A state-of-the-art multi-criteria model for drug benefit-risk analysis

Tervonen, T.; Hillege, H.L.; Buskens, E.; Postmus, D.

A state-of-the-art multi-criteria model for drug benefit-risk analysis

T. Tervonen (1,*), H.L. Hillege (2, 3), E. Buskens (2), D. Postmus (2)

(1) Faculty of Economics and Business, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands

(2) Department of Epidemiology, University Medical Center Groningen, PO Box 30.001, 9700 RB Groningen, The Netherlands

(3) Department of Cardiology, University Medical Center Groningen, PO Box 30.001, 9700 RB Groningen, The Netherlands

(*) Corresponding author. Email: t.p.tervonen@rug.nl. Tel.: +31503634495. Fax: +31503632275

Abstract

Drug benefit-risk analysis is based on firm clinical evidence related to various safety and efficacy outcomes, such as tolerability, treatment response, and adverse events. In this paper, we propose a new approach for constructing a supporting multi-criteria model that fully takes into account this evidence. Our approach is based on the Stochastic Multicriteria Acceptability Analysis (SMAA) methodology, which allows us to compute the typical value judgments that support a decision, to quantify uncertainty, and to compute a comprehensive benefit-risk profile. As an example, we constructed a multi-criteria model for the therapeutic group of second-generation antidepressants. We analyzed Fluoxetine, Paroxetine, Sertraline, and Venlafaxine according to relative efficacy and absolute rates of several common adverse drug reactions using meta-analytical data from the literature. Our model showed that there are clear trade-offs among the four drugs. Based on our experiences from this study, SMAA appears to be a suitable approach for quantifying trade-offs and decision uncertainty in drug benefit-risk analysis.

Keywords: clinical pharmacology; decision analysis; simulation methods; meta-analysis; risk communication

1 Introduction

Drug Benefit-Risk (BR) analysis is done daily by health care professionals, such as regulators, practicing physicians, and employees of insurance companies, to evaluate the safety and efficacy of different medical compounds. Popular indices of therapeutic benefit include the treatment effect, generally expressed as either the absolute change or the relative change in the rate of events, and the number of patients who need to be treated to attain one positive outcome or to prevent one adverse outcome. The harmful effects of treatment can be presented in a similar way. Although simple aggregate measures such as the numbers needed to treat and the numbers needed to harm seem easy to interpret, drug BR analysis generally includes various benefit and risk criteria and consequently must include value judgments (1-3). In such a setting, the use of Multi-Criteria Decision Analysis (MCDA) may be more appropriate as it provides a framework for analyzing complex decision problems involving value trade-offs.

The use of MCDA in the context of drug BR analysis was first proposed by Mussen et al. (4). Their work included a general framework for constructing a multi-criteria decision model for BR assessment of new drugs by regulatory authorities. Although being an important seminal work in the field, they score alternative drugs on the different benefit and risk criteria solely based on point estimates, thereby ignoring the sampling variation that is inherent in criteria measurements that are based on clinical trials and/or observational studies. In addition, the approach suggested by Mussen et al. (4) requires Decision Makers (DMs) to provide exact weights for describing the relative importance of the different criteria. In many real-life situations, however, DMs are not able to (or do not want to) restrict themselves to one particular set of

weights. Felli et al. (5) provided a similar application of MCDA in drug BR analysis. However, instead of using continuous measurements, the authors proposed to use categorical value scales for all BR attributes included in the model. Although it makes the model easier to apply in different contexts, there is a substantial risk of losing information by mapping measurements from a continuous scale to ordinal categories.

To overcome the limitations of the two previous approaches, we propose to use Stochastic Multicriteria Acceptability Analysis (SMAA) (6-8) as a new approach to drug BR analysis. Our choice of the SMAA methodology is supported by its proven applicability in risk assessment (9, 10) and published real-life analyses alike (11-17). To demonstrate its applicability in drug BR analysis, we will apply the SMAA-2 method (7) in the setting of a recently published meta-analysis considering the potential benefits and risks of several commonly prescribed second-generation antidepressants (18).

2 Methods

2.1 Stochastic Multicriteria Acceptability Analysis

SMAA-2 considers a discrete multi-criteria decision problem consisting of a set of m alternatives (such as different types of drugs) that are evaluated in terms of n criteria (such as several efficacy and safety criteria). The vector of criteria measurements corresponding to alternative i is denoted by $\mathbf{x}^i = (x_1^i, \ldots, x_n^i)$, where x_k^i represents the performance of alternative i on criterion k. It is assumed that the DM's preference structure can be represented by a real-valued utility or value function $u(\mathbf{x})$. The value function serves to rank the m alternatives by mapping their performance on the different criteria to a scalar index of preferability or value. It has the property that

alternative i is preferred over alternative j if and only if $u(\mathbf{x}^i) > u(\mathbf{x}^j)$. Although SMAA-2 can be applied with any type of value function, it is generally assumed that all criteria are mutually preferentially independent (19), which implies that the value function is additive, and the partial value functions $u_k(x_k)$ are used to normalize the criteria measurements by mapping them on a zero-to-one scale. The partial value functions can be obtained from the actual criteria measurements $x_k^1, \ldots, x_k^m$ through linear scaling, so that the worst value is 0 and the best value is 1. The additive value function is of form:

$$u(\mathbf{x}) = w_1 \cdot u_1(x_1) + \ldots + w_n \cdot u_n(x_n),$$

where the weights w_k (normalized, so that they sum to one) rescale the values of the partial value functions in such a way that the full swing (i.e. increase from the worst to the best value) in the scaled function indicates the importance of the criterion (20). For example, $w_i > w_j$ implies that if the DM had to choose between improving either criterion i or criterion j from the worst to the best value, he or she would increase the performance on criterion i.

The SMAA methodology has been developed for situations where neither criteria measurements nor weights are precisely known. So, instead of using deterministic values, probability distributions are specified for all criteria measurements x_k^i that are included in the model (in our setting, appropriate shapes for these distributions are derived from clinical trials and/or observational studies). Similarly, the DMs' unknown preferences are represented by a uniform weight distribution in the feasible weight space $W = \{\mathbf{w} \in R^n \mid w \geq 0, \sum_{j=1}^{n} w_j = 1\}$.

Instead of using the value function to rank the alternatives for an elicited weight vector $\mathbf{w} = (w_1, \ldots, w_n)$, which is the traditional approach in MCDA, SMAA computes for each alternative the weights a "typical" DM supporting this alternative would have. These so-called *central weight vectors* can be presented to the DM to help him or her understand what kind of weights favor a certain alternative, without providing any preference information. Mathematically speaking, the central weight vector of an alternative is defined as the expected center of gravity of all possible weight vectors that rank the alternative at the first place (c.f. (7)). It is expressed as a multidimensional integral over the criteria and weight distributions and can therefore be numerically computed by using Monte Carlo simulation (21). In each Monte Carlo iteration, values for the model parameters (i.e. criteria measurements and weights) are drawn from their corresponding distributions, and a ranking of the alternatives is obtained by plugging these values into the value function. For the alternative that is ranked at the first place, the current sampled weight vector is stored in a table, and the simulation proceeds with the next iteration. After all Monte Carlo iterations have been completed, an alternative's central weight vector is computed by averaging over all stored weight vectors that are associated with this alternative.

In addition to the central weight vectors, SMAA-2 defines two other types of descriptive measures: rank acceptability indices and confidence factors. The *rank acceptability index*, denoted by $b(i, r)$, describes the share of all possible values of the weight vector $\mathbf{w}$ and imprecise criteria measurements $\mathbf{x}^1, \ldots, \mathbf{x}^m$ for which alternative i is ranked at place *r*. Its value can be interpreted as the probability that alternative i is ranked at place *r*, where 0 indicates that the alternative will never obtain rank *r* and 1 indicates that alternative i will always obtain rank *r*.

The *confidence factor* of an alternative is the probability for this alternative to obtain the first rank when its central weight vector is used to scale the partial value functions. If there is no uncertainty in the criteria measurements, the confidence factor of each alternative will be equal to 1. In our setting, however, the criteria measurements are considered to be stochastic variables, so we are likely to obtain confidence factors of less than 1 for at least some of the alternatives included in the model. Just like the central weight vectors, the rank acceptability indices and the confidence factors can numerically be computed by using Monte Carlo simulation. In practice, the computation of the descriptive indices can be made more efficient than described here. For more details on the implementation of SMAA-2 based on Monte Carlo simulation, the reader is referred to (21).

So far, we considered the case when the SMAA analysis is conducted without any preference information. In practice, however, it may be possible to elicit some preference information from the DMs, such as a partial or complete ranking of the criteria. This information can easily be incorporated into the model by restricting the feasible weight space accordingly (c.f. (21)). If there is no preference information, the decision making is aided mainly through central weight vectors and confidence factors. When preference information is incorporated into the model, the rank acceptability indices can be used for finding the "best" alternative and for quantifying the risks related to uncertainty of outcomes.

2.2 Model for the therapeutic group of antidepressants

To demonstrate the application of SMAA in drug BR analysis, we constructed a model for the therapeutic group of antidepressants using efficacy and safety data from a published meta-analysis (18). We would like to stress that our model is illustrative

in nature, meaning that the results should not be interpreted as a full BR assessment of the different drugs included in the model.

2.2.1 Criteria

In the meta-analysis, the efficacy and safety of ten commonly prescribed second-generation antidepressants were compared (18). From this study, we selected the four antidepressants for which sufficient quantitative data was available: Fluoxetine, Paroxetine, Sertraline, and Venlafaxine. The criteria used to evaluate these four drugs are summarized in Table 1. We included one benefit criterion (treatment response) and five risk criteria, each corresponding to a different adverse event (diarrhea, dizziness, headache, insomnia, and nausea). There is a certain overlap between efficacy and insomnia, because improved efficacy can lead to less insomnia. For sake of simplicity, we disregarded this possible source of double-counting and assumed the criteria to be independent.

Treatment response, defined as an improvement of at least 50% on either the Hamilton Rating Scale for Depression (HAM-D) or the Montgomery-Asberg Depression Rating Scale from baseline to the end of the follow-up period, was used as a measure of efficacy in all clinical trials that were included in the meta-analysis. The authors conducted three separate analyses (effects were combined by using a random effects model) to estimate the relative efficacy (i.e. the proportion of respondents in one treatment group divided by the proportion of respondents in another treatment group) of (i) Paroxetine over Fluoxetine, (ii) Sertaline over Fluoxetine, and (iii) Venlafaxine over Fluoxetine. The frequency at which a specific adverse event occurred was reported in the meta-analysis as the mean incidence across all comparative trials and observational studies that included this event.

SMAA allows the criteria measurements to be modeled with arbitrary distributions. In our setting, the distributions follow directly from the results of the meta-analysis. In particular, it follows that the pooled incidences of the adverse events as well as the log of the pooled efficacy ratios can be considered as independently and normally distributed random variables (22). The means μ_k^i of these distributions are taken to be equal to the (log of the) pooled effect-size estimates, and the standard deviations σ_k^i are derived from the corresponding 95% confidence intervals as reported in the meta-analysis. For example, the pooled incidence of diarrhea for Fluoxetine was found to be equal to 11.7 with a 95% confidence interval of [6.8, 16.6] (18). The upper (lower) bound of this confidence interval was computed by adding (subtracting) $1.96 \cdot \sigma_{Diarrhea}^{Fluoxetine}$ to (from) the effect-size estimate of 11.7, so the estimated standard deviation of the pooled incidence of diarrhea for Fluoxetine will be equal to

$$\sigma_{Diarrhea}^{Fluoxetine} = \frac{(16.6 - 6.8)}{2 \cdot 1.96} = 2.5 .$$

2.2.2 Preference information

We performed three analyses: one with missing preference information, and two with a criteria ranking elicited from an expert in the field of antidepressants. For the latter, we explained the SMAA model and multi-attribute utility theory to the expert and asked her to consider a scenario of mild depression and a scenario of severe depression.

For both of these scenarios, we started by asking the expert to identify the criterion that she would most like to increase from the worst to the best value, given the range of the scales as depicted in Table 1. Then we asked for the second one, etc. This

process is similar to swing weighting in multi-attribute utility theory (23). However, since no exact "weights" are required, it resembles more the environment of medical decision making.

Let us denote by $\succ$ the strict preference relation. The elicitation process resulted in the following ranking for mild depression: Diarrhea $\succ$ Nausea $\succ$ Dizziness $\succ$ Insomnia $\succ$ Headache $\succ$ Efficacy. For severe depression the ranking was similar with the exception of efficacy being the most preferred criterion (i.e. Efficacy $\succ$ Diarrhea $\succ$ Nausea $\succ$ Dizziness $\succ$ Insomnia $\succ$ Headache).

2.3 Software

All analyses were conducted by using the JSMAA v0.2 software, an open source implementation of the SMAA methods in Java that is freely available from: http://www.smaa.fi.

3 Results

We completed the models with the criteria measurements listed in Table 2. The three analyses were executed with 10,000 Monte Carlo iterations, thereby giving the results sufficient accuracy (95% confidence error margins of ± 0.01) (21).

3.1 No preference information

The rank acceptability indices resulting from the analysis without preference information are listed in Table 3 and visualized as a column chart in Figure 1. These indices show that all drugs have reasonable BR profiles and should be considered for further analysis. In a situation like this, the decision can be aided through the central weight vectors (see Table 4). By looking at the central weights, we can see clear

trade-offs among the four drugs. For example, if the DM displays an *a priori* preference of Paroxetine, then based on the BR profiles expressed through the central weights, apparently nausea has the highest relative importance. If the DM accepts our model and is rational, he or she should favor lowering first nausea from the worst scale value (34%) to the best one (11.1%).

By contrasting a DM's preferences for scale swings (Table 1) with the central weights presented in Table 4, the DM can quickly decide which drug is preferable in the current situation. For example, if a DM considers the scale swing of efficacy (0.25) more important than the scale swing of dizziness (20.0, see Table 1), then he or she should prefer the BR profile of the three other drugs over Fluoxetine, because it is the only drug for which the central weight of efficacy is considerably lower than the central weight of dizziness. In addition, the confidence factors (Table 4) quantify the risk associated with the decision. For example, if a DM finds Fluoxetine's central weight vector to correspond with his or her preferences, the confidence factor (0.48) shows that the clinical data is too uncertain for making a truly informed decision.

3.2 Mild and severe depression

Rank acceptability indices for the scenario of mild (severe) depression are presented in Table 5 (Table 6) and illustrated in Figure 2 (Figure 3). Both the mild and severe depression scenarios lead to a relatively high first rank acceptability for Paroxetine. It had also a good rank profile in the analysis without preference information, and thus could be considered to have the "best" (i.e. reasonably high rank acceptabilities for the best ranks and low acceptabilities for the worst ranks) overall BR profile if no additional information is available. The rank acceptabilities of the other alternatives are more sensitive to the preferences. For example, the rank profile of Fluoxetine

depends completely on preferences. It achieves a significantly higher first rank acceptability (0.30) for the mild depression scenario than for the severe depression scenario (0.01). For Venlafaxine, the case is opposite: its first rank acceptability is a lot higher for severe depression (0.40) than for mild depression (0.21).

4 Discussion

Drug BR analysis has multiple uses, ranging from regulatory decision making to supporting decisions of a practicing physician. The MCDA-based approach suggested in this paper can be adapted for all contexts. As an example, we constructed a multi-attribute model for the therapeutic group of antidepressants by using data from a published meta-analysis. Despite the fact that the differences among the four antidepressants were mostly insignificant from a frequentist perspective, our results show that there are clear trade-offs among these drugs when the uncertainty surrounding the criteria measurements is taken into account. This could be seen from the central weight vectors of the analysis without preference information, and also from the rank acceptability indices of the analyses of the mild and severe depression scenarios that differed only in preference of the efficacy criterion.

Instead of having a different model for each therapeutic group, one could also consider constructing a more generic model by using the dimensions of an existing utility instrument, such as the EQ-5D or the Health Utilities Index. Although such instruments are suitable for calculating QALYs in the context of cost-effectiveness analysis, there is an important drawback when using them for drug BR analysis: their dimensions are defined in terms of generic health attributes—such as physical functioning, social functioning, and vitality—and may therefore not be very sensitive and responsive to the disease of interest. So, although our results have shown that

there are clear trade-offs among the four antidepressants, the relative differences in safety and efficacy may not be large enough to significantly change a patient's health status when this is measured in terms of generic health attributes.

Compared to the MCDA-based approaches proposed by Mussen et al. (4) and Felli et al. (5), the use of SMAA has two main advantages. The first advantage of the SMAA methodology is the possibility to include the sampling variation that is inherent in criteria measurements that are based on clinical trials. Mussen et al. (4) and Felli et al. (5), in contrast, do not explicitly include parameter uncertainty into their models: they solely rely on point estimates when assessing the performance of each of the alternatives against the different BR criteria. The capability of our model to propagate uncertainty to the results (in terms of rank acceptability indices and confidence factors) allows us to quantify the risks that are associated with any decision that is based on the results of the BR analysis.

The second advantage of the SMAA methodology over the two existing approaches is the possibility to characterize typical trade-offs supporting a drug BR profile without knowing or eliciting the preferences beforehand. The possibility to use our model without any preferences as well as with scenario-based ordinal preferences lowers the effort required to apply the model in different situations, and also increases the transparency of the decision making process.

The scenario-based rank acceptabilities can be used in operational support of decisions depending on drug BR analysis. If a drug has low (<0.20) rank acceptabilities for the first ranks, additional risk management practices should be used if that drug is chosen. For example, in our scenario of severe depression, Fluoxetine obtained a cumulative acceptability of only 0.06 for the first two ranks. If the BR

analysis leads to prescription decision of Fluoxetine due to external constraints (local reimbursement policy, patient profile including allergies, etc), the future patient consultancy should be sensitive to change of medication as other drugs with "better" BR profiles are available.

To conclude, we presented a new MCDA-based approach to drug BR analysis with an example application to the therapeutic group of second-generation antidepressants. In contrast to previous models, our model is based on the SMAA methodology, which allows us to take into account the sampling variation that is inherent in criteria measurements that are based on clinical trials and/or observational studies. In addition, by making the trade-offs among the four analyzed drugs explicit, we separated clinical data from subjective judgments, thereby increasing the transparency of the decision making process. The constructed model is specific to the therapeutic group of antidepressants, and future research should analyze the applicability of the SMAA methodology to other therapeutic groups.

Acknowledgements

This study was partly supported by the Escher project (T6-202), which is a project of the Dutch Top Institute Pharma. We thank Barbara van Zwieten for her expert opinion and information about antidepressants. In addition, we acknowledge Kit Roes, Hein Fennema, Hans van Leeuwen, and Marcel Hekking of Schering-Plough, as well as Petra Denig, for their valuable remarks and suggestions. The authors declare no conflicts of interest.

References

1. McGregor M, Caro JJ. QALYs: are they helpful to decision makers? Pharmacoeconomics. 2006;24(10):947-52.

2. Claycamp HG. Rapid benefit-risk assessments: no escape from expert judgments in risk management. Risk Anal. 2006;26(1).

3. Garrison LP, Jr., Towse A, Bresnahan BW. Assessing a structured, quantitative health outcomes approach to drug risk-benefit analysis. Health Aff. 2007;26(3):684-95.

4. Mussen F, Salek S, Walker S. A quantitative approach to benefit-risk assessment of medicines – part 1: The development of a new model using multi-criteria decision analysis. Pharmacoepidemiol Drug Saf. 2007;16:S2-S15.

5. Felli JC, Noel RA, Cavazzoni PA. A multiattribute model for evaluating the benefit-risk profiles of treatment alternatives. Med Decis Making. 2009;29(1):104-15.

6. Lahdelma R, Hokkanen J, Salminen P. SMAA - stochastic multiobjective acceptability analysis. Eur J Oper Res. 1998;106(1):137-43.

7. Lahdelma R, Salminen P. SMAA-2: Stochastic multicriteria acceptability analysis for group decision making. Oper Res. 2001;164(3):444-54.

8. Tervonen T, Figueira JR. A survey on stochastic multicriteria acceptability analysis methods. J Multi-Criteria Dec Anal. 2008;15(1-2).

9. Tervonen T, Linkov I, Figueira JR, Steevens J, Chappell M, Merad M. Risk-based classification system of nanomaterials. J Nanopart Res. 2009;11(4):757-66.

10. Tervonen T, Figueira JR, Lahdelma R, Almeida Dias J, Salminen P. A stochastic method for robustness analysis in sorting problems. Eur J Oper Res. 2009;192(1):236-42.

11. Hokkanen J, Lahdelma R, Salminen P. A multiple criteria decision model for analyzing and choosing among different development patterns for the Helsinki cargo harbor. Socio-Economic Planning Sciences. 1999;33:1-23.

12. Tervonen T, Hakonen H, Lahdelma R. Elevator planning with stochastic multicriteria acceptability analysis. Omega. 2008;36(3):352-62.

13. Kangas A, Kangas J, Lahdelma R, Salminen P. Using SMAA-2 method with dependant uncertainties for strategic forest planning. Forest Policy and Economics. 2006;9:113-25.

14. Kangas J, Hokkanen J, Kangas A, Lahdelma R, Salminen P. Applying stochastic multicriteria acceptability analysis to forest ecosystem management with both cardinal and ordinal criteria. Forest Science. 2003;49(6):928-37.

15. Hokkanen J, Lahdelma R, Miettinen K, Salminen P. Determining the implementation order of a general plan by using a multicriteria method. J Multi-Criteria Dec Anal. 1998;7(5):273-84.

16. Kangas J, Kangas A. Multicriteria approval and SMAA-O method in natural

resources decision analysis with both ordinal and cardinal criteria. J Multi-Criteria Dec Anal. 2003;12:3-15.

17. Lahdelma R, Salminen P, Hokkanen J. Locating a waste treatment facility by using stochastic multicriteria acceptability analysis with ordinal criteria. Eur J Oper Res. 2002;142(2):345-56.

18. Hansen RA, Gartlehner G, Lohr KN, Gaynes BN, Carey TS. Efficacy and safety of second-generation antidepressants in the treatment of major depressive disorder. Ann Intern Med. 2005;143:415-26.

19. Keeney R, Raiffa H. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. New York: John Wiley & Sons; 1976.

20. von Nitzch R, Weber M. The effect of attribute ranges on weights in multiattribute utility measurement. Management Sci. 1993;39:937-43.

21. Tervonen T, Lahdelma R. Implementing stochastic multicriteria acceptability analysis. Eur J Oper Res. 2007;178(2):500-13.

22. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. Psychol Methods. 1998;3(4):486-504.

23. Belton V, Stewart TJ. Multiple Criteria Decision Analysis - An Integrated Approach: Kluwer Academic Publishers, Dordrecht; 2002.

Table 1: Criteria characteristics.

Name	**Measurement unit**	**Preference direction**	**Scale range**
Efficacy	Relative value compared with Fluoxetine	↑	[0.98, 1.23]
Diarrhea ADE's	Absolute %	↓	[1, 20.6]
Dizziness ADE's	Absolute %	↓	[4.4, 24.4]
Headache ADE's	Absolute %	↓	[8, 31.3]
Insomnia ADE's	Absolute %	↓	[3.4, 21.3]
Nausea ADE's	Absolute %	↓	[11.1, 34]

Table 2: Criteria measurements. The values are given as mean ± standard deviation. The measurement units are as presented in Table 1.

Drug	Ln(Efficacy)	Diarrhea	Dizziness	Headache	Insomnia	Nausea
Fluoxetine	0 ± 0	11.7 ± 2.5	7.2 ± 1.45	16.6 ± 3.27	13.7 ± 1.89	18.6 ± 1.79
Paroxetine	0.086 ± 0.056	9.2 ± 1.86	10.6 ± 1.58	21.2 ± 5.15	14.3 ± 2.93	18.3 ± 3.7
Sertraline	0.095 ± 0.044	15.4 ± 2.65	7.5 ± 1.48	20.2 ± 3.78	15 ± 3.21	19.5 ± 2.6
Venlafaxine	0.113 ± 0.048	5.5 ± 2.32	15.7 ± 4.44	12.8 ± 2.45	11.2 ± 3.98	31 ± 1.68

Table 3: Rank acceptability indices from the analysis without preference information. In columns of Rank 2 and 3 also the cumulative indices from Rank 1 are presented.

Drug	Rank 1	Rank 2 (cum. 1+2)	Rank 3 (cum. 1+2+3)	Rank 4
Fluoxetine	0.20	0.28 (0.48)	0.30 (0.78)	0.22
Paroxetine	0.25	0.29 (0.54)	0.27 (0.81)	0.19
Sertraline	0.17	0.25 (0.42)	0.29 (0.71)	0.30
Venlafaxine	0.39	0.18 (0.57)	0.15 (0.72)	0.29

Table 4: Central weights and corresponding confidence factors from the analysis without preference information.

Drug	Conf. factor	Central weight vector					
		Efficacy	Diarrhea	Dizziness	Headache	Insomnia	Nausea
Fluoxetine	0.48	0.08	0.14	0.23	0.18	0.16	0.22
Paroxetine	0.45	0.18	0.17	0.15	0.13	0.15	0.22
Sertraline	0.34	0.21	0.10	0.22	0.13	0.15	0.20
Venlafaxine	0.74	0.18	0.21	0.12	0.21	0.19	0.09

Table 5: Rank acceptability indices from the scenario of mild depression.

Drug	Rank 1	Rank 2 (cum. 1+2)	Rank 3 (cum. 1+2+3)	Rank 4
Fluoxetine	0.30	0.35 (0.65)	0.26 (0.91)	0.08
Paroxetine	0.45	0.33 (0.78)	0.17 (0.95)	0.05
Sertraline	0.04	0.10 (0.14)	0.26 (0.40)	0.60
Venlafaxine	0.21	0.23 (0.44)	0.30 (0.74)	0.26

Table 6: Rank acceptability indices from the scenario of severe depression.

Drug	Rank 1	Rank 2 (cum. 1+2)	Rank 3 (cum. 1+2+3)	Rank 4
Fluoxetine	0.01	0.05 (0.16)	0.23 (0.39)	0.71
Paroxetine	0.42	0.31 (0.73)	0.20 (0.93)	0.07
Sertraline	0.18	0.31 (0.49)	0.37 (0.86)	0.14
Venlafaxine	0.40	0.32 (0.72)	0.20 (0.92)	0.08

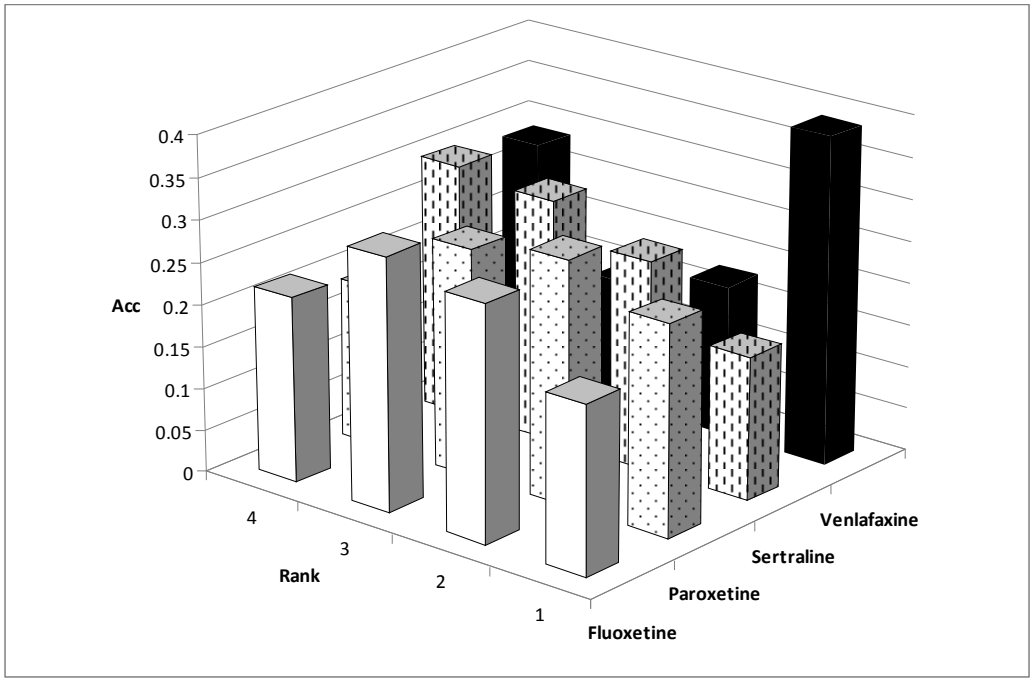

Figure 1: Rank acceptability indices for the model without preference information.

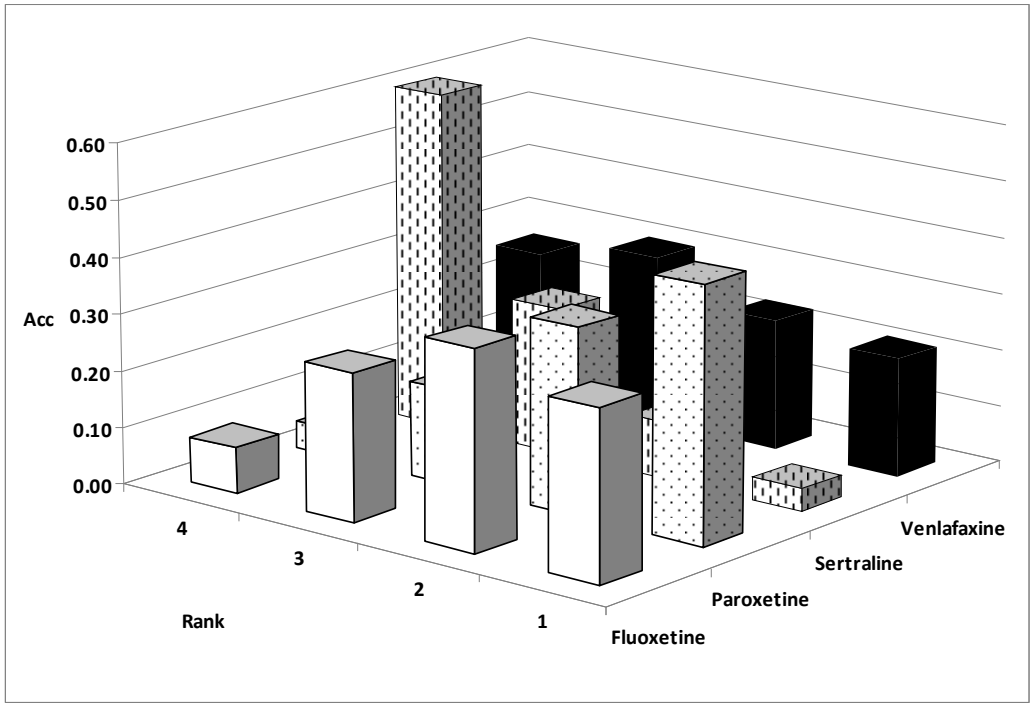

Figure 2: Rank acceptability indices for the scenario of mild depression.

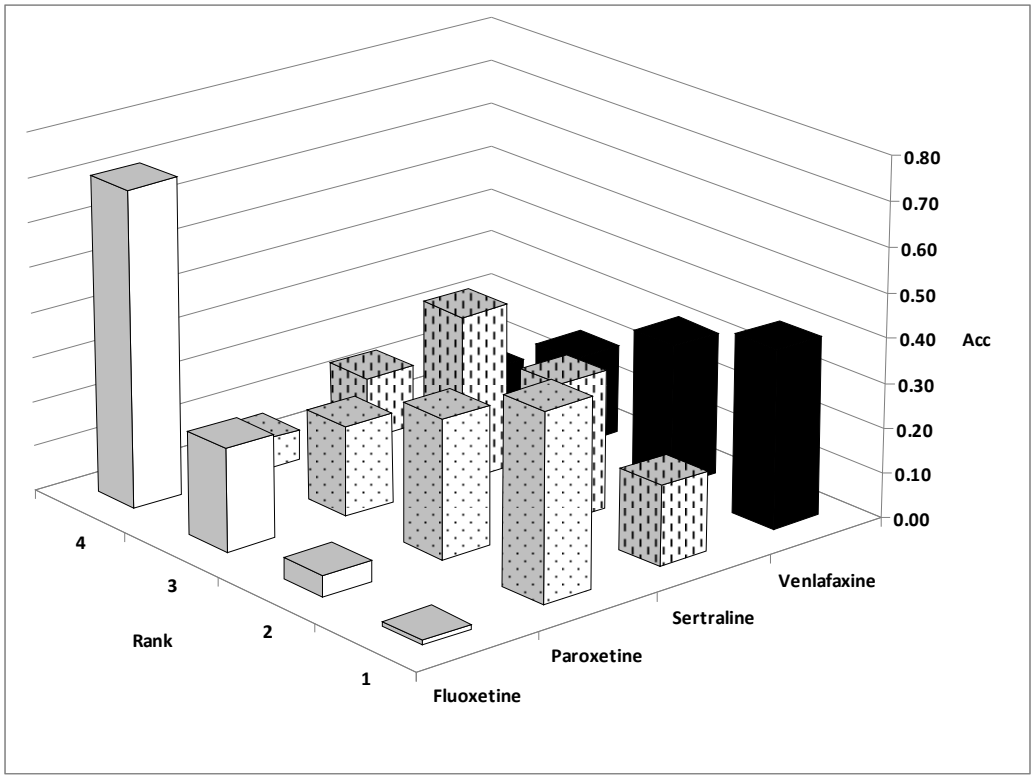

Figure 3: Rank acceptability indices for the scenario of severe depression.