Second, several sources state that 100 mg sertraline is a sufficient dose.[1,2] Moreover, the dose is a valid representation of usual practice in Iran, as there is reluctance to increase the dose given findings that 'often, adequate clinical activity, and saturation of the 5-HT transporters, are achieved at starting dosages. As a rule, higher dosages do not increase antidepressant efficacy, but may increase the risk of adverse effects'.[2]

Third, the difference in the amount of attention given is an inherent aspect of comparing behavioural activation and TAU in routine practice. Adjusting for this difference would lead to an invalid comparison in an effectiveness study. The question whether extra attention given to the TAU group would reduce the difference between behavioural activation and TAU is a legitimate one, but goes beyond the scope of this study.

Fourth, last observation carried forward was not used – this is a misinterpretation of the paper; intention-to-treat analysis was used, as it is the gold standard. Analysing only completers leads to biased conclusions. We used mixed regression analyses that use all available data and yield valid estimates under certain assumptions in the light of missing data.[3] The suggestion is that a therapy-completers analysis would yield different conclusions. However, the effects are quite similar when only treatment-completers are analysed – Hamilton Rating Scale for Depression: time × condition, $F(1,78.02) = 10.05$, $P = 0.002$; time squared × condition, $F(1,78.40) = 7.94$, $P = 0.006$; Beck Depression Inventory: time × condition, $F(1,78.02) = 6.84$, $P = 0.011$; time squared × condition, $F(1,78.35) = 5.37$, $P = 0.023$.

Fifth, the influence of referral type was analysed, and tables with statistics are available online.[4] It is difficult to understand that this was missed (e.g. 'referral did not change the condition × time and condition × time squared effects', p. 207). Moreover, if anything, the differences between conditions were stronger in participants who were referred by healthcare professionals.

Finally, all patients were capable of understanding the information about the offered treatments and making the necessary decisions. All individuals were seen by a psychiatrist to check eligibility, including capacity to consent to participate in the study, as part of the good clinical practice guidelines we applied.

1 Browne G, Steiner M, Roberts J, Gafni A, Byrne C, Dunn E, et al. Sertraline and/or interpersonal psychotherapy for patients with dysthymic disorder in primary care: 6-month comparison with longitudinal 2-year follow-up of effectiveness and costs. *J Affect Disord* 2002; **68**: 317–30.

2 Sadock BJ, Sadock VA. *Kaplan and Sadock's Synopsis of Psychiatry: Behavioral Sciences/Clinical Psychiatry* (10th edn): 1085. Lippincott Williams & Wilkins, 2007.

3 Schafer JL, Graham, JW. Missing data: our view of the state of the art. *Psychol Methods* 2002; **7**: 147–77.

4 Moradveisi L, Huibers MJH, Renner F, Arasteh M, Arntz A. Behavioural activation v. antidepressant medication for treating depression in Iran: randomised trial. *Br J Psychiatry* 2013; **202**: 204–11.

Latif Moradveisi, Department of Clinical Psychological Science, Faculty of Psychology and Neuroscience, Maastricht University, The Netherlands, and Department of Psychiatry, Faculty of Medicine, Kurdistan University of Medical Science, Sanandaj, Iran; **Marcus J. H. Huibers**, Department of Clinical Psychology, VU University, Amsterdam, The Netherlands; **Fritz Renner**, Department of Clinical Psychological Science, Faculty of Psychology and Neuroscience, Maastricht University, The Netherlands; **Modabber Arasteh**, Department of Psychiatry, Faculty of Medicine, Kurdistan University of Medical Science, Sanandaj, Iran; **Arnoud Arntz**, Department of Clinical Psychological Science, Faculty of Psychology and Neuroscience, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands. Email: arnoud.arntz@maastrichtuniversity.nl

doi: 10.1192/bjp.202.6.466a

Effect of 9/11 on suicide: appropriateness of a time series model

Although the paper by Claassen *et al*[1] investigates an exciting issue, I have some concerns about the model identification. It seems that the authors identified the appropriate model of the time series only by using the Akaike Informations Criterion (AIC), which has certain limitations. For example, the selected ARMA (15,0) and ARMA (0,6) models are of high order and long memory. In general, the AIC suggests such models of high order only when a trend or seasonality is present in the analysed time series. Usually, if a time series is stationary, a model of an order below three is found.[2] A more complex method for model identification that avoids relying only on the AIC was introduced by Box & Jenkins.[2] Their algorithm includes several acquisition parameters in the process of model identification, which are:[4] 0, make the series stationary, consider differencing; 1, choose a provisional model; 2, estimate the model parameters; and 3, check the adequacy of the model.

One key aspect is the requirement of stationarity. If the time series is not stationary, an ARIMA model should be considered instead of a mere ARMA model. The ARIMA model enables one to include terms for a trend or seasonality, respectively, directly in the model. The high order of the chosen model makes it likely that the time series in the paper indeed possesses a trend or seasonality. Furthermore, as the ultimate assessment of a correct model, Box & Jenkins demanded non-significant auto-correlations of the residuals, which were apparently also not checked in the paper. As these important aspects were not respected, the chosen model might not be correct.

Figure 1 below displays a time series with an underlying trend. When an ARMA model is assumed, the AIC suggests an ARMA (6,0), which does not fulfil the requirement of non-significant autocorrelations of the residuals on a significance level of $\alpha = 0.05$. Nevertheless, the simple differentiation of the time series leads to a straightforward ARIMA (1,1,2), which, in contrast to the previous case, meets this requirement.

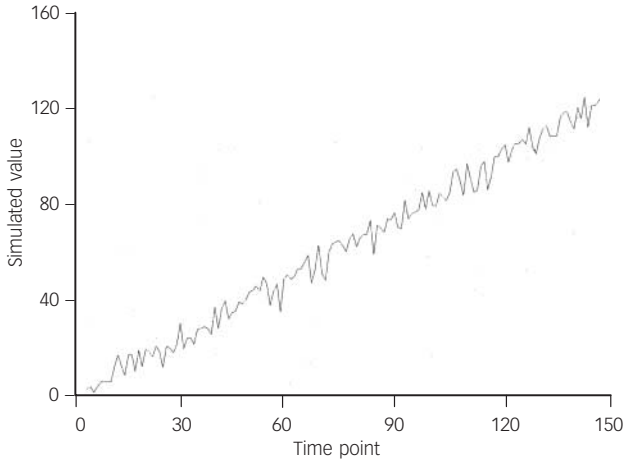

Fig. 1 Exemplary time series with an underlying trend.

The consequence of a non-fitting model would be a falsely estimated standard error, which would directly lead to insufficient statistical tests and thus incorrect *P*-values.[2,4–6] When the control group of suicides in 1998 was regarded, an even larger post-9/11 effect over a period of 180 days was found than in the group of interest (suicides in 2001). This effect was rejected because of non-significant statistical tests, which is, as shown above, not

https://doi.org/10.1192/bjp.202.6.467 Published online by Cambridge University Press

appropriate under the performed model identification. Therefore, it would be necessary to re-evaluate the considered time series in terms of model identification by the Box & Jenkins method and apply them again to the time series. I expect a notable change of results.

1 Claassen CA, Carmody T, Stewart SM, Bossarte RM, Larkin GL, Woodward WA, et al. Effect of 11 September 2001 terrorist attacks in the USA on suicide in areas surrounding the crash sites. *Br J Psychiatry* 2010; **196**: 359–64.

2 Rinne H, Specht K. *Zeitreihen: Statistische Modellierung, Schätzung und Prognose [Time Series: Statistical Modelling, Estimation, Prognosis]*. Vahlen, 2002.

3 Box GEP, Jenkins GM. *Time Series Analysis: Forecasting and Control*. Holden-Day, 1976.

4 Helfenstein U. Box–Jenkins modelling in medical research. *Stat Methods Med Res* 1996; **5**: 3–22.

5 Montgomery DC, Weatherby G. Modeling and forecasting time series using transfer function and intervention methods. *AIIE Transactions* 1980; **12**: 289–307.

6 Box GEP, Tiao GC. Intervention analysis with applications to economic and environmental problems. *J Am Stat Assoc* 1975; **70**: 70–9.

Daniel Medenwald, Institute of Medical Epidemiology, Biostatistics and Informatics, University of Halle-Wittenberg, Halle, Germany. Email: daniel.medenwald@uk-halle.de

doi: 10.1192/bjp.202.6.467

Little evidence for the usefulness of violence risk assessment

Troquete and colleagues report a cluster randomised trial of the effect of violence risk assessment on future offending.[1] They found that people in the risk assessment group were non-significantly more likely to re-offend than those in the control group. We welcome this analysis of the practical value of risk assessment. There are now literally thousands of published violence risk assessment studies, most of which claim validity for their risk assessment method on the basis of statistical discrimination between violent and non-violent groups using measures such as the area under the curve (AUC) or other indicators of effect size.[2] Recent criticism of the AUC as an outcome measure has emerged because it does not reflect the accuracy of predictions in the real world, and even high AUC values are associated with a low positive predictive value (PPV) for rare events. However, the PPV of a risk assessment is only a proxy for the usefulness of a risk assessment. A risk assessment alone is not valuable unless it leads reasonable interventions that can reduce future harm. Therefore, the utility of a risk assessment must ultimately be judged by its ability to contribute to harm reduction. In contrast to the large number of papers about the statistical aspects of risk assessment, there may be as few as four published controlled studies of the ability of risk assessment to reduce harm.[2]

The *British Journal of Psychiatry* has published two earlier studies of the utility of risk assessment. Abderhalden *et al* reported a cluster randomised trial of risk assessment among in-patients that found that intervention wards had a reduction in violence. However, interpretation of this study is difficult because the intervention wards had high rates of violence pre-trial and post-trial rates of violence in the experimental and control wards did not differ.[3] Also in the *Journal*, van de Sande and colleagues reported a cluster randomised trial that found that risk assessment was associated with a reduction in violence but not seclusion among in-patients.[4] In the nursing literature, Kling *et al* reported a study in in-patient settings that found that risk assessment was not helpful in reducing violence.[5]

Risk assessment has become the dominant paradigm in mental health practice, policy and legislation in most high-income countries. It should therefore trouble colleagues who support 'evidence-based practice' to know that there is so little evidence for the effectiveness of risk assessment.

1 Troquete NAC, van den Brink RHS, Beintema H, Mulder T, van Os TWDP, Schoevers RA, et al. Risk assessment and shared care planning in out-patient forensic psychiatry: cluster randomised controlled trial. *Br J Psychiatry* 2013; **202**: 365–71.

2 Wand T. Investigating the evidence for the effectiveness of risk assessment in mental health care. *Issues Ment Health Nurs* 2012; **33**: 2–7.

3 Abderhalden C, Needham I, Dassen T, Halfens R, Haug H-J, Fischer JE. Structured risk assessment and violence in acute psychiatric wards: randomised controlled trial. *Br J Psychiatry* 2008; **193**: 44–50.

4 van de Sande R, Nijman HLI, Noorthoorn EO, Wierdsma AI, Hellendoorn E, van der Staak C, et al. Aggression and seclusion on acute psychiatric wards: effect of short-term risk assessment. *Br J Psychiatry* 2011; **199**: 473–8.

5 Kling RN, Yassi A, Smailes E, Lovato CY, Koehoorn M. Evaluation of a violence risk assessment system (the Alert System) for reducing violence in an acute hospital: a before and after study. *Int J Nurs Stud* 2011: **48**: 543–9.

Timothy Wand, Emergency Department, Royal Prince Alfred Hospital and Sydney Nursing School, University of Sydney, NSW, Australia. Email: timothy.wand@sswhas.nsw.gov.au; **Matthew Large**, School of Psychiatry, University of New South Wales, Randwick, NSW, Australia

doi: 10.1192/bjp.202.6.468

Authors' reply: We agree with Wand & Large that there currently is very limited support for the use of structured risk assessment instruments as a method for violence prevention. So far only a small number of studies, four including our own, examined this issue. It is troubling that most research efforts seem to focus on the development of new risk assessment instruments and establishing their psychometric properties, rather than on testing the effectiveness of existing instruments. Although identification of predictors and development of instruments are crucial steps in the maturation of both risk assessment and forensic psychiatry, the field needs to move beyond these issues.

The most important risk and protective factors associated with recidivism have by now been established and are agreed on by the research community. There is no disputing the existence of correlations between mental illness, substance misuse, client well-being, quality of life and recidivism. That is why all, or a considerable selection of these factors, are commonly included in risk assessment instruments.[1–3] It seems it is time to move forward and start investigating the benefits of risk assessment instruments and their contribution to more effective treatment interventions in terms of reduction of criminal and violent behaviour. As we ourselves have experienced, introducing randomised trials in clinical practice is difficult, but it can be done, and is an essential step before implementation can be advocated.

A definitive answer about the contribution of structured risk assessment to violence prevention cannot be given at this time. The first signs are not good. The four available studies find either no significant reduction of violent outcome, or the interpretation of their findings is problematic due to differences between study groups at baseline. Differences in clinical setting of the various studies further complicate the integration of findings. Our own data were collected in a community-based forensic mental health setting. In contrast, the other three studies were completed in acute psychiatric (admission) wards. These two settings service different populations, making comparisons less straightforward. It is too early for a proper systematic review on this subject, but the overall picture is not yet convincingly in favour of changing treatment policies by systematically employing structured risk assessment in clinical care.

https://doi.org/10.1192/bjp.202.6.467 Published online by Cambridge University Press