# Statistical Spin: Linguistic Obfuscation—The Art of Overselling the CBT Evidence Base

Farhad Dalal

## Abstract

This paper takes a close look at two interlinked studies on mindfulness based cognitive therapy (MBCT), Teasdale and colleagues, 2000 and Ma and Teasdale, 2004. The second study corroborates the findings of the first study to claim that MBCT is a cost-effective treatment that prevents the recurrence of depression in 50% of the population. A close reading of the statistical and linguistic manipulations reveals that the outcomes are closer to 25%, and then too for a very limited population. In sum, the paper argues that the evidence for MBCT is being vastly oversold, as is the evidence for the effectiveness for cognitive behavioural therapy (CBT) more generally.

*Key words*: mindfulness, mindfulness based cognitive therapy (MBCT), cognitive behavioural therapy (CBT), evidence-based therapy, statistical evidence, critique

## Introduction

Many within academia and the helping professions have come to believe that the scientific credentials of cognitive behavioural therapy (CBT) as a treatment have been empirically established beyond reasonable doubt. The same is true of mindfulness based cognitive therapy (MBCT) as a treatment for depression. Not only is MBCT regularly prescribed by GPs for their depressed patients, the practice of mindfulness itself has

become popularised and promoted as a panacea for the ills of the modern condition.

The website dedicated to MBCT tells us that

> The UK *National Institute for Clinical Excellence* (NICE) has recently endorsed MBCT as an effective treatment for prevention of relapse. Research has shown that people who have been clinically depressed 3 or more times (sometimes for twenty years or more) find that taking the program and learning these skills helps to reduce considerably their chances that depression will return. The evidence from two randomized clinical trials of MBCT indicates that it reduces rates of relapse by 50% among patients who suffer from recurrent depression. (MBCT)

The evidence for this belief are two trials; the first being "Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy" (Teasdale et al, 2000), and the second "Mindfulness-based cognitive therapy for depression: replication and exploration of differential relapse prevention effects" (Ma & Teasdale, 2004).

The fact that the second study reproduced the findings of the first study generated much excitement within CBT communities as this allowed them to claim that they had reached the holy grail of empirical science. This being that because their findings had been shown to be repeatable, they constituted scientifically validated objective facts. In the literature, both studies are spoken of and thought to be exemplars of good rigorous empirical scientific research.

In this paper I will subject both these papers to a close reading and raise questions about the scientific basis of the claims being made in them. The form of the critique might be described as a discursive analysis, or deconstruction. And because the research protocols followed in these two studies are conventional norms within CBT research and academic psychology, the critique also raises deeper and more serious questions about research practices in general within this field. I need to emphasise that what follows is in no sense a critique of mindfulness per se, but it is a critique of the way that "data" is being generated, used, and promulgated within the cognitivist paradigm.

MBCT is said to be a part of the third wave of CBT treatments. The first was behaviourism itself. The second wave began with

Farhad Dalal

the bolting of cognitivism on to behaviourism. The ethos of the second wave (cognitive behaviourism) is primarily that of control, thought control. As Layard has put it, "[The purpose of CBT is to train us to] directly address our bad feelings and *replace* them by positive feelings" (Layard, 2005, p. 188; italics added).

In contrast, the third wave is much "softer"; it draws on Eastern philosophies, and so in these treatment protocols we find an abundance of terms like attachment, compassion, mindfulness, and acceptance. The ethos here is that of acceptance. Given that the ethos of one sharply contradicts that of the other, control *vs*. acceptance, the first question to be asked is: what characteristics do the "waves" share that allows them all to be called forms of CBT?

It seems to me that their shared characteristics have little to do with content (as the forms of CBT are so divergent), and more to do with the ways in which the "treatment" is legitimated and promulgated. To count as CBT, the treatments have to be "tested" and found to be efficacious by the conventions of positivist empirical science. Next, the treatments themselves should be manualised so that all practitioners closely follow in the footsteps of the study itself. Clearly there are critical questions to be asked of this way of proceeding, but this is not going to be my focus here. Instead, I am going to take issue with them on the grounds of the very paradigm that they situate themselves—positivist empirical science—and show that in Ben Goldacre's pithy phrase, that they are examples of *Bad Science* (Goldacre, 2012).

But before I get into the science itself, I want to say something about the relationship of the ethos of the third wave to the Eastern philosophies that it makes use of. CBT strips out certain meditative practices from the philosophies and meaning worlds that have generated them, and reduces them into sets of techniques and skills to be learnt. In proceeding in this way, not only have these practices have been instrumentalised, but further, the end to which they have been instrumentalised constitutes a perversion of those very philosophies that generated them. For example, in the very same breath as mentioning the Buddha, Layard, and Clark celebrate the fact that, "The Resilience Programme [which utilises mindfulness] is now being used for every soldier

in the US Army, with the aim of reducing the incidence of post-traumatic stress disorder after traumatic experiences on the battlefield" (Layard & Clark, 2014, p. 230). In other words, the Buddhist practice of mindfulness is being used to turn human beings into more efficient, more resilient, killing machines. Even so, despite the perversion, if CBT did in fact achieve what it claims to achieve, to the extent to which it claims to achieve them, then in one sense that would be all right. Whatever works. But does it work? And if it does work, then to what degree?

## Research culture

The activity of science is imbued by politics as is all human activity and so is never value free. The directions taken are very often the result of struggles and rivalries in the power-relational field between and within professions, university departments, individuals, and so forth rather than in the disinterest pursuit of truth. In other words the choices made and directions advocated are never as objective and innocent as they are made out to be.

With this in mind, the first interesting anomaly to think about is the consistency with which CBT treatments are researched and promoted as *adjuncts* to psychiatric medication, not as *alternatives* to it. At the very least this is curious given that CBT positions itself as the equivalent of a drug and uses language to that effect, speaking of "minimal effective doses", "inoculation", and so forth. It seems to me that one reason that CBT has not directly contested the efficacy of psychiatric medication is in order not to provoke the ire of the powerful pharmaceutical industry.

Instead of challenging conventional psychiatric discourse, CBT has made an alliance with it. In proceeding in this way, CBT has both stepped into, as well as helped create and sustain, a very particular kind of scientific research culture in which a number of assumptions have become unquestionable norms.

First in line is the *DSM* and its psychiatric premise that a) *there exist* discreet objective "mental illnesses" that have a chemical or organic basis, and further b) there are magic-bullet drug treatments that can target the specific organic cause—some form of chemical imbalance—and fix it (Whitaker, 2010).

Over the last few years it has increasingly become apparent that the research culture within "evidence-based medicine" has

largely been driven in directions that have served the fiscal interests of the pharmaceutical industry with the connivance of their champions bought and paid for within academia, "Ninety per cent of published clinical trials are sponsored by the pharmaceutical industry, [which means that they] . . . dominate this field, they set the tone, and *they create the norms*" (Goldacre, 2012, italics added). One norm being that of publication bias. Studies that fail to demonstrate the efficacy of a treatment being tested, simply do not see the light of day. This fact has been known about for a long time.

> [This was] first formally documented by a psychologist called Theodore Sterling in 1959. He went through every paper published in the four big psychology journals of the time, and found that 286 out of 294 reported a statistically significant result. This, he explained, was plainly fishy: it couldn't possibly be a fair representation of every study that had been conducted, because if we believed that, we'd have to believe that almost every theory ever tested by a psychologist in an experiment had turned out to be correct . . . In 1995, at the end of his career, the same researcher came back to the same question . . . and found that almost nothing had changed [Sterling, 1959, Sterling et al., 1995]. (Goldacre, 2012)

How has this situation come about?

First, standards are set abysmally low. The regulatory authorities have agreed that all it need take is for *two* clinical trials to show a positive outcome, for them to grant the treatment a licence. And all that is meant by "positive outcome" is that the treatment is shown to be better than placebo. This is the case for MBCT. The critical catch is this, "There is no limit to the number of trials that can be conducted in search of these two significant trials" (Kirsch, 2011, p. 195). How important is this, and why does it matter? One way of answering this is by recalling how the other primary treatment for depression—antidepressants— came to be accepted as the treatment of choice.

A group of researchers (Turner et al., 2008, cited in Goldacre, 2012, p. 427) tracked all the clinical trials registered with the US Food and Drug Administration (FDA) for all the antidepressants that were launched between 1987 and 2004. Thirty-eight of the trials concluded that the treatment being tested worked, and thirty-six found that the treatment did not. Thirty-seven of the

thirty eight trials with positive results were published, while only three out of the thirty-six negative trials were. Of the remaining thirty-three trials with negative outcomes, twenty-two never saw light of day again, and astonishingly, the remaining eleven "were written up as if the drug were a success". Consequently, when doctors looked to the published "research" to help make an informed decision, they found forty-eight trials apparently demonstrating the efficacy of antidepressants, set against three that did not. Any sensible person would conclude on this basis that the evidence for the efficacy of antidepressants is overwhelmingly conclusive.

In sum, over the last few decades, as the social scientist John Abrahams has concluded, there has been "a consistent trend towards deregulation, for the benefit of the [pharmaceutical] industry" (Goldacre, 2012). This then is the research culture in which the science of CBT is cultivated, a culture of low standards and at times outright deceit. And as is bound to be the case, the "science" of CBT is contaminated by the culture in which it is grown and cultivated".

## Research practice: how a statistical truth comes to be camouflaged as an absolute truth

It would be helpful to review what a clinical trial looks like. Say you want to test the efficacy of a drug called Zon for disease X. You start by finding (for the sake of argument) a hundred people suffering from X. However, you discover that many of them are also suffering from Y or/and Z, and so you remove them from the study. Similarly, many more are removed from the study for a range of other reasons. Eventually, you are left with say, twenty. They are now randomly allocated to one of two groups of ten. Members of one group are given a placebo, and the other the drug Zon. Double-blind protocols are utilised. Six years later, you look at the progress of the disease in each individual. You discover that seven out of the ten who were on Zon are in remission from disease X (70%); while in the control group, only two are still well (20%). A result! So the claim is made: The drug Zon is now scientifically proven to cure X! But actually the claim should be that the drug Zon works more often than not—it works some of the time with some people. In other

words, it is *guaranteed* that it will *not work* on a certain percentage of the population.

What happens next is that it is conveniently "forgotten" that the drug has been tested on a particular skewed population, on just twenty per cent of those known to be suffering from X, and that it has been helpful to about three quarters of them. When Zon is prescribed to members of the 80% not included in the study, then it is a form of off-label prescribing.

The mythical story of Zon is not all that mythical.

One study found that just 6% of representative asthma sufferers would have been eligible to participate in asthma treatment trials (Travers et al., 2007). And quite astonishingly, out of every eight people suffering from depression who had *volunteered* to take part in treatment trials, only one was deemed eligible to participate (Keitner & Posternak, 2003).

Further, the language used in promotional literature and journal abstracts are mostly without any caveats suggesting that Zon is the foolproof treatment of choice for *all* those suffering from X—the entire hundred—rather than as something that is likely to increase your chances of recovery if you happened to belong to a population akin to that of the 20% that the treatment was tested on.

The "evidence-base" for CBT is generated in exactly this manner. Here is a recent example. The press release for a CBT trial (Wiles et al., 2013) says:

> The CoBalT team, comprising researchers from the Universities of Bristol, Exeter and Glasgow, recruited 469 patients . . . with treatment-resistant depression . . .

> At six months, 46 per cent of those who received CBT in addition to usual care had improved, reporting at least a 50 per cent reduction in symptoms of depression, compared to 22 per cent of those who continued with usual care alone. This beneficial effect was maintained over 12 months.

> The findings demonstrate that CBT provided in addition to usual care including antidepressant medication is an effective treatment that reduces depressive symptoms, and improves the quality of life in patients whose depression has not responded to the most common first-line treatment for depression in primary care. (University of Bristol, 2012)

Although this CBT treatment did not fare anywhere near as well as that of our mythical Zon, it is nevertheless being claimed that the research has scientifically demonstrated that it is an "effective treatment". Let us look a little more closely at the figures.

The fact that 46% said that they felt better—means *more than half did not*! In their own terms then, over half the population is "CBT resistant".

However, the important figure is not the 46% that they highlight, but 24%—a figure never mentioned. This being the number that remains once we remove the 22% who had improved without the help of CBT within the control group. We have to assume that a comparable number in the CBT group might also have similarly improved—*that after all is the point of a control group*. In other words at the end of the study, just 24% *more* of the people in the CBT group improved when compared to the numbers improved in the control group. In other words, the research showed that *about two out of ten more people* came to feel better because of CBT. And this improvement has taken place in combination with the treatment as usual, presumably, drug therapy. Or to put it the other way, the treatment will be *ineffective for eight out of ten patients*.

I am sure that some people have been helped and are being helped, and this surely is a good thing. But it is not an overwhelmingly good thing, at least not to the degree to the inflated claim that is trumpeted in the rhetoric.

## The objectification of subjectivity

CBT prides itself in producing objective knowledge. How is it derived? For example, notice, even the "successful" 24% were by no means *cured* of depression. What the subjects reported was that their symptoms of depression *were reduced* by about 50%.

But what does it mean to say that "the symptoms of depression are reduced by 50%"? While the number of people at a bus stop is countable, tangible, and therefore objective, this is not the case for the 50%. The 50% is arrived at by asking people to state on a scale of one to five how depressed they are, and other questions of that ilk. The fact that the answers are numbers, and the mix of numbers are amalgamated into other numbers,

renders the impression that what is being spoken of is objective. But this is not the case. The answers to these sorts of questions about subjective experiences, although couched in numbers, remains fundamentally subjective; it is in this way that the illusion of objectivity comes to be manufactured.

So the "results" of the trial actually *ought* to be announced in this way:

> About two out of ten *more* people came to feel somewhat better because of having received CBT; however, although better, they are still depressed, only less depressed. Meanwhile, eight out of ten people were not helped by the treatment.

The culmination of this kind of exaggeration and distortion, repeatedly published in prestigious scientific journals has meant it is now become established scientific "fact" that CBT is unquestionably an *entirely* effective therapy. For example, the statement on the website of the prestigious (scientific) Royal College of Psychiatrists, is without any cautions or caveats. It says:

> CBT has been shown to help with many different types of problems. These include: anxiety, depression, panic, phobias (including agoraphobia and social phobia), stress, bulimia, obsessive compulsive disorder, post-traumatic stress disorder, bipolar disorder and psychosis. CBT may also help if you have difficulties with anger, a low opinion of yourself or physical health problems, like pain or fatigue. (Royal College of Psychiatrists, 2013)

## MBCT

### *Prevention and reduction*

I will now attend more directly to the two studies on MBCT, and for the purposes of this paper, I will remain primarily with the first. (The results of the second trial are marginally better than those of the first, but are broadly similar).

The first study is entitled "Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy". The very first word in the title is compelling: "Prevention"; it suggests that the treatment will *halt* the recurrence of major depression. This is exactly what the abstract proudly claims has

been achieved, "MBCT offers a promising cost-efficient psycho-logical approach to *preventing* relapse/recurrence in recovered recurrently depressed patients" (Teasdale et al., 2000, p. 615, my italics). But when we dip into the paper itself, we find that their research question is in fact much more modest in its aspirations, as it speaks not of prevention, but *reduction*, "Does this inter-vention, when offered in addition to TAU, *reduce rates of relapse* and recurrence compared to TAU alone?" (Teasdale et al., 2000, p. 617, my italics).

In the body of their paper the authors freely bounce from "prevent" to "reduce" and back again as it suits them, "The finding that MBCT *prevented* relapse . . ." (p. 622), "can signifi-cantly *reduce risk* of future relapse/recurrence . . ." (p. 623), "MBCT offers a promising cost-efficient psychological approach to *preventing* relapse/recurrence . . ." (p. 615), "*reduce* future risk of relapse and recurrence of depression . . ." (p. 618), and so on (all italics mine).

Well, which is it and does it matter? Is the intention of the research to prevent (a very strong and compelling claim), or is it merely to reduce (worthy, but not so sexy). This confusion allows these researchers to have their cake and eat it too, which is illustrated in the very first sentence of the paper, "Relapse and recurrence following successful treatment of major depressive disorder (MDD) *is common* and often carries massive social cost" (Teasdale et al, 2000, p. 615, my italics).

Consider, if relapse and recurrence "is common", then in what sense is it being claimed that the prior treatment (what-ever it was) was successful? If relapse is indeed *common* then surely what this demonstrates is that the treatment is not all that successful. I am reminded of the old joke, "operation successful, patient dead". This state of affairs—where relapse is *common* following treatment—would be completely unacceptable in other arenas of medicine. It seems to me that in revealing this reality in their very first sentence, the authors have blown CBT's cover—its claim that the beneficial effects of CBT treatment are sustained over many years.

### What is MBCT?

In this study, MBCT training was to be given to those who had suffered previously from depression and recovered it, but were

not currently depressed. The hope of the researchers being that MBCT training would "inoculate" the patients and prevent them relapsing into further episodes of depression.

> MBCT is a manualized group skills-training program . . . It is designed to teach patients in remission from recurrent major depression . . . to relate differently to, their thoughts, feelings, and bodily sensations . . .

> After an initial individual orientation session, the MBCT program is delivered by an instructor in eight weekly 2-hr group training sessions involving up to 12 recovered recurrently depressed patients. (Teasdale et al., 2000, p. 618)

Over the next year each patient had four follow up sessions. The question that the researchers had to answer at the end of the sixty week period was: are the people who received MBCT better off than those that did not receive MBCT, and if so, by how much?

## *The evidence*

The trial started out with an intent to treat sample of 145. At the end of the sixty weeks they had complete data on 137. Of the people that dropped out they say:

> Of the 13 patients allocated to MBCT not included in the per-protocol sample, 6 failed to attend any training sessions and 7 . . . dropped out after attending fewer than four sessions. (Teasdale et al., 2000, pp. 618–619)

As is the convention in this kind of research protocol, the thirteen are not included in any of the calculations. It is worth pausing to think about this for a moment. The conventional rationale for not including the thirteen in the analysis would be that as they did not take part in "the treatment", we have no data on them.

But in any analysis that is anything more than a simplistic arithmetic one, we would have to inquire *why* these six *adults decided* not to continue having attended the "orientation session". Was it a meaningful rational choice based on the information they gleaned in the orientation session, or were they simply disorganised, unmotivated, and therefore "CBT resistant" (this

being a new diagnostic category developed for those who are not sufficiently persuaded to participate in CBT treatments)?

So rather than saying that "six *failed* to attend any training sessions", one might equally say that "six *elected* not to go any further because they were not persuaded by the information they were given". One might say the same of the seven who "dropped out" having experienced "the treatment" itself. If this were indeed the case, then surely the thirteen should be included among those for whom CBT will not work its magic.

### *Halving, doubling, and disappearing*

The study divided the participants into two groups, one to receive MBCT as well as treatment as usual (TAU) (seventy-one people), and the other just TAU (sixty-six people). TAU can be whatever people usually do, go to their GPs, take anti-depressants, go to therapy, meditate, and so on.

At the end of the sixty weeks they found that out of the seventy-one people in the MBCT group, thirty-one had relapsed (44%); and out of sixty-six in the TAU group, thirty-eight people relapsed (58%). It is worth representing the findings in simple pictures, as the conventional means of representing the infor-mation—consisting as it does of a dense mix of statistical nota-tion and highly digested claims—obscures more than it reveals (see Figure 1).
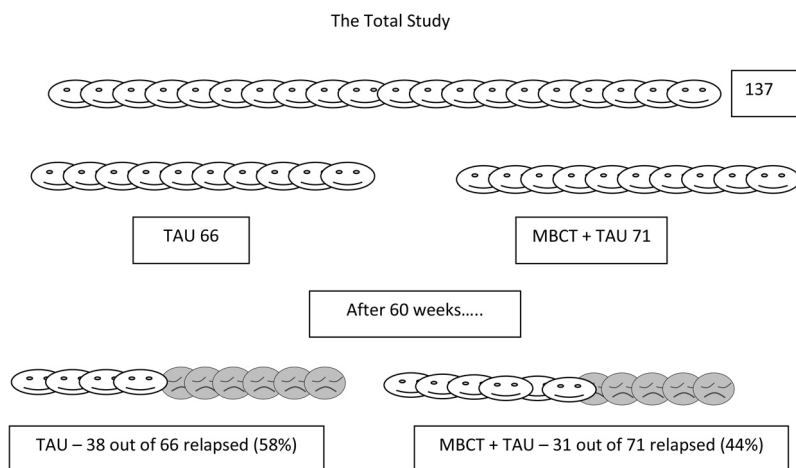
The Total Study



Figure 1

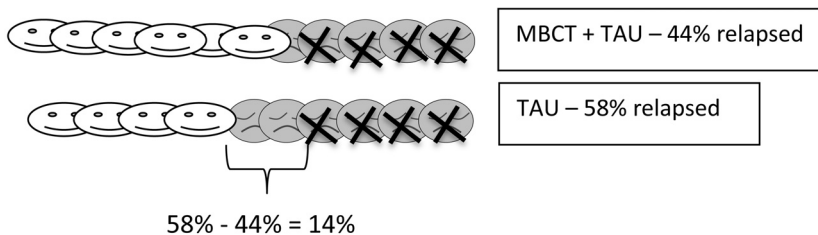Farhad Dalal

58% - 44% = 14%

Figure 2

When we remove the numbers from the control group, we are left with just 14% fewer people relapsing in the MBCT group (see Figure 2).

Interestingly, these figures are *never stated* in this form in the paper. In one place they tell us that of the 105 who had suffered three or more previous episodes of depression, "40% (22/55) of MBCT participants experienced relapse/recurrence compared with 66% (33/50) of TAU participants" (Teasdale et al., 2000, p. 620). And in another place they tell us that of the thirty-two who had suffered two or less previous episodes of depression, "56% (9/16) of MBCT participants experienced relapse/recurrence compared with 31% (5/16) of TAU participants" (Teasdale et al., 2000, pp. 620–621). So, the total numbers that relapsed after MBCT treatment amounted to (twenty-two plus nine) thirty-one, out of a total of (fifty-five plus sixteen) 71—44%; and the figures for those who underwent TAU: thirty-eight out of sixty-six relapsed—58%.

So why are they never cited in this form? Presumably, it is because these figures are not very impressive. They show that *just one or two more* people out of every ten people (14%) in the MBCT group fared better. In other words there is no statistical difference in the outcomes between those who received the treatment, and those who did not. The null hypothesis is still true—statistically speaking, *the treatment makes no difference*.

And if we add in the thirteen who elected not to take part in the study—as I would be inclined to do—it would make the outcome even more invidious.

### *Success with those who have suffered at least three previous episodes of depression*

However, it turns out that the figures for relapse rates are much better for a subset of the total population—those that had had three or more previous episodes of depression. They discovered that for some (unknown) reason, fewer members of this group relapsed in comparison to those who had previously recovered from two or less previous episodes of depression. It is the findings for this grouping that the paper highlights to such a degree that it obscures everything else, including the fact that *in toto* the treatment does not do much at all.

For those who suffered three or more previous episodes of depression (105 out of 137), they say, "Over the total study period, in the intent-to-treat sample, 40% (22/55) of MBCT participants experienced relapse/recurrence compared with 66% (33/50) of TAU participants" (Teasdale et al., 2000, p. 620).

On this basis they tell us that that this constitutes a "39% reduction in risk" following MBCT treatment. It was not self-evident (at least to me) how they arrive at this figure from those that they have allowed us to know about. Let us therefore look at the situation diagrammatically (see Figure 3).
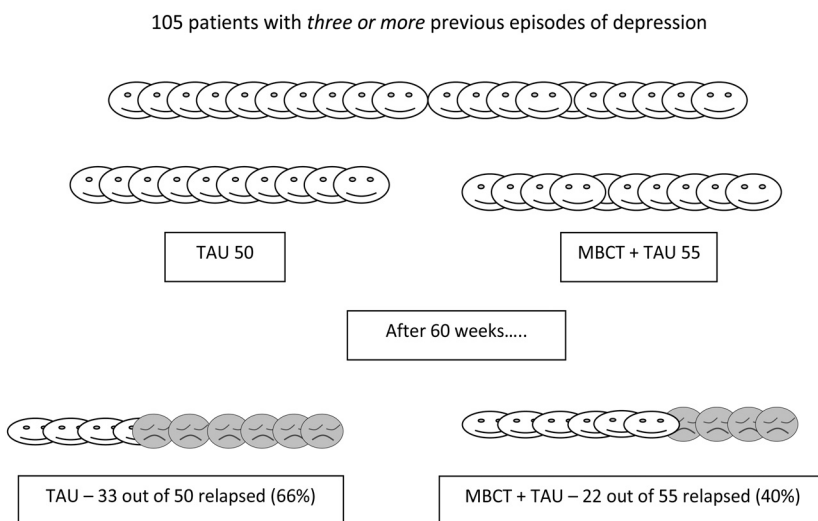
105 patients with *three or more* previous episodes of depression



TAU 50

MBCT + TAU 55

After 60 weeks…..

TAU – 33 out of 50 relapsed (66%)

MBCT + TAU – 22 out of 55 relapsed (40%)

Figure 3

Farhad Dalal

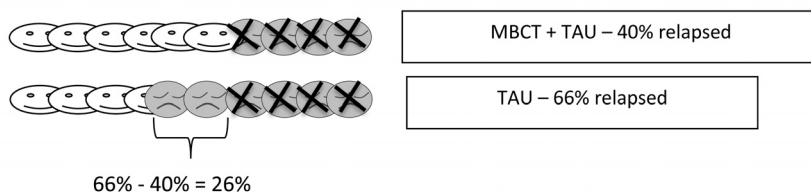MBCT + TAU – 40% relapsed

TAU – 66% relapsed

66% - 40% = 26%

Figure 4

To my mind, the obvious way of reading the results is as before, which is say that 26% (66 – 40 = 26) of those in the MBCT group were better off than those in the TAU group (see Figure 4).

Meanwhile, they make no mention of the 26% and instead proffer the more impressive figure 39% and call it "the reduction in risk". What is going on?

Let us go back to basics to think about this situation in regards to the three or more situation.

The study shows us that if we did nothing new (TAU), then about 66% of this group (three or more) are likely to relapse over the year. This is our base line—call this number X.

On this basis, the number of people relapsing *after* treatment (Y) becomes the measure of the success or not of the treatment.

- If the same numbers relapse with and without treatment $(Y = X)$, then $Y/X = 1$, and the treatment is useless.
- If fewer people relapsed after treatment $(Y < X)$, then $Y/X < 1$, and the treatment is helpful.
- If more people relapse after treatment $(Y > X)$, then $Y/X > 1$, and the treatment is harmful.

Now, there are two ways of reading the relationship between Y and X.

To my uneducated eye, the obvious way of reading these figures is to say:

- The number of people who relapsed with no treatment (X) is 66%.
- The number of people who relapsed after treatment (Y) is 40%.
- Therefore, the people in the treatment group were better off by $(X–Y) = 66\% – 40\% = 26\%$.

I have subtracted one figure from the other.
They, instead, elect to divide one figure into the other.

$$Y/X = 40/66 = 0.61 \ (61\%)$$

$$100\% - 61\% = 39\%$$

To be sure they have performed a calculation, but what is it actually describing? They have elected *to name the result* a "reduction in risk". But is this what this calculation really means? Does this figure *really* describe a "reduction in risk"?

Here is a way to understand the difference in these two calculations, and what this is likely to mean in practice. I could for example, try to impress you by telling you that I have discovered the secret of quadrupling your chances of winning the lottery; *my* secret will increase your chances of winning by a dizzy 400%. How? Buy four tickets. True, I have increased your chances of winning by 400% but only *relative* to your original chance (one ticket). But in absolute terms, all I have done is to increase your chances of winning from one out of 14 billion, to four out of 14 billion. Not very impressive really.

This is exactly the difference between the 26% and the 39%; the 26% reduction being the absolute reduction in risk, and the 39% being the relative reduction in risk, *relative to the 66%*!

To my mind, as a sufferer of depression, what I would want to know is: what difference will joining the MBCT group make to *my well-being*. Answer: if I join the MBCT group, I will *reduce my chances of relapse* by 26% (but only if I have relapsed on *at least* three previous occasions). That is it. That is the big finding. Statistically speaking, I am much more likely not to feel any benefit from going through MBCT training—the likelihood of *not benefiting* being around 75%. Nevertheless, the 26% speaks directly and concretely to my experience and so I am able to put its meaning to me into words. The 39% meanwhile, is an arithmetic mystification, by which I mean that I find it impossible to put into words what this means to *my experience* in any meaningful way.

But of course, 26% is not an impressive a number as 39%.

I was helped to see the deeper connection between the two figures by the mathematician Roger Porkess, who pointed out that 39% of sixty-six is twenty-six! It begins to seem that we are

in the territory of numerology masquerading as statistical science.

It is part of the art of obfuscation not to trouble the reader by informing them that "the reduction in risk" they are citing is relative rather than absolute, nor to trouble then with what the absolute reduction actually is, nor to reveal the actual calculations used to derive the figures in the first place. These practices are normative conventions within this research field.

### The art of amplification

Even so, let us allow them their 39%, and look at what they do with it. They say, "in participants with three or more previous episodes of depression . . . an 'adequate dose' of MBCT *almost halved* relapse/recurrence rates over the follow-up period compared to TAU" (Teasdale et al., 2000, p. 621, my italics). In this sentence, the 39% (less than four out of ten) has been morphed into "almost half". To be even more precise, to call it "almost a half" is to increase 39% by 11%, and this 11% is more than 25% of the original figure of 39%. That is a very, very big "almost"—25% bigger.

The stealthy amplification continues a couple of paragraphs later, where we find that the caveat "almost" is no longer deemed necessary, "*the halving of relapse* . . . rates in a group of high risk . . . would appear to be a clinically useful outcome" (Teasdale et al., 2000, p. 621, my italics). In a few paragraphs we have gone from 26% (which does not even get a mention) to 39%, to almost a half, to a half. Notice the transmutation process: we start with *a number* 39%; we then change the form of representation from numbers to words; this allows 39% to be rendered as "almost half"; next, because "almost half" is so close to "half", we may as well call it "half". Having got to "half" the discourse reverts from the domain of words back to numbers and we find ourselves faced with 50%, which after all is simply the numerical representation of half (see Figure 5).

To my mind the claim of 50%, is deceitful and deeply unethical.

This then is the message that is broadcast to all and sundry: MBCT halves the relapse rates of all those prone to depression. And sure enough, this is exactly what we find on the official MBCT website, "The evidence from two randomized clinical
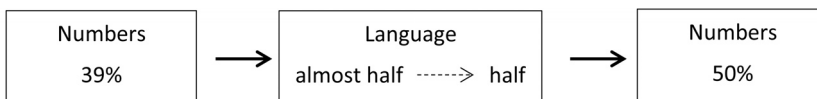
| Numbers | | Language | | Numbers |
|---|---|---|---|---|
| 39% | → | almost half ------≫ half | → | 50% |

Figure 5

trials of MBCT indicates that it reduces rates of relapse by 50% among patients who suffer from recurrent depression" (MBCT).

### *The economic argument*

The economic rationale for this treatment also falsely empha- sises "prevention" in lieu of "likelihood" and "reduction". They say that the, "*Preventative* effect of MBCT was achieved for an average investment of less than 5 hours of instructor time per patient, suggesting [this is] . . . a cost-efficient strategy for *prevention*" (Teasdale et al., 2000, p. 622, my italics).

Fund holders are bound to be excited at the prospect of *preventing* depression for hardly any cost at all—for five hours work per patient. They would be less excited if they were told that what was being achieved was not "prevention", but simply a 26% *reduction in the likelihood of relapse*, and this too for a very narrow population. Fund holders are even less likely to be impressed if their attention is drawn to the fact that *the treatment will make no difference to 75%* of those who had suffered three or more previous episodes of depression, who will continue to relapse as before.

Neither is it clear as to how they arrive at the figure of less than five hours. To calculate the amount of instructor time invested, one needs to add up the orientation sessions, one per patient (twelve hours), the training itself (8x2 = 16 hours), and the four follow up sessions per patient (4x12 = 48 hours). This amounts to a total of seventy-six hours of instructor time, which divided between twelve patients, results in 6.3 instructor hours per patient.

To be sure this is not the most important part of the critique, but because they do not tell us how they calculated this figure, it could constitute yet another inflation of the actual results of the study.

Farhad Dalal

### Therapy or training?

But there is another element of note in the statement above: *there is no therapist present*; instead we find an "instructor". Ma and Teasdale (2004) inform us that MBCT was previously referred to as "attentional control training", and they too refer to "instructors" delivering the training. Given that there is no therapist present, surely rather than calling this therapy, it would be more correct to call MBCT an educational programme or a "training". Nothing wrong with that; we should just call a spade a spade.

It seems to me that two reasons lie behind the change in title from training to therapy. First, therapy carries more kudos than training. This is because therapy evokes connotations of cure, while training evokes connotations of symptom management. And second, the notion of training is better suited to the ethos of the second wave (control), rather than that of the third wave (acceptance). But as Shakespeare once almost said, a training by another name remains a training.

### Two or less episodes of depression

But what of those who had two or less previous episodes of depression? It would seem that MBCT did not help them. The researchers say, "MBCT appeared to have *no prophylactic effects* in those with only two previous episodes" and "MBCT prevented relapse and recurrence in patients with a history of three or more episodes of depression, *but not in patients with only two previous episodes*" (Teasdale et al., 2000, p. 622, my italics). The mild neutral tone is repeated in the follow up paper four years later, "The earlier finding that a group of patients with two previous episodes of depression . . . showed *no evidence of benefit* from MBCT was also replicated" and "MBCT can be *relatively unhelpful* for a particular group of patients (Ma & Teasdale, 2004, p. 38, my italics).

But what are the actual figures regarding relapse for the two-or-less group? "Over the total study period . . . 56% (9/16) of MBCT participants experienced relapse/recurrence compared with 31% (5/16) TAU participants" (Teasdale et al., 2000, pp. 620–621).

Do look at the numbers again. *More* of the people who received MBCT relapsed, compared to those who did not get the

treatment. Twenty-five per cent more of them. Yet the researchers do not say that the treatment made them ill, all they say is that there is *no evidence of benefit* of MBCT to this group.

Surely on the same grounds that they claim that MBCT is efficacious for the three-or-more group, 26%, they ought to say that MBCT is positively harmful to those who have suffered two-or-less, 25%. Sure the numbers are small, but the differentials are close to identical.

If we put the figures for the two-or-less group through the same arithmetic procedure as they did to arrive at 39%, then for the two-or-less group we find that the treatment will *increase the chances of relapse by 80%* relative to the original chance.

It is clear that for the two-or-less group the MBCT treatment is iatrogenic—it is making these people *more ill*. This embarrassing fact is obscured and packaged in bland neutral statements of "no benefit", and that they have determined that the *p* value shows no statistical significance, "As in Teasdale et al.'s (2000) study, these patients [two-or-less] showed a nonsignificantly greater tendency to relapse following MBCT" (Ma & Teasdale, 2004, p. 38).

### p *values,* p-*hacking, and effect size*

This is where the term *p*-hacking becomes important. Statisticians use something called a "*p*" value to say whether their findings are statistically significant or not. The convention within the research community is that if *p* is less than 5%, then the results are considered statistically *significant*.

The first thing to be said is that the 5% is by no means an objective measure of some natural phenomenon—be it Avogadro's number, or the weight of something. The 5% is arrived at through a *professional consensus* that this is where one will draw the line between statistically significant and not significant. However, many statisticians might well have issues with whether this is a reasonable place to draw the line or not. I stress this in order to underline that the claim that something is "statistically significant" is also, despite numbers and calculations, in many senses something subjective rather than objective information.

*p*-hacking must be rife, because almost every published study magically ends up just under this threshold. As Theodore

Sterling has previously noted 286 out of 294 papers reported a statistically significant result.

*p*-hacking is easy because researchers do not share their raw data, nor do they declare the way that they calculate *p*; all they do is to state its value embedded in a mix of dense hieroglyphics that immediately saps the uninitiated of their will to live, and can look like this:

$$\chi^2(1, N = 99) = 8.49, p < 0.005$$

It is on this sort of basis, that Teasdale and colleagues are able to dismiss the fact that this subgroup is positively harmed by MBCT.

## Further issues

There are a number of other elements in the paper that fall short of the self-aggrandising rhetoric endemic to CBT discourse, its claim that its research methods are scientifically *rigorous*.

If the authors of this study were truly rigorous, then they would curtail their claims even further; this is because more than three quarters of the participants were women and close to 100% of them were white.

So properly speaking, the findings as such are limited to the category "white women"; to utilise the treatment with other social groupings is a form of "off-label prescribing" (recall the Zon situation).

It is also the case as the authors of both papers tell us that "MBCT was most effective in preventing relapses *not preceded by life events*" (Ma & Teasdale, 2004, p. 31, my italics), and more effective when the depression was driven by "autonomous ruminative-thinking" (Ma & Teasdale, 2004, p. 32). This then constitutes another major constraint that further curtails their claims on the efficacy of MBCT.

And finally, there is no way of emphatically knowing how and why those in the MBCT group relapsed less than those in the TAU group. Because apart from anything else, over the sixty week study period, those in the MBCT group accessed other forms of help *more* than the TAU group. The MBCT group visited their GP for a depression related issue 6% more than the TAU group (58% MBCT *vs.* 52% TAU); the MBCT group reached

out for counselling 15% more than the TAU group (49% MBCT *vs*. 34% TAU); and the MBCT group used medication 5% more than did the TAU group (45% MBCT *vs*. 40% TAU).

Therefore the study cannot unequivocally demonstrate that the reduction in relapse rates in the MBCT group was due to MBCT or because of the additional help they accessed over the sixty week period.

The researchers make tokenistic gestures towards the limitations of their study, seeking to lend their position an air of scientific humility, saying things like, "the sample sizes in the two groups mean that estimates of risk have appreciable margins of error" (Teasdale et al., 2000, p. 622). They even confess, "The relapse/recurrence rate in patients . . . *was clearly substantially above* the expected annual incidence rate . . . it is clear that the intervention did *not reduce risks of major depression to the 'normal' range*" (Teasdale et al., 2000, p. 621, my italics).

This is all forgotten in the triumphalism we find in the journal abstract (which is all a busy professional is likely to have the time to read).

## In conclusion

In conclusion, I want to underline the following issues.

The evidence base for CBT is of the statistical kind that speaks to likelihoods but it is presented as though it were generating certainties no different to the more objective disciplines of mathematics or physics.

This next point is quite important. It is the case that both the supporters and detractors of CBT seem to be in agreement that in round figures, CBT is of benefit to about 50% of the population. It is on this sort of figure that the Exchequer has been persuaded to part with eye-watering amounts of money. But if it is found to be the case that the figures we have come across in this study (that the efficacy is not 50% but 25%) are representative of the CBT research base in general, then something much more worrying is going on. Then we would have to ask: is the whole CBT research base infected and corrupted by these kinds of statistical malpractices?

I doubt that the Exchequer would be all that keen to empty the public purse on the more realistic but much more meagre

Farhad Dalal

promises of *reduction in likelihood* rather than *prevention*, and a likelihood of reduction of 25% rather than 50%.

When all is said and done, then ethically, the findings of this research should in fact be announced in this way:

> Over the period of a year, MBCT is likely to reduce the chances of relapse into depression for something between two to three out every ten people within a very specific group of patients—white women who are not currently depressed, but who have suffered and recovered from at least three previous episodes of depression. Further, their depressions ought not to be caused by actual life events, but fed by "autonomous ruminative thinking". Therefore, about seven out of ten people within this group (three or more)—"almost three quarters"—will continue to relapse despite the treatment.

These exciting claims, it should be remembered are for the grouping that has responded best to the treatment. Ethically, the abstract should also highlight the fact that:

> Of those who have suffered and recovered from two-or-less previous episodes of depression, 80% are *more likely to become ill because of the treatment*. However, they will be relieved to know that their suffering caused by MBCT although personally prob-lematic, is statistically not significant.

> And if the entire population of those who had previously suffered depression (but were currently in remission) were put through this treatment, then only about one or two people out of every ten are likely to reduce their chances of relapse.

So despite the range of caveats supplied by the researchers themselves, despite the fact that over three quarters of those who the treatment is meant to help will not be helped, despite the fact that those with two-or-less episodes will actually be made to feel considerably worse, despite the fact that this treatment is deemed inappropriate for anyone who is depressed because of a life event (most of human kind surely), the researchers feel able to end their abstract with the rousing statement, entirely devoid of qualifications, "MBCT offers a promising cost-efficient psy-chological approach to *preventing* relapse/recurrence in recov-ered recurrently depressed patients". But even worse, many of the great and the good—governments and policy makers—have been gullible enough to believe them.

## References

Goldacre, B. (2012). *Bad Pharma*: *How Drug Companies Mislead Doctors and Harm Patients*. London: Fourth Estate.

Keitner, G. I., & Posternak, M. A. (2003). How many subjects with major depressive disorder meet eligibility requirements of an anti-depressant efficacy trial? *The Journal of Clinical Psychiatry*, *64*(9): 1091–1093.

Kirsch, I. (2011). Antidepressants and the placebo response. In: M. Rapley, J. Moncrieff, & J. Dillon (Eds.), *De-Medicalising Misery* (pp. 189–196). London: Palgrave Macmillan.

Layard, R. (2005). *Happiness*. London: Penguin.

Layard, R., & Clark, D. (2014). *Thrive*: *The Power of Evidence-Based Psychological Therapies*. London: Allen Lane.

Ma, S. H., & Teasdale, J. D. (2004). Mindfulness-based cognitive therapy for depression: replication and exploration of differential relapse prevention effects. *Journal of Consulting and Clinical Psychology*, *72*(1): 31–40.

Mindfulness-based Cognitive Therapy (MBCT). Does it work? Available at: http://mbct.com/about/does-mbct-work/ (accessed May 2015).

Royal College of Psychiatrists (2013). *Cognitive Behavioural Therapy*. London: Royal College of Psychiatrists. Available at: http://www.rcpsych.ac.uk/expertadvice/treatments/cbt.aspx (accessed May 2015).

Sterling, T. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *American Statistical Association Journal*, *54*: 30–34.

Sterling T., Rosenbaum, W., & Weinkam, J. (1995). Publication decisions revisited—the effect of statistical tests on the decision to publish and vice versa. *American Statistician*, *49*: 108–112.

Teasdale, J. D., Williams, J. M. G., Soulsby, J. M., Segal, Z. V., Ridgeway, V. A., & Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology*, *68*(4): 615–623.

Travers, J., Marsh, S., Williams, M., Weatherall, M., Caldwell, B., Shirtcliffe, P., Aldington, S., & Beasley, R. (2007). External validity of randomised controlled trials in asthma: to whom do the results of the trials apply? *Thorax*, *62*(3): 219–223.

Turner, E., Matthews, A., Linardatos, E., Tell, R., & Rosenthal, R. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *The New England Journal of Medicine*, *358*: 252–260.

Farhad Dalal

University of Bristol (2012). Cognitive behavioural therapy is effective in reducing depression in people whose symptoms have not responded to treatment with antidepressants. Press release issued: 7 December 2012. Available at http://www.bristol.ac.uk/news/ 2012/8986.html (accessed May 2015).

Whitaker, R. (2010). *Anatomy of an Epidemic*. New York: Broadway Books.

Wiles, N., Thomas, L., Abel, A., Ridgway N., Turner, N., & Campbell, J. (2013). Cognitive behavioural therapy as an adjunct to pharma-cotherapy for primary care based patients with treatment resistant depression. *The Lancet*, *381*(9864): 375–384.