Scientific U-turns: eight occasions when science changed its mind

Michael Hal Sosabowski and Paul R. Gard

ABSTRACT The Scientific Method is the series of processes by which hypotheses, ideas and theories are shown to be true beyond a reasonable scientific doubt. Most science 'fact' is expressed in terms of probabilities rather than certainties. Thus, by means of statistical calculations, researchers aim to determine whether an observed association between two events or characteristics may have occurred by chance (coincidence), whether they frequently occur together (correlation) or whether they occur together because one causes the other (causative relationship). In this article we review the Scientific Method and consider the statistical tests that are applied. We then focus on the occasions when science changes its mind and review eight such occurrences.

The role of science can be argued as providing new knowledge that consists of facts, which in turn come from evidence generated when testing hypotheses. But during the quest for new knowledge it sometimes happens that existing knowledge, or well-founded theories that fall short of being absolute fact but nevertheless have much evidence to bear them out, needs correcting. This is when science does a U-turn. This is one of the defined rules that any competent scientist will hold dear, and yet it goes against human nature. This is because of the normal need to fulfil one's ingrained self-serving confirmation bias – the tendency to search for evidence to confirm pre-existing beliefs, especially those in which the individual has an investment, emotional or otherwise:

Faced with the choice between changing one's mind and proving that there is no need to do so, almost everyone gets busy on the proof. (John Kenneth Galbraith; https://en.wikiquote.org/wiki/John_Kenneth_Galbraith)

And so the ability to change one's mind in the light of new evidence is one that develops during the career of the aspiring scientist, and those that have reached the dizziest heights of our trade are often the most disarmingly honest:

In science it often happens that scientists say, 'You know that's a really good argument; my position is mistaken,*' and then they actually change their minds and you never hear that old view from them again. . . . I cannot recall the last time something like that has happened in politics or religion.* (Carl Sagan, 1987 CSICOP address; https://en.wikiquote.org/wiki/Carl_Sagan)

When my information changes, I alter my conclusions. What do you do, sir? (John Maynard Keynes; https://en.wikiquote.org/wiki/John_Maynard_Keynes)

This skill attribute is probably more finely developed in those who have it woven into their professional ethical personae and it may be less well understood by the layperson. Society often cannot understand it when science changes its mind, and that which hitherto was stated as fact becomes fiction and vice versa.

During the 2016 discussion concerning childhood obesity, Public Health England had to step in and describe as 'irresponsible' the advice given by the National Obesity Forum that eating fat could help combat obesity and type 2 diabetes (BBC News, 2016). The sight of a public body and a charity engaged in two diametrically opposed scientific views, each apparently having evidence to support their case, appeared to add to the pre-existing confusion and wavering confidence in science, causing the then-MP for Leicester East and Chair of the Home Affairs Select Committee, Keith Vaz, to opine:

Ordinary people are now caught in a whirlwind of conflicting advice at a time when they

desperately need clarity, consistency and straight talk. Quite simply they don't know where to turn. (Keith Vaz; https://hansard.parliament.uk/Commons/2016-05-23/debates/1605231000001/DietaryAdviceAndChildhoodObesityStrategy)

To some degree, Mr Vaz had a point: to anyone unfamiliar with the Scientific Method (Figure 1), the periodic occasions when science changes its stance appear to be confidence-sapping and serve to undermine the public's trust in science. Once, science claimed, it was healthy to drink a glass of red wine a day; then science later recanted and said that it wasn't after all. Understanding of the Scientific Method as well as the knowledge it produces may be a key to offering the public the reassurance it seeks, that science really does know what it's talking about.

The Scientific Method, certainty and probability

Scientific research is not presented in terms of certainties but in terms of probability. Thus, by means of mathematical calculations, researchers aim to determine whether an observed association between two events or characteristics may have occurred by chance (coincidence), whether they frequently occur together (correlation) or whether they occur together because one causes the other (causative relationship).

Let us consider two human conditions. The first condition is schizophrenia, a severe psychiatric condition in which people affected frequently report delusions (for example that their neighbours are spying on them) and hallucinations (for example hearing a voice or voices). It affects up to about 13 per 1000 individuals during

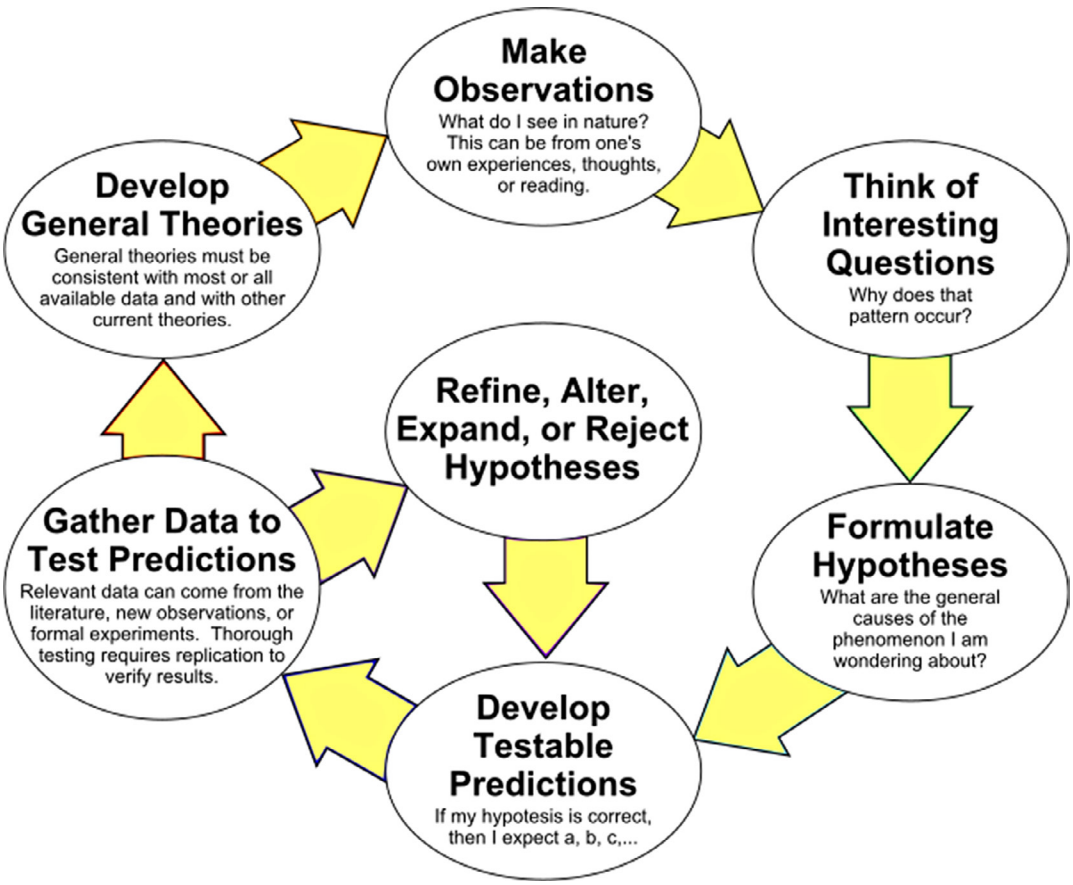

Figure 1 The Scientific Method as an ongoing process; reproduced with permission from ArchonMagnus under a Creative Commons CC BY-SA 4.0 licence

their lifetime (Perälä *et al*., 2007). The second condition is tonic–clonic (or grand mal) epilepsy, in which people affected experience seizures during which, in the tonic phase, the body becomes entirely rigid and, in the clonic phase, there is uncontrolled jerking. These may last a few seconds or several minutes. Approximately 5 per 1000 of the population experience tonic–clonic seizures (Joint Epilepsy Council, 2011). Mathematically, the chance of a single individual being affected by *both* schizophrenia and tonic–clonic seizures is 6.5 per 100 000 individuals (or 0.065 per 1000; 0.5% × 1.3%). Early in the last century, this statistic led doctors to believe that patients with tonic–clonic epilepsy were protected against schizophrenia because they had never seen a patient with both. This belief was implemented by attempting to treat schizophrenia by inducing epilepsy, initially by injections of chemicals such as camphor and later by passing an electric current across the head (electroconvulsive therapy, ECT). Today, ECT is still used, but usually for the treatment of depression-like illnesses that are quite different from schizophrenia; rarely is ECT used for the treatment of schizophrenia. The observed relationship between epilepsy and schizophrenia was not preventive but purely 'coincidental'. If the probability of *ever* seeing a patient with both epilepsy and schizophrenia is 0.0065%, the probability of *never* seeing a patient with both is 99.9935%. This is an example of an observation and speculated relationship that turned out to be wrong. Today, this fact would be revealed quickly by considering the relative prevalences of the two conditions. The doctors' belief, however, was not without value as it did lead to the development of a treatment for another illness.

Let us consider another, non-medical, example. If you were to take a shuffled deck of 52 playing cards and deal four cards off the top of the deck, you would probably suspect that the deck was 'rigged' if the first four cards were all aces. You would know that, mathematically, such an event is possible, but in reality it is highly unlikely. The chance of the first four cards dealt from a shuffled deck being the four aces is 1 in 270 725 ($\frac{4}{52} \times \frac{3}{51} \times \frac{2}{50} \times \frac{1}{49}$) or 0.00037%. It could happen by chance, but it is highly unlikely. You would conclude that the deck of cards was fixed. In the case of the schizophrenia and epilepsy belief, the probability of the lack of association of the two conditions occurring by chance was 99.99% – it should never have been questioned. In most scientific research, if the probability of an observed event occurring by chance is less than 5%, it is concluded that the event has not occurred randomly but that something has been done to make it happen. In scientific terminology, a results is termed 'significant' if the probability of the observation occurring by chance is less than 5%.

To put this into context, when dealing from a shuffled deck of cards, the probability that the first card will be red (diamond or heart suit) is 26/52 (50%); the probability that the second card is also red is 25/51 (49%) and that the third card is also red is 24/50 (48%). Thus, dealing three cards of the same colour 'off the top' could occur on 11.8% (0.5 × 0.49 × 0.48) of occasions. The probability of dealing four cards of the same colour off the top would be 5.5%. By scientific convention, if someone deals four cards of the same colour off the top of a shuffled deck of cards, you must accept that this is a statistical possibility and therefore not question the honesty nor integrity of the dealer. But if they deal five cards of the same colour off the top (the probability is 2.53%), you might choose to challenge them; in scientific terms, you would state that the observed pattern of dealt cards is unlikely to have occurred by chance, and therefore the deal has been influenced by the dealer. When making your challenge, you would have to accept that the sequence of cards could have been dealt, randomly, by an honest dealer, on 2.53% of deals, i.e. about 1 in 40 deals. There is a 1 in 40 chance that your challenge is incorrect and that, rather than receiving compensation from a crooked dealer, you lose the case and have to pay them compensation for defamation of character.

In scientific research, the 5% probability hurdle is used because it raises suspicions that the observed results have not occurred by chance at a relatively low threshold (1 in 20); it encourages further research. If the observed results have actually occurred by chance, the probability of a second experiment producing the same significant results, again by chance, is 5% times 5%, i.e. 0.25%. This highlights the importance of repeating research in an attempt to replicate results. Typically, researchers present their findings in terms of probability, not certainty. They would not claim that a card dealer was crooked; they would claim that the probability

(p) of the card dealer being honest is 0.0253 (or 2.53%), and therefore the findings suggest crookedness. The smaller the p value, the greater the likelihood that the observations are real, true or significant.

The concept of probability versus certainty has led to several incorrect scientific conclusions throughout history and, importantly, the popular press have frequently promulgated incorrect theories without appreciating the statistical assumptions underlying those theories.

Cases

1. The cause of cholera

Cholera was at one time the cause of death of millions of people, with pandemics routinely sweeping continents. Before 1854, it was believed to be spread via 'malodorous airs' (https://en.wikipedia.org/wiki/John_Snow). By charting specific cases within a narrowly defined area of London, John Snow developed a theory that it was transmitted by infected water. He removed the handle from a water pump in Broad Street, an action that is generally accepted to have halted the outbreak and at the same time to have proven how the disease is transmitted. It transpired that the pump had been situated too close to a disused cesspit into which a nappy from a cholera-infected baby had been thrown. Interestingly, the scientific community later rejected Snow's assertions, partly because society wasn't prepared to accept the distasteful faecal–oral route proposed. Only when William Farr, hitherto one of Snow's antagonists, scrutinised Snow's dot maps and other data, and his own gathered data, was Snow's hypothesis vindicated.

2. Vitamin C and the common cold

The apocryphal link between vitamin C and resistance to the common cold is one of the science folklores that seems still to be propagated within society in spite of science's best efforts. Somewhat surprisingly, it was originally proposed by Linus Pauling (1976), the only scientist to win two unshared Nobel prizes. He is largely the cause of the mistaken belief that vitamin C affects one's propensity to catch the common cold. He also suggested that large doses of vitamin C caused the majority of 100 terminal cancer patients to survive 3–4 times as long as would otherwise have done. Subsequent double-blind studies of both of these claims showed them to be false and yet the myth that vitamin C is a cure for the cold is still pervasive (www.quackwatch.com/01QuackeryRelatedTopics/pauling.html). It has been shown that Pauling cherry-picked from the literature, citing the articles that supported his assertion and ignoring those that did not.

3. Red wine, antioxidants and longevity

For some years, society was given the definite impression that drinking a glass of red wine a day afforded verifiable health benefits. This so-called 'French Paradox' considered the low incidence of coronary heart disease in France and attributed it to red wine in general and one of its contents, resveratrol, which is an antioxidant, in particular. The underlying theory about resveratrol, polyphenols and antioxidants in general is that they mop up free radicals. Free radicals are atoms or molecules with an unpaired electron which are very reactive and so will bind to almost any molecule (including DNA). This will lead to cell death and, over time, to premature signs of ageing. It might be logical to draw the conclusion that a molecule that sweeps up free radicals might reduce ageing, or its appearance at least. This belies the fact that sometimes the body produces its own free radicals: leucocytes (white blood cells) produce them to kill off pathogens. Ipso facto, sometimes free radicals help rather than hinder. In the 1980s, there appeared to be a rush of research suggesting the positive health effects of antioxidants, both in terms of longevity and in reducing ageing (or at least the appearance of it). For example, it was proposed that there is a positive relationship between high uptake of β-carotene (an antioxidant) and reduced risk of cancer. While on the subject of carotenes, we can put the other carotene myth to bed: eating carrots does not do anything for your eyesight (although carotenoids are involved in the pathway for sight perception) – this was a deliberate myth promoted by the War Office during the Second World War to explain the increased success rate of Allied fighters, who always appeared to know where the Luftwaffe were. It was said that the pilots had sharp eyesight because they ate many carrots. It was nothing to do with carrots and everything to do with newly discovered radar, a secret of which it was critical to keep from the Axis powers.

The French Paradox appears to be another case where correlation does not necessarily mean causation (NHS Choices, 2014):

this prospective study of nearly 800 older community-dwelling adults shows no association

between urinary resveratrol metabolites and longevity. This study suggests that dietary resveratrol from Western diets in community-dwelling older adults does not have a substantial influence on inflammation, cardiovascular disease, cancer or longevity.

4. MMR and autism

The MMR and autism controversy has been described as '*perhaps, the most damaging medical hoax of the last 100 years*' (Flaherty, 2011). The root cause was a single individual researcher with his own agenda, which, when added to the shrill support of the tabloids, created fear and uncertainty among parents. In 1998, a now-retracted article (Wakefield *et al*., 1998) was published in *The Lancet* which claimed that colitis and autism spectrum disorders are linked to the combined measles, mumps and rubella (MMR) vaccine. The media pounced on this proclamation and, as a result, thousands of children either underwent three separate vaccinations or no vaccinations at all. Only as a result of an investigation by a *Sunday Times* journalist, Brian Deer (http://briandeer.com/mmr/lancet-paper.htm), was the research debunked (although one might argue that it was never 'bunked' in the first place). In hindsight, the flaws in the article appear glaringly obvious: the 'sample' size was 12 children and the reported correlation between the vaccination and bowel or behavioural problems in eight of the 12 was anecdotal. The study did not constitute a case series, cohort study or case–control study, as would normally be required (Goldacre, 2009). Moreover, the fact that all 12 gathered in the same place (the Paediatric Gastroenterology Clinic at the Royal Free Hospital in London) meant that they were effectively a self-selecting, non-random sample, each with bowel disorders. It later transpired that Andrew Wakefield, the principal investigator, had failed to disclose a conflict of interest inasmuch as he had an interest in a patent relating to a new single measles vaccine. The 12 children were preselected by the researchers and not sequential referrals as suggested in the article. A curious facet of this case is that the media stood behind the researcher, holding him up as the courageous lone voice fighting the system. Predictably, vaccination rates dropped, with commensurate results in terms of preventable deaths, and any number of parents whose children developed autism that they would have developed with or without the vaccinations went through the needless guilt of believing they were responsible for their child's affliction.

5. The prosecutor's fallacy

This example is not so much about how science changed its mind as about how statistics did (albeit statistics presented as science and by a scientist). The phrase and concept were first coined in 1987 by William Thompson and Edward Schumann (Thompson and Schumann, 1987) and yet the fallacy caused mother-of-two Sally Clark to be found guilty in 1999 of murdering her two babies in succession, primarily because of the flawed interpretation of statistics by the now-discredited Professor Sir Roy Meadow, who was considered an expert in children who came to harm by their parent's hand.

The basic fallacy results from misunderstanding conditional probability and neglecting the prior odds of a defendant being guilty before that evidence was introduced. When a prosecutor has collected some evidence (for instance a DNA match) and has an expert testify that the probability of finding this evidence if the accused were innocent is tiny, the fallacy occurs if it is concluded that the probability of the accused being innocent must be comparably tiny. (https://en.wikipedia.org/wiki/Prosecutor's_fallacy)

Professor Meadow famously quoted the chance of two babies in the same household dying of SIDS (sudden infant death syndrome) as being 1 in 73 000 000 and cited Meadow's Law:

One sudden infant death in a family is a tragedy, two is suspicious and three is murder unless proven otherwise. (https://en.wikipedia.org/wiki/Meadow%27s_law)

He arrived at this figure by simply multiplying the chance of *one* baby dying of SIDS by itself (i.e. by squaring 1 in 8543); in other words, he assumed that both events were totally independent of each other and the chance of having a second death by SIDS was vanishingly small if a first death by SIDS had occurred. He was effectively stating that a second SIDS death is less likely than the first when in fact it is more likely since there is likely to be some underlying cause (e.g. genetic predisposition or some environmental factor). This needs to be considered in the light of knowledge

existing at the time, that the probability of a woman developing schizophrenia immediately after the birth of her baby is approximately 1 per 1000, but that the risk of developing it after the birth of a subsequent child is 50% if a previous birth has been affected (Jones, Chandra, Dazzan and Howard, 2014). This contrasts with the 1 per million that might have been calculated using the logic presented above.

The press at the time made much of Professor Meadow's number, effectively stating that 1 in 73 000 000 was the chance of the death of the two babies being accidental and therefore, crucially, the chance of Sally Clark being innocent. The jury appeared to share that view and she was found guilty, because they too, as well as the defence team, and the judge of the original trial *and* the judge of the first appeal (Goldacre, 2009) fell victim to the prosecutor's fallacy. In this context, the fallacy can be explained as follows. Two babies have died, either by double SIDS or by double murder, both of which have a low probability of having taken place. But, having taken place, both are now far more likely because it has to be one or the other (or, one supposes, one of each). The jury should therefore have been directed to consider the *relative* likelihood of double murder and double SIDS and not just the probability of double SIDS (even if the 1 in 73 000 000 figure was accepted, which it shouldn't have been).

Although double SIDS is very rare, double infant murder is likely to be rarer still, so the probability of Clark's innocence was quite high. Hill calculated the odds ratio for double SIDS to double homicide at between 4.5:1 and 9:1. (Hill, 2004)

Furthermore, Professor Meadow's original figure for a single SIDS death of 1 in 8543 was also flawed, and should have been closer to 1 in 1300. He had arrived at the figures by cherry-picking characteristics of the Clark family that seemed to bolster the prosecution case by making SIDS seem less likely. Such factors included the fact that the babies came from a non-smoking household, the parents were affluent, middle class and in a stable relationship. He also ignored those characteristics that made SIDS more likely (the two babies were boys, and boys have a higher predisposition to SIDS).

When considering the fact that the number of 1 in 73 000 000 would have been uppermost on the jury's mind, the judges at Sally Clark's second appeal said:

We rather suspect that with the graphic reference by Professor Meadow to the chances of backing long odds winners of the Grand National year after year it may have had a major effect on [the jury's] thinking notwithstanding the efforts of the trial judge to down play it. (R. v. Clark, [2003] EWCA Crim 1020, 11 April 2003)

6. MRSA superbugs in hospitals

The MRSA scare of 2005 arose when newspapers in general and tabloids in particular reported a sharp and chilling increase in MRSA (methicillin-resistant *Staphylococcus aureus*) in UK hospitals. MRSA is a bacterial infection that just doesn't respond to normal antibiotics, and understandably these reports generated public disquiet. The manner in which the reports were covered didn't help, with the *Sunday Mirror* famously describing '*the mop of death*' on one front page splash. The unusual aspect to this episode is that the reportage seemed to come before the research; and when hospital microbiologists attempted to replicate the results being screamed by the red tops (www.mirror.co.uk/news/world-news/killer-bug-trains-buses-1599712) they came back negative. Ben Goldacre, probably the UK's best-known debunker of faked or flawed science, was instrumental in uncovering the reality (Goldacre, 2009). It transpired that all the positive MRSA results emanated from one testing laboratory, Chemsol Ltd, run by a commercial company consisting of one employee, a 'Dr' Christopher Malyszewicz. The full toe-curling account can be read in Goldacre's book *Bad Science*, but the short story is that Dr Malyszewicz's lab was described by the journalists as '*the lab that always gives positive results*'. The *News of the World* described him as a '*respected MRSA specialist*' and *The Sun* called him '*the UK's top MRSA expert*'; however, it became apparent that he was qualified only inasmuch as he held a BSc from Leicester Polytechnic and a non-accredited correspondence course PhD but, crucially, no training in microbiology. His microbiology laboratory was a shed in his garden that used kitchen units as laboratory benchware. Somewhat curiously, the tabloids appeared to view him as a solitary heroic crusader attempting to turn the tide of an indifferent system, much like Dr Wakefield of MMR/autism infamy. Astonishingly, when

challenged about their articles by accredited, renowned and world-class microbiology laboratories such as University College Hospital's, the tabloids maintained their positions, implying that their position was concrete based on Malyszewicz's '*research*' even in the face of overwhelming evidence to the contrary from respected institutions.

7. Thalidomide

Thalidomide became infamous in the late 1950s and 1960s for being identified as the cause of birth defects in children born to mothers who had taken it as an anti-emetic (anti-nauseant) during pregnancy (https://en.wikipedia.org/wiki/Thalidomide). It was originally made and sold in 1953 by Chemie Grünenthal as a sedative or hypnotic after its toxicity had successfully been tested on several animal models. Once it came into general use, it was noticed by Chemie Grünenthal researchers that it had anti-emetic properties. This, plus the fact that in Germany it was being sold as an over-the-counter (OTC) drug as opposed to prescription-only medicine (POM), led to its use by pregnant women as an anti-emetic. The company marketed the drug thus, claiming that the reduction in morning sickness in expectant mothers made it a 'wonder drug'. Times have clearly changed and nowadays no drug could be breezily given to pregnant women without rigorous and painstaking testing. However, in the 1950s compliance protocols were very different and pharmacovigilance less diligent. Instances of babies being born with phocomelia (malformation of limbs) started appearing and were eventually traced back to the use of thalidomide. In this regard, thalidomide is said to be teratogenic (from Greek for 'monster forming'); this is also an example of how times have changed in terminology. Many scientific terms, when considered in isolation, would now be considered to be hurtful or offensive: cyclopia, cyclopamine, and red, white and brown dwarves (classes of imploded stars). Thalidomide was marketed in 46 countries, including most of Europe and the UK, and by the early 1960s 10 000 children had been affected with phocomelia.

The precise mechanism by which thalidomide causes its teratogenicity is not fully understood but it has two enantiomers (non-superimposable mirror images, not unlike left and right gloves). One of these acted as a sedative and anti-emetic, and the other as a teratogen (Figure 2). Separating

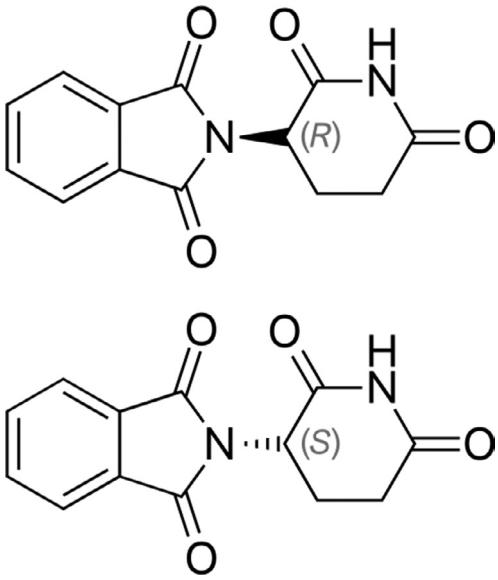

Figure 2 The *R*- and *S*- enantiomers of thalidomide

the enantiomers in the hope of using only the non-harmful one would not achieve the desired result: thalidomide racemises (changes from an optically pure compound that consists of only one enantiomer into a mixture of two enantiomers). This can happen *in vivo* due to the acidic hydrogen on the chiral centre.

The one country where thalidomide was not licensed for use was the USA, and this was due to the diligence of an FDA (Food and Drug Administration) worker called Frances Oldham Kelsey. Kelsey ran some animal model tests on rats and found that thalidomide acted as an abortifacient in rats. In humans, in the doses used, it was teratogenic and not an abortifacient. Kelsey prevented probably thousands of birth defects by meticulous lab work.

8. Beta blockers and hypertension

In some cases, 'cause and effect' for diseases and treatments have been misunderstood, despite appropriate understanding of the statistics. Let us consider the example of drug treatment for high blood pressure (hypertension). The heart acts as a pump to move blood around the body; in doing this, it generates pressure within the fluid of the blood. Blood pressure is measured in millimetres of mercury (mm Hg), i.e. the height to which the pressure in the blood vessel can push a column of mercury: 1 mm Hg is equivalent to 133.3 Pa. Unlike a domestic central heating system, however,

the veins and arteries through which the blood flows are not rigid, they are able to contract and stretch. In an average healthy adult, as the heart contracts (systole), it generates a blood pressure of 120 mm Hg and, as the heart relaxes (diastole), the pressure falls to 80 mm Hg. High blood pressure, above about 140 mm Hg systolic and 90 mm Hg diastolic, is known to be a major risk factor for stroke, heart failure, dementia and premature death. For each 2 mm Hg above normal, there is a 10% increase in the risk of mortality from stroke. It has been recognised for a long time, therefore, that lowering blood pressure is a logical approach to decreasing the risk of stroke, heart failure, dementia and premature death. At times of stress, the adrenal gland secretes adrenaline that acts to increase the rate and force of contractions of the heart and to dilate the lungs to allow better gas exchange. As part of this 'fight or flight' response, adrenaline causes constriction of the blood vessels supplying the gut and skin, and dilation of the blood vessels supplying muscle and the brain. This is achieved by adrenaline acting via two different sub-types of receptor, alpha and beta receptors. Alpha receptors are responsible for the vasoconstriction that decreases blood flow to the gut and skin while beta receptors are responsible for the vasodilation of the blood vessels supplying muscle. Beta receptors cause the heart to increase the rate and force of its contractions and cause bronchodilation in the lungs. In 1965, Goodman and Gilman's textbook *The Pharmacological Basis of Therapeutics* (commonly referred to as the *Blue Bible*; originally Goodman and Gilman (1965), now Brunton, Knollman and Hilal-Danden (2017)) discussed a newly emerging class of drugs called the beta blockers. These drugs were seen to block the effects of adrenaline on the beta receptors and therefore prevented the increase in rate and force of cardiac contraction. There was a consequent decrease in blood pressure.

In a later edition of the same textbook, in 1975, the beta blocker propranolol was reported to be effective at treating moderate hypertension when used alongside other drug therapies. By the time of the publication of the eighth edition of the textbook in 1990, the relevant part of the chapter on antihypertensive agents read:

[Beta blockers] *were not thought to have antihypertensive effects when they were first investigated* … [the] *antihypertensive effect was subsequently demonstrated for propranolol and all other* [beta blockers] … *The* [beta blockers] *provide effective therapy for many cardiovascular and other diseases, and they are useful for all grades of hypertension.*

In the UK, the National Institute for Health and Clinical Excellence (NICE) published its guidelines for the treatment of essential hypertension in 2004 (Clinical Guideline 18, see Williams *et al*., 2004) in which it stated:

Drug therapy should normally begin with a low dose thiazide-type diuretic. If necessary second line add a beta blocker unless the patient is at risk of new-onset diabetes.

In younger patients, aged under 55, with moderately raised blood pressure and who may be managed on one drug, consider beginning with a beta-blocker.

Pharmacology textbooks of the time listed beta blockers as first- or second-line treatment for hypertension, stating that millions of patients worldwide were having their hypertension controlled by beta blockers alone. By 2006, however, NICE guidelines were updated with the phrase:

One class of drugs that caused particular debate was the beta-blockers. In head-to-head trials, beta-blockers were usually less effective than a comparator drug at reducing major cardiovascular events, in particular stroke.

In the guidelines of 2011 (www.nice.org.uk/guidance/cg127/chapter/1-guidance), data gathered from seven major clinical trials were presented to indicate that, although beta blockers were effective at lowering blood pressure, they did not necessarily prolong life or prevent development of other diseases. In terms of all causes of death among patients with hypertension, those being treated with beta blockers had a risk of dying of 94% compared with untreated patients, i.e. those receiving placebo. The probability of this marginal decrease in risk having occurred by chance was determined as being 16%, i.e. beta blockers did not significantly decrease mortality in patients with hypertension and two of the seven trials suggested that patients receiving beta blockers were more likely to die. Furthermore, beta blockers only decreased the risk of heart attack (myocardial infarction) to 92%; again the

probability that this marginal decrease was due to chance was assessed as 76%, with two of the trials showing greater risk for beta blocker users. For stroke, the reduction was to 81%, but this was also seen to be non-significant, and the probability that the reduction occurred due to chance was 8%, though no trial showed beta blockers to be worse than placebo.

The lesson learned from this sequence of events is that high blood pressure is associated with increased risk of heart disease, stroke, dementia, heart failure, etc. The perceived logic was that drugs that lowered blood pressure would therefore decrease the risks of the other problems. The search was thus on to find drugs to lower blood pressure: beta blockers lower blood pressure and therefore became a popular form of antihypertensive medication. What was forgotten was that the aim of the treatment was actually to decrease the risk of heart disease, stroke, dementia, heart failure, etc., not to lower blood pressure. When the effects of the beta blockers on these other conditions were investigated, they were found to be small. The most recent guidelines for the treatment of hypertension state that beta blockers are not the preferred therapy for hypertension – they are only recommended for particular patients (such as pregnant women) and as a final resort when all other treatments have failed.

References

BBC News (2016) Public Health England: advice to eat more fat 'irresponsible'. *BBC News*, 23 May. Available at: www.bbc.co.uk/news/health-36345768.

Brunton, L., Knollman, B. and Hilal-Dandan, R. (2017) *Goodman and Gilmans' The Pharmacological Basis of Therapeutics*. 13th edn. New York: McGraw-Hill Education.

Flaherty, D.K. (2011) The vaccine-autism connection: a public health crisis caused by unethical medical practices and fraudulent science. *Annals of Pharmacotherapy*, **45**(10), 1302–1304.

Goldacre, B. (2009) *Bad Science*. 1st edn. London: Fourth Estate.

Goodman, L.S. and Gilman, A. (1965) *Goodman & Gilman's The Pharmacological Basis of Therapeutics*. 3rd edn. New York: Macmillan.

Hill, R. (2004) Multiple sudden infant deaths – coincidence or beyond coincidence? *Pediatric and Perinatal Epidemiology*, **18**(5), 320–326.

Jones, I., Chandra, P.S., Dazzan, P. and Howard, L.M. (2014) Bipolar disorder, affective psychosis, and schizophrenia in pregnancy and the post-partum period. *The Lancet*, **384**(9946), 1789–1799.

Joint Epilepsy Council (2011) *Epilepsy Prevalence, Incidence and Other Statistics*. Joint Epilepsy Council. Available at: www.epilepsyscotland.org.uk/pdf/Joint_Epilepsy_Council_Prevalence_and_Incidence_September_11_%283%29.pdf.

NHS Choices (2014) Resveratrol's health benefits 'overstated'. *NHS Choices*, 13 May. Available at: www.nhs.uk/news/heart-and-lungs/resveratrols-health-benefits-overstated/.

Perälä, J., Suvisaari, J., Saarni, S.I. Kuoppasalmi, K., Isometsa, E., Pirkola, S., Partonen, T., Tuulio-Henriksson, A., Hintikka, J., Kieseppa, T., Harkanen, T., Koskinen, S. and Lonnqvist, J. (2007) Lifetime prevalence of psychotic and bipolar I disorders in a general population. *Archives of General Psychiatry*, **64**(1),19–28.

Pauling, L. (1976) *Vitamin C and the Common Cold*. San Francisco: W.H. Freeman.

R. v. Clark, [2003] EWCA Crim 1020, 11 April 2003.

Thompson, W.C. and Schumann, E.L. (1987) Interpretation of statistical evidence in criminal trials: the prosecutor's fallacy and the defense attorney's fallacy. *Law and Human Behavior*, **11**(3), 167–187.

Wakefield, A.J., Murch, S.H., Anthony, A., Linnell, J., Casson, D.M., Malik, M., Berelowitz, M., Dhillon, A.P., Thomson, M.A., Harvey, P., Valentine, A., Davies, S.E. and Walker-Smith, J.A. (1998) Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *The Lancet*, **351**(9103), 637–41 [retracted].

Williams, B., Poulter, N.R., Brown, M.J., Davis, M., McInnes, G.T., Potter, J.F., Sever, P.S. and McG Thom, S. (2004) Guidelines for management of hypertension: report of the Fourth Working Party of the British Hypertension Society, 2004-BHS IV. *Journal of Human Hypertension*, **18**(3), 139–185.

Michael Hal Sosabowski is Professor of Public Understanding of Science in the School of Pharmacy and Biomolecular Sciences at the University of Brighton. Email: M.H.Sosabowski@brighton.ac.uk

Paul R. Gard is Professor of Experimental Therapeutics in the School of Pharmacy and Biomolecular Sciences at the University of Brighton.