

Bayesian Meta-analyses of the Tolerability of Selective Serotonin Reuptake Inhibitors and Tricyclic Antidepressants for Treating Patients with Depression*

Chan Siew Pang *MSc (Management), MSc (Medical Statistics)*, Poh Kim Leng *PhD*¹

National Disease Registries Office, Health Promotion Board

¹Department of Industrial & Systems Engineering, National University of Singapore

ABSTRACT

Background. This study aimed to compare, via meta-analysis, the tolerability of selective serotonin reuptake inhibitors (SSRIs) with tricyclic antidepressants (TCAs) for treating depression. A Bayesian model was developed and the results were compared with the classical models.

Methods. The outcome of interest was the combined odds ratio of premature withdrawal from treatments due to drug-related side-effects. Two separate meta-analyses — one for primary-care and one for the general setting — were conducted. Unlike the classical approach, the proposed Bayesian model allows researchers to combine their expert opinions with published data. The iterative Markov Chain Monte Carlo method (written in Stata 7.0) was applied for generating the posterior distributions for Bayesian analyses.

Results. The classical models showed that significantly fewer patients receiving SSRIs withdrew prematurely due to drug-related side effects. The Bayesian models based on SSRI-favoured priors gave similar results, although the interpretation was philosophically different. Moreover, the likelihoods for the analyses were rather weak, so one must interpret the results with extra care.

Conclusion. The Bayesian analyses with SSRI-favoured priors revealed that SSRIs might be better tolerated than TCAs in the general setting. However, as the posteriors were strongly influenced by the priors, more studies need to be conducted before any conclusion can be made. The Bayesian models provided more insights to the problem and the nature of data selected for meta-analyses.

Keywords: Bayesian meta-analysis, depression, Gibbs sampler, Markov Chain Monte Carlo, SSRIs, TCAs

INTRODUCTION

Depression is a common disorder that is becoming better understood as an illness that can be chronic, recurrent, and refractory to treatment.¹ Generally defined as a mood disorder that impairs normal functions, depression may be caused by many factors. People with the condition typically have difficulty regulating certain brain chemicals called neurotransmitters. One is at risk if there is a family history of depression. Biochemical factors are also important and certain medications and hormones may

affect one's mood. The onset of depression might be related to the occurrence of medical illnesses. Last but not least, losing a loved one, financial concerns, work stress, or relationship problems may all contribute to depressive disorders.

There is a wide class of antidepressants available for treating all forms of depression regardless of cause. By working within the brain to increase the levels of either noradrenaline, serotonin or both, antidepressants help to reduce the symptoms of anxiety and negative thoughts usually experienced by sufferers. However, they do not act immediately and the lifting of moods typically takes up to 2 weeks or longer.

* Presented at the International Conference on Evidence-based Medicine, Singapore, held on 18 to 19 October 2003.

The most common antidepressants are selective serotonin reuptake inhibitors (SSRIs) and tricyclic antidepressants (TCAs). Widely regarded as one of the most effective antidepressants, the latter had been the first-choice medication for treating depression for over 30 years. TCAs work by preventing the uptake of norepinephrine and serotonin, thus building up the concentration of these transmitters and improving the communication between neurons. However, there are a number of problems with TCAs as their effects in the brain are not restricted to alleviating depression. They also interact with a number of other brain receptors, thus causing side-effects like dry mouth, drowsiness, dizziness, blurred vision, constipation, urinary difficulties, tremor and tachycardia. In addition, they tend to lower blood pressure and consequently cause a feeling of faintness. Morbidity and mortality caused by TCA overdose are also widely reported.^{2,3} As such, TCAs are usually prescribed in severe cases of depression.

On the other hand, SSRIs work only on the serotonin system. Introduced in the 1980s, SSRIs possess an improved side-effect profile over TCAs with their selective mode of action.⁴ While retaining good clinical efficacy, they have few of the anticholinergic, antihistaminergic and cardiotoxic effects and are probably safer in overdose than TCAs.^{5,6} As a result, SSRIs are recognised to be better tolerated and more acceptable to patients.⁷⁻¹¹ However, there are other well-documented side-effects associated with the use of the drugs. These include nausea, vomiting, diarrhoea, decreased appetite, fatigue, increased sweating, sleep disturbances and impotency.¹⁰ Moreover, SSRIs may also interact with other drugs so extra care must be taken in prescription.

Both TCAs and SSRIs achieve similar efficacy, with 60 to 80% of patients responding adequately.^{12,13} However, their side-effect profiles vary substantially, so the choice of medication for treating depression depends primarily on patients' tolerability. As such, the main objective of this paper was to conduct meta-analyses on the tolerability of SSRIs and TCAs. In addition, a Bayesian model was developed. This may help to provide biomedical researchers with a basic understanding on the workings of Bayesian statistical methods, which are increasingly becoming attractive to researchers in many fields. The results of the proposed Bayesian model were compared with those of the classical model.

METHODS

Bayesian Model

As opposed to the familiar classical approach, the proposed Bayesian meta-analysis model allows prior information — in the form of expert opinion — to be incorporated into analysis. To facilitate discussion, let the overall effect size (for example, odds ratio, rate difference, risk ratio or mean difference), the study-specific effect size, the between-study precision and within-study precision be denoted as θ , φ_i , τ and ϕ_i , respectively. The first 3 quantities are unknown and unobserved. The term “data” refers to observed individual effect sizes of the studies considered for meta-analysis. Based on the celebrated Bayes' theorem, the proposed model may be expressed as:¹⁴

$$p[\theta, \varphi_i, \tau, \phi_i | \text{data}] \propto L[\text{data}; \varphi_i | \phi_i] \times L[\varphi_i | \theta] \times g[\theta] \times g[\tau]$$

where $L[\text{data}; \varphi_i | \phi_i]$ and $L[\varphi_i | \theta]$ are the likelihood functions, $g[\theta]$ the prior distribution for θ , $g[\tau]$ the prior distribution for τ and $p[\theta, \varphi_i, \tau, \phi_i | \text{data}]$ the joint posterior distribution where analysis is made. This model allows observed effect sizes (data) to vary around their individual study-specific effect sizes φ_i , which in turn belong to a distribution characterised by the overall effect size θ . The priors, as opposed to the likelihood, express the analyst's expert opinion about θ and τ .

The first and foremost task of conducting Bayesian analysis is to derive the joint posterior based on suitably-chosen likelihoods and priors. The form of the likelihood depends on the nature of data, while the choice of priors is usually restricted to the conjugate family of distributions. When the posterior has the same distributional family as the priors, one says that the priors and the likelihood distributions are conjugate. The concept of conjugacy ensures that the posterior distribution is mathematically tractable. Once the distributional form of the posterior is deduced, it may be generated and the overall effect size θ can be determined.

The above formulation illustrates several important differences between the Bayesian and classical models. First, no prior distribution of θ is allowed in the classical approach and the analysis is solely based on the likelihood function(s). Second, the overall effect size θ is considered as a random quantity in Bayesian analysis. As readily seen, all information required for

Bayesian analysis is described by the posterior. In practice, one may not have prior information before analysis and the advice is to specify a flat (non-informative) distribution for the priors so that they have little influence on the posterior. Not surprisingly, the posterior is then dominated by the likelihood and the result will be identical to the classical analysis. However, the interpretation is philosophically different.

In most instances, normal distribution is appropriate for the observed effect sizes. If normal distribution is not immediately appropriate, one may perform logarithmic transform on the observed effect sizes. The prior distributions are chosen within the related conjugate family such that θ belongs to a normal distribution with parameters μ (location) and ν (scale), while τ belongs to a gamma distribution characterised by quantities λ (shape) and η (scale).

By fixing 4 parameters in the set-up, the model becomes extremely complicated (Annex 1). Theoretically, inferences about θ should be made from this posterior. However, the complicated form of the posterior makes computation extremely difficult and one may resort to simulation techniques, such as the Markov Chain Monte Carlo (MCMC), for generating the posterior.¹⁵⁻¹⁷ Via the Gibbs sampler algorithm, the technique works directly with the conditional posterior distributions, such as $p[\theta | \varphi, \tau, \phi, \text{data}]$ and $p[\tau | \theta, \varphi, \phi, \text{data}]$, instead of the joint posterior (Annex 2).¹⁸ The advantage is that conditional posteriors have simpler structures as they are derived from the full posterior by treating other parameters as fixed. By drawing a large number of values from these conditional posterior distributions, one obtains the full posterior of interest.

Since the procedure is iterative in nature, the parameters generated depend on the initialising values. As such, the number of burn-ins must be decided before analysis. This is the beginning set of runs that are discarded under the assumption that they are not representative of the joint posterior distribution. The eventual analysis is based on the updated set of values after burn-ins.

The posterior distribution of θ could be summarised by means of a 95% posterior interval (PI). In classical statistics, the random variables in a confidence interval (CI) are the limits and not the overall effect size, θ . With a 95% CI constructed, one may claim that the unknown θ is contained in 95% of all possible intervals obtained with the same sampling method. In the case of Bayesian analysis, the PI provides a more natural

interpretation as θ may be stated to have a 0.95 probability of being within the interval.

A user-friendly Stata 7.0 (Stata Corporation, Texas, USA) programme was written for facilitating the computation through Gibbs sampling, which is the basis of the proposed Bayesian model. The programme allows users to specify the prior parameters and the number of burn-ins and updates, with options for a graphical display of the iterative history.

Selection of Studies

The above-mentioned Bayesian model was applied to compare the tolerability of SSRIs and TCAs. There were 2 sets of meta-analyses. The first was concerned with primary-care patients' premature discontinuation from treatments due to drug-related side effects. It is important to investigate patients' tolerability in the primary care setting since the majority of depression cases are first seen and treated in general practice.¹⁹⁻²² Randomised controlled trials investigating the efficacy and tolerability of SSRIs (fluoxetine, fluvoxamine, paroxetine, sertraline and citalopram, etc.) against a TCA or an antidepressant with identical mechanism of action (amitriptyline, imipramine, dothiepin, clomipramine and lofepramine, etc.) in patients with depressive disorders were identified through MEDLINE and Cochrane Library search up to May 2004, previous meta-analyses and literature review.^{13,23-28} The patients' depressive orders were assessed by means of the Research Diagnostic Criteria (RDC), Diagnostic Systems (DSM-III), Hamilton Rating Scale for Depression (HRSD), Clinical Global Impression Score (CGI), Clinical Anxiety Scale (CAS) and Montgomery-Asberg Depression Rating Score (MADRS), and so on. No language restriction was imposed in the search. Studies were excluded from analysis if there were insufficient information on study design, description of treatments and tolerability and source of subjects, and so on.

In the second analysis, the Bayesian model was compared with a reported result published in the Cochrane Database of Systematic Reviews.²⁸ This analysis served to provide a more complete assessment of the comparative tolerability of the antidepressants in the general setting. In addition, it also helped to reveal more details about the model's theoretical properties. The published result was based on a search on the Cochrane Collaboration Depression, Anxiety and Neurosis Controlled Trials Registers (1977-1999), MEDLINE (1966-1999), EMBASE (1974-1999), specialist journals, previous systematic reviews,

Table 1. Selected clinical trials with primary-care patients.

Study	Patient Selection	Antidepressants (SSRI/TCA)	Duration	Profile of Patients
Corne 1989	Met RDC; HAM-D \geq 17	Fluoxetine/dothiepin	6 weeks	Gender: 70% female Mean age: 41.7
Stott 1993	MADRS \geq 16; CAS \geq 11	Paroxetine/amitriptyline	8 weeks	Gender: 66.5% female Mean age: 42.8
Rosenberg 1994	HAMD \geq 14	Citalopram/imipramine	6–22 weeks	Gender: 70% female Age range: 19–65
Doogan 1994	Met DSM-III-R; MADRS \geq 22; CGI \geq 4	Sertraline/dothiepin	6 weeks	Gender: 70.5% female Mean age: 47.1
Moon 1996	Met DSM-III-R; MADRS \geq 18	Paroxetine/lofepramine	6 weeks	Gender: 71.3% female Mean age: 43.7
Christiansen 1996	HAMD \geq 15	Paroxetine/amitriptyline	8 weeks	Age range: 18–65
Ravindran 1997	MADRS \geq 20; CAS \geq 11	Paroxetine/clomipramine	12 weeks	Gender: 73.5% female Mean age: 42.6

Table 2. Primary-care patients discontinued from treatments due to side effects.

	SSRIs (No./Total)	TCAs (No./Total)	OR (95% CI)
Corne 1989	7/49	2/51	0.29 (0.07–1.12)
Stott 1993	35/243	49/262	1.36 (0.85–2.18)
Rosenberg 1994 *	43/380	16/92	1.74 (0.87–3.46)
Doogan 1994	5/83	2/96	0.35 (0.08–1.60)
Moon 1996	5/60	4/62	0.76 (0.20–2.94)
Christiansen 1996	9/71	9/73	0.97 (0.36–2.59)
Ravindran 1997	54/500	84/502	1.65 (1.15–2.36)
Total	158/1386	166/1138	

* Evaluated at 22 weeks

Table 3. Meta-analyses of the tolerability of SSRIs and TCAs in primary care.

Models	Combined OR (95% Interval Estimate)
Classical Fixed-Effect Model	1.35 (1.06–1.73)*
Bayesian Model in favour of SSRIs (Prior OR: 1.50)	1.47 (1.21–1.77)**
Bayesian Model in favour of SSRIs (Prior OR: 1.25)	1.23 (1.02–1.48)**
Bayesian Model with “Non-informative Indifferent” Prior (Prior OR: 1.00)	0.53 (0.11–2.52)**
Bayesian Model with “Non-informative” Prior in favour of TCAs (Prior OR: 0.75)	0.53 (0.11–2.52)**

* 95% Confidence Interval

** 95% Posterior Interval

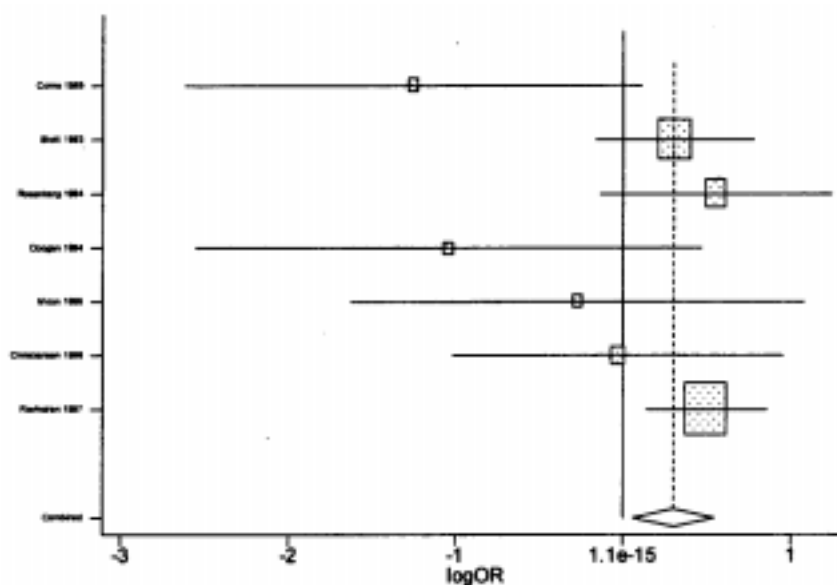


Fig. 1. Individual study results based on the classical model.

conference abstracts, government documents and reference lists of relevant papers.²⁸

For the abovementioned analyses, the effect size of interest was the overall or combined odds ratio (OR) of premature withdrawal from treatments due to drug-related side effects. The combined and study-specific ORs were computed such that a value above unity suggests that SSRIs were better tolerated than TCAs. To allow direct comparison with published results, the odds ratios were computed by Peto's Observed-Expected (O-E) method.²⁹ In order to provide more insights into the issues, several priors reflecting different views of the comparative tolerability of the antidepressants were applied in the proposed Bayesian analyses. The data were entered into Stata for analysis and all statistical tests were conducted at 5% level of significance.

RESULTS

Discontinuation from Primary-Care Treatment Due to Side Effects

The data were extracted from 7 randomised double-blinded clinical trials.³⁰⁻³⁶ The trials, mostly conducted at multi-centres, involved 2524 patients (with 1386 received SSRIs and 1138 received TCAs). Information concerning drug treatments, inclusion criteria and basic demographics of the patients can be found in Table 1. Several studies, including those reported in a similar meta-analysis, were omitted from analysis because of insufficient information.^{26,37} In addition, 2 trials which

recruited only elderly patients (65 years and above) were also excluded.^{21,38}

Only 3 of the selected studies, notably those with large sample sizes, were in favour of SSRIs (Table 2 and Fig. 1). Combining all 7 studies, the classical model produced an overall OR of 1.35 (95% CI: 1.06–1.73) (Table 3), thus suggesting that SSRIs were better tolerated than the TCAs. The fixed-effect model was applied in this context because the assumption of homogeneity was not discarded (p-value: 0.09).

However, the result must be interpreted with care. The combined OR was largely influenced by one trial that favoured SSRIs strongly.³⁶ Omitting this trial would reduce the combined OR to 1.15 (95% PI: 0.83–1.60), thus suggesting that SSRIs were not better tolerated than TCAs.

The Bayesian models were built next. Recall that 4 parameters must be fixed for the prior distributions for θ (combined odds ratio) and τ (between-study precision). Different prior values for θ reflect the different beliefs of the comparative tolerability of SSRIs and TCAs. To induce normality, the observed ORs were transformed and the prior for θ refers to combined log OR. The final results were reported as OR by performing the necessary back-transformation. Next, the prior for τ was standardised as Gamma[λ : 0.01, η : 0.01]. The choice of this distribution reflected the lack of prior information regarding between-study precision. Also, the number of burn-ins was set a priori at 500 and the Markov chain would thereafter be run

another 1000 times before the final analyses were conducted.

In the first attempt, the prior for θ was chosen as Normal[μ : 0.4055, ν : 100]. This reflects a highly-concentrated normal distribution with the overall OR believed to be 1.50, i.e., the SSRIs were better tolerated than TCAs. The selection of this prior was based on the general beliefs that SSRIs were associated with a significantly lower risk of toxicity.² The combined OR turned out to be 1.47 (Table 3). Since the 95% PI did not contain unity, SSRIs were found to be significantly better tolerated than TCAs.

In the second attempt, suppose an expert reported that SSRIs were better tolerated than TCAs, but the prior combined OR was fixed at 1.25, i.e. Normal[μ : 0.2231, ν : 100]. The combined posterior OR was 1.23 (Table 3). This is somewhat lower than that reported in the previous analysis.

In the next exploratory Bayesian analysis, a prior suggesting that SSRIs and TCAs were identical in terms of their tolerability was fitted. In this case, the “indifferent” prior for θ was chosen as Normal[μ : 0, ν : 0.000001]. Due to the low precision of 0.000001, this flat prior resembled that of a uniform distribution, thus suggesting that there was little prior information regarding θ . As shown in Table 3, the combined OR turned out to be 0.53 and the associated 95% PI was 0.11–2.52. As a result, one may interpret that SSRIs were not significantly better tolerated than TCAs. The fairly wide interval estimate was a result of the inclusion of 2 non-informative priors.

To further illustrate the properties of the proposed Bayesian model, a prior suggesting that TCAs were better tolerated than SSRIs was fixed next, i.e., Normal[μ : -0.2877, ν : 0.000001]. This was a “non-informative” prior OR fixed at 0.75 (in favour of TCAs). The result was identical to the previous analyses based on “indifferent” prior (Table 3).

The above analyses based on Bayesian models shared a very important common feature. The posteriors were dominated by their respective priors. This was the result of a “weak” likelihood, which was due to the small sample of selected trials. Moreover, Table 2 also shows that almost all individual ORs contain unity in the 95% CIs. With a relatively “weak” likelihood, the posterior result would be strongly influenced by the prior, i.e., the combined posterior ORs were largely similar to the prior ORs. However, this point was not explicitly highlighted by the classical model. In this case, the Bayesian analysis revealed more details of the data. It

must also be emphasised that the above analyses were conducted for exploratory purposes. One should fix a confirmed prior before conducting the analysis.

In passing, note that the resultant posteriors based on the 4 different sets of priors were fairly normal and the Markov chains exhibited no obvious pattern of divergence after the burn-in values had been discarded (figure not shown).

Discontinuation from Treatment Due to Side Effects in the General Setting

It was not feasible to perform sub-group meta-analysis by drug class in the primary-care example as there were few studies included for analysis. In the next example, SSRIs were compared with 3 tertiary TCAs (amitriptyline, imipramine and clomipramine) separately. For all analyses, the priors for θ and τ were fixed as Normal[μ : 0.2231, ν : 100] and Gamma[λ : 0.01, η : 0.01] respectively. The choice of priors reflected the belief that SSRIs were better tolerated than TCAs.^{39,40} In addition, both the number of burn-ins and updates were set a priori at 1000.

As shown in Table 4, the SSRIs were indeed better tolerated than TCAs in the general setting. As in the case of primary-care analysis, the combined ORs of the Bayesian models were largely influenced by the priors.

In the case where SSRIs were compared with amitriptyline, 31 studies were selected. The classical fixed-effect model was chosen to compare with the Bayesian model because statistical tests suggested no strong evidence of study heterogeneity (p-value: 0.79). Both models showed favourable results for SSRIs but the Bayesian model’s combined OR was substantially lower (Table 4).

A similar observation was made for comparison with imipramine where 29 studies were included for analysis. The Bayesian model’s posterior combined OR was almost identical to the prior OR. The random-effect model was chosen for classical analysis because there was strong evidence of study heterogeneity (p-value<0.01). In passing, note that the lower ends of the 95% CIs and PIs for both classical and Bayesian analyses were close to unity.

Last but not least, there was again a large disparity in results when SSRIs were compared with clomipramine. Dominated by the prior, the posterior combined OR of the Bayesian model reported a less favourable effect for SSRIs when compared with the fixed-effect classical model. With only 9 studies selected, there was no

Table 4. Meta-analysis of the tolerability of SSRIs and TCAs in the general setting.

Models	SSRIs vs amitriptyline	SSRIs vs imipramine	SSRIs vs clomipramine
Classical Model	OR: 1.57 (1.27–1.95)	OR: 1.48 * (1.09–2.01)	OR: 1.68 (1.24–2.26)
Bayesian Model in favour of SSRIs (Prior OR: 1.25)	OR: 1.34 (1.11–1.61)	OR: 1.26 (1.05–1.53)	OR: 1.29 (1.09–1.53)

* Based on random-effect model

strong evidence of study heterogeneity according to the classical analysis (p -value: 0.18). The Markov chains exhibited no obvious pattern of divergence in the above-mentioned Bayesian analyses (figure not shown).

DISCUSSION

Replication of experimental results has long been a central feature of scientific inquiry, and it raises questions concerning how to combine the results obtained. Meta-analysis is often defined as the statistical analysis of a collection of results from individual studies for the purpose of integrating the findings.⁴¹ It involves the combination of quantitative evidence from studies that have investigated a common question.

The theoretical details of the classical model for meta-analysis are well known.^{42,43} Following the rationale of conventional statistical theory, the effect size of interest is considered as an unknown but fixed quantity that can be accurately estimated from data obtained from a proper literature search. Motivated by the current need for evidence-based medicine, a Bayesian model was developed in this paper. It differs from the classical approach in 2 aspects. First, it allows prior information — in the form of expert opinion — to be incorporated into analysis. Though subjective in nature, such information, it is argued, may provide a more realistic approach in data analysis. Many biomedical researchers may have accumulated a large amount of experience through practice and it is costly to ignore such information. Second, the analysis is conducted on the posterior distribution which summarises all the information, both prior- and data-based, that the analysts have about the unknown parameters.

As described, the proposed Bayesian model allows observed effect sizes (data) to vary around their individual study-specific effect sizes, which in turn belong to a distribution characterised by the combined effect size θ . This is essentially a random-effect model designed for situations where there is substantial

heterogeneity among studies. It was preferred over a fixed-effect model for the abovementioned analyses because there were differences in treatments (for example, types of antidepressants, dose of drugs and treatment duration), types of patients (for example, inclusion criteria, culture of drug compliance), experimental designs (for example, with or without a placebo arm) and type of statistical analyses applied. Consequently, it is naïve to assume that study heterogeneity does not exist even with the support of formal statistical tests. Moreover, such statistical tests may lack power in detecting the underlying differences among studies. It is very rare for medical studies of the same objective/nature to be exactly comparable.

The proposed Bayesian model may also be called a hierarchical model because, loosely speaking, more than one level of prior and likelihood is specified. In this case, a particular observed quantity depends on an unknown parameter, which in turn follows a second-stage prior. This sequence of priors and parameters constitute a model with an extended or hierarchical data structure.

This paper aimed to analyse the tolerability of SSRIs and TCAs in patients with depressive disorders. As one of the most common illnesses that affects a large number of individuals in all countries, depression is a “whole-body” disorder affecting the nervous system, moods, thoughts and behaviour. As both SSRIs and TCAs are effective in treating depression, the choice of medication depends mainly on patients’ tolerability.

There was no clear evidence from the meta-analysis on patients receiving primary care that SSRIs were better tolerated than TCAs. Of the 7 clinical trials considered, only 1 favoured SSRIs significantly. The Bayesian models demonstrated that the resultant posteriors were strongly influenced by the priors fixed before analysis. As such, the primary-care physicians must be vigilant when using SSRIs (fluoxetine, praxetine, citalopram and sertraline in particular).

In the general setting, however, the result was slightly more optimistic. Based on current findings, amitriptyline, imipramine and clomipramine were not as safe as SSRIs. This result conformed with the general beliefs. However, as in the analysis with primary-care patients, the posteriors of the Bayesian analyses were strongly dominated by the priors. Consequently, the safety of antidepressant therapies should be monitored carefully as patients who suffer from depression may experience different tolerability profiles.

The proposed Bayesian model provides biomedical researchers an alternative framework for conducting meta-analysis. For future research, one may consider conducting meta-analysis on other newer antidepressants and different types of patients such as the elderly.

Finally, it is worthwhile to make some suggestions regarding the elicitation of prior distributions for future research. In practice, one may form a “community” of prior distributions with inputs from a number of experts.⁴⁴ Inferences may then be based on a consensus of the posterior results or a “final” prior may be formed by averaging the prior distributions elicited from the experts.⁴⁵ Alternatively, one may also consider constructing a primary prior distribution and a number of similar priors that belong to the same class.⁴⁶ More information concerning elicitation methods can be found in references.⁴⁷⁻⁴⁹

CONCLUSION

The Bayesian models based on SSRI-favoured priors suggested that SSRIs might be better tolerated than TCAs in patients. However, as the posteriors were largely influenced by the priors, more studies need to be conducted before any confirmatory remark could be made. As there was no overwhelming evidence that SSRIs were indeed better tolerated than TCAs, physicians should exercise extra care in using antidepressants.

In evidenced-based biomedical research, expert opinions and published results form a rich source of information. By offering an alternative perspective in meta-analysis, the proposed Bayesian model is very useful when prior information regarding the unknown parameter of interest is available. It allows published results from literature search to be combined with expert opinions. As such, one expects the Bayesian model to provide more insight to biomedical problems with reliable prior information. In fact, one may view the classical model as a special case within the broader framework of Bayesian statistics. Thus, the

revolutionary Bayesian model is found to be more versatile than the conventional model.

ACKNOWLEDGEMENTS

We would like to express our appreciation to all the delegates who attended our presentation at the International Conference on Evidence-based Medicine held in Singapore on 18 to 19 October 2003. Their enlightening suggestions during the Q&A session helped to improve the final draft. The authors also wish to thank Dr Khin Lay Wai, Epidemiologist, Clinical Trials & Epidemiological Research Unit, Singapore, for reviewing the Bayesian model when it was first developed in 2001. Special thanks also go to the anonymous reviewers for their advice and insightful critiques.

REFERENCES

1. Sampson SM. Treating depression with selective serotonin reuptake inhibitors: a practical approach. *Mayo Clin Proc* 2001; 76:739-44.
2. Barbey JT, Roose SP. SSRI safety in overdose. *J Clin Psychiatry* 1998; 59 (15 Suppl):42-8.
3. Cheeta S, Schifano F, Oyefeso A, Webb L, Ghodse AH. Antidepressant-related deaths and antidepressant prescriptions in England and Wales, 1998-2000. *Br J Psychiatry* 2004; 184:41-7.
4. Goldstein BJ, Goodnick PJ. Selective serotonin reuptake inhibitors in the treatment of affective disorders-III. Tolerability, safety and pharmacoeconomics. *J Psychopharmacol* 1998; 12 (3 Suppl B):S55-87.
5. Pacher P, Ungvari Z, Nanasi PP, Furst S, Kecskemeti V. Speculations on difference between tricyclic and selective serotonin reuptake inhibitor antidepressants on their cardiac effects. Is there any? *Curr Med Chem* 1999; 6:469-80.
6. Stokes PE. Fluoxetine: a five-year review. *Clin Ther* 1993; 15:216-43.
7. Zohar J, Westenberg HG. Anxiety disorders: a review of tricyclic antidepressants and selective serotonin reuptake inhibitors. *Acta Psychiatr Scand Suppl* 2000; 403:39-49.
8. Katona C. Managing depression and anxiety in the elderly patient. *Eur Neuropsychopharmacol* 2000; 10 (4 Suppl):S427-32.
9. Baldwin DS. Unmet needs in the pharmacological management of depression. *Hum Psychopharmacol* 2001; 16(2 Suppl):S93-9.
10. Wagstaff AJ, Cheer SM, Matheson AJ, Ormrod D, Goa KL. Paroxetine: an update of its use in psychiatric disorders in adults. *Drugs* 2002; 62:655-703.
11. Linda VM, Ramesh S. Role of selective serotonin reuptake inhibitors in psychiatric disorders: a comprehensive review. *Prog Neuropsychopharmacol Biol Psychiatry* 2003; 27:85-102.
12. Lader MH. Tolerability and safety: essentials in antidepressant pharmacotherapy. *J Clin Psychiatry* 1996; 57 (2 Suppl):39-44.
13. Steffans DC, Krishnan KR, Helms MJ. Are SSRIs better than TCAs? Comparison of SSRIs and TCAs: a meta-analysis. *Depress Anxiety* 1997; 6:10-8.
14. Berry DA, Stangl DK. Bayesian methods in health-related research. In: Berry DA, Stangl DK, editors. *Bayesian Biostatistics*. New York: Marcel Dekker, 1996;3-66.

15. Chib S, Greenberg E. Understanding the Metropolis-Hastings algorithm. *Ann Math Stat* 1995; 49:327-35.
16. Brooks SP. Markov chain Monte Carlo method and its application. *Stat* 1998; 47:69-100.
17. Carlin BP, Gelman A, Neal RM. Markov chain Monte Carlo in practice: a roundtable discussion. *J Am Stat Assoc* 1998; 52:93-100.
18. Casella G, George EI. An introduction to Gibbs sampling. *Ann Math Stat* 1992; 46:167-74.
19. Thompson C, Thompson CM. The prescribing of antidepressants in general practice. II: a placebo-controlled trial of low-dose dothiepin. *Hum Psychopharmacol* 1989; 4:191-204.
20. Kernick DP. Which antidepressant? A commentary from general practice on evidence-based medicine and health economics. *Br J Gen Pract* 1997; 47:95-8.
21. Kyle CJ, Petersen HE, Overo KF. Comparison of the tolerability and efficacy of citalopram and amitriptyline in elderly depressed patients treated in general practice. *Depress Anxiety* 1998; 8:147-53.
22. Mahendru RK, Mahendru S. Selection of antidepressant drugs in general practice. *J Ind Med Assoc* 2001; 99:54-5.
23. Montgomery SA, Henry J, McDonald G. Selective serotonin reuptake inhibitors: meta-analysis of discontinuation rates. *Int Clin Psychopharmacol* 1994; 9:47-53.
24. Montgomery SA, Kasper S. Comparison of compliance between serotonin reuptake inhibitors and tricyclic antidepressants: a meta-analysis. *Int Clin Psychopharmacol* 1995; 4 (9 Suppl):33-40.
25. Anderson IM. Selective serotonin reuptake inhibitors versus tricyclic antidepressants: a meta-analysis of efficacy and tolerability. *J Affect Disord* 2000; 58:19-36.
26. MacGillivray S, Arroll B, Hatcher S, Ogston S, Reid I, Sullivan F, et al. Efficacy and tolerability of selective serotonin reuptake inhibitors compared with tricyclic antidepressants in depression treated in primary care: a systematic review and meta-analysis. *BMJ* 2003; 326:1014-7.
27. Yildiz A, Pauler DK, Sachs GS. Rates of study completion with single versus split daily dosing of antidepressants: a meta-analysis. *J Affect Disord* 2004; 78:157-62.
28. Barbui C, Hotopf M, Freemantle N, Boynton J, Churchill R, Eccles MP, et al. Treatment discontinuation with selective serotonin reuptake inhibitors (SSRIs) versus tricyclic antidepressants (TCAs). *Cochrane Database Syst Rev* 2004;(2).
29. Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomised trials. *Prog Cardiovasc Dis* 1985; 27:335-71.
30. Corne SJ, Hall JR. A double-blind comparative study of fluoxetine and dothiepin in the treatment of depression in general practice. *Int Clin Psychopharmacol* 1989; 4:245-54.
31. Stott PC, Blagden MD, Aitken CA. Depression and associated anxiety in primary care: a double-blind comparison of paroxetine and amitriptyline. *Eur Neuropsychopharmacol* 1993; 3:324-5.
32. Rosenberg C, Damsbo N, Fuglum E, Jacobsen LV, Horsgard S. Citalopram and imipramine in the treatment of depressive patients in general practice. A Nordic multicentre clinical study. *Int Clin Psychopharmacol* 1994; 9 (1 Suppl):41-8.
33. Doogan DP, Langdon CJ. A double-blind, placebo-controlled comparison of sertraline and dothiepin in the treatment of major depression in general practice. *Int Clin Psychopharmacol* 1994; 9:95-100.
34. Moon CAL, Vince M. Treatment of major depression in general practice: a double-blind comparison of paroxetine and lofepramine. *Br J Clin Pract* 1996; 50:240-4.
35. Christiansen PE, Behnke K, Black CH, Ohrstrom JK, Bork-Rasmussen H, Nilsson J. Paroxetine and amitriptyline in the treatment of depression in general practice. *Acta Psychiatr Scand* 1996; 93:158-63.
36. Ravindran AV, Judge R, Hunter BN, Bray J, Morton NH. A double-blind, multicentre study in primary care comparing paroxetine and clomipramine in patients with depression and associated anxiety. Paroxetine study group. *J Clin Psychiatry* 1997; 58:112-8.
37. Thompson C. Sertraline in a primary care setting. In: Racagni NB, Fukuda T, editors. *Biological Psychiatry*. Amsterdam: Elsevier, 1991;863-5.
38. Hutchinson DR, Tong S, Moon CAL, Vince M, Clarke A. Paroxetine in the treatment of elderly depressed patients in general practice: a double-blind comparison with amitriptyline. *Int Clin Psychopharmacol* 1992; 6 (4 Suppl):43-51.
39. Swinkels JA, De Jonghe F. Safety of antidepressants. *Int Clin Psychopharmacol* 1995; 9 (4 Suppl):19-25.
40. Kasper S, Hoflich G, Scholl HP, Moller HJ. Safety and antidepressant efficacy of selective serotonin re-uptake inhibitors. *Hum Psychopharmacol* 1994; 9:1-12.
41. Glass GV. Primary, secondary and meta-analysis of research. *Edu Res* 1976; 5:3-8.
42. DerSimonian R, Laird N. Meta-analysis in Clinical Trials. *Control Clin Trials* 1986; 7:177-88.
43. Fleiss JL. The statistical basis of meta-analysis. *Stat Methods Med Res* 1993; 2:121-45.
44. Kadane JB. Progress toward a more ethical method for clinical trials. *J Med and Philosophy* 1986; 11:385-404.
45. Spiegelhalter DJ, Freedman LS, Parmar MK. Bayesian approaches to randomised trials. *J Royal Stat Soc (Ser A)* 1994; 157:357-416.
46. Greenhouse JB, Wasserman L. Robust Bayesian methods for monitoring clinical trials. *Stat Med* 1995; 14:1379-91.
47. Chaloner K. Elicitation of prior distributions. Bayesian methods in health-related research. In: Berry DA, Stangl DK, editors. *Bayesian Biostatistics*. New York: Marcel Dekker, 1996;141-56.
48. Kadane JB, Wolfson LJ. Experiences in elicitation. *Stat* 1998; 47:3-19.
49. O'Hagan A. Eliciting expert beliefs in substantial practical applications. *Stat* 1998; 47:21-35.

DERIVATION OF THE JOINT POSTERIOR DISTRIBUTION

Let y_i : observed effect size in study i , φ_i : study-specific effect size in study i , θ : overall effect size, \mathbb{S}_i^2 : within-study variance of y_i , ϕ_i : within-study precision of y_i ($1/\mathbb{S}_i^2$), σ^2 between-studies variance, τ : between-studies precision ($1/\sigma^2$) and k : total number of studies included for meta-analysis. Note that only 2 quantities, namely y_i and \mathbb{S}_i^2 , are observed (ϕ_i is derived), while φ_i , θ and σ^2 are unobserved and unknown. Since the selected studies for meta-analysis are usually independently conducted, it is reasonable to assume that $y_i \sim \text{i.i.d.}[\varphi_i, \mathbb{S}_i^2]$ and $\varphi_i \sim \text{i.i.d.}[\theta, \sigma^2]$. Collecting terms and based on the sequential nature of Bayes' theorem, one yields the more explicit form of the joint posterior distribution:

$$p[\theta, \varphi, \tau, \phi | \mathbf{y}] \propto \prod_{i=1}^k p[y_i; \phi_i | \varphi_i] \times \prod_{i=1}^k p[\varphi_i | \theta] \times g[\theta] \times g[\tau]$$

where $\mathbf{y}=(y_1, y_2, \dots, y_k)'$ is the vector for the observed sample, $\phi=(\phi_1, \phi_2, \dots, \phi_k)'$ the vector of observed within-study precision, $\varphi=(\varphi_1, \varphi_2, \dots, \varphi_k)'$ the vector for study-specific effect sizes which belong to a distribution characterised by θ , $g[\theta]$ the prior for θ , $g[\tau]$ the prior for τ , $p[y_i; \phi_i | \varphi_i]$ the individual sample distribution for y_i and $p[\varphi_i | \theta]$ the individual distribution for study-specific effect sizes. With $\theta \sim \text{Normal}[\mu, \nu]$ and $\tau \sim \text{Gamma}[\lambda, \eta]$, the joint posterior may be explicitly written as:

$$p[\theta, \varphi, \tau, \phi | \mathbf{y}] \propto \prod_{i=1}^k \sqrt{\frac{\phi_i}{2\pi}} \exp\left[-\frac{\phi_i}{2}(y_i - \varphi_i)^2\right] \times \prod_{i=1}^k \sqrt{\frac{\tau}{2\pi}} \exp\left[-\frac{\tau}{2}(\varphi_i - \theta)^2\right] \\ \times \sqrt{\frac{\nu}{2\pi}} \exp\left[-\frac{\nu}{2}(\theta - \mu)^2\right] \times \frac{\eta^\lambda}{\Gamma(\lambda)} \tau^{\lambda-1} \exp[-\eta\tau]$$

where $\Gamma(\bullet)$ is a gamma function.

MCMC ALGORITHM VIA THE GIBBS SAMPLER

With a complicated joint posterior, it is more efficient to work on the conditional posterior distributions. The conditional posterior distributions are relatively simple in structure. For example, φ_i occurs in only 2 terms and its posterior conditional on other parameters is:

$$p[\varphi_i | \varphi_{-i}, \theta, \tau, \phi, \mathbf{y}] \propto \sqrt{\frac{\phi_i}{2\pi}} \exp\left[-\frac{\phi_i}{2}(y_i - \varphi_i)^2\right] \times \sqrt{\frac{\tau}{2\pi}} \exp\left[-\frac{\tau}{2}(\varphi_i - \theta)^2\right]$$

where φ_{-i} represents the vector of all other study-specific effects in studies other than i . This is a product of 2 normal distributions. Similarly, the posterior of major concern

$$p[\theta | \varphi, \tau, \phi, \mathbf{y}] \propto \prod_{i=1}^k \sqrt{\frac{\tau}{2\pi}} \exp\left[-\frac{\tau}{2}(\varphi_i - \theta)^2\right] \times \sqrt{\frac{v}{2\pi}} \exp\left[-\frac{v}{2}(\theta - \mu)^2\right]$$

is again a normal distribution. Last but not least, the conditional posterior for τ (between-studies precision) is:

$$p[\tau | \theta, \varphi, \phi, \mathbf{y}] \propto \prod_{i=1}^k \sqrt{\frac{\tau}{2\pi}} \exp\left[-\frac{\tau}{2}(\varphi_i - \theta)^2\right] \times \frac{\eta^k}{\Gamma(\lambda)} \tau^{\lambda-1} \exp[-\eta\tau]$$

The central idea of MCMC is to cycle through the conditionals and randomly draw the parameters. The following procedure is based on the concept of Gibbs sampling:

1. with starting values $\varphi_i^{(0)} = y_i$, $\theta^{(0)} = \sum_{i=1}^k \varphi_i / k$ and $\tau_b^{(0)} = k / \sum_{i=1}^n (\varphi_i - \theta)^2$
2. draw each φ_i randomly using its conditional posterior and the current values of θ and τ
3. draw θ randomly using its conditional posterior and the current values of φ and τ
4. draw τ randomly using its conditional posterior and the current values of φ and θ
5. record the current values of φ , θ and τ
6. repeat steps 2 to 5 for a sufficiently large number of times, say 1000
7. the parameters generated represent a sample from the full joint posterior
8. summarise θ from the generated sample of posteriors by computing its mean, variance and interval estimates

If the procedure is run sufficiently long, one may eventually reach the true posterior distribution of interest.