CHAPTER 1
What is pharmacovigilance and how has it developed?

Origins and definition of pharmacovigilance

In the beginning, there was thalidomide. It can be argued that the history of pharmacovigilance goes back further but, for practical purposes, the story of modern pharmacovigilance begins there.

In the late 1950s there was little, if any, regulation of medicines outside the USA (where thalidomide was not marketed), and their testing and development was almost entirely in the hands of pharmaceutical companies. In the case of thalidomide, unjustified claims of safety in pregnancy were made and its use as a sedative was targeted at pregnant women. The drug turned out to be a teratogen, producing a variety of birth defects but particularly limb defects known as phocomelia (see Figure 1.1). Worldwide, about 10,000 fetuses were affected, particularly in Germany where the drug was first marketed. Since phocomelia was otherwise a very rare congenital abnormality, the existence of a major increase in its incidence did not go unnoticed in Germany but the cause was initially thought to be environmental. In 1961 a series of just three cases associated with thalidomide was reported in The Lancet, the problem was finally recognised and the drug withdrawn from sale.

At the beginning of the 1960s, publication of possible adverse effects of drugs in the medical literature was effectively the only mechanism for drawing attention to them. Thalidomide produced a non-lethal but visible and shocking adverse effect, leading people to ask why so many damaged babies had been born before anything had been done? This question is central to subsequent

An Introduction to Pharmacovigilance. By Patrick Waller.
developments. It is unlikely that we will ever be able to predict and prevent all the harms which may be caused by medicines but limiting the damage to much smaller numbers is now achievable. Today we would expect to be able to identify an association between drug and outcome analogous to thalidomide and phocomelia after the occurrence of less than 10 cases, i.e. at least three orders of magnitude more effectively than five decades ago.

The overriding lesson learnt from thalidomide was that we cannot just wait until a drug safety problem, quite literally in this case, hits us between the eyes. So thalidomide led directly to the initial development of the systems we now have, although it is only quite recently (i.e. since the early 1990s) that the term pharmacovigilance has become widely accepted.

Pharmacovigilance has been defined by the WHO as ‘The science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other drug-related problems’. There are other definitions but this very broad one seems to be the most appropriate since there is a clear implication that the process is one of ‘risk management’. This is a concept which is applicable to many
What is pharmacovigilance and how has it developed?

Aspects of modern life but, surprisingly, its explicit use in relation to pharmaceuticals is quite a recent development.

Thalidomide is not merely of historical interest since in the last few years it has made something of a comeback. The reasons for this exemplify the point about risk management since the risk of fetal malformation can be successfully managed by avoidance of the drug during pregnancy. It also demonstrates another concept which is central to the practice of pharmacovigilance – the balance of benefit and risk. Thalidomide appears to have benefits in some diseases that are otherwise difficult to treat conditions, e.g. refractory multiple myeloma – these appear to outweigh the risk of fetal malformation if there is an effective pregnancy prevention scheme in place. A further point which thalidomide illustrates well, and is relevant to many other drug safety issues, is that not everyone is at the same risk of a particular adverse effect. In this case, a substantial part of the population i.e. women who are not of childbearing capacity, are not at risk at all.

**Main lessons from thalidomide**

- The need for adequate testing of medicines prior to marketing.
- The need for government regulation of medicines.
- The need for systems to identify the adverse effects of medicines.
- The potential relationship between marketing claims and safety.
- Avoidance of unnecessary use of medicines in pregnancy.
- That some risks can be successfully minimised.

The ramifications of the thalidomide tragedy were many-fold but the key lesson for the development of pharmacovigilance was that active systems for detecting hazards are needed. Within a few years this had been taken forward with the introduction of voluntary (or ‘spontaneous’) schemes for reporting of suspected adverse drug reactions (ADRs). These have stood the test of time as an alerting mechanism or ‘early warning system’ and will be covered in more detail in Chapter 3.

**Scope and purposes of pharmacovigilance**

In the past, the process of pharmacovigilance has often been considered to start when a drug is authorised for use in ordinary practice. Nowadays, it is more commonly considered to include all safety-related activity beyond the point at which humans are first exposed to a new medicinal drug.
The ultimate purpose of pharmacovigilance is to minimise, in practice, the potential for harm that is associated with all active medicines. Although data about all types of ADRs are collected, the main focus is on identifying and preventing those which are defined to be serious. This means an ADR which meets at least one of the following criteria:

- Fatal
- Life-threatening
- Causes or prolongs hospitalisation
- Results in long-term disability

Additionally, all congenital abnormalities are considered serious and the definition of ‘serious’ allows the application of medical judgement such that a reaction may be considered serious, even if there is not clear evidence that one of the above criteria is met.

Non-serious reactions are important to individual patients and health professionals involved in their treatment but they can usually be managed clinically and they impact much less on the balance of benefit of risk and the public health. Thus, pharmacovigilance may be seen as a public health function in which reductions in the occurrence of serious harms are achievable through measures which promote the safest possible use of medicines and/or provide specific safeguards against known hazards. Pregnancy prevention in users of thalidomide is an example of such a safeguard; monitoring white blood cell counts to detect agranulocytosis (absent white blood cells) in users of the antipsychotic drug clozapine is another.

In order to minimise harms there is first a need to identify and assess the impact of unexpected potential hazards. For most medicines, serious ADRs are rare; otherwise their detection would result in the drug not reaching or being withdrawn from the market. For products which do reach the market, serious hazards are seldom identified during pre-marketing clinical trials because sample sizes are almost invariably too small to detect them. In addition, the prevailing conditions of clinical trials – selected patients, short durations of treatment, close monitoring and specialist supervision – almost invariably mean that they will underestimate the frequency of ADRs relative to what will really occur in ordinary practice.

During pre-marketing clinical development, the aims of pharmacovigilance are rather different to the broad public health function described above. In volunteer studies and clinical trials there
is an overriding need to protect individuals being exposed. There is also a need to gather information on harms which occur in order to make a provisional assessment of safety and to plan for post-marketing safety development.

**Development of pharmacovigilance since the 1960s**

In the early 1970s another drug safety disaster occurred – this was the multi-system disorder known as the oculo-mucocutaneous syndrome caused by practolol (Eraldin) – a cardioselective beta-blocker used to treat angina and hypertension. As in the case of thalidomide, several thousand individuals were permanently damaged before the association was recognised. The fundamental problem in this instance was a failure of timely identification despite having an early warning system in place. Ultimately the system was dependent on doctors suspecting an association between drug and disease. Probably because of the unusual nature of the syndrome – dry eyes, skin rash and bowel obstruction – and a long latency period (averaging almost two years in respect of the onset of the most serious bowel manifestations), relevant cases were not reported until the association was identified in the medical literature. Around 3,000 cases were then retrospectively reported to the UK ‘Yellow Card’ scheme, an example of the potential effect of publicity on ADR reporting. Subsequent attempts to develop an animal model of practolol toxicity failed, indicating that the problem could not have been predicted from pre-clinical studies.

**Main lessons from practolol**

- Some adverse effects are not predictable from pre-clinical studies.
- Spontaneous reporting schemes are not invariably effective.
- Long latency effects and clinical manifestations not known to be related to other drugs may not be suspected as ADRs by doctors.
- Additional, more systematic methods of studying post-marketing safety are needed.

The overriding message from practolol was that spontaneous ADR reporting alone is insufficient as a means of studying post-marketing safety. Thus, in the late 1970s various schemes designed to closely monitor the introduction of new drugs were suggested, but most of them were not implemented. The basic idea was that initial users of new drugs would be identified through prescriptions and monitored
systematically rather than waiting for someone to recognise a possible adverse effect. The concept did come to fruition in the UK in the early 1980s with the development of ‘prescription-event monitoring’, a method which is still in use today (see Chapter 3).

The first drug studied by prescription-event monitoring was benoxaprofen (Opren), a non-steroidal anti-inflammatory drug (NSAID) which frequently produced photosensitivity reactions, i.e. rashes in light-exposed areas. A published case series of five deaths related to hepatic and renal failure led to withdrawal of the drug in 1982, even though some doubts were expressed as to whether they were caused by the drug, particularly as prescription-event monitoring did not reveal any indication of these effects. Many of the patients who experienced serious ADRs with benoxaprofen were elderly; this was due to reduced excretion of the drug as a consequence of renal impairment. Even though it is well-recognised that many patients who use NSAIDs are elderly, benoxaprofen had not been adequately studied in this population prior to marketing. A reduction in the dosage recommendations for the elderly was implemented briefly but it was too late to save the drug. Because the usage of benoxaprofen took off rapidly after launch and an important adverse effect – photosensitivity reactions – was common, a large number of spontaneous reports were received in a short period of time, swamping the primitive computer systems then used and pointing up the need for purpose-designed databases. The issue also illustrated the need for patients to be properly informed about possible ADRs and how to minimise the risk – in this case by avoiding exposure to the sun. It was therefore influential in moving us towards the introduction of patient information leaflets – these became compulsory in the EU during the 1990s.

Main lessons from benoxaprofen

- Uncertainty about cause and effect from individual case reports – further impetus to the need for formal post-marketing studies.
- The need to study a drug in the population that will use it (e.g. the elderly).
- The need for purpose-designed computer systems to handle ADRs more promptly and effectively.
- The concept of intensive surveillance of new drugs, achieved in the UK by the introduction of the Black Triangle scheme (see Glossary).
- The need for patients to be informed about possible ADRs.
As it turned out, benoxaprofen was just the first of a series of NSAIDs withdrawn for various safety reasons in the 1980s. During this decade, pharmaceutical companies started to conduct their own post-marketing surveillance studies and UK guidelines related to their conduct were drawn up in 1987. However, initially, the value of such studies turned out to be limited because they usually lacked comparator groups and often failed to meet the planned sample-size. The UK guidelines were revised in 1993 with the aim of improving the quality of studies. The principles of the revised, so-called Safety Assessment of Marketed Medicines or SAMM, guidelines also became a blueprint for the first EU level guidance on the topic.

During the mid-1980s, the term pharmacoepidemiology was first used to mean the scientific discipline of the study of drug use and safety at a population level. The discipline developed strongly during the 1990s with the increasing use of computerised databases containing records of prescriptions and clinical outcomes for rapid and efficient study of potential safety hazards. In some instances prescription records are held in a separate database to clinical events, and linkage between the two databases needs to be achieved through some common identifier in the two sets of data in order to study adverse events at an individual patient level.

Towards the end of the 1980s pharmacovigilance eventually recognised and started to deal with the problem of dependence on benzodiazepines – so-called ‘minor tranquillisers’ such as chlordiazepoxide (Librium) and diazepam (Valium) that had been introduced in the 1960s. Advice was issued to limit the dose and duration of such treatments although, even today, such recommendations are widely ignored. The issue brought into focus the problems faced in dealing with the misuse and abuse of prescription drugs. This is another example of a situation where spontaneous ADR reporting failed to highlight an important concern, the issue eventually coming into focus as a result of pressure from advocates for groups of affected patients.

As well as the problem of delayed identification of real hazards, pharmacovigilance has suffered from the reverse, i.e. apparent identification of hazards which turn out not to be real. To some extent this is inherent in a system which relies much on clinical suspicions – sometimes these will be wrong. The consequences are that sometimes a drug may be unnecessarily withdrawn or people become too scared to use it. For example, Debendox
(or Bendectin), a combination product containing an antihistamine doxylamine, was widely used for the treatment of nausea and vomiting in pregnancy in the 1970s. It was withdrawn in the early 1980s on the basis of concerns that it might cause fetal malformations, a concerted campaign against the drug and impending litigation. At the time, the evidence of a hazard was very weak but it was not possible to exclude a significant risk to the fetus. Subsequently, many studies of this potential association were performed and collectively they provided no evidence of an increased risk of fetal malformations. This example illustrates the intrinsic difficulty of disproving the existence of a hazard once concern has been raised. A more recent, very high profile example illustrating the same point was the suggestion made in late 1990s that combined measles, mumps and rubella (MMR) vaccine might be a cause of autism in children. Despite there being little credible evidence for this suggestion, it was impossible to completely disprove it and hard to convince worried parents. Vaccine campaigns were damaged and a significant number of cases of measles occurred in the UK for the first time in many years.

The mother of all drug safety scares occurred with oral contraceptives (OCs) in 1995. It was not the first ‘pill’ scare – this story began in the late 1960s when it was discovered through spontaneous ADR reporting and confirmed in formal studies that combined OCs (containing an oestrogen and a progestagen) increased the risk of venous thromboembolism (VTE). This led to a reduction in the dose of oestrogen to 20–30 µg of ethinyloestradiol which lessened (but did not abolish) the risk without compromising efficacy. Nevertheless, when the risk of thrombosis became public knowledge many women were scared and stopped taking OCs. It is important to recognise that most women using OCs are relatively young and healthy – this impacts considerably on their perception of the risk. When OCs are stopped abruptly by sexually active women without immediate use of an effective alternative, unwanted pregnancies occur and abortion rates increase. The have been several ‘pill’ scares over the years related to VTE and also to other safety issues – e.g. a possible association with myocardial infarction and a small increase in the risk of breast cancer. In each instance, many women who stopped using OCs later returned to using OCs but the public health impact of each of these scares in terms of unwanted pregnancies was considerable. This has been particularly unfortunate since pregnancy itself is fundamentally
riskier than using any OC and there may also be compensating health benefits from using them.

In 1995 a WHO study of OCs unexpectedly found a two-fold increase in the risk of VTE when use of so-called ‘third-generation’ (3G) OCs was compared to ‘second-generation’ (2G) OCs. The difference between these pills was the progestagen component – desogestrel or gestodene for 3G OCs and levonorgestrel for 2G OCs. This was surprising as it had always been considered that VTE risk was simply related to the dose of the oestrogen component of the pill. Another multinational study which could address the relative safety of 3G and 2G OCs was ongoing and a further study was quickly conducted using a UK database. Within about three months the results of three studies were available and their findings were all quite similar. Arguments were put forward that the associations seen in these studies were not necessarily causal and also that it was possible that 3G OCs might have benefits which would compensate for the increase in VTE risk. There was general agreement that the absolute level of risk – VTE is quite rare in healthy young women, even if they take the pill – was not such that 3G OCs should be withdrawn from the market but nevertheless the UK’s expert regulatory committee felt that doctors and women needed to know. Despite a clear message being provided that no one should stop taking OCs, many women did, presumably because the media coverage scared them. It did not help that the principal investigator of one of the studies flew from Canada to London to give a press conference criticising the committee’s advice because the public get more worried when experts disagree. At the time, the European Medicines Agency had recently been formed but co-operation on nationally authorised products was in its infancy. Various authorities in Europe and around the world adopted different positions and it was not until 2001 that the EU reached an agreed position on the issue.

Over a period of several years, more studies were done and the effects of the various progestagens on blood clotting investigated. Ultimately, it was shown that there were plausible differential effects of these agents on clotting and there was enough consistency in the risk data to convince most scientists that the observed association was causal. But, despite good intentions all round, it was hard to escape the feeling that more harm than good had been done and that the communication tools used were inadequate. In 1997 the WHO convened a meeting of experts to specifically
consider how communication in pharmacovigilance could be improved (see Chapter 5).

Main lessons learned from the OC safety issues

- Drugs are sometimes marketed at the wrong dose.
- There may be differences in safety between drugs of the same class.
- Harm may result from safety warnings.
- Uncertainty and debate about risks may fuel public concern.
- The power of the media to influence users is much greater than the authorities.
- The need for greater international co-operation in pharmacovigilance.
- There is a need to develop more effective communication tools.

One important point about the OC issues discussed above is that the data on which they were based did not (after the initial signal in the 1960s) come from spontaneous ADR reporting. Despite that, causation was debatable because the studies were not randomised trials but ‘observational’. VTE is a sufficiently rare outcome in young women that it would be extremely difficult to conduct a large enough clinical trial to detect a doubling of risk.

Later in life, women have also been prescribed female sex hormones – in lower doses and as replacement therapy (HRT). In this age group the baseline risks of VTE, arterial cardiovascular disease and various cancers are much greater and therefore, it is more feasible to study them in clinical trials although they do need to be large and long-term. Therefore observational studies of these outcomes were performed first and, in general, they appeared to show that HRT reduced the risk of arterial disease outcomes, i.e. myocardial infarction and stroke. HRT was not authorised for the purpose of reducing cardiovascular risk but in the 1980s and 1990s it was quite widely used for this purpose. The fundamental problem in performing such studies is that women using HRT may be healthier to start with, although it is possible to address this, at least to some extent, in the design and analysis. Another important point is that the outcome in question is a benefit (i.e. a reduction in risk) and, because of such biases, observational studies rarely provide convincing evidence of benefit. It is generally accepted that randomised trials are needed to establish efficacy and benefit.

Eventually, large randomised trials were set up but they had to be stopped early because they tended to show the opposite of what was
What is pharmacovigilance and how has it developed?

expected – i.e., an increase in cardiovascular risk. Warnings were issued and, because there is no major downside to suddenly stopping HRT, communication was intrinsically easier than with OCs. Indeed, the intended effect of the warnings was that women who were inappropriately using long-term HRT should stop taking it. However, conveying the right messages was not straightforward because there were multiple risks involved, and they are time-dependent and cannot simply be expressed as a proportion (e.g. 1 in 100).

HRT, like the last two issues I am going to cover here, came to a critical point in the first three or four years of the new millennium. However, history is not yet ‘complete’ on any of these issues, indeed one often wonders whether it ever can be – e.g. with the return of previously withdrawn drugs like thalidomide and clozapine. The latter is an antipsychotic drug which was first introduced in the 1970s and then withdrawn following reports of agranulocytosis, i.e. absence of white blood cells. It was reintroduced with compulsory blood monitoring around 1990.

Selective serotonin re-uptake inhibitors (SSRIs) are antidepressants which were brought to the market in the late 1980s and have since largely replaced older, ‘tricyclic’ antidepressants such as amitriptyline. The main reason why they have done so – apart from effective marketing – is that they are less toxic to the heart in overdose, i.e. there is a greater margin of safety in relation to dose. Depressed patients are at risk of taking an overdose and therefore this is potentially an important advantage.

There have been two controversial issues with SSRIs – withdrawal reactions and a possible increase in the risk of suicide. Problems experienced by patients when they stop treatments are often quite difficult to assess because they could possibly be related to recurrence of the disease. Nevertheless, the potential for SSRIs to produce withdrawal reactions was identified during their development, and when spontaneous reports were received post-marketing it was hardly a new ‘signal’. There were very large numbers of such reports received but few were serious and the level of usage of the drugs was high. Over a period of years it became clear that the problem was occurring much more commonly than initially thought, particularly in users of paroxetine (Seroxat), a fairly short-acting drug. Ultimately, greater care was needed in withdrawing patients more gradually from these drugs. Suggestions have been made that SSRIs are drugs of dependence but most scientists do not accept this because features such as craving and dose-escalation
are generally absent. Importantly, it emerged that the nature of some of the more unpleasant symptoms patients experienced – e.g. so-called ‘electric shock’ sensations in the head was being lost in the data processing systems. This was due to inadequate coding such cases often became ‘paraesthesia’, something that hardly conveys how unpleasant such sensations can be. Thus it was recognised that we need better ways to capture unusual patient experiences and this gave considerable impetus to allowing patients to report their adverse reactions to the authorities. That approach had been used in the USA for many years but hardly at all in Europe until the early years of the new millennium.

The possibility that any drug might increase the risk of an outcome associated with the disease it is being used to treat is invariably difficult to evaluate. Suicidal feelings and actions are relatively common in depressed patients and it is not surprising when they occur in a patient who has recently started treatment. Nevertheless, around 1990 a clinician in the USA saw several patients treated with fluoxetine (Prozac) who had suicidal thoughts and he published a case series suggesting that the drug might be responsible. This prompted a review of all the clinical trial data for the drug which did not support the proposition but it was never completely refuted. Over the years more clinical trial data accumulated for various drugs in the class and studies were conducted in children and adolescents, the latter being a high-risk group for suicide. Even in severely depressed patients, completed suicides are rare in clinical trials and therefore the evidence that is available relates mostly to attempted suicide (also uncommon in trials) and thoughts of suicide measured on various scales. Trials of paroxetine in children produced some potentially worrying findings that for some time were known only to the manufacturer. When the regulatory authorities eventually received the data, they issued warnings against the use of this drug in children. The company was investigated and prosecution considered but the law was found to be insufficiently clear that they were obliged to immediately submit concerning clinical trial data to the authorities when a trial was being conducted outside the authorised indication. This issue again pointed to the potential importance of clinical trials to the assessment of safety and raised concern about a lack of transparency with clinical trial data. Already, considerable steps have been taken towards making clinical trial data publicly available through mechanisms other than publication in the literature
What is pharmacovigilance and how has it developed?

which is slow and selective. The jury is still out on whether SSRIs directly increase the risk of suicide but there is general agreement that the early phase of treatment is a high-risk period and that careful monitoring of patients is required.

Finally, what is probably the most important drug safety issue of recent years? The answer is the increased risk of cardiovascular outcomes associated with selective COX-2 inhibitors (coxibs). This possibility was first uncovered in basic research but not followed through; the first clinical indication of a problem came from a trial known as VIGOR which was published in 2000. At the time, two drugs in the class – rofecoxib and celecoxib – had just been authorised. The VIGOR study was a randomised comparison of rofecoxib and naproxen (a standard NSAID) designed to establish whether or not there was a difference in the rates of serious gastrointestinal adverse effects of these two drugs. In that respect, rofecoxib was clearly preferable and the trial results led to rapid uptake of coxibs – on the basis that they were supposedly safer. The VIGOR study also found an important difference in the rate of cardiovascular events such as myocardial infarction – these were five-fold more common in patients taking rofecoxib, compared to naproxen. This information was included in the original publication but lacked prominence and was presented as a five-fold reduction with naproxen rather than an increase with rofecoxib. The paper has since been the subject of extensive criticism.

Over the years there have been suggestions that standard NSAIDs might reduce the risk of cardiovascular outcomes (as aspirin does) and one explanation for the finding in the VIGOR study put forward was that naproxen is ‘cardioprotective’ whereas rofecoxib is not. Ultimately, it took a large clinical trial comparing rofecoxib with placebo to establish beyond any doubt that this was an adverse effect of rofecoxib (rather than a lack of benefit) and the findings of that study led to the drug being withdrawn from the market in late 2004. This event sent shockwaves around the world that are still reverberating leading people to question why such a trial had not been done much earlier, i.e. before millions of people had used the drug. It also left a big cloud hanging over the remaining drugs in the class – some have been withdrawn and some remain in the market. At one stage, the proposition that coxibs might be given to people at high risk of gastrointestinal and low risk of cardiovascular disease seemed reasonable but it has since been discovered that, to a considerable extent, risk factors
for these problem overlap in individual patients. To make matters even more complicated, it appears that some standard NSAIDs might also increase the risk of cardiovascular events and, at the present time, our ability to assess the relative safety of drugs in the same class remains rather limited.

**Main lessons learned from recent major safety issues**
- The need for vigorous follow-up of safety signals with appropriate studies.
- The difficulty of assessing outcomes which are related to the drug indication.
- The potential value of clinical trials in assessing safety and the importance of the choice of comparator drug(s).
- Important safety data may emerge from clinical trials performed for other purposes.
- The need for greater openness about clinical trial data.
- The potential importance of off-label use (e.g. in children) to safety.
- There is a need to evaluate medicines properly in children.
- The need for greater patient involvement in drug safety.
- The complexity of evaluating and communicating multiple risks (and benefits).
- The need for regulatory authorities to have powers to ensure that companies adequately investigate potential risks with marketed products.

**Conclusion**

The issues discussed above are necessarily selective and my narration of them is broad. The intention is primarily to illustrate that pharmacovigilance has experienced many teething problems and that most of its developments have been in response to quite specific lessons learned from landmark safety issues. In this chapter, I have tried to illustrate what pharmacovigilance is and how it has progressed over a period of nearly half a century. Despite that progress, no one should doubt that there is a long way to go yet. The current limitations of the discipline and how we might overcome them are considered in Chapter 8.