guidelines for reporting trials. In this case, when the light boxes were modified to 100 lux, the disparity in intensity was very obvious and we did not feel that the study would conform to the important double-blind aspect of the design. It would have been very clear to any patient who received the 100 lux box that they had been assigned to the low-intensity arm of the trial. We therefore modified the boxes to administer 2000 lux at 20 min in the low-intensity arm. The boxes appeared bright, but literature on seasonal affective disorder indicates that this would not be a therapeutic dose within this time frame, whereas 10 000 lux at 20 min would be a therapeutic intensity/dose.

As we stated in the introduction to our study, the primary outcome measure for this trial was seizure control. We have reported these results separately[1] and that paper is fully referenced in our study. Although it is possible that bright light therapy may result in an increase in seizures for some patients, this was not a statistically significant finding in our previous study and, as yet, the risk remains theoretical. Clinicians will be aware that seizure control should be carefully monitored following the introduction of any new treatment offered to people with epilepsy.

In presenting the results of our study for publication we have sought to provide as clear an account of the data as possible. The results are by no means clear-cut or definitive. However, there are some interesting aspects to the data that suggest that this may not be a dead end in terms of a treatment option for some people with epilepsy. This study stands as a guide for future research. We hope that its limitations, which we fully acknowledge and have set out at length in the Discussion, will serve as a useful guide for future research in this area.

1 Baxendale S, O'Sullivan J, Heaney D. Bright light therapy as an add on treatment for medically intractable epilepsy. *Epilepsy Behav* 2012; **24**: 359–64.

Sallie Baxendale (on behalf of the authors), Department of Neuropsychology (Box 37), National Hospital for Neurology & Neurosurgery, Queen Square, London WC1N 3BG, UK. Email: s.baxendale@ucl.ac.uk

doi: 10.1192/bjp.202.6.465b

Results for behavioural activation are overstated

The study by Moradveisi *et al*,[1] which is applicable to both secondary mental health and primary care, looks at the prospect of using minimally trained staff in delivering behavioural activation against pharmacological intervention in the treatment of severe depression. We would like to highlight the following points for further clarification.

First, an obvious problem of the study was the lack of a placebo arm, which would have lent credibility. As the cultural avoidance of antidepressants in Iran has been highlighted, adding a placebo group would have removed some bias such as paying for medication in the treatment as usual (TAU) group after 3 months and also in the analysis.

Second, sertraline was used at a suboptimal dose and was slowly titrated, against prevailing practice. A meta-analysis shows an optimum dose for sertraline between 100 and 150 mg/day – doses below the therapeutic range were significantly less effective, i.e. by 7%.[2] Sertraline reached its lowest therapeutic dose of 100 mg at 6 weeks. All drop-outs occurred before the mid-point assessment and only three were as a result of medication side-effects.

Third, there was a significant difference in the amount of attention that participants received in each group. Participants in the behavioural activation group received 50% more face-to-face sessions than the TAU group. The study did not adjust for this in the analysis.

Fourth, last observation carried forward (LOCF) was used in the study. However, 5% of drop-outs occurred in the behavioural activation group as opposed to a significant 30% from the TAU group. Last observation carried forward is used frequently in intention-to-treat studies but standard errors and confidence intervals from LOCF underestimate uncertainty.[3] As there are no strategies for universal use, reasons for the choice of a certain method have to be provided when designing and analysing clinical trials.[4] Last observation carried forward analysis seems to have favoured the behavioural activation group.

Many other limitations of the study are cited in the paper itself. Significant numbers of participants were recruited via advertisement or word of mouth, which seemed to have attracted more women and perhaps more psychologically minded individuals. It would have been helpful to include these advertisements as a supplement to the paper in order to identify any bias.

Finally, we wondered whether an ethics committee would allow this type of study to go ahead in the UK as it included individuals with severe depression. In England and Wales, before recruitment to a trial, potential participants must be assessed under the Mental Capacity Act 2005; in Scotland, the Adults with Incapacity (Scotland) Act 2000 (para. 72) must be used.[5] Since the authors of the study state that 'the study's aim was to investigate whether a simple psychological treatment [. . .] would be a viable alternative to antidepressant medication [. . .] in a non-Western country', we are unsure of an equivalent law in Iran and whether this criterion was met.

1 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.

2 Bollini P, Pampallona S, Tibaldi G, Kupelnick B, Munizza C. Effectiveness of antidepressants. Meta-analysis of dose–effect relationships in randomised clinical trials. *Br J Psychiatry* 1999; **174**: 297–303.

3 Mallinckrodt C, Clark W, David S. Accounting for dropout bias using mixed-effects models. *J Biopharm Stat* 2001; **11**: 9–21.

4 Unnebrink K, Windeler J. Intention-to-treat: methods for dealing with missing values in clinical trials of progressively deteriorating diseases. *Stat Med* 2001; **20**: 3931–46.

5 General Medical Council. *Consent: Patients and Doctors Making Decisions Together*. GMC, 2008.

Mukesh Kripalani, Adult Crisis Resolution & Intensive Home Treatment, Tees, Esk and Wear Valleys NHS Foundation Trust, Middlesbrough, UK. Email: drmukesh@doctors.org.uk; **Muhammad Suleman**, Adult Crisis Resolution & Intensive Home Treatment, Adult Affective Team, Tees, Esk and Wear Valleys NHS Foundation Trust, Middlesbrough, UK

doi: 10.1192/bjp.202.6.466

Authors' reply: We thank Kripalani & Suleman for their critical remarks. Before addressing them point by point, a general remark is required. Our trial was an effectiveness, not an efficacy, trial. We compared a new treatment previously tested elsewhere (behavioural activation) with treatment as usual (TAU) (antidepressant medication) in Iran. An effectiveness trial aims to assess outcomes in usual care, not to test specific mechanisms, which affects the type of control condition(s). Some criticisms make sense from an efficacy study point of view, not from an effectiveness study point of view. Also of note is that the initial response to TAU was quite good, and that the longer-term response of behavioural activation accounted for its superiority.

We do not see how a placebo arm could have assessed cultural influences on TAU. To study this interesting topic, both a placebo and a natural course condition are needed to see whether placebo in Iran does worse than in other cultures compared with doing nothing.

https://doi.org/10.1192/bjp.202.6.466a Published online by Cambridge University Press

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.

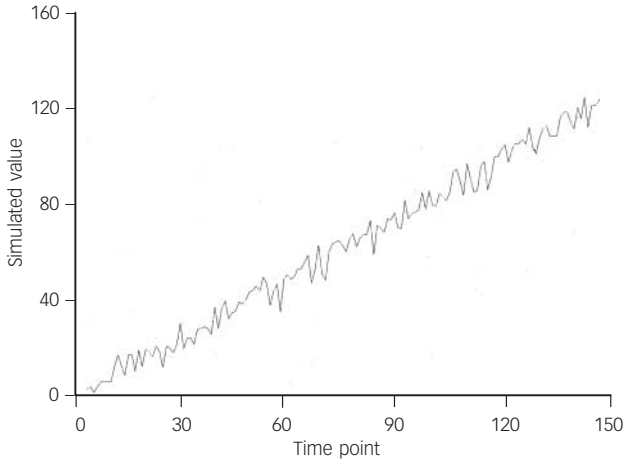

Fig. 1 Exemplary time series with an underlying trend.

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