NEARLY EVERYBODY knows what a placebo is; the concept is fairly easy to grasp. The placebo effect, however, is a different story and will require some explanation.

Placebos, by definition, have no effect. A placebo may take any form, from a sugar pill to an elaborate procedure; as long it can not and does not have any specific therapeutic effect, it is a placebo. A lack of a specific therapeutic effect means that it doesn’t do anything to directly physically change the person in any way that might provide some improvement in their underlying condition. The killing of disease-causing bacteria by antibiotics is a specific therapeutic effect, as is the lowering of blood sugar by insulin. These are simple to understand, but, as
you will find out later, sometimes we guess or make up theoretical specific therapeutic effects in order to explain the perceived response to a drug. We need to test those theoretical effects by comparing the treatment to a placebo. This concept is simple, and the process is often necessary in order to demonstrate true effectiveness. Despite the simplicity and apparent necessity of placebo tests, there is still a reluctance to put medical treatments to the placebo test, despite how often treatments actually fail that test.

Placebos can take the form of injections, operations, and other physical treatments, as long as they don’t have any specific therapeutic effect. Active placebos are also used, but they are still placebos. Active placebos have some noticeable effect on the patient, but as long as the ‘active’ component doesn’t directly affect the underlying condition, it is still a placebo. For example, a placebo pill might be used that creates tingling of the tongue, because the drug it is being tested against has such an effect. This helps with ‘blinding’ the patient, so that they cannot tell whether or not they had the placebo, but, importantly, active placebos still don’t have a direct effect on the underlying condition.

So if a placebo is inert and has no effect on the condition being treated, how then do we have something called the ‘placebo effect’? The answer lies in the fact
that what is happening to our bodies on a purely physical level doesn’t correlate with our perception – how we feel. We know from some very interesting observations and experiments that our perception of our own pain, well-being, health and happiness is poorly correlated to the objective state of our bodies. Sometimes just having some reassurance and knowing that someone is looking out for us makes us feel better, even if the pill that the person just gave us has no active ingredients. But the problem goes deeper than that. People can be quite convinced that symptoms attributed to the specific condition being treated have improved after they have been given treatment that we know to be ineffective. There are many reasons why we might perceive ourselves to be better, which will be covered later.

The answer then, to the problem of the placebo effect lies in the difference between what is actually happening to us, and the way we perceive ourselves. It is the difference between the objective and the subjective; between any specific therapeutic effect and the perceived therapeutic effect.

However, it is not that black and white. A treatment may have some specific therapeutic effect (it may actually work on the underlying condition and physically change you) while simultaneously having some placebo effect in addition to the real effect. This added effect makes the
overall perceived effect (from the patient’s point of view) greater than the real or specific effect.

By doing some adding and subtracting, you can see that the placebo effect is basically the perceived therapeutic effect, minus any specific therapeutic effect (even if there is no specific effect). The placebo effect is the extra benefit explained by the perception of improvement. To put it mathematically:

\[
\text{specific therapeutic effect} + \text{placebo effect} = \text{perceived therapeutic effect.}
\]

So if a drug (or any other form of treatment) is a pure placebo (no specific therapeutic effect), then the entire perceived effect will be equal to the placebo effect. The formula for this would be:

\[
\text{zero (specific effect)} + \text{placebo effect} = \text{perceived therapeutic effect}
\]

which is the same as saying:

\[
\text{perceived therapeutic effect} = \text{placebo effect.}
\]

Similarly, if a drug has a real (specific) effect on the patient, and is delivered (say) while the patient is unconscious or without their knowledge (where there cannot be a placebo effect), then the specific effect of the drug and its perceived effect (once they wake up) will be the same. In this last case, the formula would be:
specific therapeutic effect + zero (placebo effect) = perceived therapeutic effect

or:

perceived therapeutic effect = specific therapeutic effect.

This rule applies to any therapy: any treatment, whether it be a pill, a lotion, an injection, a physical treatment, psychological treatment, or an operation. And teasing out the placebo effect – the difference between the specific (real, physical, pathological) effect and the perceived effect of any given therapy – is what we will be trying to achieve, because that is how we separate what really works from what we think works.

So what are some of the placebos out there? Obviously, things that are labelled ‘placebo’ are placebos. These are drugs or devices that have been carefully manufactured so that they have no meaningful physical interaction with the body, and everybody knows that they are placebos. In reality, though, just about any treatment can have a placebo effect. In fact, treatments that are not known to be placebos can have a much stronger placebo effect. But this makes sense: if you are given a tasteless pill and told that it is a placebo, with no possibility of having any effect on you, you are not likely to feel much better after taking it.
Homeopathy provides a good example of a pure placebo. Homeopathy uses ingredients that are given in such massive dilutions that it is highly unlikely that any molecules of the original substance remain. Scientifically, the final product can have no active ingredients. Yet it is clear, if only from the fact that homeopathy is still in common use after so many years, that it has a perceived effect. It might not work for everyone, but there are people out there who swear by it – not just the practitioners. Those people think it works, even though it does not have any specific therapeutic effect.

You will see from the following chapters that the same thing applies to many surgical procedures. The difference is that the proponents of any particular operation usually have a scientific explanation to justify the treatment – one that can’t be ruled out as easily as homeopathy. However, you will also learn that a scientifically plausible explanation is no guarantee that the treatment works; it only gives an explanation as to how it might work.

Why, then, do people think that their condition has improved when nothing has happened to them, physically? There are a number of reasons, and the answer in any individual case may be a combination of these.

If we look at it logically, there are really only three reasons why someone’s condition would improve after receiving treatment that had no specific effect:
The placebo effect

1. they did get better, but it wasn’t due to the treatment
2. they did not get better, we just think they did, or
3. they did not get better, they just think they did.

For explanation number 1, that they got better anyway, there are several explanations. The most obvious one is that most conditions are self-limiting. This is because our bodies have evolved over a pretty long time to handle most conditions that might harm us – that’s what evolution is for. It is surprising how much people underestimate the natural healing processes of the human body. My pets live long and happy lives with virtually no veterinarian involvement, yet humans apparently need constant maintenance? I don’t think so. This explanation refers to the ‘natural history’ of the condition – what happens when it is left alone. Most people who have taken antibiotics for their stubborn cold and then improved fall into this category. As Voltaire said: ‘the art of medicine consists of amusing the patient while nature cures the disease.’

For doctors, we often come to appreciate the often favourable natural history of many conditions later in our careers, after seeing what happens to the few patients who missed out on treatment or refused; we are often surprised at how well they recover.

One of the early proponents of evidence-based
Surgery, the Ultimate Placebo

medicine, who was critical of accepting treatments on face value without proper scientific trials, was Archibald Cochrane, after whom the Cochrane Collaboration – the ‘mothership’ of evidence-based medicine – is named.

As a prisoner of war in Germany during World War II, Cochrane was the medical officer overseeing 20,000 prisoners of war, all of whom had diarrhoea, with frequent epidemics of typhoid, diphtheria and other infections. With no medicine (except for aspirin and antacids), he expected hundreds to die. He remembers his shock at the reply from one of his captors to his request for doctors: ‘Ärzte sind überflüssig’ (doctors are superfluous). In his six months at the camp, only four prisoners died, three as a result of being shot by their captors.

On returning to Britain, Cochrane began to question many of the (then) standard medical practices, practices that were later shown to be ineffective (like bed rest after a heart attack). His observation that so much medical treatment relied on ‘amusing the patient while nature cures the disease’ led him to call for more randomised trials (which were only then starting to be used in medicine) to properly (scientifically) test ‘standard’ medical treatments.

Another reason why people get better without any treatment is a phenomenon known as ‘regression to the mean’. For example, on average (the ‘mean’), a person
The placebo effect

will have a certain number of episodes of back pain in their life, of varying severity. If you select a group of patients that currently have back pain (out of a population of average people) to test your new treatment, it is very likely that many of them will not have back pain in (say) six weeks, because they will ‘regress to the mean’. Back pain fluctuates, and selecting patients who all currently have back pain is setting yourself up to show good results with whatever you do to them because it is very unlikely that 100 per cent of them will still have back pain when you examine them later. Similarly, if you select people based on having high blood pressure at the time of selection, when you test them again later (after your treatment) their blood pressure will (on average) be closer to the mean (lower than before) because blood pressure varies each time you take it. As you can see, this can make any treatment look pretty good, and this is a particular problem in clinical research.

Daniel Kahneman refers to this phenomenon in his book *Thinking Fast and Slow*, in which he describes a flight instructor who considered his method of berating poorly performing pilots after an exercise to be particularly effective because they perform slightly better next time. He found this to be more effective than praising well performing pilots, because they often did worse next time. All this occurs because in a group of people
(or pilots) who are repeatedly tested, there is a natural, random variation so that the performance of these people will not always be the same. If you pick those at the end of the spectrum at one point in time, they are unlikely to all remain at the end of the spectrum (among the best or the worst) on later testing. To pick those at the extremes and then attribute their fall back towards the average to your intervention is cheating – it is not proof of cause and effect. This is the same phenomenon as the often quoted underperformance of sports stars after appearing on the cover of *Sports Illustrated* (when they are at their peak). This is not causation – it is regression to the mean.

Regression to the mean is one of the many reasons that scientific studies of treatments should always include a ‘control’ arm – a group of patients treated exactly the same in every way except that they did not have the test treatment. Patients in a control arm can be given a placebo, or simply not given the treatment, but the presence of a control arm is the most important thing in any test of a treatment. For years, doctors have been watching patients get better and attributing the improvement to their treatment. Only the use of a control group can properly test (in the scientific sense) the effectiveness of a treatment, and placebos make the best ‘controls’.

Another reason why people improve after receiving inactive treatment is because they are receiving other
The placebo effect

treatment at the same time: concomitant treatment. This is overlooked surprisingly often in our rush to attribute cause-and-effect to any association we see.

In one scientific study of an expensive, high-tech, new, genetically engineered drug, BMP (bone morpho-genetic protein – something designed to make bone) was compared to old-fashioned bone grafting (using bone taken from the patient’s own pelvis) in patients with unhealed leg (tibia, or shinbone) fractures. The results showed that both treatments, when placed between the unhealed ends of the bone, worked equally well. This was great for the manufacturers of BMP: it looked like BMP could replace bone grafting, which is painful and time-consuming. But patients in both groups also received surgery to refix their fracture; they had their tibia reamed out and had a metal rod inserted to stabilise the bone ends, a recognised treatment for unhealed fractures. This is concomitant treatment, and without a group of patients who did not receive any other intervention (or better still, a group of patients that received placebo BMP), we don’t know if either treatment (the BMP or the graft) made any difference to the healing rate – it could have all been due to the concomitant treatment. Sneaky? I think so, but most people don’t look into it that much: they just see that a scientific experiment was done, it was reported in a journal, and the results of the
BMP and the bone graft treatments were similar. It was certainly good enough for the US FDA (Food and Drug Administration), who approved the drug. The drug was a massive hit for a while, but has since fallen out of favour.

The example above highlights the problem of comparative effectiveness research – comparing one treatment to another. When two treatments are shown to be equally effective, nobody considers the possibility that the two treatments are equally ineffective and that the benefit was due to something else, like concomitant treatment. Or time. I expand on this in Chapter 7.

When testing new treatments, the best way to account for things like the natural history of the disease, regression to the mean, concomitant treatment and any other reason people get better is to test it side by side against a placebo. If the only difference between your treatment group and the control group is your treatment, then any improvement in the treatment group over the control group is likely to be a result of the treatment. This difference will represent the specific therapeutic effect.

Now, to address explanation number 2: where the patients did not get better, but we (the observer or doctor) thought they did.

First of all, patients generally want to please their doctor, so if the doctor asks how the patient feels after an
The placebo effect

operation, the patient is more likely to say that they are a little better (even if they are not) than if an independent assessor asks them.

Improvement can also be manufactured simply by measuring or reporting things differently. We can measure outcomes in many different ways, and then choose to only use (report) the combination that showed improvement. By random chance, even if there was no real change in a patient’s condition, when using multiple scoring tools, surveys and measurements, some would go up and some would go down, because these things aren’t perfectly stable. Just measuring a quality of life score a few days apart can change the score up or down, even if nothing has happened to the patient. Like blood pressure, it is not exactly the same all the time; it fluctuates. Cherry-picking the outcomes that look good is called selective outcome reporting and is one way of generating a biased report.

Selective outcome reporting is common in medical research, but I believe that it is not normally deliberate; it is just human nature to see what you want to see; to sift through the data until you see what you are expecting to find, and then to use the outcome that fits your idea of what the results should be. If you believe your treatment works, you are more likely to believe the favourable results and discard the others, and then find a reason to explain
why the unfavourable results might be ‘wrong’. This is called confirmation bias, which is also called ‘believing is seeing’.

Another reason why we might think the patient got better when they didn’t is because we asked the wrong questions. Doctors and patients use different criteria to judge the success of surgery, so outside the confines of a structured research study with validated, predefined outcome tools, what do we use to judge success? One common measure of surgical success, used in many specialties, is the ‘revision rate’. A revision is a second operation – a re-operation, usually required because the first one failed at some stage. In fact, many national surgical registries use revision as their main outcome in order to monitor the quality of surgery performed, often because it is a ‘hard’, objective outcome that can be reliably measured. No one would argue that this isn’t important, and this tool has been used successfully to improve surgical quality and safety, but for some, the revision rate equates to the failure rate of the surgery. Similarly, some may consider a procedure that doesn’t require revision to be a success.

Now, let’s take the patient’s perspective: other factors are also important in rating the ‘success’ of surgery, for example the degree of pain relief and improvement in function. The problem with using the revision rate in
The placebo effect

isolation as a measure of success is that it does not give
the whole picture, from the patient's perspective.

This can lead to anomalies, for example where the
success of knee replacement surgery, based on the need
for revision surgery, is over 95 per cent after five years,
and around 99 per cent after one year. Yet the rates of
ongoing pain and patient dissatisfaction one to two years
after this procedure are around 10 to 25 per cent. This
can lead to a mismatch in the ratings of success between
surgeons or administrators and patients.

Similarly, we previously surveyed patients with limb
fractures and their surgeons to measure their satisfaction
with the outcome of treatment. Overall, the surgeons
rated the success as much higher than the patients, and
the surgeons' ratings depended mainly on whether or not
the bone healed. Patient satisfaction, on the other hand,
was unrelated to fracture healing or the need for further
surgery, and depended on factors such as whether or not
they were able to return to work.

In general, surgeons tend to dwell on the technical
aspects of surgery (what the X-rays looks like, the
amount of tumour removed, the restoration of blood flow
and the avoidance of complications, for example). These
may be relevant, but they are not necessarily aligned
with a measurable clinical benefit for the patient (relief
of pain, prolonged survival, etc.). In spine surgery, for
researcher concentrating on the favourable findings and discounting the information that doesn’t fit with their expectations. This is supported by research around placebos where greater expectation of success leads to greater success – but this is the topic of the next chapter.

It should be noted that the placebo effect can be strong enough that it can override harmful specific effects of treatment. All of the factors contributing to the perceived benefit of treatment may be strong enough to support treatment that actually causes harm. This makes sense, because the overall perceived effect is really just the addition of any direct (specific) effect and the indirect (placebo) effect. If the placebo effect is sufficiently positive, it can more than make up for small or moderate negative specific effects, so that the sum of both effects is still positive.

Certainly, surgery has some negative effects. It causes pain and exposes patients to other risks and discomforts. If the positive direct effects of a truly effective operation can lead patients to feel an overall benefit, then so can the indirect, placebo effects.

In this chapter we have learned that a placebo has no specific, direct effect on the underlying condition being treated, although it can have effects on the patient’s perception. Treatments that do have a specific (true) effect on the underlying condition can also have a placebo effect.
The placebo effect however, is the perceived or measured effect beyond that of any specific effect.

To recap, there are several reasons for a placebo effect: the patient may have improved anyway (natural history, regression to the mean, concomitant treatment, etc.) or the improvement was only perceived by the patient (through psychological or indirect physiological mechanisms) or perceived by the observer (biased observer, selective reporting, measurement error, etc.).

I will persist with this definition of the placebo effect (any effect that contributes to the total perceived effect, beyond the specific effect of the treatment), even though I have described other reasons (that is, reasons other than the placebo) for the extra effect seen with placebos, such as the natural history of the condition. I have chosen to do so because it is easy to understand, and because that is how many people understand it. I accept that the excess benefit seen is not directly due to the placebo, and this is where confusion can arise between the placebo (which has no effect), and the placebo effect. This is why this excess effect (beyond the true effect) has been labelled a 'meaning response' rather than a placebo effect. For a full explanation of this concept, I refer you to Daniel Moerman’s book Meaning, Medicine and the Placebo Effect.

One problem with using the term ‘placebo effect’ to describe an effect that is not necessarily due to the
placebo treatment is that treatments are often justified based on this ‘effect’. This is addressed in Chapter 8.

Now that we have a basic understanding of the placebo effect, let’s see how this fits with the science of medicine.