Clinical Topics

Identification of adverse reactions to new drugs. III: Alerting processes and early warning systems

GEOFFREY R VENNING

Abstract

The process of discovery for 18 important adverse drug reactions was analysed in terms of study designs and other features which determined whether published reports of suspected adverse reactions were convincing or provided valid evidence of causality. First alerts were almost always anecdotal reports.

For earlier alerting during clinical trials event recording is needed, associated with dose-response data and the use of patients as their own controls.

Introduction

In part I¹ an attempt was made to identify a list of important adverse drug reactions to form the basis for a review of the discovery process as it had existed in recent years, mainly since thalidomide. In part II^{2,3} the history of the discovery process and "avoidable delays" were reviewed in respect of 18 important reactions identified in part I. This paper attempts to analyse the processes leading to alerting for the adverse reactions on the basis of the characteristics of the 18 published reports identified as providing the first alerts.

Classification of discovery processes—The studies leading to identification are classified in two ways—firstly, according to study design and, secondly, according to the presence or absence of specific features which may affect the validity of any study irrespective of design.

Designs of reported studies

An adverse drug reaction may be regarded as a clinical event or illness in someone taking a drug as a result of an unwanted action of the drug. That adverse clinical events may occur from other causes, or for no known reason, highlights the need for an epidemiological approach to validation of suspected adverse reactions. Even when a drug apparently produces a unique or unusual clinical syndrome it is important to verify the frequency of occurrence of the defined syndrome in non-users. If this can be confirmed as very low then causality for a suspected reaction may be established on the basis of a small number of reports. Even under these conditions, and certainly in all other circumstances, an important factor in validation is the presence or absence of data on frequency. This may take the form of data on the incidence of the reaction among users of the drug or alternatively data on the prevalence of drug usage among patients with the particular clinical event defined as the suspected reaction.

These considerations prompted the choice of the following classification of study designs.

Firstly, a cohort of drug users may be reviewed. Any such study gives data on the incidence of the reaction, as the number of users is known and constitutes the denominator. Included in this category of study will be clinical trials, whether controlled or uncontrolled, post-marketing surveillance studies, and epidemiological cohort studies.

Secondly, a group of patients may be reviewed who experience a particular unwanted clinical event and the prevalence noted of the prior usage of one or more drugs. Such a study gives no data on the incidence of the reaction, as the denominator (the number of drug users) is unknown. Studies of this type could be defined as "valid case studies" and would include case-control studies, in which a separate group of patients or normal subjects is also reviewed for comparison of the prevalence of drug usage. In assessing case studies it is important to know whether the group of patients with the observed clinical event is an unselected group, representative of all patients with such an event in a defined population, or whether it is selected and thus liable to bias, in which case the observed prevalence of drug usage may be misleading.

For the prevalence data to be useful it is necessary that a case study should represent a consecutive series of cases, including non-users as well as users of the drug. Included in this class will be studies based on special disease registries, hospital surveillance schemes—for example, the Boston Collaborative Drug Surveillance Program⁴—and reviews of national mortality and morbidity statistics, as well as case-control studies and acceptable case studies as defined above. In all of these types of studies data are available on the denominator—the number of clinical events—to form the basis for assessing the prevalence of drug usage, which may then be compared with the known level of drug usage in the general population. The classification selected does not cover all types of controlled epidemiological study. Some cross sectional studies may look simultaneously at the prevalence of drug usage and of defined clinical events—for example, in a hospital population. In such studies the selection bias in hospital admission may limit the validity of both the denominators, and other factors may be required to establish validity. (In one such study in the survey phenacetin nephropathy was verified, primarily through the evidence on dose response rather than through the data on incidence of the reaction among users and of prevalence of drug usage among patients with the reaction, both of which lacked validity, as explained.)

Thirdly, there are studies not included in the above classes. These are clearly the least useful for verification of adverse reactions to drugs and include anecdotal reports such as single case reports and case studies reporting a number of patients taking a drug who have been noted to experience a particular clinical event—the suspected adverse reaction. These will also include studies based on selective registries—for example, of patients taking oral contraceptives who develop endometrial cancer, or of children with brain damage who have been given pertussis vaccination. These should be regarded as biased registries. Studies based on voluntary adverse reaction reporting systems—for example, yellow cards—fall into this category. Most anecdotal reports of suspected reactions also lack information on each of the two denominators described. There is no information on the number of clinical events in a population including both users and non-users.

Having considered these three categories of study it is natural to consider whether any method of study might be possible which would provide information on both denominators, thus making it possible

University Department of Community Medicine and General Practice, Oxford

GEOFFREY R VENNING, BM, FRCP, visiting scientist (now director of research and development, Janssen Pharmaceutical Limited, Marlow, Bucks SL7 1ET)

to estimate the incidence of a suspected adverse reaction among users of a drug and also the prevalence of drug users among patients who experience the clinical event which is suspected of being an adverse reaction. The only type of study with these characteristics is a record linkage system. Apart from pilot studies such as that in Oxford by Skegg *et al*⁵ this approach has not been applied, and studies in this survey are therefore classed as (a) cohort studies, (b) valid case studies, and (c) other studies.

Other factors affecting validity

In evaluating the discovery processes for adverse reactions defined criteria must be used for accepting or rejecting evidence of causality. The following definitions are used here. (a) *Convincing*: This is the level of strength of evidence accepted for defining an adverse reaction as verified. Evidence has not always constituted complete scientific proof but has been accepted where the author is convinced that the drug is a cause of the reaction without any reasonable doubt. (It is also noted that occasionally a suspected reaction is so serious and widespread that regulatory action may be justifiable on suspicion even when the evidence does not meet the criteria defined as convincing.) (b) *Strong*: Such evidence still requires verification but indicates a probability of causality. (c) *Weak*: This type of evidence does no more than raise the possibility of causality.

The following features of clinical studies have been accepted as providing evidence to these standards.

CONVINCING EVIDENCE

Rechallenge data, particularly data on more than one rechallenge.

Significant dose-response data may provide convincing evidence, even from studies with poor design features, with lack of information on the denominators defined above.

Control data—Here study design is critical. Data from prospective cohort studies are in theory stronger than data from case-control studies, but where due attention is paid to possible confounding factors the latter may be convincing.

Experimental data on mechanisms of pathogenesis—Rawlins and Thompson⁶ classified adverse reactions into type A where the adverse effect is a consequence of the pharmacological action of a drug, often predictable and dose dependent, as distinct from type B, which are idiosyncratic. Reports of type A reactions are in general, almost by definition, convincing.

Evidence from close association in time and space—Although not encountered in this study, it must be recognised that local reactions at the site of injection and immediate adverse reactions such as acute anaphylaxis at the time of drug administration can normally be accepted as convincing.

STRONG EVIDENCE

Uniqueness of adverse event—This would include, for example, special clinical or histological features.

Extreme rarity of adverse event in the absence of drug usage. Self evident.

Data on minor variations of adverse reaction or formes frustes—(Here the problem arises that one cannot assume that a drug that consistently causes minor changes in organ function can necessarily be held responsible for occasional severe dysfunction.)

WEAK EVIDENCE

Association of drug with common or only moderately rare events relative to extent of usage of drug—This would include, for example, the first reports of myocardial infarction in association with oral contraceptives.

Demonstration of pharmacological actions of drug not established as relevant to pathogenesis of adverse reaction—This would, for example, include the argument that oral contraceptives were likely to cause thromboembolism because effects on clotting factors had been detected; there was no evidence that these factors were relevant to the pathogenesis of thromboembolism.

Improvement after withdrawal, when not supported by rechallenge data, offers only weak evidence of causality.

Results

Table I shows the categories and descriptions of the studies alerting to the 18 adverse reactions. In three of the 18 instances the studies were formal cohort studies, and in two were valid case studies; in the remaining 13, however, alerting arose from anecdotal reports, including single case reports in four instances. Table II shows the strength of the evidence in these 18 studies: the evidence was convincing in one report, strong in 12, and weak in five.

TABLE I—Study designs in publications leading to first alerts of 18 important adverse drug reactions

Adverse reaction	Drug	Description of study
<i>Cohort studies (n = 3)</i>		
Lactic acidosis	Phenformin	Clinical pharmacology
Tardive dyskinesia	Phenothiazines	Therapeutic trial
Retroperitoneal fibrosis	Methysergide	Therapeutic trial
<i>Valid case studies (n = 2)</i>		
Vaginal cancer	Stilboestrol	Survey of unusual histological type of cancer
Nephropathy	Phenacetin	Survey of unusual histological type of nephropathy
<i>Anecdotal studies (n = 13)</i>		
Skin rash	Practolol	Single case report
Keratitis		Anecdotal report
Peritonitis	Oral contraceptives	Single case report
Pulmonary embolism		Anecdotal report
Myocardial infarction	Aerosols	" "
Deaths from asthma	Clioquinol	" "
Subacute myelo-optic neuropathy	Chloramphenicol	" "
Aplastic anaemia	Halothane	" "
Jaundice	Antibiotic	" "
Colitis	Lincomycin	" "
"	Clindamycin	" "
"	Phenylbutazone	" "
Aplastic anaemia		Single case report

TABLE II—Strength of evidence and factors affecting validity at time of alerting; 18 important adverse drug reactions

Adverse reaction	Drug	Nature of evidence	Type*
<i>Convincing evidence</i>			
Lactic acidosis	Phenformin	Experimental data on mechanism of pathogenesis	A
<i>Strong evidence</i>			
Vaginal cancer	Stilboestrol	Special histological features	A
Nephropathy	Phenacetin	Special histological features	B
Keratoconjunctivitis	Practolol	Special clinical features	B
Deaths from asthma	Aerosols	Special clinical features	B
Sclerosing peritonitis	Practolol	Rarity of event	B
Retroperitoneal fibrosis	Methysergide	Rarity of event	B
Aplastic anaemia	Chloramphenicol	Rarity of event	B
Aplastic anaemia	Phenylbutazone	Rarity of event	B
Pseudomembranous colitis	Antibiotics	Rarity of event	B
Tardive dyskinesia	Lincomycin	Data on variant of adverse drug reaction	A
	Clindamycin		
	Phenothiazines		
<i>Weak evidence</i>			
Pulmonary embolism	Oral contraceptives		B
Myocardial infarction			B
Subacute myelo-optic neuropathy	Clioquinol		B
Skin rash	Practolol		B
Jaundice	Halothane		B

*Rawlins and Thompson's classification.⁶

Discussion

ALERTING

In considering the role of formal controls incorporated into the design of studies several points emerge. At the alerting stage five out of 18 reports came from studies meeting the defined criteria for classification as cohort studies or valid case studies, three being cohort studies and two valid case studies. *The design*

features, however, were not the main reasons for the validity in the three cohort studies, but they were important in the two case studies.

In the phenformin study the strength of the evidence rested primarily on the experimental evidence for the mechanism of the adverse effect; the study design, which made it possible to assess the incidence of the reaction, naturally added to the strength of the evidence. In the methysergide report two patients were observed with an unusual adverse reaction. The physician published information on all patients he had treated with the drug (500), thus giving data on incidence of the reaction. This made the report somewhat more convincing than an anecdotal report of two cases of the suspected reaction with no information on the denominator of drug usage. Data on incidence from a series of this kind, however, may be biased by selection for publication; other users with equal experience uncomplicated by adverse reactions may be less likely to publish. The same applies to the phenothiazine report. In the phenacetin and stilboestrol studies unusual histological features were identified; the study designs ensured that all patients with these features were reviewed, so that the prevalence of drug usage among patients with the reaction could be assessed.

The remaining 13 publications were anecdotal reports, often of single cases; in some of these, however, the evidence was strong. The practolol keratoconjunctivitis report identified special clinical features, as did the Australian report of deaths from asthma associated with aerosols which identified a cardiac type of death. (The story of this adverse reaction is complex. Deaths from asthma in childhood and subsequently shown to be attributable to aerosols were not usually cardiac.) In other anecdotal reports the strength of the evidence depended on the rarity of the clinical event in the absence of drug use. This is, of course, relative to the extent of drug usage in the population. A few cases of aplastic anaemia with a drug of limited use would be strong evidence. The first anecdotal reports associated with chloramphenicol were less convincing, as the drug was in very widespread use at the time. Further verification was needed (and was obtained quite promptly). Five of the anecdotal reports lacked good evidence of causality. I have shown elsewhere⁷ that first alerts of this kind may turn out to be false alarms and not form a logical basis for regulatory action. When, however, the suspected adverse reaction is as serious as fatal pulmonary embolism or myocardial infarction in a young woman or hepatic necrosis with a new anaesthetic or the widespread occurrence of subacute myelo-optic neuropathy, which reached epidemic proportions in Japan, then verification or refutation becomes an urgent matter. The other first alert which failed to provide good evidence was the practolol dermatitis report. Of itself, this was not a very serious reaction. Skin reactions to drugs are common, and perhaps this should be taken into account in assessing the strength of the evidence. Two years later different authors verified causality by patch testing and rechallenge. If this had been done in the earlier study this adverse reaction would not have needed further verification.

The present state of alerting to adverse reactions to drugs is that anecdotal reports provide the main source of information. Many discussions of the problem in recent years have considered the possible value of national voluntary reporting systems, the British yellow card system in particular, hospital surveillance schemes such as the Boston Collaborative Drug Surveillance Program, extension of phase IV (postmarketing) trials into monitored release schemes for postmarketing surveillance, and monitoring of national morbidity and mortality statistics. It is salutary to note the contribution these procedures made to the alerting process for these 18 important adverse reactions. The only actual first alert from any of these mechanisms was the Australian Drug Regulatory Agency warning on deaths from asthma associated with aerosols. This may have been fortuitous

in so far as the alert related to a cardiac mechanism of death in adults, whereas subsequent studies eventually verified a real problem confined to children and not characterised by cardiac death. Monitoring of mortality statistics contributed to solving this problem and was the first alert in Britain, one year after the Australian warning, which appears to have been overlooked elsewhere. The British yellow card system, phase IV clinical trials, and hospital schemes such as the Boston programme did not contribute to alerting to these 18 reactions, but the yellow card system began only in 1964 and the surveillance programme in Boston in 1966. From then on, however, these systems failed to provide early warnings for the eight remaining reactions—three effects of practolol and the reactions to aerosols, stilboestrol, lincomycin, and clioquinol.

LESSONS TO BE LEARNT

Lessons for the clinical stage—Study design features such as randomisation and control data are now almost universally recognised as critical for establishing the efficacy of new drugs. Although these particular features have not proved to be critical for early alerting to suspected adverse reactions, the importance of event recording has been identified, in the study which proved capable of showing the relation between practolol and eye complaints in a series including only 71 patients. Observations on control periods incorporated into the study were important. Possibly the combination of event recording with the use of patients as their own controls may in the future lead to earlier alerting at the clinical trial stage, through identification of significant excesses of minor organ specific complaints which may provide an early warning of less frequent but more serious adverse reactions affecting the same organ.

Lessons for regulatory agencies for alerting—So far as the Committee on the Safety of Medicines is concerned it is salutary to note that the yellow card system made no contribution to first alerting for any of the important adverse reactions identified. (The value of the system for verification will be considered in part IV.) I have shown elsewhere that some anecdotal reports could form a satisfactory basis for regulatory warnings.⁷ They should also trigger specific action to ensure early verification.

I thank Sir Richard Doll, Professor Abe Goldberg, Dr G Jones, Dr David Skegg, and Professor Martin Vessey for helpful comments and criticisms.

Part IV of this series will be published next week.

References

- 1 Venning GR. Identification of adverse reactions to new drugs. I—What have been the important adverse reactions since thalidomide? *Br Med J* 1983;**286**:199-202.
- 2 Venning GR. Identification of adverse reactions to new drugs. II—How were 18 important reactions discovered and with what delays? *Br Med J* 1983;**286**:289-92.
- 3 Venning GR. Identification of adverse reactions to new drugs. II (continued)—How were 18 important reactions discovered and with what delays? *Br Med J* 1983;**286**:365-8.
- 4 Jick H, Miettinen OS, Shapiro S, et al. Comprehensive drug surveillance. *JAMA* 1970;**213**:145-60.
- 5 Skegg DCG, Richards SM, Doll R. Record linkage for drug monitoring. *J Epidemiol Community Health* 1981;**35**:25-31.
- 6 Rawlins MD, Thompson JW. Pathogenesis of adverse reactions. In: Davies DM, ed. *Textbook of adverse drug reactions*. Oxford: Oxford University Press, 1981.
- 7 Venning GR. The validity of anecdotal reports of suspected adverse reactions—the problem of false alarms. *Br Med J* 1982;**284**:249-52.

(Accepted 11 November 1982)