**Mixed-effect model with two-level**
Simulation Design1 ($E(p_1) = 0.05$, *steady*)	$E(p_1) = E(p_2) = E(p_3) = E(p_4)$	0.36	0.02	0.13
	$E(p_1) = E(p_2) = E(p_3) < E(p_4)$	0.35	0.04	0.17
	$E(p_1) = E(p_2) < E(p_3) = E(p_4)$	0.71	0.86	0.87
	$E(p_1) = E(p_2) < E(p_3) < E(p_4)$	0.74	0.90	0.9
	$E(p_1) < E(p_2) < E(p_3) < E(p_4)$	1	1	1
Simulation Design2 ($E(p_1) = 0.05$, *sharply*)	$E(p_1) = E(p_2) = E(p_3) = E(p_4)$	0.36	0.02	0.13
	$E(p_1) = E(p_2) = E(p_3) < E(p_4)$	0.36	0.05	0.17
	$E(p_1) = E(p_2) < E(p_3) = E(p_4)$	0.94	0.99	0.99
	$E(p_1) = E(p_2) < E(p_3) < E(p_4)$	0.94	0.99	1
	$E(p_1) < E(p_2) < E(p_3) < E(p_4)$	1	1	1
Simulation Design3 ($E(p_1) = 0.1$, *steady*)	$E(p_1) = E(p_2) = E(p_3) = E(p_4)$	0.14	0.03	0.06
	$E(p_1) = E(p_2) = E(p_3) < E(p_4)$	0.16	0.05	0.08
	$E(p_1) = E(p_2) < E(p_3) = E(p_4)$	0.74	0.95	0.96

Simulation design	Scenario	Power*		
		MDRM	Mixed-effect model with two-level randomly assigned**	Mixed-effect model with two-level
	$E(p_1) = E(p_2) < E(p_3) < E(p_4)$	0.97	1	1
	$E(p_1) < E(p_2) < E(p_3) < E(p_4)$	1	1	1
Simulation Design4 ($\boldsymbol{E(p_1) = 0.1}$, ***sharply***)	$E(p_1) = E(p_2) = E(p_3) = E(p_4)$	0.14	0.03	0.06
	$E(p_1) = E(p_2) = E(p_3) < E(p_4)$	0.17	0.08	0.09
	$E(p_1) = E(p_2) < E(p_3) = E(p_4)$	0.98	1	1
	$E(p_1) = E(p_2) < E(p_3) < E(p_4)$	0.98	1	1
	$E(p_1) < E(p_2) < E(p_3) < E(p_4)$	1	1	1
Simulation Design5 ($\boldsymbol{E(p_1) = 0.15}$, ***steady***)	$E(p_1) = E(p_2) = E(p_3) = E(p_4)$	0.16	0.04	0.06
	$E(p_1) = E(p_2) = E(p_3) < E(p_4)$	0.21	0.04	0.06
	$E(p_1) = E(p_2) < E(p_3) = E(p_4)$	0.60	0.83	0.78
	$E(p_1) = E(p_2) < E(p_3) < E(p_4)$	0.59	0.85	0.81
	$E(p_1) < E(p_2) < E(p_3) < E(p_4)$	1	1	1
Simulation Design6	$E(p_1) = E(p_2) = E(p_3) = E(p_4)$	0.19	0.04	0.06
	$E(p_1) = E(p_2) = E(p_3) < E(p_4)$	0.22	0.06	0.09
	$E(p_1) = E(p_2) < E(p_3) = E(p_4)$	0.90	0.99	0.99

Simulation design	Scenario	Power*		
		MDRM	Mixed-effect model with two-level randomly assigned**	Mixed-effect model with two-level
$(\boldsymbol{E(p_1) = 0.15}, \boldsymbol{sharply})$	$E(p_1) = E(p_2) < E(p_3) < E(p_4)$	0.93	0.99	1
	$E(p_1) < E(p_2) < E(p_3) < E(p_4)$	1	1	1
Simulation Design7 $(\boldsymbol{E(p_1) = 0.2}, \boldsymbol{steady})$	$E(p_1) = E(p_2) = E(p_3) = E(p_4)$	0.24	0.03	0.06
	$E(p_1) = E(p_2) = E(p_3) < E(p_4)$	0.22	0.04	0.06
	$E(p_1) = E(p_2) < E(p_3) = E(p_4)$	0.53	0.68	0.62
	$E(p_1) = E(p_2) < E(p_3) < E(p_4)$	0.58	0.73	0.66
	$E(p_1) < E(p_2) < E(p_3) < E(p_4)$	1	1	1
Simulation Design8 $(\boldsymbol{E(p_1) = 0.2}, \boldsymbol{sharply})$	$E(p_1) = E(p_2) = E(p_3) = E(p_4)$	0.24	0.03	0.06
	$E(p_1) = E(p_2) = E(p_3) < E(p_4)$	0.27	0.06	0.08
	$E(p_1) = E(p_2) < E(p_3) = E(p_4)$	0.85	0.98	0.97
	$E(p_1) = E(p_2) < E(p_3) < E(p_4)$	0.86	0.99	0.98
	$E(p_1) < E(p_2) < E(p_3) < E(p_4)$	1	1	1

* It means type I error when $p_1=p_2=p_3=p_4$
** In the case of a combination drug, the combination pairs are assigned within a cluster randomly

4.3.2 Signal detection comparison

To determine if a signal was significant, we calculated the false positive rate and true positive rate by comparing the bootstrap and the directed calculated prediction interval obtained through the two-level mixed effect model and compared the results to the LFDR, a signal detection indicator in MDRM.

The process of setting the signal is described in detail in Section 4.2.2. The process of setting the signal in the two-level mixed effect model is described below. After setting one cluster, select all cases that include the cluster we set among 177 cases, and set all clusters included in the case to signal. At this time, all cases that contain the cluster we set are themselves set as signals in the MDRM. In this way, the denominators of the signals of the two methods are set differently. Therefore, we tried to compare TPR and FPR for the cases where 11 clusters and 17 clusters were set as signals, respectively.

At first, the case where there was no signal was examined based on the entire cluster, and based on the Bootstrap PI, only about 0.04 of the totals were falsely detected as having a signal. The directed calculated PI, was 0.04, showing no significant difference from the method obtained by bootstrapping. In the case of MDRM, a relatively high type 1 error of 0.23 was confirmed.

When the number of signals was set as 11, the case where only the ratio of p_4 was varied, TPR was 0.32 and 0.13 in MDRM and two-level mixed effect model, respectively, showing no significant difference compared to type 1 error. When the ratio of $E(p_3)$ was

increased, it was confirmed that it increased to 0.5 in the case of the two-level mixed effect model. In addition, it was confirmed that the degree increased from 0.51 to 0.58 when $E(p_3) < E(p_4)$ with 0.15 compared to when $E(p_3) = E(p_4)$ was set to 0.1. In contrast, in MDRM, it was found that the values were not very high, i.e., 0.28 and 0.3. This confirms that the power of MDRM is not very high when the difference in the ratio is not large; rather, the two-level mixed effect model works better in this case.

However, when the differences in all cases except p_1 were widened, MDRM's TPR was 0.65 and 0.66, which was higher than the two-level mixed effect model's TPR with 0.4 and 0.60, confirming that MDRM's detection rate was also high. However, the FPR is also 0.15 and 0.12, which is higher than the FPR of two-level mixed effect, 0.09, so the overall value is highly estimated. However, in the directed calculated PI, it was difficult to confirm with a comparative index because it could be seen that the TPR was generally low. This tendency was similar even when more than half of the 30 cases with 17 were signals. When only the ratio of p_4 was varied, the TPR was 0.26 and 0.09 in the MDRM and two-level mixed effect model, respectively. However, the ratio was lower than that of 11 signals.

When the ratio of $E(p_3)$ was increased, it was confirmed that it increased to 0.61 in the case of the two-level mixed effect model. In addition, it was confirmed that the degree increased from 0.61 to 0.68 when $E(p_3) < E(p_4)$ with 0.15 compared to when $E(p_3) = E(p_4)$ was set to 0.1. In contrast, in MDRM, it was found that the values were not very high, i.e., 0.26 and 0.27. This confirms that the power of MDRM is not very high when the difference in the ratio is small, the two-level mixed effect model works better in this case.

However, when the differences in all cases except p_1 were large, MDRM's TPR was 0.79 and 0.85, which was higher than the two-level mixed effect model's TPR with 0.53 and 0.73, confirming that MDRM's detection rate was also high. In conclusion, it can be concluded that MDRM works well when the ratio is large enough, but FPR is high, and when the comparison deviation is not large, our model, the two-level mixed effect model, works well.

Table 4.4 Average detected proportion on MDRM LFDR, bootstrap and direct PI based on two-level mixed effect model for iteration 50

Number of True signals	p_1	p_2	p_3	p_4	Detected proportion* by LFDR of MDRM	Detected proportion* by bootstrap PI of two-level mixed effect model	Detected proportion* by directed PI of two-level mixed effect model
No signal	$E(p_1)=0.05$	$E(p_2)=0.05$	$E(p_3)=0.05$	$E(p_4)=0.05$	33.20/147 0.23	1.08/30 0.04	1.15/30 0.04
11 (Signal)	$E(p_1)=0.05$	$E(p_2)=0.05$	$E(p_3)=0.05$	$E(p_4)=0.1$	22.14/139, 2.53/8 (0.16, 0.32)	1.88/19, 1.47/11 (0.10, 0.13)	1.23/19, 1.17/11 (0.07, 0.11)
	$E(p_1)=0.05$	$E(p_2)=0.05$	$E(p_3)=0.1$	$E(p_4)=0.1$	18.42/139, 2.22/8 (0.13, 0.28)	2.71/19, 5.60/11 (0.14, 0.51)	1.24/19, 2.18/11 (0.07, 0.20)
	$E(p_1)=0.05$	$E(p_2)=0.05$	$E(p_3)=0.1$	$E(p_4)=0.15$	24.83/139, 2.40/8 (0.18, 0.3)	2.71/19, 6.43/11 (0.14, 0.58)	1.13/19, 2.71/11 (0.06, 0.25)
	$E(p_1)=0.05$	$E(p_2)=0.05$	$E(p_3)=0.1$	$E(p_4)=0.2$	18.74/139, 2.42/8 (0.14, 0.30)	2.89/19, 6.40/11 (0.15, 0.58)	1.17/19, 2.51/11 (0.06, 0.23)
	$E(p_1)=0.05$	$E(p_2)=0.1$	$E(p_3)=0.15$	$E(p_4)=0.2$	20.67/139, 5.20/8 (0.15, 0.65)	1.74/19, 4.34/11 (0.09, 0.40)	1.00/19, 4.00/11 (0.05, 0.36)
	$E(p_1)=0.05$	$E(p_2)=0.1$	$E(p_3)=0.2$	$E(p_4)=0.3$	16.31/139, 5.29/8 (0.12, 0.66)	1.72/19, 6.64/11 (0.09, 0.60)	-
17 (Signal)	$E(p_1)=0.05$	$E(p_2)=0.05$	$E(p_3)=0.05$	$E(p_4)=0.1$	25.14/132, 3.92/15 (0.19, 0.26)	1.27/13, 1.59/17 (0.10, 0.09)	1.30/13, 1.13/17 (0.1, 0.07) -

Number of True signals	p_1	p_2	p_3	p_4	Detected proportion* by LFDR of MDRM	Detected proportion* by bootstrap PI of two-level mixed effect model	Detected proportion* by directed PI of two-level mixed effect model
	$E(p_1)=0.05$	$E(p_2)=0.05$	$E(p_3)=0.1$	$E(p_4)=0.1$	20.86/132, 3.88/15 (0.16, 0.26)	3.60/13, 10.4/17 (0.28, 0.61)	1.25/13, 3.22/17 (0.10, 0.19)
	$E(p_1)=0.05$	$E(p_2)=0.05$	$E(p_3)=0.1$	$E(p_4)=0.15$	22.31/132, 4.05/15 (0.17, 0.27)	3.69/13, 11.56/17 (0.28, 0.68)	1.17/13, 3.56/17 (0.09, 0.21)
	$E(p_1)=0.05$	$E(p_2)=0.05$	$E(p_3)=0.1$	$E(p_4)=0.2$	25.85/132, 4.5/15 (0.20, 0.3)	3.17/13, 11.15/17 (0.24, 0.66)	1.20/13, 3.08/17 (0.092, 0.18)
	$E(p_1)=0.05$	$E(p_2)=0.1$	$E(p_3)=0.15$	$E(p_4)=0.2$	27.24/132, 11.88/15 (0.21, 0.79)	2.347/13, 8.98/17 (0.18, 0.53)	1.00/13, 5.20/17 (0.08, 0.31)
	$E(p_1)=0.05$	$E(p_2)=0.1$	$E(p_3)=0.2$	$E(p_4)=0.3$	26.37/132, 12.76/15 (0.2, 0.85)	2.878/13, 12.38/17 (0.22, 0.73)	-

* the number of detected signals/ the number of not true signals (False Positive Rate), the number of detected signal/ the number of true signal (True Positive Rate)

5. Application

We apply our proposed method to the MOACDM database, which is based on the EHR data from hospitals whose database is based on the Korea Institute of Drug Safety & Risk Management. In this case, there are possible to approach for the adverse event drug signal. Also, this kind of database if based on the common data model. It can be possible to expand the overall data set. If this data structure shares the same structure. So, it can have good characteristics, if necessary. However, our study is for based on the severance hospital for CDM database approach.

Our approach can be used to predict drug-induced hepatotoxicity from datasets. We include a listing of drugs previously known to be hepatotoxic and FDA Sentinels which are meloxicam, celecoxib, valproic acid, and lamotrigine. Nine drugs which are azilsartan, candesartan, eprosartan, fimasartan, irbesartan, losartan, lomesartan, telmisartan, and valsartan. We applied to our model to these drugs to detect the multiple drugs signal.

A total of 13 drugs were used as the basic drug group, the case of combined use of the basic drug and additional doses was examined, and a total of four-drug combinations were used. As there was no case of lomesartan being taken, and there was no case of taking a single drug alone in the case of azilsartan and valproic acid, they are excluded from the target group.

The case of taking the 13 drugs that became the baseline drug within 60 days before and after taking the drug was fixed as a combination drug, and the event for the case of hepatotoxicity up to 60 days based on the date of taking the baseline drug was reported. When considering the most complex drug based on the baseline drug, if the drug was consumed after an adverse event hepatotoxicity occurred, it was excluded from the target group and analyzed.

In this application, patients who have undergone at least two aspartate transaminase (AST) and alanine transferase (ALT) lab tests due to the relevant medication are targeted, and patients whose baseline test results fall within the normal range are targeted. The definition of hepatotoxicity was based on the Mini-sentinel's diagnosis-based definition of hepatotoxicity, and an event was defined as any of the following criteria being met: ALT increasing by more than five times the upper limit of normal, alkaline phosphatase (ALP) increasing by more than two times, ALT increasing by more than three times, and total bilirubin increasing by more than two times compared to baseline (Patel et al., 2016).

We considered additional doses based on baseline medications. Based on the same list, the combination risk of related drugs was confirmed. We considered the case of taking a combination with a drug known to have hepatotoxicity and focused on patients who took up to four-drug combinations.

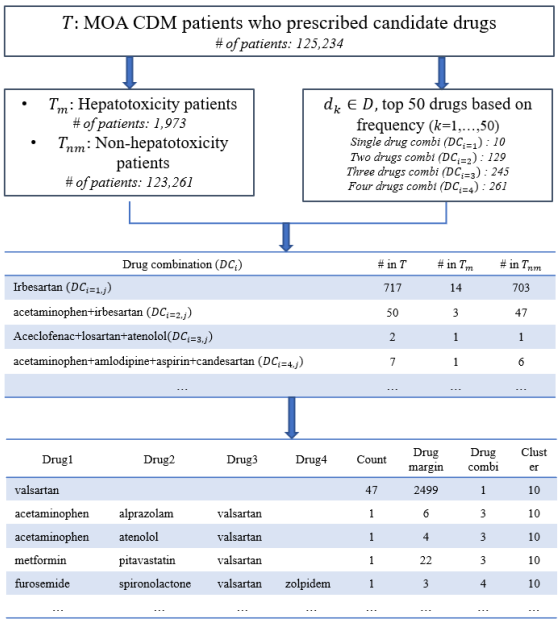
In the group of drug exposure for the total set, we consider the drug component name for the PO, which is the medication is taken by mouth ‘bid’ or twice a day. At based on this

list, we include our candidate for the top50 drugs based on the number of drug frequency. The list below is the top10 drug component list for including our data analysis.

Table 5.1 Top10 Combination drug list based on frequency

Drug	Records
Aspirin	22782
Amlodipine	21624
Clopidogrel	21418
Acetaminophen	18787
Magnesium	17871
Atovarstatin	14667
Rosuvastatin	13727
Furosemide	13427
Prednisolone	13025
Nicorandil	10488

We consider hepatotoxicity as the outcome for our application. To apply the proposed method, we used the MOACDM database based on hepatotoxicity as an example. As the representative diagnosis is known to have a certain level of risk due to drugs, the evidence for applying this diagnosis may be clear. All drugs having data reported four-drug combinations were considered to identify the top 50 drugs among them, and local FDR was used as a signal detection index.



Drug combination (DC_i)	# in T	# in T_m	# in T_{nm}
Irbesartan ($DC_{i=1,j}$)	717	14	703
acetaminophen+irbesartan ($DC_{i=2,j}$)	50	3	47
Aceclofenac+losartan+atenolol($DC_{i=3,j}$)	2	1	1
acetaminophen+amlodipine+aspirin+candesartan ($DC_{i=4,j}$)	7	1	6
…	…	…	…

Drug1	Drug2	Drug3	Drug4	Count	Drug margin	Drug combi	Cluster
valsartan				47	2499	1	10
acetaminophen	alprazolam	valsartan		1	6	3	10
acetaminophen	atenolol	valsartan		1	4	3	10
metformin	pitavastatin	valsartan		1	22	3	10
furosemide	spironolactone	valsartan	zolpidem	1	3	4	10
…	…	…	…	…	…	…	…

Figure 5.1 Overall dataset flow

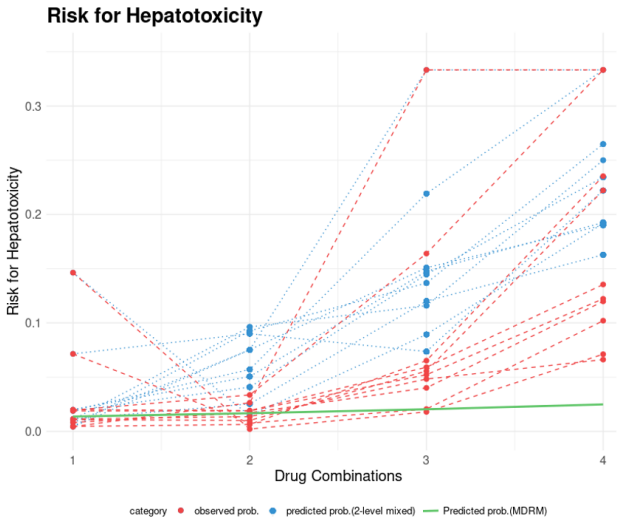
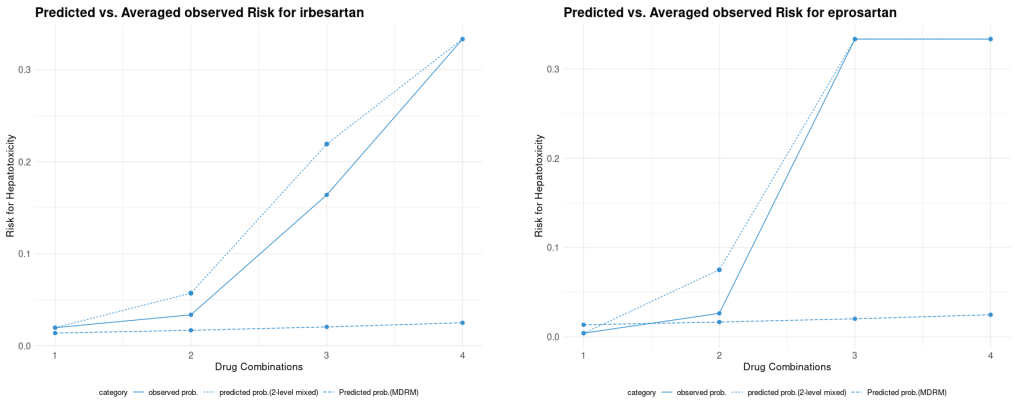


Figure 5.2 Risk Prediction graph for overall clusters

* Thick green line is for predicted risk of MDRM

Based on the real-word database, we showed the absolute value of differences between predicted and observed values. The graph below shows how the absolute difference between observed and predicted data, respectively, differs as the drug combination increases. As a result, when the drug combination was low, there was no significant difference between MDRM or other methods. Therefore, it can be said that MDRM does not properly reflect the observed risk even after the optimization process, and it can be confirmed that the risk in the case of the generalized linear mixed-effect model without duplication has the smallest difference compared to the existing observed risk.



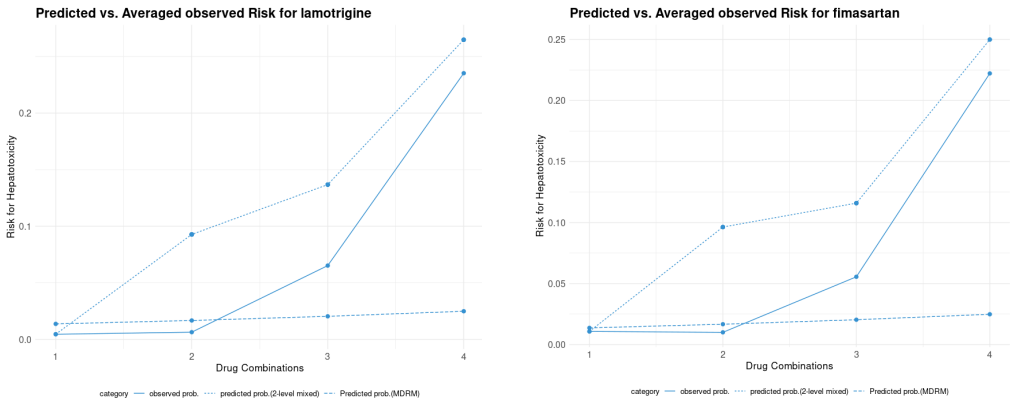


Figure 5.3 Difference between observed and predicted risk by drug combinations

According to Figure 5.3, in the case of the predicted risk of MDRM always depends only on all drug combinations and shows the same risk in other cases. In the case of randomly assigned to one cluster, it could be seen that the prediction was relatively like the observed risk. In the case of metoprolol and metformin, it was possible to observe that virtually less bias exists, and in the case of metformin and methotrexate, the predicted risk in the MDRM case was estimated to be higher than the observed risk. The actual mixed-effect model worked well. Thus, we could demonstrate that our mixed-effect model works well fit with real data.

To obtain prediction intervals for each cluster based on fitting with a two-level mixed effect model, we determined whether it was a signal or not depending on whether the prediction interval included 0 or not. As a result, it was confirmed that meloxicam and telmisartan were not significant. It was confirmed that it was a significant signal in the

other eight clusters. When we confirmed that the upper bound of the risk for the lamotrigine-related cluster composition rose to 2.12, we were able to confirm that the risk for the lamotrigine-related compound composition had the strongest effect on hepatotoxicity.

Table 5.2 Result of signal detection based on the two-level mixed effect model

Main drug	Bootstrap lower bound (2.5% quantile)	Bootstrap upper bound (97.5% quantile)
Candesartan	0.53	1.23
Celecoxib	0.71	1.25
Eprosartan	0.62	1.66
Fimasartan	0.59	1.81
Irbesartan	0.59	1.66
Lamotrigine	0.87	2.12
Losartan	0.46	1.37
Meloxicam	-1.18	2.78
Telmisartan	-0.35	2.07
Valsartan	0.48	1.31

In contrast, when the same data was analyzed in MDRM, the result of signal detection according to each combination was obtained. The list below has been found to be significant. Notably, in the case of the meloxicam related cluster, it was determined that it was not significant in the two-level mixed effect model, but here it was determined as a significant combination. Clinically, meloxicam may cause hepatotoxicity, but the extent is

not severe, so there is a question as to whether the meloxicam-related combination detected through the LFDR index can be used as a clinical judgment criterion.

Table 5.3. Result of signal detection based on MDRM

Drug1	Drug2	Drug3	Drug4	count	Drug margin	drug combinations	LFDR
candesartan				36	3151	1	<0.0001
acetaminophen	candesartan			8	122	2	<0.0001
celecoxib	oxycodone hydrochloride			5	60	2	<0.0001
celecoxib	furosemide	spironolactone		8	28	3	<0.0001
celecoxib	furosemide	magnesium	spironolactone	3	7	4	<0.0001
celecoxib	furosemide	spironolactone	ursodiol	5	30	4	<0.0001
irbesartan				14	717	1	<0.0001
losartan				43	2150	1	<0.0001
acetaminophen	losartan			13	157	2	<0.0001
Meloxicam				12	82	1	<0.0001
Meloxicam	metoclopramide	telmisartan		14	50	3	<0.0001
Meloxicam	metoclopramide	telmisartan	ursodiol	11	25	4	<0.0001
Telmisartan				19	266	1	<0.0001
Amlodipine	telmisartan			19	1453	2	<0.0001
valsartan				47	2499	1	<0.0001
acetaminophen	valsartan			17	150	2	<0.0001

6. Discussion and Conclusion

Our proposed method has fewer false positives compared to the conventional method (MDRM) in the high-dimensional DDI structure, even if the power was estimated to be higher. What this means is that the structure of multiple combination drug considering the cluster variation. For SRS data among existing high-dimension DDIs, our approach can be a good alternative. MDRM has the advantage of estimating the parameters by estimating the MLE through the EM algorithm, but it is dependent only on the number of drug combinations and assumes that it linearly increasing. As it was not possible to connect the estimation according to the type, it is thought that this study has with sufficient advantages in this area.

As this probability itself is different for each single drug, it is better than the existing MDRM method, which estimated the same probability regardless of single drug status. MDRM method provides a formula that can determine signal detection through LFDR. We presented two signal detection indices: Directed calculated PI and Bootstrap PI. In the case of Bootstrap PI, although there were differences, it showed higher power compared to LFDR when the difference was not large, and it was proven to be a useful indicator for detecting signal clusters through our two-level mixed effect model's bootstrap PI, even if the difference was not noticeable. This can be presented as a meaningful indicator in the context of polypharmacy. It is a novelty in this study.

However, this two-level mixed effect model also has limitations. As only two-level is currently considered, the same probability is still estimated among those who consumed the same number of drugs within the cluster. This aspect needs to be improved, and it seems to be a task to be able to apply the part to know the actual drug intake rate or the indication that the drug was taken.

In the future, we will design a three-level mixed-effect model based on the drug combination. One limitation is that tests based on a two-level mixed-effect model only verify for the level base drug and combination. This makes it difficult to predict adverse events for the specific combinations. Consequently, we cannot verify the specific combination. In a future study, we will focus on verifying the specific risk prediction for the drug combination set. We need to develop the signal detection method for combination drugs more than three-level mixed model.

References

Adam, A. and H. Blockeel (2015). Dealing with overlapping clustering: a constraint-based approach to algorithm selection. Proceedings of the 2015 International Conference on Meta-Learning and Algorithm Selection - Volume 1455. Porto, Portugal, CEUR-WS.org**:** 43–54.

Alomar, M., et al. (2020). "Post marketing surveillance of suspected adverse drug reactions through spontaneous reporting: current status, challenges and the future." Therapeutic advances in drug safety **11**: 2042098620938595-2042098620938595.

Efron, B. and R. Tibshirani (2002). "Empirical bayes methods and false discovery rates for microarrays." Genetic Epidemiology **23**(1): 70-86.

Harrison, X. A., et al. (2018). "A brief introduction to mixed effects modelling and multi-model inference in ecology." PeerJ **6**: e4794.

Ho, T. B., et al. (2016). "Data-driven Approach to Detect and Predict Adverse Drug Reactions." Curr Pharm Des **22**(23): 3498-3526.

Ibrahim, H., et al. (2021). "Signal Detection in Pharmacovigilance: A Review of Informatics-driven Approaches for the Discovery of Drug-Drug Interaction Signals in Different Data Sources." Artificial Intelligence in the Life Sciences **1**: 100005-100005.

Martin, I., et al. (2019). "The mixed model for the analysis of a repeated-measurement multivariate count data." Statistics in Medicine **38**(12): 2248-2268.

Nash, J. C. and R. Varadhan (2011). "Unifying Optimization Algorithms to Aid Software System Users: optimx for R." Journal of Statistical Software **43**(9): 1 - 14.

Nguyen, D. A., et al. (2019). "A survey on adverse drug reaction studies: data, tasks and machine learning methods." Brief Bioinform.

Ventola, C. L. (2018). "Big Data and Pharmacovigilance: Data Mining for Adverse Drug Events and Interactions." P & T : a peer-reviewed journal for formulary management **43**(6): 340-351.

Wang, X., et al. (2020). "Propensity score-adjusted three-component mixture model for drug-drug interaction data mining in FDA Adverse Event Reporting System." Statistics in Medicine **39**(7): 996-1010.

Wang, X., et al. (2018). "Mixture drug-count response model for the high-dimensional drug combinatory effect on myopathy." Stat Med **37**(4): 673-686.

Patel, U. D., N. C. Hardy, and D. H. Smith. "Validation of acute kidney injury cases in the mini-sentinel distributed database." Validation of health outcomes details. Accessed October 27 (2013): 2016.

Mixed Model Theory . (2022). Retrieved from http://www2.compute.dtu.dk/courses/02429/enotepdfs/eNote-10.pdf.

Best practice issues in inference . (2022). Retrieved from https://www.middleprofessor.com/files/applied-biostatistics_bookdown/_book/best-practices-issues-in-inference.html.

Appendix Table 1. Baseline table for the simulation

Drug combinations	Drug Margin Approximate range	Event Proportion Approximate range	Observed pair	Possible pair
1	1000 ~ 30000	0.0008 ~ 0.11	30	30
2	60 ~ 1000	0.007 ~ 0.17	110	435
3	10 ~ 2000	0.02 ~ 0.35	30	4060
4	1 ~ 120	0.1 ~ 0.5	7	27405

국 문 요 약

약물 조합에 따른 부작용 탐지를 위한 일반화 선형 혼합 모형을 이용한 위험 예측

복합약물 복용에 따라 유발된 약물 부작용에 대한 연구는 약물-약물 상호 작용과 관련된 중요한 이슈로 떠오르고 있다. 노인 연령일수록 약물에 대해 복합 약물을 복용하는 경우는 점점 더 높아지고, 최근 시기일수록 3개 이상의 복합 약물을 복용하는 경우도 증가하고 있다. 그러나 기존의 통계방법론에 따라 주로 집중되고 있는 연구는 두개의 약물 조합의 부작용에 집중되어 있기 때문에, 고차원 약물 상호작용의 위험 예측 방법에 대한 연구가 부족한 실정이다. 기존에 복합약물에 대한 혼합 약물 반응 모델이 존재하고 이의 경우에는 여러 개의 복합 약물의 경우를 고려하지만, 이 또한 약물의 개수에 따라 선형적으로 위험을 가정하여 추정할 뿐, 약물 자체의 특성에 대한 고려를 포함하여 모델링 하고 있지 않다.

따라서 우리는 여러 개의 약물 조합에 따른 특성을 반영한 약물 부작용을 고려하기 위해 우리는 단일 약물을 기반으로 여러 개의 약물 조합의 위험을

예측하기 위한 2레벨 일반화 선형 혼합 모형을 제안한다. 이는 약물의 조합에 대한 부분이 하나 존재하고, 또 하나의 조합은 기본 약물을 기반으로 묶인 집단에 의해 약물의 변동성이 존재하는 것을 고려하여, 이에 대한 경우에 약물 부작용을 적절히 예측하기 위한 모형이다.

따라서 본 논문에서는 기본이 되는 약물 단위로 그룹화하여 집단 간의 변동을 추정하고자 했다. 우리는 이러한 특성을 반영하지 않은 모형은 위양성을 증가시키는 것을 제시하고자 했다. 시뮬레이션 과정을 통해 우리의 2레벨 일반화 선형 혼합 모형이 기존 혼합 약물 반응 모형에 비해 성능이 우수한 것을 살펴볼 수 있었고, 실제 공통데이터 모델에서 간독성에 적용하여 임상적으로 유의한 조합을 실제 시그널로 잘 예측하는지를 비교하여 결과를 제시하고자 하였다.

핵심되는 말: 약물 부작용, 부작용 감지, 자발적 보고 시스템, 혼합 효과모형, 약물-약물 상호작용, 공통 데이터 모델, 고차원 중첩 구조, 복합약물