Psychopharmacology in Turmoil: an ethical or scientific crisis?
David Healy: Columbia October 20th 2005
(Lightly edited for the purpose of posting)

Thank you very much for this invitation.

Preamble
What I'm going to do first is to whip through a few slides and leave copies of them there. These first slides have to do with the conflicts of interest that you may wish to pick up with me at the end of the talk.

As you'll see from the first slide (1), when I accessed under F.O.I. provisions the material pharmaceutical companies have on me, you find things like this from Eli Lilly who it seems are happy to organise for a person to be here in the audience to keep an eye on the things that I say and check to see if I can be sued on the basis of the things said.

The second slide is from a posting by Pfizer (2) just before the FDA hearing on antidepressants and suicidality in children last year. This is a 41 page piece on the FDA site, largely about me. The third slide (3) gives you a flavour of the things Pfizer had to say about me. 'Dr Healy is pretty clearly from this, an unsavoury customer of some sort who before he got involved as an expert witness in the SSRI trials', (I have actually been involved in very few trials), 'said completely different things about this group of drugs to the kind of things that he says now'. This is pretty well completely untrue. I do not think you would be able to get a hair between the views that I had prior to being involved as an expert witness and the views that I have on this group of drugs now.

But this talk was initially designed to be a debate between Jim Coyne and I. The kinds of things that Pfizer had been saying previously were things that Dr Coyne has also been saying at much the same time that Pfizer began to say similar things.

One of your handouts is an article by Dr Coyne who I understand is due to talk to some of you in a weeks time on the martyrdom of David Healy. The claims made in this article I find are very hard to pin down. I think some of his issues come through clearer on a listserve he contributes to and the next set of slides give you some of Dr Coyne's claims about what I've got up to (4,5,6). I first heard of Dr Coyne through the Globe and Mail letter (5), after I'd lost my job at the University of Toronto. In this letter Dr Coyne goes through his views which are that it's unclear why the University of Toronto had thought to offer me a post in the first instance. I'd done no research, he claims, and where I had done research on healthy volunteers it was flawed.

The only other people at this point that were making the same claims that you see here from Dr Coyne were Pfizer. I am very happy to field any questions that you may have on this piece of healthy volunteer research at the end of the talk. But the points Dr Coyne and Pfizer claim such that I did this research on 20 underlings that worked with me are untrue. In fact pretty well all the things that are being said here are untrue.

One of the other points Dr Coyne makes is 'Dr Healy put the first reports of the research that 2 healthy volunteers had gone on to become suicidal in an obscure journal'. It is the case that, at the point that, Dr Coyne began to say the things that he began to say first, it was hard to access the article that outlined what had happened to the healthy volunteers that had become suicidal taking Zoloft. One of the reasons

this was hard to access was that it had been offered first to the BMJ who didn't want to have anything to do with it.

This leads on to a further letter from the BMJ earlier this year after they'd had a brief piece from me that covers some of the points you're going to hear during the course of the next hour (slide 7). After this article had been through the peer review process and had been approved by the peer reviewers - the reviews were longer than the article – the article was held up and remains held up by BMJ's lawyers, even though everything in the article is in the public domain. So what you need to take into account from this is that in many cases some of our best journals refuse to take articles on the basis of sheer terror about what pharmaceutical companies might do to them rather than on the basis of any concerns about the scientific merits of the work.

Now I would concede that when he first made the points he made about the healthy volunteer study, there may have been some basis for Dr Coyne to make these points in that it was hard to find the actual work. He might have come to me to ask more, but didn't. But there was no excuse for Pfizer to make exactly the same points as they had every single piece of paper from this particular study and they knew the things reported in the media as being claimed by them were incorrect.

You'll see Dr Coyne also says (8) that you need to note that Dr Healy gets paid for work as an expert witness and takes considerable amounts of money from drug companies. You're going to hear all about the cash that I take from drug firms in the course of this talk.

Dr Coyne has a vigorous style as you see here (slide 9). One of the people who takes issue with him gets told that they have the critical faculties of a broken lawn chair. Carl Elliott gets criticised as you see here. And Sarah Boseley, one of the journalists that he took issue with, responded in the way you see that I think gives you a feel for the way Dr Coyne tackles issues (10, 11). She makes it clear she is being found guilty by association. Just because work of hers had ended up on a website run by the Church of Scientology that she was equivalent to being a scientologist.

Two more bits before we get into the main body of the talk. These slides were all from three or four years ago but here are comments from a few weeks ago (12, 13). Dr Coyne says that 'David Healy show up with junk science for fees greater than 30,000 euros'. Just for the record I haven't taken a single euro in legal fees ever – you really can't take an action against a pharmaceutical company in Europe. And 'he's armed with junk science and he cooks data'.

Now just to make things easier for you the research that I'm going to talk about here isn't my own research. There is only one bit of research that you're going to hear about during the course of the next hour that I have had a direct hand in and I'll clearly flag that up so you can be duly wary of it.

Problem 1

The next slide (14) shows an electric chair powered by wind power. This aims to make the point that there are an awful lot of good things that people intend that end up being put to a purpose that may be quite at odds with the original intention. I don't think that you can really at the end of the day infer from what people do what their motives may have been. So in the course of this talk you don't want to think that I'm making inferences about or that I think I know why particular people may have done

things they did. What I'm going to be talking about is a system that I think we're all caught up in, from which it's very hard for any one to emerge untainted.

Lets begin with this excerpt from a review of one of the drugs that we use now to treat people who have got schizophrenia (15). This is a review by the FDA of Risperdal. The FDA are here saying that if the pharmaceutical company that produces Risperdal were in any way to claim that it was superior to any of the older drugs in the field that they would be breaking the law. There is no evidence that FDA have seen to support such a claim. Now if you work for the pharmaceutical company that produces Risperdal, you might think this would give you a problem in that Risperdal costs 30 to 40 times more than the older drugs. So if it's not superior to the older drugs which hospital, which pharmacy, which physician is going to use it?

Well the answer for a pharmaceutical company faced with a problem like this is easy. You get into evidence-based medicine. You hire the experts. You get the Professors of Psychiatry and others, in this case, from the state of Texas. You convene them and get them to read the evidence. You get them to read the clinical trial material on Risperdal compared with older drugs and compared to placebo, in order to establish just how good Risperdal is or not compared to the older drugs. You convene these talking heads in a room and ask them to weigh things up and devise a set of protocols, guidelines, and algorithms, which outline how you would treat people who've got schizophrenia (16).

These algorithms will say that on the basis of the evidence you should use drugs from group A first, then drugs from group B second, as in this algorithm here, which is the algorithm for treating people who are depressed from TMAP – the Texas Medication Algorithm Project (17).

You have to of course pay these experts to read all the clinical trials and you've got to pay the experts to convene in a room and experts don't come cheap. So does this exercise just prove that, if you pay your experts you get the answer you want? In the case of schizophrenia, the experts in Texas endorsed Risperdal over the older drugs and later went on to endorse Zyprexa and all of the more recent drugs to treat people who've got schizophrenia. And when it comes depression they endorsed the SSRIs over the older drugs. They always seem to be able to recommend for first line use the drugs that are on patent.

Well just to confuse you, look at the next slide (18). There are 2 bits to note. Down at the end you'll see a hint of an algorithm that goes on to argue that if you treat people who have schizophrenia use drug A first, in this case Risperdal, then drug B and drug C. Janssen Pharmaceuticals, who convened and created TMAP, did the same exercise in the UK, once Risperdal was licensed there. They convened a bunch of experts, paid those experts to read the evidence, and to spend time in the room etc. The experts came up with the views that 'Well yes, if you are going to treat people who have got schizophrenia you use Risperdal before you use the older drugs'. I know all about this because I was paid by Janssen to do it. I was in the room. I read the evidence. One of the authors on the authorship line you see here is me.

Hopefully this next slide is going to make you even more unsure about what's going on here. This is the National Institute for Clinical Excellence in the UK, NICE, who are independent of the pharmaceutical industry (19). The experts for NICE also read all of the evidence and sat in rooms and they aren't paid by any of the companies. And what conclusions do they come up with? They conclude that you should use Risperdal and Zyprexa and all of the newer drugs before you use the older drugs and you should use the newer drugs to treat people who are depressed rather than the

older drugs. So question is have the pharmaceutical companies made a rare mistake. Have they gone and paid the experts when they didn't need to?

Let's raise the bar a bit further. This next slide shows another piece of work undertaken by Janssen in the UK (20). They were keen to ensure that the experts didn't just say that Risperdal was superior to the older drug but actually if we switched everybody over from the older drugs, from drugs like Chlorpromazine, to Risperdal that the NHS would save money. So what do you do? You convene a bunch of experts in a room and you get them to look at the evidence and you get them to outline treatment pathways. You look at the costs of the treatment pathway if patients take Risperdal versus one of the older drugs. And what do you know, the experts conclude that even though Risperdal costs 30 times more that the NHS would save money if we just switched people over from the older drugs to Risperdal. I know all about this because I was in that room and read the evidence. The beauty of the process is that you don't have to have the people in the room who are linked into the pharmaceutical industry. You can have people in the room who are hostile to pharmaceutical companies, and the outcome will be the same.

This exercise resulted in this peer reviewed article here (21) which is the kind of thing that I or colleagues here in the US could take along to their hospital formulary group to say 'look we really need to have Risperdal on the hospital formulary because, don't you know, even though it seems to cost more money, in actual fact we are going to save money if we just use it instead of the older drugs'.

A comparable exercise was undertaken for Zyprexa, which you see in the next slide (22). This is not an exercise that I was involved with. Again though the conclusion is that if you just switch people over from the older drugs that they've been using to Zyprexa, that the NHS would save money.

Well this issue of how good the new drugs are and whether the services would save money was looked at in a controlled trial that came out 2 years ago in JAMA. This VA trial (Rosenheck et al) compared Zyprexa and haloperidol in the treatment of people who have schizophrenia. Both drugs come out as equi-effective but the overall cost for Zyprexa was 40-50 times more than the cost of haloperidol. There was no cost saving. Clinically both drugs were just the same, which the FDA in essence had said they would be, but the new one drug cost far more than the older one.

And the recent CATIE study comes to pretty much the same conclusions. So how come the experts convened in all these rooms, me included, so clearly came to the wrong answer?

Well part of the answer may lie on the next slide (23), which shows a portfolio of articles on Zoloft being co-ordinated for Pfizer by CMD, Current Medical Directions, an agency based here in New York. The next slide (24) shows you the things that CMD said that they do, as of 2 or 3 years ago when I checked out their website. The website has now changed. So CMD write up articles, review articles, abstracts, journal supplements, product monographs, expert commentaries etc.

What you find on the inside of the portfolio is things like you see in the next slide which is the PTSD page of the portfolio (25). You see over here on the right that 2 PTSD clinical trials that have been completed. They've been written up. One of them will go to the New England Journal of Medicine and one will go to JAMA. In fact one does end up with JAMA and the other comes out in the Archives of General Psychiatry. Over in the left hand corner you will see 'author TBD' - Author To Be Determined.

Now this gave us the opportunity to look at all of the articles during the period that CMD were coordinating for Pfizer that ultimately ended up in print versus all the articles on Zoloft that weren't listed in that particular portfolio (26). So if you compare the CMD articles versus the non-CMD Zoloft articles, the next slide shows you that the impact factor of the journal in which the CMD articles ends up in is 2 times greater than the impact factor of the journal in which the non CMD articles end up in. If you look at the numbers of previous articles that authors on the authorship line of the CMD articles have, you find that these people have at least 2 times more articles to their name than those who are authors on the group of articles that aren't being coordinated by CMD.

So if a clinician gets an article that seems to come from Harvard and is in the New England Journal of Medicine, but this article is clearly linked to a pharmaceutical company, what do you think the clinician will do when faced with this? Do they think that this is just an advert for one of the pharmaceutical companies or do they treat this as the genuine scientific thing? Well the answer is here on the far right. This column shows that the CMD group of articles are 3 times more likely to be cited than the articles on Zoloft not being coordinated by CMD. In other words we treat these articles as the real scientific thing.

The British Journal of Psychiatry usually has 2 peer reviewers. I got at least 6 reviews from 6 different peer reviewers of this article and it went back to all of the reviewers after we'd made all the changes. It then went through the legal department of the British Journal of Psychiatry and was finally copy edited to death before it was published. The British Journal of Psychiatry just like BMJ is very scared of pharmaceutical companies.

This next slide (27) gives you a flavour of the way that pharmaceutical companies view articles, which you and I might think are clinical studies (26). You see here this communication from Pfizer shows company personnel deciding who ends up on the authorship line. Which senior people in France, in this case, would suit the company's interests.

So this all leads us to this next point, which is data on completed suicides and suicidal acts in the trials of antipsychotics sent in to FDA for registration purposes. Now remember, people like me and like all of you in this audience I think if asked to review the evidence on Risperdal compared with the older drugs would have come to the same conclusion that we should be using Risperdal first. But if you remember FDA didn't come to that conclusion. The difference is that the published articles aren't showing the same thing as FDA has been seeing. What the FDA has been seeing, which isn't in any of the published articles, are figures of this sort here which are the completed suicides and suicidal acts on Risperdal and Zyprexa (28).

If you look at the Zyprexa suicide figures, these come from 2500 people who were in 5 clinical trials. You repeatedly hear that one of the problems we've all got is that the pharmaceutical companies hide trials. Well far from hiding these trials, these 5 clinical trials have given rise, when I checked last, to 234 different publications, of which 41 are major publications. These trials haven't been hidden. There are 8 articles or abstracts per trial. But in all of these scientific publications you aren't able to find the fact that this is the drug with the highest suicide rate in clinical trial history. You aren't able to find either from any of the articles that have been published or from the FDAs reviews of the drug how many people went on to suicidal acts. And, given that these figures suggest Zyprexa has the highest suicide rate in clinical trial

history, you would have to think from the fact even FDA aren't prepared to show what the figure of suicidal acts is that it may be rather high.

This next slide shows the greatest divide in all of medicine between what the scientific data show on the one hand and what the actual published articles show on the other hand (29). These are the articles on Paxil and Zoloft in the treatment of children who are depressed. You see repeated down through the slide that all of these authors say that these drugs are safe and effective. Safe and effective. Safe and effective etc. Even though on the right hand side there you see that I've put in for the first 3 or 4 studies the rates at which children who are depressed in these trials go on to a suicidal act, 9% in the first case. This drug is safe and effective?

The third article in this series, by Keller et al, deserves to be one of the most famous articles in medicine (30). It's a good symbol for something. This is trial 329. These 20 authors include some of the biggest and best known names in the field. Now remember here, I don't know what any of these authors will have known about the raw data. They may well have been just as caught up in a process as I was when I said Risperdal is the drug we should be using before the older drugs. But this article comes out in 2001. A few years beforehand SmithKline faced with the fact that at this point they'd done 2 trials in children who are depressed, trial 329, which you've just seen there and one further one, had concluded as the next slide (31) shows, that Paxil does not work for children who are depressed. The article that you've just seen however says that it does work and that in fact it works awfully well. How could this happen?

Well if you read the next slide (32), it seems that GSK recognise that they have a problem that needs to handled. Because if word gets out that the drug may be unhelpful or may fail to work for children it isn't just going to mean that it won't be given to children, it may well reduce the overall use of the drug. The slide shows that they recognise they are going to have to be careful about what they show the FDA. Ultimately how they're going to move this forward is they are going to get positive data from study 329 and they are going to publish that.

One of the things you need to know is that even if a drug hasn't been approved by the FDA for use, if you have experts writing articles that say that this drug is useful, and if experts at meetings where there are 300 or 400 people in the audience say this drug is useful – as experts did with Study 329 - clinicians will use it off label. And in fact where it might be a crime for Janssen to put in their adverts that Risperdal is superior to the older drugs, if there is anything that I have written which says it is superior, they can put that quote from me into the advert, even though the FDA says they can't make the claim. Because you see FDA don't regulate me or any of the other academics you see on the authorship lines of any of these articles.

The next slide (33) shows a 2-page spread from the BMJ. Most people when they see this see the problem as being on the left. If only the BMJ and other journals didn't take adverts like this and if only people like me didn't get free pens and free cups and small little things like trips to the Caribbean for meetings, and if only RCTs were all that clinicians depended on, such as the RCT which you see on the right, then every thing would be okay in medicine. I hope I've brought home to you is that the problem isn't on the left, the problem is the advert on the right.

The companies want you to think the problem is on the left. They hope you'll focus all your efforts on thinking no free lunch is the answer to the problem. While you're doing that, the pharmaceutical companies are in the business of producing the evidence to look like the kind of evidence that clinicians say influences them. When I

get ambushed by the media in the Caribbean where I'm playing golf and some journalist pops out and say's 'Doctor, aren't you influenced by all this?' and I say 'No.' He or she says ' Well what are you influenced by?' I say 'Well I'm influenced by the evidence.' The industry have been taking clinicians at their own word on this point for a long time.

Problem 2

Now I want to change gears and introduce you to a completely different problem. This is a slide (34) you've seen before – the greatest divide in all of medicine - but despite all you've just seen, at this point I want you to assume that nothing here has been ghost written, that all of these authors saw all of the data and wrote all of these articles.

The American College of Neuropsychopharmacology will have assumed just this. When the crisis first blew up as to whether the SSRI group of drugs could cause problems in children, the ACNP, as you see in this slide (35), convened a Task Force to look at the issue. The group they convened included authors on almost all of the articles that you've just seen, and investigators in the RCTs for most of these drugs. They conclude that the SSRI group of drugs does not increase the risk of suicidal thinking or suicide attempts. They came to this conclusion in part on this basis that as you see the results are not "statistically significant".

But as you see in the next slide the data from ACNP's own report shows a very clear increase in risk (36). It may be that ACNP said that there is no increase in risk because of ACNP didn't write their own report. It was written by GYMR, a PR agency based in Washington DC. Maybe this is why this mistake was made. Or perhaps there is a deeper problem in the field.

Let me introduce you to a group of curves, which you'll see repeatedly through the next few slides (37). If you're looking at whether a drug does good or does harm, what you do is you plot the data from an RCT in the form of a p value function curve like this. What you've got at the top of the curve is the point estimate - that's the best bet as to what the data from this particular clinical trial has shown.

Further down you've got the confidence interval and the p value. Now if the p value, as here, is above the 95% confidence interval the findings are often portrayed as not being statistically significant. But even though in this case the result is not statistically significant, you can see that almost all of the data falls to one side of the line that goes through 1.0, which is the axis for no effect.

Let's see how this plays out. Here in the next slide (38) are 2 groups, Hommes et al and Messori et al, who are in the business of trying to look at whether subcutaneous Heparin works or not. As you see the broken and unbroken curves here are almost precisely the same. But the conclusion from the 2 groups is the opposite to this. You see in this next slide (39) that Messori et al say that Hommes et al had things all wrong. If you look at line 2 or paragraph 2 where they say 'the result when we looked at the data was an odds ratio of 0.61 and Hommes had an odds ratio of 0.62'. They go on to claim that 'we have precisely the opposite findings to Hommes et al' and they claim this on the basis that the confidence interval in their case includes 1.0 while in the Hommes et al case it doesn't.

The people in this audience who are best placed to realise how absurd this is are probably the people who hate figures and say they know nothing about statistics. You're the ones who can see that this claim is totally ridiculous. But this in fact is the

basis of all of the arguments about whether the SSRIs cause suicide or not. Those who say the SSRIs do not cause suicide are taking a Messori et al position.

Worse again, in the next two slides (40, 41) you will see that we are currently training all of our juniors to make exactly the mistake you've just seen from Messori. Look at this - if the odds ratio includes 1.0 what do you conclude? You conclude that the figures are not significant.

Well that's from a UK text but people get things wrong in the UK. So here's JAMA for you saying if when you look at the data the upper end of confidence interval lies on the opposite side of the 1.0 value compared to data from the lower end of the confidence interval, you should not conclude that anything has been shown.

Let's put this to the test and ask people which of the 2 drugs in the next slide (42) would you take, drug A or drug B? Drug A is a drug that has 2 times increased risk that if you are on it you are going to drop dead, and this is a statistically significant increase. Drug B has an 8 fold increase in risk that you're going to drop dead if you're on it but this is not statistically significant. Which drug should you take? Well the answer is you should take drug A because it is 4 times safer than drug B.

The next slide (43) shows the Beasley et al article in the BMJ on Prozac and suicide. Beasley et al work for Lilly who make Prozac. This came out when the fuss about Prozac blew up. Here is all of their data and this is what their data looks like. The increased risk was 1.9 times greater on Prozac compared to placebo and you see the confidence interval there. What do they conclude and what do the BMJ happily let them conclude? Well, that there is no statistically significant difference between fluoxetine treated and placebo treated patients. This is reasonable, but you should not go on from to the conclusion which is 'data from these trials do not show that fluoxetine is associated with an increased risk of suicidal acts or the emergence of substantial suicidal thoughts among depressed patients'. If you look at the confidence interval for this data, the scientific data from this article are consistent with the risk being 16 times greater, and the best guess as to what the increased risk is on Prozac compared to placebo points to a figure of 1.9 times greater.

You can't believe all the data in the Beasley article, because Lilly and GSK and Pfizer have handled the data from their clinical trials in a manner that is very interesting. The next slide (44) shows you a schema of the distribution of suicidal acts as they happened in the clinical trials for Fluoxetine, Sertraline and Paroxetine. A group of suicidal acts happened when you whipped the person off whatever treatment they'd been on before, and before they went into the randomised phase of the trial. Then further acts happened when they were on the active treatment or on the placebo and further acts happened after the trial was over. But if you read the published data and FDA reviews of these drugs you see that some pre-randomised and post-treatment suicidal acts have been coded under the heading of placebo. There has been a misleading manipulation of the data as outlined in the next slide (45).

If you go back to the Beasley data that you just saw which gave an increased odds ratio of 1.9 – this risk was computed on the basis that there was one placebo suicide but this placebo suicide shouldn't be there. It didn't happen on placebo. If you take this into account, the risk in fact was infinitely greater on Prozac than placebo.

There's a further trick being played right now with MHRA in the UK even though this issue is in the full glare of publicity. In this case GSK have taken suicidal acts that happened during the post-treatment phase and added these into the mix. What they have done can be seen in an article published in the BMJ in February. I have to put

it to you that academics, journalists and others interested in these issues should be asking the regulator just when GSK's suicides and suicidal acts happened after treatment stopped and for instance whether in the case of some acts coded under the heading of placebo patients hadn't just been started on another antidepressant for instance, and whether this is a reasonable approach to the data.

Lilly went onto look at the rate at which people who have an eating disorder went on to suicidal acts. In this article by Wheadon et al you'll see they claim (46) - 'that if you look at all of the data from our eating disorder trials there isn't any increased risk in people with eating disorders being treated with Prozac compared to placebo'. When in actual fact the data from their own article shows that the risk was 1.5 times greater on Prozac compared to placebo.

Then there is this article by Khan et al who have had access to the files that the FDA have on all of these drugs and have put together data from 40,000 people who had been randomised to one of the SSRIs or to one of the other antidepressants or to placebos (47). This article claims that there isn't any increase in risk on SSRIs. Khan et al haven't taken into account the fact that not all of the placebo suicides happened in people taking placebo, but even so their data shows a 1.4 times increased risk on treatment compared to placebo. Yet they conclude 'there is no statistical difference in the suicide rate among patients taking SSRIs, other antidepressants or placebo'. We can all live with this but I would suggest to you that you ought not be able to live with the next statement that "the only possible conclusion supported by the present data is that prescription of SSRI antidepressants is not associated with a greater risk of completed suicide". This is the wrong interpretation of this data.

The next slide (48) again comes from Khan et al, who also looked at suicide rates in trials submitted to FDA seeking license for use in anxiety. They conclude that they were surprised to find that people who are anxious are much more at risk of going on to kill themselves then they had thought. 'There is a 10-fold risk that people who are anxious will go on to commit suicide compared with the general population'. But if you look at the data they use, from almost 13,000 people randomised to active drug there were 11 suicides, while in almost 4,000 people randomised to placebo there were no suicides. How the authors conclude that data like this points to an increased risk from anxiety rather than it's treatment is hard to understand.

The problem we have here goes back to 2 men. R A Fisher who is the person who is responsible for the idea that we want things to be statistically significant at the 95% level, $p < 0.05$. Fisher thought he was making a claim about what we can say about what's real in the world out there. These days usually we wouldn't view data from RCTs that way. The kind of tests, he used, we would view as means to describe or summarise the data as opposed to tests from which we would make inferences about the real world.

Jerzy Neymann on the other hand said 'Look Fisher's wrong'. If you do an RCT or if you do any piece of work at all it's a good chance that the data you'll end up with will not be statistically significant at the $p< 0.05$ level. That does not mean that you should throw the data out - you should instead use confidence intervals. And just because the confidence interval goes through the 1.0 axis as in this next slide (49) doesn't mean that you throw the data out either. Data of this sort can reveal things about the things that you ought to do next. If this confidence interval here was say the data for people going on to a suicidal act after they have been exposed to one of the antidepressants, you might conclude from this that what we need are clinical trials with a greater degree of power. Or you might conclude, well actually no we can get by with much smaller clinical trials and just put into the trials rating scales

sensitive to the emergence of suicidal ideation. You shouldn't stop thinking about the data just because the data doesn't reach the p< 0.05 point or breaches the axis through 1.0.

Let me give you an update on Fisher's and Neyman's positions. This is Paul Anthony speaking on behalf of PhRMA in Princeton in June (50). What Mr Anthony is concerned about is that if people don't stick to the rigid p<0.05 level of significance that there is a big risk that we may go on to slur good drugs and good pharmaceutical companies. It's a bit like the person next door to you saying that I think he's a child molester. And even though it's ultimately proven that you aren't, the slur sticks. We in the industry don't really want this to happen, so we want you guys to get back to doing good science. You want to do clinical trials that will give you outcomes that are clearly statistically significant or not.

The problem here of course is this leaves us all hostage to power in the statistical sense (51). In this case if pharmaceutical companies simply don't do big enough trials, you are not going to find that the problem is statistically significant at the 0.05 level. And if you don't find that, you are being invited to say the problem isn't real. And if the problem isn't real, you overestimate the good that drugs do. And if you overestimate the good the drugs do, you over use them. And if you overuse the drugs, you kill or injure people if in fact the problem is real.

The opposite point of view, the Neymann point of view, is presented here by David Graham of FDA who says 'you know, FDA's position on this is unreal. It's a bit like stating that if you've got a 100 chambered gun, FDA are saying that the gun isn't loaded until all but 5 of the gun barrels are filled. One of the interesting things here to think about is how many actual bullets would need there be in this 100-chambered gun before you'd be prepared to take the risk with it? (52).

This comes back to the real world in which both clinicians and patients have to deal with uncertainty the whole time - we're in the business of best bets. And if you recall what I actually said earlier was the best bet is the point estimate. In contrast in terms of the SSRI and suicide data we're in a world where we're being invited to say that there isn't any risk even when there is very clear clinical trial data that points to a risk.

Now the next slide (53) shows work that I have been involved in, so you'll have to take great care here, because if you recall Dr Coyne had concerns about me cooking the data. This slide shows the results from analysing all of the clinical trials in which SSRIs have been involved. Something like 680 odd trials from 87,000 patients. (You'll be pleased to know that I'm not the only author on this article so perhaps the risks aren't too great). What you see here is the point estimate or the odds ratio for an increased or decreased risk of suicidal acts from the SSRI group of drug from 1982 when the first clinical trials get reported through to the present. The first clinical trials reported were on 2 drugs we had in Europe that were never licensed in America. Now over in Europe, we had the Medici and we know how to hide corpses just as well as you do so that's perhaps why the data looks the way it does here for the early years – seemingly indicating a reduced risk on treatment compared to placebo.

Now the way the SSRI story has played out since has been that you had case reports, which were picked up by the media in 1990 - 15 years ago now - where on the basis of 6 people who had gone on to Prozac, people up in Harvard, where they think all sorts of odd things, claimed that these 6 people had become suicidal taking Prozac. What Eli Lilly did in response was to look at all of the data from the clinical trials that were done which you've just seen and concluded that there is no evidence

of an increase in risk that would be consistent with the claims being made. And how the story has played since has been that you have case reports on one hand versus the science on the other hand.

But in fact, if you go back and look at the science from 1988, 2 years before the first report that anybody had become suicidal taking Prozac, you see that if all the RCT data is added up on people who had gone on to a suicidal act on active treatment versus placebo that there was the increased risk on treatment. A 2.9 times greater risk and that actually the scientific data was consistent with that increase being 18.9 times greater. This increased risk remains clearly there in the scientific data ever since. So I put it to you that it should have been obvious to us right from the start that there was an increased risk. It is clearly ridiculous to wait until the error bar here doesn't include 1.0 before you conclude that it has been shown that there is an increase in risk.

But that is exactly what the regulator is still doing. The next slide (54) shows the data on completed suicides in adults from the MHRA working paper put out at the end of 2004. This is available to anyone but I warn you its deeply misleading. I'll be happy to explain to anyone just where the risk of being mislead or deceived lies. But if you work through the data as I do, you should find the increased risk of completed suicides for these drugs (which does not include Prozac because the makers of Prozac did not give all of the clinical trial data to the regulator in the UK) that is 2.66 times greater on active treatment versus placebo. The confidence interval indicates that something like 92 or 93 barrels of our 100 chambered gun have bullets in them. Yet based on data of this sort the FDA and MHRA are still saying that there isn't even a signal of an increase in risk to adults from these drugs.

Problem 3

Let me move on to another problem. Let's assume people are handling the data the way I think it should be handled and conceding that it shows an increase in risk. In point of fact, FDA cracked a year ago in the case of children with the help of people here from Colombia and said 'well, actually the children's data now does look like it's statistically significant and based on this we have to do something'. And they began to warn (55).

But even after that does the American Psychiatric Association pay any heed to either Colombia or FDA? Well the answer is 'No'. APA, as of this year, is still saying that there is no evidence that antidepressants increase the risk of suicide (56). They even go further. Here you see a spokesperson for APA saying 'you know, what could be going on here is just that children who go on to a suicidal act on one of these drugs are happier to actually tell people afterwards that this is what they've done'. And from David Fassler's point of view this is a good thing - you can intervene to help the child who tells you this (57). I'm not sure what you can do for the adults who go on to complete suicide.

APA however still believes that antidepressants save lives (58). But, there isn't any evidence that any of the antidepressants reduce the risk of people going on to commit suicide. So why do the APA say this? Well they may be saying it on the basis that if the antidepressants work and people keep on having them that perhaps you therefore reduce the risk of suicide.

Now when it comes to looking at whether the antidepressants work or not, this is where having statistical significance set at $p < 0.05$ comes into play (59). On this issue its a good idea to insist that companies produce data indicating unequivocally or beyond any reasonable doubt that the drug does something because we don't

want drugs out there that don't work. We really do want to hold pharmaceutical companies to a relatively high standard to prove that their drugs work, and indeed companies can gain by this as it keeps inferior products off the market.

Or do we really wish this? Well, look at this quote from Paul Leber from the approval process for Zoloft (60). Pharmaceutical companies don't have to prove that their drugs work in all of the trials that they do. They could conceivably do 100 clinical trials and the drug might only be shown to work in 2 of them. But as Dr Leber indicates here on the basis of the law as it stands at the moment Pfizer would actually get a license for Zoloft based on data like that. In fact of the first 16 Zoloft trials that were done, the drug can only clearly be shown to work in 5. So in roughly 1 in 3 of the trials the drug worked in, in the other 2 out of the 3 trials there isn't really good evidence that the drug beats placebo. Where does this leave us?

I want to tackle this problem from a slightly different angle. What happens when one of the antidepressants 'works' is that you find that in a group of people who get put on the drug compared to a group of people taking placebo that 50% of those put on the drug show some improvement on a rating scale compared with 40% on placebo, if you only look at published trials, or 45% on placebo if you have access to all trials (61).

In this context the idea that an antidepressant 'works' does not mean that people do not go onto suicide or that they get back to work, it means some change has been produced on a rating scale. But let's assume this really means working. Based on the idea that we want things that work, the culture and the money in health care follows evidence of this sort. It follows the blue bar and not the red bar. We don't do the red things we do the blue things. But let me put it to you that RCTs in this case don't prove that the pill works in this sense. What they do is allow us to quantify the contribution that the pill may make to the therapeutic response.

If you agree with me on this point, then all of a sudden the figures can be recast as they are in the next slide (62), where 80% of the therapeutic response has come from a variety of factors that have to do with the natural history of the problem that you have got. Which is that on average most people who are depressed have a condition that lasts 12 to 16 weeks. They'll usually take a month or two to seek help at all, so by the time they come to me they're almost half way through the problem that they've got and there's a good proportion of them that are just going to get well anyway. And the simple fact that they bring themselves for help means a good deal, as may the fact that I problem solve things with them about their work and their marriage and perhaps offer them some reasonable advice as regards the foods they are eating or their level of fitness and make recommendations about jogging or whatever. These things all make a difference.

But what we can't do is we can't tease these out and quantify the contribution that each of these separate components to the placebo component of the therapeutic response make, whereas we can quantify the contribution the pill makes. But if the money and the culture in health care is going to follow the evidence, surely it ought to follow the red bar which is worth 4 times more than the blue bar. We should be ensuring that people get good advice about the things that they're eating and the lifestyle that they have and we should problem solve things with them and they should be told that in the natural history of the problem they've got is that it will often get well of it's own accord anyway. And yes they can have pills but there are risks to the pills and if you are actually going to have them you need to know what these risks are.

In fact we treat the data from the antidepressant trials as though it looks like the data in the next slide (63), which is the kind of data you might get from a trial giving penicillin to GPI, tertiary syphilis, where the placebo response will be almost nil but the response to the drug will be awfully good. All of us would probably say in this case 'yes, the money in the culture in health care should follow the blue bar here'. We should screen. We should teen screen or whatever, if the evidence looks like this. Who cares about the bedside approach of the clinician who gives the treatment, the key thing is to ensure that the person who has the illness gets the treatment. But that isn't the kind of data we have for the antidepressants. Given the kind of data we have, there is a real issue about whether we should screen for depression in teenagers and if we do screen what do we do then?

Coda

To conclude. Thirty years ago, if you were ill your clinician would say to you you should have this or that and patients weren't encouraged to ask any questions. That's changed now partly as a consequence of the fact that we all know there is RCT evidence for treatments that supposedly work or not. We are not at the mercy of authoritative but idiosyncratic clinical judgements the way we once were. Patients will now come and ask what the evidence shows?

In terms of the treatments we have in psychiatry, however, the evidence does not in fact show what we should do. It just isn't clear what the right way is to treat people who've got schizophrenia or people who are depressed. The trials show limited short term effects but nothing more, and we do not know whether we are going to be better in the longer run opting for the limited benefits.

A further problem is perhaps apparent now after this talk, which is that even in the case of those limited benefits, the evidence is less believable than once thought.

This limited and debatably believable evidence gives rise to a further problem. Thirty odd years ago, in the case of pills like the antidepressants, a paternalistic physician working in primary care, usually male, would not have used new drugs as quickly as they might now do. If they put any of you on a new antidepressant and you came back after a week and had turned blue and had grown feathers, they'd say to you 'well hey you know the only thing we've done differently here recently is you've just been put on this new pill so maybe we should halt it'. Now we didn't care for that world because the variation from physician to physician was huge and we don't like variation like this. Part of the point in moving into an evidence based world has been to cut down on variation.

So, if you go along to a primary care physician in the UK now it will usually be a younger physician, who will stick to the evidence, and who will usually be female as opposed to male - because only the brightest people get into medicine these days - and she will put you on one of the newer drugs first. She won't be slow to move from the older drugs that she may know well to newer drugs – because this is what the evidence and expert assessment of the evidence point to. After a week on this pill, if you come back and you've turned blue and grown feathers, what does she say? Well she'll check the computer screen and look at the evidence base and say there isn't any evidence that this drug will turn you blue or that you'll grow feathers, but you must be awfully anxious about what's happened so why don't you double the dose.

We are in a world where health care is being segmented and in the case of illnesses like hypertension, where screening is easy, and where the treatment can be readily standardised through the use of protocols and guidelines, healthcare delivery is being handed over to nursing staff. The same can happen for depression. Giving a

Hamilton rating scale is easy and if antidepressants save lives as opposed to physicians saving lives, then physicians need to be very careful they are not going to be replaced.

I happen to think that other healthcare professionals could be just as good as physicians, and perhaps more sensitive to the broader picture. But in our brave new evidence based world, nursing staff and clinical psychologists will have their freedom for manoeuvre restricted just as completely as physicians. They too will be trapped into having to give what the guidelines say. So we all have a problem.

This leads to my last slides – a quote from Chuang Tzu from 323 BC (64, 65). He says its obvious when you're faced with a problem of a thief who pinches purses, robs and things like that, that what you need to do is you need to fasten up your property with ropes, lock it up with locks, bolt with bolts. This is elementary good sense. Except when a strong thief comes along. He'll pick up the whole lot, put it on his back and go on his way with only one fear - that your ropes and locks and bolts will give way. We have tried to control pharmaceutical companies with evidence-based medicine and their biggest fear at the moment is that we will stop believing in evidence-based medicine.