

‘Comparative efficacy and acceptability of 21 antidepressant drugs’, meta-analysis: critical observations and context

P. K. Gillman. PsychoTropical Research, Bucasia, Queensland, Australia.
Address for correspondence: kg@matilda.net.au

Abstract

The Cipriani et al. study in the *Lancet*, of ‘21 anti-depressants’ is critically appraised and the likelihood that it will be applied inappropriately in clinical practice, and without due regard for the extensive caveats which were expressed by the authors and must be observed, is discussed.

RCTs, and the resulting proliferation of meta-analyses they generate, do not deserve to be regarded as having the elevated evidential status that they presently receive. They do not have the superior epistemic validity, or reliability, or extrapolability, that appears to be assigned to them, even if they are carried out to exacting and ideal designs. Most are significantly flawed methodologically, and ghost writing, deceit, and bias are inadequately accounted for by many commentators.

This is especially important in relation to the small overall degree of improvement that these drugs engender. It is problematic to justify assigning to most of these drugs the epithet of ‘anti-depressants’, since in most cases the small ‘standardized mean difference’ of 0.3 indicates a modest reduction of symptoms, but not wellness or a normalisation of function.

The predominance, of meta-analysis in the landscape of clinical decision-making is such that there is a methodological monoculture. Other methodologies and clinical experience have been mistakenly relegated to inferior status. The pronouncements resulting from M-A may be simplistically interpreted and may assume a dictatorial attire and be mis-used by non-medical decision-makers. They foster a blinkered and inflexible approach to therapeutics which stifles innovation.

Introduction

This ‘Lancet Cipriani 21 antidepressants meta-analysis’ paper published in Feb 2018 in the *Lancet* (1), is yet another ‘**meta-analysis (MA)**’, adding to more than 200 already existing M-A studies about AD drugs (2). Professor Ioannidis (one of the authors) recently published ‘*The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses?*’. Does this one, perhaps the most thorough and prestigious thus-far, usefully advance the cause? When one looks at his previous comments on Cipriani’s last M-A (3) one might well wonder what happened to persuade Professor Ioannidis that yet another M-A might be worthwhile? It will become a much referred-to work, so it is important for it to be seen in a meaningful clinical treatment perspective.

I predict the caveats and limitations rightly expressed by the authors about the validity and generalisability of their data and results will soon, as usual, be forgotten in the oversimplified implementation resulting from the insufficiently informed or critical use of ‘guidelines’.

Address for correspondence. Dr Ken Gillman: kg@matilda.net.au

We must remember that the trials considered include only a highly selected and small sub-group of the general population who actually get treated with these drugs: they do not include those with suicidal depression (perhaps the most important group to treat effectively), the old, those with medical comorbidities, those on other medications, and that large group who fail to achieve remission (~70%) with the first-offered treatment (often an SSRI).

In effect, such analyses are germane to only around 10% of ‘real-world’ patients who are given antidepressant drug treatment, and none of those seen in specialist practice.

This paper clearly entailed a great deal of work, even if there are uncertainties concerning its strength and usefulness. These uncertainties are highlighted in this commentary.

One can only express admiration for the industry and application of those concerned in researching and writing this paper: the following comments are not intended to be dismissive or disrespectful of these eminent researchers.

One of my [recent commentaries](#) is a detailed discussion of the many problems with guidelines, which emanate largely from over-reliance on meta-analysis and, therefore inevitably, on randomised controlled trials.

The subject of guidelines and M-A is a complex issue (not amenable to a ‘postage-stamp’ précis) and those ‘coming to it’, with less knowledge and experience of science, will benefit from appreciating that there are extensive, serious, and relevant problems, in the practice and publishing of medical research, only some of which are mentioned here.

These are expanded-on in my commentary [‘Guidelines: problems aplenty’](#) and in other commentaries in the menu heading above ‘Bias in science’. These cover the considerable complications caused by scientific fraud, ghost-writing, hiding and distorting of ‘raw’ [un-coded] patient data, and more.

That commentary, ‘Guidelines: problems aplenty’, extensively cited professor Ioannidis (one of the authors of this paper). I cannot immediately think of many other researchers whose work I admire and respect more, nor ones whose work I have cited more frequently. I wonder how he will feel about his contribution to this paper in years to come?

Hackneyed as this old computer programmer’s phrase may be, it is obligatory to start by repeating it “garbage in, garbage out”. In layman’s language, you cannot make a silk purse out of a sow’s ear, nor build a castle on sand.

“On two occasions I have been asked, “Pray, Mr. Babbage, if you put into the machine wrong figures, will the right answers come out?” ... I am not able rightly to apprehend the kind of confusion of ideas that could provoke such a question.

Charles Babbage, *Passages from the Life of a Philosopher*”

Key problems

What exactly is being assessed? how is it assessed? who assessed it? are they competent? honest? trustworthy? can we see the actual patient record (clinical notes) or the original [un-coded] data? The answer to those questions is often ‘no’, or, ‘we cannot be sure’.

Science requires scrupulous honesty and objectivity, as well as accuracy and reproducibility: otherwise, it is just ‘a castle built on sand’.

Address for correspondence. Dr Ken Gillman: kg@matilda.net.au

These and other questions are key issues to be dealt with before even the most tentative treatment recommendations can be considered based on such meta-analyses, especially when they show only marginal degrees of improvement or benefit in symptoms, and no convincing evidence at all of improvements in medium to long-term real-life outcomes viz. reduction in suicide, less days off work, less time on state benefits for disability, improved leisure and social functioning etc.

It is also relevant to point out that the recommendations from this Cipriani analysis have huge implications for financial expenditure on health care, where billions of dollars are expended on AD drugs. It would be good to be confident in the quality of the evidence and certain about the extent and usefulness of the benefit.

A fundamental unaddressed question is whether all these drugs are meaningfully assigned the epithet of ‘antidepressant’ (labelling a drug as an AD has more to do with the ‘sales & marketing’ department than it has to do with agreement among pharmacologists, or proven benefit, as I have illustrated with the example of mirtazapine (4)). Next, a more critical evaluation of how well they actually work, and on what symptoms is needed. Also, their effect on the long-term course and illness outcome is not even being considered.

I will not even mention the failure to assess or discuss long-term adverse effects.

One needs to focus on the indisputable fact that the average improvement, compared to placebo, in the **Hamilton rating scale for depression (HAM-D)** in these trials is small, ~3 points only out of fifty compared to placebo. Most patients come no-where near getting completely better. Yet, to hear the talk about these results, one could be excused for thinking most patients were getting ‘better’. In fact, the ‘standardised mean difference’ (SMD) for the HRSD scores from Cipriani was 0.3, which equates to an HRSD score of around two points, see below for what that really means (5, 6).

Anyone paying for such ineffectual treatment might justifiably demand their money back.

Moncrieff’s comment is about the only ‘non-laudatory’ published response relating to Cipriani that I have found whilst updating this comment of mine (from its first iteration in Feb, a couple of weeks after Cipriani’s article was available). Moncrieff’s comment was rebuffed by Prof Young (7, 8), with the ‘interesting’ comeback ‘*In focusing the argument on change in total HRSD score, Dr. Moncrieff appears unaware that **the scale was never intended to measure change.** A more robust way of analysing it was recently demonstrated, using the rating of subjective mood (item 1 on the HRSD), which would be akin to the CGI. This avoided the influence of antidepressant side-effects on the scale and found clear benefits for paroxetine and citalopram over placebo.*’ So, never mind M-A, just hang your hat on whichever measurement you prefer!

Really? In that case, if ‘*it was never intended to measure change*’ why has it been ubiquitously used for that purpose over and over again, for decades? And then he informs us in a perverse assertion that using only 1 item from the whole HAM-D scale suddenly becomes better — miraculous! *I requiem meam doleat* (I rest my case), as Cicero might have said. Or as Sir Humphrey might have said ‘a refreshingly original and imaginative interpretation minister’. The goal-posts are being moved so fast in this discussion that I am becoming dizzy.

Address for correspondence. Dr Ken Gillman: kg@matilda.net.au

The HAM-D does not adequately assess the central symptoms of the illness, which are anergia and anhedonia. This key question of defining ‘biological’ drug-responsive depression is covered in various of my other commentaries, especially in relation to professor Parker’s work on ‘CORE’ symptoms (9-12).

The [Hamilton] rating-scale question is discussed in my above-mentioned [commentary on guidelines](#). To compound the poor rating-scale problem, especially the significant influence of ‘sedative’ effects on scores (independent of any AD effect), there is a minimal assessment of **functional capacity** related to anergia [and anhedonia] — the assessment of social activities, work activities, and leisure activities, as corroborating measures of lack of motivation and drive. These features are fundamental to understanding and assessing depression, and yet they are relatively poorly assessed and rated, or not assessed at all, by the rating scales on which these RCTs & meta-analyses are predicated (13-15).

Look at this [online version](#) of the HAM-D to see what I mean. Qs 4, 9,10,11 & 12 might all be improved by any anxiolytic/sedative — a one gradation change in each of those produces a 5-point improvement of your score, more than double that needed to get a drug approved by the FDA as an AD.

The question of inter-rater reliability is not even mentioned any more (15).

Cipriani et al. specifically state ‘we were not able to quantify some outcomes, such as global functioning’.

I do not intend to sound scathing, but not assessing anergia or functioning more fully is like not asking a patient with angina how many stairs they can manage without stopping because of pain. It is hard politely to describe one’s astonishment at this basic error and omission.

And, the very use of the word antidepressant is a misrepresentation. You would not call a drug an antibiotic if it only slightly slowed down the growth of bacteria, without killing them. To label drugs as antidepressants when many of them merely produce a small change in symptoms, which are not even central to what we presume is the core of the illness, is an assumptive misconception and misrepresentation. At this point we should remind ourselves that the epithet ‘antidepressant’ was generally assigned at the behest of the sales and marketing division — it bears little relationship to pharmacology or effectiveness. [I commented about mirtazapine](#) in this regard — the first paper about this drug in humans, using it as a preoperatively a sedative, was entitled ‘A double-blind group comparative study using the **new anti-depressant** Org 3770 ...’ (16). Needless to say, at that time there was absolutely no evidence that it was an effective antidepressant, but the principle of primacy, being so powerful, caused the idea and the label to stick — if that was a purposeful manoeuvre it constituted brilliant marketing, or was it just intellectual laziness.

For those who favour using Bayesian reasoning to advance hypotheses, there is an obvious, unaddressed, pharmacological contradiction that demands adequate explanation. The frequently unjustified epithet of ‘antidepressant’ glosses over the fact that various of these drugs work, or do not work, in different ways, and have different effects on neurotransmission.

Therefore, to suppose, *prima facie*, that ‘antidepressants’ working via different mechanisms and neurotransmitters are equally effective is improbable, implausible, and without justification.

Address for correspondence. Dr Ken Gillman: kg@matilda.net.au

Therefore, a methodology that appears to show that such is the case is inevitably suspect. Note also, various of these drugs are hardly used in clinical practice — e.g. moclobemide (despite having the most benign SE profile), trazodone, reboxetine, showing clinicians have given up on them. The widespread perception of these drugs effectiveness amongst practising psychiatrists is substantially discrepant with the supposed scientific evidence from RCTs. Since these two kinds of observations do not possess a qualitative epistemological superiority one over the other, this indicates that something is seriously wrong. I certainly gave up on those three drugs pretty quickly, I regard moclobemide (and mirtazapine) as an ‘active placebo’.

Incidentally, what happened to moclobemide? It is absent from the Cipriani MS: I asked the authors why this was but received no reply.

Questions a Bayesian would wish to address are various, prominent among them being: why is there an [apparent] significant difference between different SSRIs when they are all equally effective as SRIs? That applies especially to citalopram and escitalopram (its isomer), and desvenlafaxine & venlafaxine; bupropion (a pro-drug) has hydroxy-bupropion as the dominant active molecule. Yet when hydroxy-bupropion was tested, it failed, and was successfully buried (it was called Radafaxine, but you will not have even heard of it). Why is the anti-histamine, mirtazapine (aka 6-azamianserin), [apparently] so good, but not its pharmacologically identical twin-brother, mianserin (4)? The mirtazapine vs. mianserin mystery is yet another forgotten issue concerning [deceitful data](#). What about the huge sibutramine AD trial? Only ever reported as a (very brief) conference abstract. Sibutramine was the first ‘dual-action’ AD drug (again ‘buried’). Then there is the initial tranche of trials on duloxetine which all failed, leading to it being ‘shelved’ (it is hard to see from the data presented if the authors ‘found’ those older studies), but it was resurrected later.

What most doctors and psycho-pharmacologists do not seem to understand is that the list of drugs currently regarded as antidepressants are in fact a somewhat random selection from a vast array of drugs which have been used or tested over the last few decades. There are simply no reliable or meaningful criteria for defining which of them are really antidepressants.

These anomalies require explanations and cast a shadow over the validity of MA results. A parsimonious initial supposition must be that some/many of them do not work as true ADs at all [except via non-depression-specific mechanisms, such as sedation/anxiolysis].

And, to discuss extrapolating the putative optimal AD agent[s] for the general depressed population, based on the results of such meta-analyses, as it appears this paper, and the discussion around it, suppose to do, despite the caveats given (17), is ‘heroic optimism’ and seems to ignore other methodologies, and decades of clinical experience, and thereby to assume or insinuate that they are of lesser value.

Evidently, an emphatic reminder that there is no epistemological justification for assigning qualitatively different validity to those two domains of evidence (clinical observation vs. RCTs) is required.

Indeed, an exaggerated view of the epistemic virtues of RCTs currently dominates thinking (18, 19) — Sir Austin Bradford Hill himself made a point of endorsing Claude Bernard's view that there is ‘no qualitative epistemic difference between experiment [RCTs] and observation’ [clinical experience].

Hill also opined: ‘you need neither randomisation, nor statistics, to analyse the results, unless the treatment effect is very small (20)’.

Address for correspondence. Dr Ken Gillman: kg@matilda.net.au

And, for most patients, it is small.

And, the issue of mis-coding data (and hiding the ‘raw’ data — something Healy, Gotzsche, and others have written about (21, 22): and, fraud and deceit, are not even mentioned or addressed (see also my commentary about the forgotten story of [deceitful data presented on mirtazapine](#)). There is no doubt that doctors are naive about these matters and tend to avoid calling-out scientific misconduct or fraud, even if they do recognise it.

The reality that fraud and deceit issues do seriously affect all meta-analyses receives insufficient consideration, for many different reasons.

Indeed, I wrote to one of the authors (Leucht) of this study, about his previous meta-analysis of antipsychotic drugs (23) which failed to mention the seminal paper by Houston et al (24) detailing the blatant deceit involved in the work published about risperidone. Leucht was specifically discussing the question of bias (and cf. mirtazapine) and stated in reply to me they ‘did not know about Huston’.

It is concerning to predict that it is inevitable that prestigious publications such as this will tend to dictate the kinds of treatment than ordinary doctors will use. This [excessive reliance on guidelines] has already produced a blinkered and stultified approach to antidepressant treatment, which virtually excludes, *inter alia*, the use of drugs like tranylcypromine and clomipramine. I know few experienced psycho-pharmacologists who do not agree that clomipramine is superior to amitriptyline.

In my career as a psycho-pharmacologist I have seen hundreds of patients who have alternated between amitriptyline and clomipramine. I cannot remember many who responded better on amitriptyline, but I would estimate that 19/20 would have stated unequivocally that they were better on clomipramine, having had only a partial response to amitriptyline (the difference being sufficiently great that any increased burden of side-effects was usually accepted with little complaint). That methodology, A-B-A trial, represents a powerful methodology for comparing drugs, especially when subjects are already known to suffer from a ‘biological’ depression; by virtue of previous response to an established antidepressant (or ECT): they represent a ‘better’ and ‘purer’ (more homogeneous) sample.

And, that brings us to another glossed-over but vital question. It is inevitable that the samples in these trials, frequently done in outpatients, are comprised of significant proportion of people who are never going to respond to ADs because they do not have a ‘biological’ depression. It is a frequently forgotten fact that randomised controlled trials rest, for their methodological validity, on the presupposition that the sample in question is homogenous. When that is palpably not the case the power of the RCT is much reduced, some would say invalidated, even more so when the treatment effect is small (cf. Hill).

And, these trials of 8 weeks duration are assessed with only subjective interim surrogate outcome measures (rating-scales).

Interim surrogate outcome measures have various major limitations, especially for predicting long-term treatment of chronic illnesses (cf. lithium, not a potent AD, but reduces suicide more than ‘ADs’ (25)).

Furthermore, concerning interim surrogate outcome measures, cf. the story of anti-hypertensives drugs, where early reliance on short-term BP reduction failed to show various nuances of outcome, such as the superior benefit of ACE-inhibitors on preserving kidney function (26, 27).

Address for correspondence. Dr Ken Gillman: kg@matilda.net.au

And, tranylcypromine is not even mentioned (nor moclobemide, which some might consider to be an ‘active placebo’). There are of course few trials in the modern era that used it. Why is this? At least partly because the RCTs & meta-analyses that feed guidelines do not mention it, so nobody uses it, and nobody does trials with it. The classic example of my point above about how guidelines foster a circular, self-fulfilling, restricted and blinkered approach to the subject.

As the Australian professor, Gordon Parker has said, there are major limitations to ‘level 1’ evidence derived from RCTs, which are no longer producing meaningful clinical results’ (28).

No longer? Did they ever?

Remember, the recognition of the undoubted powerful antidepressant effect of amitriptyline, clomipramine, and tranylcypromine owes nothing to double-blind trials. Nor did the discovery of penicillin, or the effectiveness of a host of other general medical drugs. Remember what Sir Austin Bradford Hill said about statistics being unnecessary if the effect was obvious.

My take?

Meta-analysis amplifies the limitations discussed above into serious problems and mis-directions, made worse still because M-A gives results a false imprimatur of authority and because prevailing ethos in the profession strongly discourages ‘non-guideline’ treatment. I have no doubt that doctors feel pressured to follow the guidelines, and liable to threat if they do not. Bureaucrats and managers who know next to nothing about medicine use guidelines to further their own (financially focussed) agendas.

The initial reaction to this paper from ‘the profession’ has been laudatory and uncritical. The quality and trustworthiness of the data, and the techniques and methodology, simply do not bear the degree of interpretation and extrapolation that they are being subjected to.

Attempting to make fine distinctions between drugs, when none of them are much better than placebo, is unrealistic and unscientific; this is even more so because ‘depression’ is not a homogeneous entity.

The very fact that such intricate manoeuvres are required to ‘show’ these minor benefits is itself clear proof that their effect is minimal — it really is that simple.

If they had a good effect and produced remission (wellness) in a majority of patients, we would not be occupying our time with this fruitless nit-picking and bickering.

Worse still, these ‘results’ (which are actually neither new nor definitive, as Gotzsche and others have also pointed out) will be accorded a degree of reliability and authority they in no way deserve. The lazy and unthinking will justify their practice by stating they ‘followed the guidelines’. The policy-makers and providers will decline to pay for, or make available, drugs like nortriptyline, clomipramine, tranylcypromine etc. It is already happening. Good doctors will cease clinical practice because their practice is unreasonably constrained and they are harassed and bullied by non-medical ‘managers’

The promulgation of guidelines has reached the point where it is producing major negative consequences and patients are suffering as a result.

Address for correspondence. Dr Ken Gillman: kg@matilda.net.au

One has to wonder if the authors of this study will remain comfortable with such consequences. They have assumed an onerous burden of 'prophecy' and history may not judge them kindly.

My parting suggestion is that some researchers get organised and produce some decent research aimed at answering useful question and expend less time endlessly comparing one 'me-too' drug with another, to little avail. A trial demonstrating that tranlycypromine treats psychotic depression, even when ECT has failed would be a start. The irony there is, of course, that such a trial is unnecessary because the effect is so decisive that randomisation, blinding, assessment with rating scales, statistical analysis etc. are all unnecessary. Remember Sir Austin Bradford Hill?

Science does not change many people's minds, and so I offer this poignant story about a [professor of psychiatry with psychotic depression was made-well, cured, restored to full functioning, by tranlycypromine.](#)

Addendum: key points from Cipriani paper

The Cipriani analysis was based on 522 double-blind studies (116 477 patients) involving 21 antidepressants (maximum 8 weeks duration).

The 21 antidepressants

agomelatine, amitriptyline, bupropion, citalopram, clomipramine, desvenlafaxine, duloxetine, escitalopram, fluoxetine, fluvoxamine, levomilnacipran, milnacipran, mirtazapine, nefazodone, paroxetine, reboxetine, sertraline, trazodone, venlafaxine, vilazodone, and vortioxetine, [but not, oddly, moclobemide, imipramine, or nortriptyline].

NB remember, there are many other ADs that might have been considered, but were not: e.g. sibutramine and hydroxy-bupropion.

Verbatim extracts

'Our assessment overall found few differences between antidepressants when all data were considered'.

All antidepressants were [a tiny bit] more effective than placebo, with ORs ranging between 2.13 (95% credible interval [CrI] 1.89-2.41) for amitriptyline and 1.37 (1.16-1.63) for reboxetine.'

'In our analyses, funding by industry was not associated with substantial differences in terms of response or dropout rates. **However, non-industry funded trials were few and many trials did not report or disclose any funding.**'

'We did not cover important clinical issues that might inform treatment decision making in routine clinical practice (eg, specific adverse events, withdrawal symptoms, or combination with non-pharmacological treatments). Additionally, because of the paucity of information reported in the original studies, **we were not able to quantify some outcomes, such as global functioning.** It should also be noted that some of the adverse effects of antidepressants occur over a prolonged period, meaning that positive results need to be taken with great caution, because the trials in this network meta-analysis were of short duration.'

'Given the modest effect sizes, non-response to antidepressants will occur. Our information unfortunately cannot guide next-step choices after failure of such a first step (ie, they do not apply to treatment resistant depression), for which well performed trials are scarce.'

'Notwithstanding these limitations, the findings from this network meta-analysis represent the most comprehensive currently available evidence base to guide the initial choice about pharmacological treatment for acute major depressive disorder in adults. All statements comparing the merits of one antidepressant with another must be tempered by the potential limitations of the methodology,³² the complexity of specific patient populations, and the uncertainties that might result from choice of dose or treatment setting. We

Address for correspondence. Dr Ken Gillman: kg@matilda.net.au

hope that these results will assist in shared decision making between patients, carers, and their clinicians.'

'In total, 87 052 participants were randomly assigned to an active drug and 29 425 were randomly assigned to placebo. The mean age was 44 years (SD 9) for both men and women; 38 404 (62.3%) of 61 681 of the sample population were women. The median duration of the acute treatment was 8 weeks (IQR 6–8). 243 (47%) of 522 studies randomly assigned participants to three or more groups, and 304 (58%) of 522 were placebo controlled trials. 391 (83%) of 472 were multi-centre studies and 335 (77%) of 437 studies recruited outpatients only. 252 (48%) of 522 trials recruited patients from North America, 37 (7%) from Asia, and 140 (27%) from Europe (59 [11%] trials were cross-continental and the remaining 34 [7%] were either from other regions or did not specify). The great majority of patients had moderate-to-severe major depressive disorder, with a mean reported baseline severity score on the Hamilton Depression Rating Scale 17-item of 25.7 (SD 3.97) among 464 (89%) of 522 studies.'

'The current report summarises evidence of differences between antidepressants **when prescribed as an initial treatment**. Given the modest effect sizes, non-response to antidepressants will occur. **Our information unfortunately cannot guide next-step choices** after failure of such a first step (ie, they do not apply to treatment-resistant depression), ...'

I confidently predict that the cautions and caveats in the section I have placed in bold above, which are correct and appropriate, will be lost in the implementation. This is what frequently happens with diagnostic guidelines like the DSM, and with treatment guidelines: thus, they become diktats, rather than advice to be implemented with caution, judgement, and consideration of individual circumstances.

Formula 1 Anti-Depressant starting grids: 2009 vs 2018

[Or the 'Cipriani stakes' for fillies?]

NB This is rather esoteric humour

Front row of the grid

2009 mirtazapine, escitalopram, sertraline, venlafaxine

2018 mirtazapine, escitalopram, sertraline, paroxetine, agomelatine

So, my *bête noire*, venlafaxine, has suffered 'dual-traction' issues in the rear differential and slipped off the front row of the grid! Paroxetine makes a comeback after several poor seasons (although poor hydraulic pressure in the central ram remains a concern over the full race distance). Several teams have appeals for using illegal slicks in qualifying before the stewards. Team cipramil mysteriously failed to qualify (there are rumours about sandbagging and patents). Team Moclobemide appear to have lost all fans and all sponsors after the revelation that the engine supplier grossly exaggerated the power output. Team Clomipramine just never got noticed in F1 circles, after starting on the wrong test-circuit, where they just went round-and-round for days and couldn't stop themselves, despite scrubbing the tyres down to the canvass. When they did turn up they had no fan base, and were overlooked, despite having the most powerful engine. Team 'Tranyl-Tyrrell', with their 6-wheel car, were disqualified on technical grounds long ago for needing special fuel, and using too high turbo-boost pressures -- but they are still consistently winning on the alternative 'non-F1' circuit, due, some say, to their inherently superior design and performance.

References

Address for correspondence. Dr Ken Gillman: kg@matilda.net.au

1. Cipriani, A, Furukawa, TA, and Salanti, G, Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet*, **2018**. 391(10128): p. 1357-1366
[http://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(17\)32802-7/fulltext](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)32802-7/fulltext).
2. Ioannidis, JP, The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. *Milbank Q.*, **2016**. 94(3): p. 485-514.
<https://www.ncbi.nlm.nih.gov/pubmed/27620683>
3. Ioannidis, JP, Ranking antidepressants. *Lancet*, **2009**. 373(9677): p. 1759-60; author reply 1761-2.
<https://www.ncbi.nlm.nih.gov/pubmed/19465221>
4. Gillman, PK, A systematic review of the serotonergic effects of mirtazapine: implications for its dual action status. *Hum Psychopharmacol*, **2006**. 21(2): p. 117-25.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16342227
5. Kirsch, I and Moore, TJ, The Emperor's New Drugs: An Analysis of Antidepressant Medication Data Submitted to the U.S. Food and Drug Administration. *Prevention & Treatment*, **2002**. 5.
<http://journals.apa.org/prevention/volume5/pre0050033r.html>
6. Moncrieff, J, What does the latest meta-analysis really tell us about antidepressants? *Epidemiology and Psychiatric Sciences*, **2018**(28 May): p. <https://www.cambridge.org/core/journals/epidemiology-and-psychiatric-sciences/article/what-does-the-latest-metaanalysis-really-tell-us-about-antidepressants/90020F9E608E60DE0B6AFF2932F9A6B9>.
7. Moncrieff, J, Antidepressants are not antidepressants – an alternative approach to drug action and implications for the use of antidepressants.
BJPsych Bull, **2018**. 42: p. 42-4.
8. Jauhar, S and Young, AH, Ideology over evidence? *BJPsych Bull*, **2018**. 42: p. 42-4.
9. Parker, G, Defining melancholia: the primacy of psychomotor disturbance. *Acta Psychiatr Scand Suppl*, **2007**(433): p. 21-30.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17280567
10. Parker, G and McCraw, S, The properties and utility of the CORE measure of melancholia. *J Affect Disord*, **2017**. 207: p. 128-135.

Address for correspondence. Dr Ken Gillman: kg@matilda.net.au

<https://www.ncbi.nlm.nih.gov/pubmed/27721186>

11. Parker, G, Roy, K, Hadzi-Pavlovic, D, Mitchell, P, et al., Subtyping depression by clinical features: the Australasian database. *Acta Psychiatr. Scand.*, **2000**. 101(1): p. 21–8.

12. Soria, V, Vives, M, Martinez-Amoros, E, Galvez, V, et al., The CORE system for sub-typing melancholic depression: Adaptation and psychometric properties of the Spanish version. *Psychiatry Res.*, **2016**. 239: p. 179–83.

<https://www.ncbi.nlm.nih.gov/pubmed/27010187>

13. Trajković, G, Starčević, V, Latas, M, Leštarević, M, et al., Reliability of the Hamilton Rating Scale for Depression: A meta-analysis over a period of 49 years. *Psychiatry Res.*, **2011**. 189: p. 1–9.

14. Tabuse, H, Kalali, A, Azuma, H, Ozaki, N, et al., The new GRID Hamilton Rating Scale for Depression demonstrates excellent inter-rater reliability for inexperienced and experienced raters before and after training. *Psychiatry Res.*, **2007**. 153(1): p. 61–7.

<https://www.ncbi.nlm.nih.gov/pubmed/17445908>

15. Bagby, RM, Ryder, AG, Schuller, DR, and Marshall, MB, The Hamilton Depression Rating Scale: has the gold standard become a lead weight? *Am J Psychiatry*, **2004**. 161(12): p. 2163–77.

<https://www.ncbi.nlm.nih.gov/pubmed/15569884>

16. Sorensen, M, Jorgensen, J, Viby-Mogensen, J, Bettum, V, et al., A double-blind group comparative study using the new anti-depressant Org 3770, placebo and diazepam in patients with expected insomnia and anxiety before elective gynaecological surgery. *Acta Psychiatr. Scand.*, **1985**. 71(4): p. 339–46.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=4003099

17. Parikh, S and Kennedy, S, More data, more answers: picking the optimal antidepressant. *Lancet*, **2018**: p.

[http://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(18\)30421-5/fulltext](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(18)30421-5/fulltext).

18. Healy, D, Trussed in evidence? Ambiguities at the interface between clinical evidence and clinical practice. *Transcultural psychiatry*, **2009**. 46(1): p. 16–37.

19. Worrall, J, Causality in medicine: getting back to the Hill top. *Prev. Med.*, **2011**. 53(4–5): p. 235–8.

<https://www.ncbi.nlm.nih.gov/pubmed/21888926>

Address for correspondence. Dr Ken Gillman: kg@matilda.net.au

20. Hill, AB, The Environment and Disease: Association or Causation? *Proc. R. Soc. Med.*, **1965**. 58: p. 295–300.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14283879
21. Le Noury, J, Nardo, JM, Healy, D, Jureidini, J, et al., Restoring Study 329: efficacy and harms of paroxetine and imipramine in treatment of major depression in adolescence. *BMJ*, **2015**. 351: p. h4320.
<http://www.bmj.com/content/351/bmj.h4320.full>
22. Schroll, JB, Penninga, EI, and Gotzsche, PC, Assessment of Adverse Events in Protocols, Clinical Study Reports, and Published Papers of Trials of Orlistat: A Document Analysis. *PLoS Med*, **2016**. 13(8): p. e1002101.
<https://www.ncbi.nlm.nih.gov/pubmed/27529343>
23. Leucht, S, Tardy, M, Komossa, K, Heres, S, et al., Antipsychotic drugs versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis. *Lancet*, **2012**. 379: p. 2063 – 2071.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=22560607
24. Huston, P and Moher, D, Redundancy, disaggregation, and the integrity of medical research. *Lancet*, **1996**. 347(9007): p. 1024–6.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8606568
25. Young, AH, Review: lithium reduces the risk of suicide compared with placebo in people with depression and bipolar disorder. *Evid Based Ment Health*, **2013**. 16(4): p. 112.
<http://www.ncbi.nlm.nih.gov/pubmed/24046391>
26. Williams, B, Recent hypertension trials: implications and controversies. *J. Am. Coll. Cardiol.*, **2005**. 45(6): p. 813–27.
<https://www.ncbi.nlm.nih.gov/pubmed/15766813>
27. Wu, HY, Huang, JW, Lin, HJ, Liao, WC, et al., Comparative effectiveness of renin-angiotensin system blockers and other antihypertensive drugs in patients with diabetes: systematic review and bayesian network meta-analysis. *BMJ*, **2013**. 347: p. f6008.
<https://www.ncbi.nlm.nih.gov/pubmed/24157497>
28. Parker, G, Evaluating treatments for the mood disorders: time for the evidence to get real. *Aust NZ J Psychiatry*, **2004**. 38(6): p. 408–14.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15209831

Address for correspondence. Dr Ken Gillman: kg@matilda.net.au