The STAR*D trial: the 300 lb gorilla is in the room, but does it block all the light?

Simon Hatcher

It cost US tax payers $35 million and the results were announced on the front page of the Washington Post—so what did the STAR*D trial tell us about how to help people presenting with depression? First the name, STAR*D stands for Sequenced Treatment Alternatives to Relieve Depression. The emphasis in STAR*D was that it was a series of pragmatic trials that as closely as possible replicated what was possible in usual clinical care. The trials are divided into four groups or levels, with each level consisting of several randomised controlled trials with the participants being people with depression who hadn't responded to treatment at the previous level. The study is the largest series of randomised controlled trials ever done in psychiatry and the results are complicated and published in numerous papers in several different journals. There is however a website hosted by the funder which summarises most of the findings (http://www.nimh.nih.gov/health/trials/practical/stard/index.shtml). For this article I have chosen to focus on the results of treatment at the different levels in the trial.

LEVEL 1: CITALOPRAM IN MAJOR DEPRESSION

The first trial was not a randomised controlled trial and was designed to answer the question what happens to people with depression who are treated with citalopram in routine clinical practice and what are the important prognostic indicators.[1] The potential participants were 4041 outpatients aged 18–75 with DSM-IV defined non-psychotic major depressive disorder referred into the trial by their clinicians. Participants were recruited from 41 clinical sites in the USA which involved 18 primary care clinics (38% of participants) and 23 specialty psychiatric settings (62% of participants). People were not eligible if

Correspondence to: Simon Hatcher, Department of Psychological Medicine, University of Auckland, Private Bag 92109, Auckland, New Zealand; s.hatcher@auckland.ac.nz

they had already had an adequate trial of the treatments in the first two levels of the STAR*D study. Of the 4041 potential participants, 1165 were excluded because they either did not meet the study requirements of having "moderate" depression—that is, a score on the 17-item HAM-D of 14 or more, (931 people) or they chose not to participate, leaving 2876 in the trial. Of these, two thirds were women with an average age of 41; for 25% the current episode of depression had lasted for at least two years; and 18% had a history of attempted suicide. Of note here is that a major incentive for patients taking part in this trial is that they got free healthcare within their local health system.

The treatment these 2876 people were intended to receive was 10–60 mg of citalopram for up to 14 weeks involving five or six outpatient visits. Also, the clinicians were supported with a treatment manual, a web-based treatment monitoring system and a centralised system of feedback of participants' reported symptoms, adverse effects and medication adherence. This part of the study was open label—so everyone knew what was being prescribed.

At the end of the 14 weeks 28% of participants had achieved remission (defined as a 17-item HAM-D score of seven or less) taking a mean dose of citalopram of 55 mg/day over 12 weeks. People more likely to achieve remission were well-educated, employed, married, white and female, with few complicating problems. Factors associated with a poorer response included co-occurring anxiety, substance abuse or physical disorders, and lower quality of life.

So this meant at the end of 14 weeks' treatment there were 2086 patients who had not achieved remission who were invited to progress to the next stage of treatment.

LEVEL 2: SWITCHING OR AUGMENTATION STRATEGIES

Study two consisted of several randomised controlled trials to test whether

switching antidepressant or augmentation with a second drug was effective in those people who had not responded in the first study. Patients who had not responded were asked to choose which randomised controlled trial they would like to participate in. This allowed patients to express some preference about which treatments they found acceptable. 1439 people who had not become symptom-free (69% of non-remitters from study one) chose to continue. The level 2 trials which patients could choose were:

▶ Medication and Psychotherapy Switch: switch to sertraline, venlafaxineXR, bupropionSR, or cognitive therapy.

▶ Medication and Psychotherapy Augmentation: add to citalopram either (a) buspirone, (b) bupropionSR or (c) cognitive therapy.

▶ Medication Only Switch or Medication Only Augmentation options were available for participants for whom cognitive therapy was unacceptable.

▶ Psychotherapy Only Switch or Psychotherapy Augmentation Only options were available for participants for whom additional medication was unacceptable (participants must have been willing to continue citalopram).

Most people chose to switch or augment their medication with another drug. Fifty one per cent (727) of the patients chose options that included switching to a different medication and were randomly assigned to one of the three switch medications. Thirty nine per cent (565) chose options that included augmenting the citalopram they were already taking, and were randomly assigned to one of the two augmenting medications.

SWITCHING MEDICATION

The 727 patients who received the switch medication treatments were randomised to change to sertraline, bupropion-SR or venlafaxine-XR.[2] A quarter of these patients became symptom-free within 14 weeks; this was similar within each of the three treatment groups. Additionally, no significant differences were found in the efficacy, safety or tolerability of the three medications to which patients were switched. Interestingly, the switch to the new medication was done directly with no washout period or reduction of dose without any apparent problems.

AUGMENTING MEDICATION

The 565 patients who received the augment medication were randomised to take either bupropion-SR or buspirone in addition to the citalopram that they were already taking in level 1.[3] Within 14 weeks of using either treatment, about one third of the patients who enrolled in the augmentation study became symptom-free. Buprorion-SR had significantly fewer dropouts due to intolerance than buspirone (12.5% vs 20.6%).

COGNITIVE THERAPY

147 participants (that is, only 10% of those who chose to continue into level 2) either switched to cognitive therapy or added it as an adjunctive treatment to citalopram.[4] Participants received up to 16 sessions of CT over 12 weeks. About 25% of those who switched to cognitive therapy alone, and about 23% of those who added it, became symptom-free. The rates were not significantly different from those who were in medication-only treatment pathways in level 2. However, among those in the cognitive therapy add-on group, remission took longer to achieve, an average of 55 days, compared with an average of 26 days among those who augmented the citalopram with another medication. The time to remission among the cognitive therapy only switch group was not statistically different from those who switched to another medication. But those who switched to cognitive therapy alone were spared the side effects experienced by those who switched to another medication. Only 369 people (25% of those entering level 2) were prepared to be randomised to cognitive therapy and these participants were more likely to be more educated, have a family history of depression or bipolar disorder or a greater length of time in level one treatment compared to those who weren't willing to have cognitive therapy. Note there was no assessment of clinician preferences or their characteristics.

LEVEL 3: PARTICIPANTS WITHOUT SATISFACTORY RESPONSE SWITCHING OR AUGMENTATION AGAIN

Patients who had not achieved remission in the previous trials could now choose to enter one of two randomised controlled trials:

► Medication Switch to: (a) mirtazapine or (b) nortriptyline.
► Medication Augmentation: add either (a) lithium or (b) thyroid hormone (T3).

377 participants chose to continue onto the level 3 randomised trials. 234 patients chose to switch medications and were randomly assigned to take either mirtazapine or nortriptyline for up to 14 weeks.[5] Overall, the two medications were about equally effective with 10–20% of patients becoming symptom-free.

Of the 142 participants who chose the augmentation trials, each was randomly prescribed either lithium or triiodothyronine to add to the medication they were already taking.[6] After being on one of these new combinations for an average of nine weeks, about 20% of participants became symptom-free. Those taking T3 complained of fewer troublesome side effects than those taking lithium. In addition significantly fewer patients on T3 stopped treatment, with only 10% of people taking the T3 discontinuing treatment, while 23% of those taking lithium discontinued. However the adequacy of the lithium treatment was hard to ascertain as only 57% of patients had lithium levels done with a median concentration of 0.6 mmol/l.

LEVEL 4: PARTICIPANTS WITHOUT AN ADEQUATE RESPONSE TO LEVEL 3 SWITCHED AGAIN

Patients who had so far not responded were eligible for random assignment to two further medication switch options: tranylcypromine or mirtazapine plus venlafaxineXR.[7] In level 4, 109 participants who had not become symptom-free in any of the previous levels were taken off all other medications and switched to one of the two treatments. After an average of nine weeks, about 10% of participants became symptom-free. Those taking the venlafaxine-XR/mirtazapine combination had fewer side effects, stayed on the medications longer, and had lower dropout rates.

COMMENTARY

One conclusion from this series of studies is that treating depression is more difficult than clinicians and researchers had previously assumed. At the end of four levels of treatment about a third of patients who had stayed in the study had failed to recover from their symptoms.[8] However there are a number of difficulties with STAR*D which makes it hard to generalise the findings. These will be dealt with in turn.

THE ABSENCE OF A "PLACEBO"

There was no placebo arm in level one of the trial. This is an important omission.

In primary care support, explanation and time ("wait and see") are important interventions in the management of low mood. At the very least the third of patients recruited from primary care should have been offered the option of a wait and see or a placebo option. For an example of the importance of this, see an April 2002 *JAMA* study[9] which compared the effectiveness of sertraline, St John's wort and a placebo in depressed patients. In this study, the placebo-treated patients had a 31.9% rate of remission of symptoms—about the same as with citalopram in level one of STAR*D. (Also interestingly, the criterion for inclusion in the *JAMA* study was a HAMDRS score of 20 or higher, while in STAR*D it was only 14 or higher).

INCLUSION CRITERIA

There are two problems with the inclusion criteria. Firstly, those people who had already not responded to the treatments in the STAR*D trial were excluded from the study. It is hard to know how significant this is without knowing about the treatment history of those people not eligible for the trial compared with those who were. However, it has the potential to inflate the remission rate in STAR*D. Secondly to consent to inclusion in the trial participants had to find prescription of an antidepressant acceptable (level one). This would exclude those who would rather have a non-pharmacological option and probably explains why so few people chose to be randomised to a psychotherapy option in level two.

CHOICE OF STRATEGIES

The choice of strategies appears hard to justify. Why choose buprorion and buspirone as initial augmentation strategies when there appears better evidence[10] for augmentation with lithium or olanzapine as a next-step drug strategy after a trial of a single antidepressant has failed? It is worth noting that although the study was funded from public money the lead investigators had significant relationships with the makers of the drugs used in the study. For example Dr A John Rush and Dr Madhukar H Trivedi receive consulting fees from or served on the advisory boards for Forest Pharmaceuticals (Celexa), Wyeth-Ayerst Laboratories (Effexor) and Bristol-Myers Squibb (Buspar); GlaxoSmithKline (Wellbutrin) and Pfizer (Zoloft); and Dr Rush had an equity interest in Pfizer.

GENERALISABILITY OUTSIDE THE USA

The study was done in a variety of clinics in the USA. Given the diverse range of healthcare providers, the relative lack of primary care and the difficulty accessing care by people without insurance it is hard to know how these results generalise into non-US health settings.

CONCLUSION

STAR*D is a large and complicated trial and like most large and complicated trials it is hard to draw clear conclusions from it. It is a trial of pharmacological treatments in those people who find drug treatments acceptable. It is of little use to clinicians and patients in primary care who don't want to start antidepressants straight away. For those working in secondary care it provides some information about what happens to people who start on an antidepressant and provides some information about potential (but by no means all) next-step strategies. The STAR*D trial may be the 300 lb gorilla of clinical trials but disappointingly it only sheds a little light on how to manage depression in clinical practice.

Competing interests: None.

REFERENCES

1. **Trivedi MH,** Rush AJ, Wisniewski SR, *et al*. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry* 2006;**163**:28–40.
2. **Rush AJ,** Trivedi MH, Wisniewski SR, *et al*. Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression. *N Engl J Med* 2006;**354**:1231–42.
3. **Trivedi MH,** Fava M, Wisniewski SR, *et al*. Medication augmentation after the failure of SSRIs for depression.[see comment]. *N Engl J Med* 2006;**354**:1243–52.
4. **Thase ME,** Friedman ES, Biggs MM, *et al*. Cognitive therapy versus medication in augmentation and switch strategies as second-step treatments: a STAR*D report. *Am J Psychiatry* 2007;**164**:739–52.
5. **Fava M,** Rush AJ, Wisniewski SR, *et al*. A comparison of mirtazapine and nortriptyline following two consecutive failed medication treatments for depressed outpatients: a STAR*D report. *Am J Psychiatry* 2006;**163**:1161–72.
6. **Nierenberg AA,** Fava M, Trivedi MH, *et al*. A comparison of lithium and T(3) augmentation following two failed medication treatments for depression: a STAR*D report. *Am J Psychiatry* 2006;**163**:1519–30; quiz 1665.
7. **McGrath PJ,** Stewart JW, Fava M, *et al*. Tranylcypromine versus venlafaxine plus mirtazapine following three failed antidepressant medication trials for depression: a STAR*D report. *Am J Psychiatry* 2006;**163**:1531–41; quiz 1666.
8. **Rush AJ,** Trivedi MH, Wisniewski SR, *et al*. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry* 2006;**163**:1905–17.
9. **Hypericum Depression Trial Study G.** Effect of Hypericum perforatum (St John's wort) in major depressive disorder: a randomized controlled trial. *JAMA* 2002;**287**:1807–14.
10. **Anderson I,** Ferrier I, Baldwin R, *et al*. Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 2000 British Association for Psychopharmacology guidelines. *J Psychopharmacol* 2008;**22**:343–96.

Unravelling factor analysis

Khalida Ismail

Factor analysis is a broad term that refers to a set of statistical methods used to detect underlying patterns in the relationships among a number of observed variables. Its origins were in the large scale studies defining the dimensions of intelligence pioneered by Thurstone.[1][2] Factor analysis can appear complicated to the general reader but the main principle is relatively straightforward: what it aims to do is identify whether the correlations between a set of multiple observed variables are explicable or can be summarised in terms of a smaller number of underlying, latent, unobserved variables, also called factors. It is useful to have a basic understanding of the specific techniques when reading articles about factor analysis. There are two main approaches: exploratory factor analysis and confirmatory factor analysis.

EXPLORATORY FACTOR ANALYSIS

Exploratory factor analysis is used for the preliminary investigation of a set of observed variables, especially where there are multiple variables, such as each question on the Hopkins Symptom Checklist.[3] In a population or sample where a diverse range of symptoms is under study, as is often the case in mental health research, the advantage of this method is that it makes no a priori assumption about the composition of underlying latent variables or factors. The applications of exploratory factor analysis are wide ranging:

1. Data reduction when multiple (over 25) variables have been measured, providing a parsimonious description of the data.

2. Classification of symptoms into clinically meaningful concepts especially when symptoms are many and diverse such as medically unexplained symptoms, symptoms to describe stress, multiple health beliefs and behaviours.

3. Definition of subscales of new measures of psychological functioning—a par exemplar is the validation of the well known General Health Questionnaire[4] which was shortened to 12 items after identifying which symptoms in the 60-item version were closely related to each other and therefore could be removed.

4. Informing the development of new hypotheses.

5. Assessment of the construct validity of a scale.

6. The generation of new factors or scores which can be used as variables in themselves.[5]

CONFIRMATORY FACTOR ANALYSIS

Confirmatory factor analysis is a method for testing whether a specified factor structure remains valid with a new dataset.[6] A factor structure represents a number of factors and the variables that load onto them. When trying to identify patterns in the multiple variables, a factor structure or a model can be proposed. The model aims to describe the associations between the factors and the observed variables, and these can then be tested to see if the proposed model holds true even if the parameters are changed. A specific model is assumed, from which "predicted" values for the correlations between the observed variables can be made. Whether or not the specified model is considered to provide an adequate explanation of how the observed variables intercorrelate—that is, "fit"—is determined by how "close" the predicted correlations are to those observed. Judging the fit of confirmatory factor analysis models is not straightforward and a variety of measures of fit are usually considered. Confirmatory factor analysis is primarily used for assessing the construct validity of questionnaires or tests.

PROBLEMS WITH FACTOR ANALYSIS

As is the case with any other statistical technique, the validity of the results from

Correspondence to: Khalida Ismail, Department of Psychological Medicine, Institute of Psychiatry, King's College London, Weston Education Centre, 10 Cutcombe Road, London SE5 9RJ, UK; khalida.ismail@iop.kcl.ac.uk