Causality Assessment in Pharmacovigilance: A Step Towards Quality Care

Dr. Rohan C. Hire, Dr. Parag J. Kinage, Dr. Nilesh N. Gaikwad
Department of Pharmacology, Government Medical College, Nagpur, Maharashtra, India – 440003.

Corresponding author
Dr. Rohan C. Hire
Email: rohanhire123@gmail.com

Abstract: Adverse drug events ranges from mild to life threatening reactions which results in inconvenience or serious morbidity and mortality. Clinicians often do not recognize drug related harm. Terms used to describe these events with medication use cause much confusion. Moreover manifestations of adverse drug reactions can be non-specific making difficult to differentiate from current illness. To determine the likelihood of relationship between the drug and the event, assessment of causality is done which is rather the important task in conducting the National Pharmacovigilance Programme in each country. Despite the great number of methods proposed, assessing the causal role of a drug in the occurrence of an adverse medical event remains one of the most controversial issues. Qualifying terms for criteria, such as “compatible”, “suggestive” or “inconclusive”, have never been strictly defined, leading to low reproducibility. So, in order to have a harmony in defining the criteria of causality assessment, researchers started developing different methods for causality assessment. All these methods were classified in three broad categories: expert judgement/global introspection, algorithms and probabilistic methods (Bayesian approaches). As a result of problems of reproducibility and validity, no single method is universally accepted. Different causality categories are adopted in each method, and the categories are assessed using different criteria. Because assessment methods are also not entirely devoid of individual judgements, inter-rater reliability can be low. In conclusion, there is still no method universally accepted for causality assessment of ADRs.

Keywords: Causality assessment, Adverse drug events, Bayesian approaches, Algorithm, Pharmacovigilance.

INTRODUCTION:
Pharmacovigilance is the science and activities relating to detection, assessment, understanding and prevention of adverse effects or any other drug related problem [1]. These adverse drug reactions (ADRs) not only add to the suffering of patients but also increase morbidity and mortality along with a financial burden on society. The overall incidence of ADRs in hospitalized patients is estimated to be 6.7% (range 1.2–24.1%) and that of fatal ADRs 0.32% (0.1–0.85%) [2]. Data indicates that in patients who experience ADRs death rates are 19.18% higher and the length of hospital stay is 8.25% higher. Total medical costs for patients with ADRs are increased by an average of 19.86% [3]. However, the lack of ability of clinicians to suspect or detect such adverse events related to drugs might lead to inappropriate management of adverse events thus exposing the patient to additional drug hazards. To minimize the suffering of the patients from ADRs, it is essential though difficult to establish causal relationship between the drug and the event which is nothing but the causality assessment. By definition, causality assessment is the evaluation of the likelihood that a particular treatment is the cause of an observed adverse event [4]. It assesses the relationship between a drug treatment and the occurrence of an adverse event. It is an important component of pharmacovigilance, contributing to better evaluation of the risk-benefit profiles of medicines [5] and is an essential part of evaluating ADR reports in early warning systems and for regulatory purposes [6].

METHODS OF CAUSALITY ASSESSMENT
Many researchers developed various methods of causality assessment of ADRs by using different criteria like chronological relationship between the administration of the drug and the occurrence of the ADR, screening for non drug related causes, confirmation of the reaction by in vivo or in vitro tests, and previous information on similar events attributed to the suspect drug or to its therapeutic class, etc. to define ADRs in different categories [7]. But because there are no defined diagnostic criteria or categories, inter-rater and intra-rater variability can be large [8]. Currently, there is no universally accepted method for assessing causality of ADRs. [9] We describe here three broad categories of various methods of causality assessment: expert judgement/global introspection, algorithms and probabilistic methods (Bayesian approaches) [9]. Expert judgements are individual assessments based on previous knowledge and experience in the field using no standardized tool to arrive at conclusions regarding causality. Algorithms are sets of specific questions with associated scores for calculating the likelihood of a cause-effect relationship. Bayesian approaches use specific findings in a case to transform the prior estimate of probability into a posterior estimate of probability of drug causation. The prior probability is calculated from epidemiological information and the
posterior probability combines this background information with the evidence in the individual case to come up with an estimate of causation. As a result of problems of reproducibility and validity, no single method is universally accepted. Reproducibility ensures an identical result, regardless of who the user is, and when he uses it. Validity means the ability of the method to distinguish between cases where the drug is responsible and cases where it is not [7].

**EXPERT JUDGEMENT/GLOBAL INTROSPECTION**

ADRs are mostly suspected or recognized by either treating physicians or clinical pharmacologists [10]. This expert judgement or global introspection is a process whereby an expert expresses judgement about possible drug causation by considering all available data relevant to a suspected ADR [11], estimates their relative importance and assigns weights to deduce the probability of the role of the drug in the untoward event [10]. Assessment of ADR in this category is either done by single expert evaluator or by a group of expert evaluators. As evaluation and assessment of ADR by these experts is purely based on their respective knowledge and experience about the subject of interest, it produces disagreement and extensive inter-rater variability. We describe here two methods based on expert opinion or global introspection.

**Swedish method by Wilholm et al. [12]**

It was used by Swedish regulatory agency. The clinician evaluates the causal relationship by considering seven different factors: (i) the temporal sequence, (ii) previous information on the drug, (iii) dose relationship, (iv) response pattern to drug, (v) rechallenge, (vi) alternative aetiological candidates and (vii) concomitant drugs. Events are classified as ‘probable’ or ‘possible’ and ‘non-assessable’ or ‘unlikely’. A limitation of this method is the small number of categories into which causality can be placed, as there may be an overlap and ADRs could be wrongly evaluated.

**World Health Organization (WHO) - Uppsala Monitoring Centre (UMC) causality assessment criteria [4]**

WHO-UMC system has been developed in consultation with the National Centres participating in the Programme for International Drug Monitoring and is meant as a practical tool for the assessment of case reports. It is basically a combined assessment taking into account the clinical-pharmacological aspects of the case history and the quality of the documentation of the observation. Since pharmacovigilance is particularly concerned with the detection of unknown and unexpected adverse reactions, other criteria such as previous knowledge and statistical chance play a less prominent role in the system. This method gives guidance to the general arguments which should be used to select one category over another.

The WHO–UMC causality assessment method includes the following four criteria [1]:

a) Time relationships between the drug use and the adverse event.

b) Absence of other competing causes (medications, disease process itself).

c) Response to drug withdrawal or dose reduction (dechallenge).

d) Response to drug readministration (rechallenge).

The level of causal association is grouped into six categories which are based on a number of the above criteria being met. Causal category is ‘certain’ when all the four criteria are met. It is ‘probable’ when criteria a, b and c are met. When only criterion a is met, the event is categorized as ‘possible’ and it is ‘unlikely’ when criteria a and b are not met (Table 1). Beside these four categories, ADR can also be categorised into ‘Unclassified/Conditional’ or ‘Unassessable/ Unclassifiable’ in WHO-UMC causality assessment. The term ‘Unclassified/Conditional’ is applied when more data is needed and such data is being sought or is already under examination. Finally when the information in a report is incomplete or contradictory and cannot be complemented or verified, the verdict is ‘Unclassifiable’. For drug-drug interactions the WHO-UMC system can be used by assessing the interacting drug, which influences the kinetics or dynamics of the non-interacting drug (which has usually been taken over a longer period), in the medical context of the patient.

**Table 1: WHO-UMC causality assessment method [1]**

<table>
<thead>
<tr>
<th>Categories</th>
<th>Time sequence</th>
<th>Other drugs/disease ruled out</th>
<th>Dechallenge</th>
<th>Rechallenge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Certain</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Probable</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Possible</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Unlikely</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Dechallenge is the clinical decision to withdraw/discontinue a drug treatment after possible ADR has occurred. A dechallenge is ‘positive’ or ‘suggestive’ if the reaction abates, partially or completely when the drug is withdrawn and is considered to be ‘negative’ or ‘against’ if the reaction does not abate when the treatment is stopped. Rechallenge is nothing but the deliberate or inadvertent administration of a further dose(s) of the same medicinal product to a person who has previously
experienced an adverse event/adverse drug reaction that might be drug related. Failure of the product, when reintroduced, to produce signs and symptoms similar to those observed when the suspect drug was previously introduced implies a negative rechallenge, while recurrence of similar signs and symptoms upon reintroduction of the suspect product implies a positive rechallenge [13].

ALGORITHMS

An algorithm is a problem-specific flow chart with step-by-step instruction on how to arrive at an answer [14]. It is a clinical instrument in the form of a questionnaire that gives detailed operational criteria for ranking the probability of causation when an ADR is suspected. Algorithms give structured and standardized methods of assessment in a systematic approach to identifying ADRs based on parameters such as time to onset of the ADR or temporal sequence, previous drug/adverse reaction history and dechallenge and rechallenge. Individual cases are approached systematically, resulting in a high degree of consistency and reproducibility. Clinical judgement is, however, required at various stages to arrive at a conclusion [15]. Currently, there are many algorithmic methods of causality assessment but no single algorithm is accepted as the ‘gold standard’, because of the shortcomings and disagreements that exist between them [5]. We shortlisted few important algorithmic methods here.

Dangaumou’s french method [16]

This method has been used by the French regulatory agency since 1977. The method separates an intrinsic imputability (possible cause between drug and clinical event) from an extrinsic imputability (bibliographical data) using seven criteria (three chronological and four semiological) in two different tables. The chronological criteria are (i) drug challenge, (ii) dechallenge and (iii) rechallenge, with an overall score of four possible categories. The semiological criteria are (i) semiology (clinical signs) per se (suggestive or other), (ii) favouring factor, (iii) alternative non-drug-related explanation (none or possible) and (iv) specific laboratory test with three possible outcomes (positive, negative or no test for the event-drug pair). Scores are grouped into ‘likely’, ‘possible’ and ‘dubious’. The advantage of this method is that it allows certain drugs taken at the same time with the ‘suspect’ drug to be excluded, because each drug is imputed separately. However, this method requires more time than most other algorithms.

Kramer et al. method [17]

This algorithm applies to a single clinical manifestation occurring after administration of a single suspect drug. In cases where multiple drugs are involved, each is assessed separately. One of the advantages of this algorithm is its transparency. However, certain levels of expertise, experience and time are required to use this method effectively.

Naranjo et al. method (Naranjo scale) [18]

It is used to assess causality in a variety of clinical situations using the conventional categories and definitions of ‘definite’, ‘probable’, ‘possible’ and ‘doubtful’. It consists of ten questions (Table 2) that are answered as ‘yes’, ‘no’, ‘unknown (don’t know)’. The event is assigned to a probability category based on the total score. A total score of 1–6 is ‘definite’, ‘probable’ 7–9, ‘possible’ 1–4 and ‘doubtful’ ≤ 5. This scale is intended to assess the likelihood of an ADR associated with only one drug, not for adverse drug events resulting from interactions between two drugs. The Naranjo scale does not address the main points that are necessary in causality evaluation of potential drug interactions.

Table 2: The Naranjo scale/ questionnaire

<table>
<thead>
<tr>
<th>Questions</th>
<th>Yes</th>
<th>No</th>
<th>Don’t know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence of previous conclusive report on adverse reaction.</td>
<td>+1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Did adverse event appear subsequent to administration of suspected drug?</td>
<td>+2</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Did adverse event improve on drug discontinuation or on administration of specific antagonist?</td>
<td>+1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Did the adverse event reappear when the drug was re-administered?</td>
<td>+2</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Are there any alternative causes other than the suspected drug that could have caused the reaction on their own?</td>
<td>-1</td>
<td>+2</td>
<td>0</td>
</tr>
<tr>
<td>Did the adverse event reappear when a placebo was administered?</td>
<td>-1</td>
<td>+1</td>
<td>0</td>
</tr>
<tr>
<td>Was the incriminated drug detected in toxic concentrations in blood (fluids)?</td>
<td>+1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Did the adverse event worsen on increasing the dose or decreased in severity with lower doses?</td>
<td>+1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Past history of any similar reaction to the same or similar drugs.</td>
<td>+1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Was the adverse event confirmed by objective evidence?</td>
<td>+1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Total score 0– Doubtful 1–4 Possible, 5–8 Probable, ≥9 Definite
Balanced assessment method (Lagier et al.) [19]

It evaluates case reports on a series of visual analogue scales (VAS), according to the likelihood that each criterion is fulfilled. Its advantage is that it considers the possibility of an alternative to causation for each of the factors and not just as a separate factor. Although each case is assessed by two independent assessors, the evaluation still depends, to a large extent, on the level of assessor’s knowledge. An evaluator needs to be an expert in the particular area to make reliable evaluations.

Summary time plot (Castle et al.) [20]

It was proposed for the industrial setting to identify patterns of ADRs. The plot summarizes the time relationship between treatment and a possible adverse reaction. With sufficient information from the causality criteria, the duration of treatment and possible adverse reaction are plotted with time on the x-axis and severity of the possible adverse reaction on the y-axis. This method does not lead to a conclusion on causality, because it only summarizes the time factor alongside other factors that are relevant to the drug-event relationship. The method is, however, quick to use, saves the use of ambiguous terminology and is applicable even with minimal information.

Ciba geigy method (Venulet et al.) [21]

The ‘Ciba-Geigy method’ resulted from a number of expert consensus meetings. Experts used their clinical judgement to assess events and assign causality on a VAS. This method was updated and replaced with a checklist of 23 questions, split into three sections: (i) history of present adverse reaction, (ii) patient’s past adverse-reaction history and (iii) monitoring-physician’s experience. This updated method was found to have a high degree of agreement (62%) when compared with evaluator’s assessments. Although the level of reliability does not assure validity, this method reflects the knowledge and experience of the evaluator, and the type of ADR that is being evaluated.

Loupi et al. method [22]

It is developed to assess the teratogenic potential of drug. The first sections of the algorithm (chrono-semiological axis) allow for the drug to be excluded if not implicated in the origin of the abnormality. The second section (bibliographical axis) weights the bibliographical data. The three questions consider alternative aetiological candidates other than the drug; chronology of the suspect drug and other bibliographical data, to arrive at a conclusion on causality.

Roussel Uclaf causality assessment method (RUCAM)[7]

This method is designed for predetermined disease states such as liver and dermatological injuries. A retrospective assessment of the reproducibility of this method among four experts showed a 37–99% agreement rate. Although this method seems quite easy to use, it is organ specific. Therefore, the criteria need to be defined by a consensus of experts for each medical field and validated before it can be of any meaningful use in ADRs other than hepatic or dermatological injuries.

Maria and Victorino (M and V) scale [23]

Maria and Victorino developed this scale for diagnosing drug induced liver injury (DILI). Probability was expressed as a score between 6 and 20, divided into five causality degrees (score of > 17, definite; 14 - 17, probable; 10 -13, possible; 6 – 9, unlikely; < 6, drug hepatotoxicity excluded). Diagnosis of DILI is complex and requires experienced clinicians in order to be accurate. Though famous, it is not free from lacunae. In cases where more than one drug is suspected, the scale needs to be computed for individual drugs. Some questions on the M&V scale apply only to immunological hepatitis, making it difficult for scores to be generated for other hepatic injuries.

Drug Interaction Probability Scale (DIPS) [24]

It was proposed by Horn et al. Drug Interaction Probability Scale (DIPS) is used to evaluate drug interaction cases. The DIPS uses ten questions that are answered ‘yes’ or ‘no’ to yield a score estimating the likelihood of drug interaction. The questions concern the pharmacological properties of the drug, the possible role of other drugs and specific patient information. The method was developed to assist users in the assessment of drug-interaction-induced adverse outcomes and also to serve as a guide for the further study of potential drug interactions. Only requirement is the adequate knowledge of either the drugs involved and/or the basic mechanisms of interaction.

PROBABILISTIC OR BAYESIAN APPROACHES

Bayesian methods for causality assessment make use of specific findings in a case to transform a prior into a posterior probability of drug causation [25]. The prior probability is calculated from epidemiological information and the posterior probability combines this background information with the evidence in the individual case. It is open-ended with no limit to the amount of case details that can be assessed. This method allows the simultaneous assessment of multiple causes [26].

Australian method [27]

It is one of the first probabilistic methods used. Conclusions are drawn from internal evidence, such as timing, and laboratory information from case reports. Previous knowledge on the suspect-drug profile is deliberately excluded in the assessment. Likelihood decisions are made only on the likelihood of a causal relationship.
Bayesian Adverse Reactions Diagnostic Instrument (BARDI):

Bayesian Adverse Reactions Diagnostic Instrument (BARDI) was developed to overcome the numerous limitations associated with expert judgements and algorithms [28]. This BARDI is used to calculate the odds in favour of a particular drug causing an adverse event compared with an alternative cause. These odds are referred to as the posterior odds. The posterior odds factor is calculated by considering six assessment subsets: one deals with background epidemiologic or clinical trials information (the prior odds) and the other five deal with case specific information (the likelihood ratios). The prior odds (PrO) factor is the ratio of the expected drug-attributable risk and the background risk of a certain adverse event in a population sharing basic characteristics with the patient being considered (such as medical condition). The five likelihood ratios (LRs) deal with any information of differential diagnostic value under the categories of patient history (Hi); timing of the adverse event with respect to drug administration (Ti); characteristics of the adverse event (Ch); drug dechallenge (De), which refers to any signs, symptoms, or occurrences after drug withdrawal; and drug rechallenge (Re) or readministration of the suspected causal drug(s).

The product of these factors is the posterior odds (PsO) [14]:

$$\text{PsO} = \text{PrO} \times \text{LR(Hi)} \times \text{LR(Ti)} \times \text{LR(Ch)} \times \text{LR(De)} \times \text{LR(Re)}$$

The Bayesian approach can be implemented as a spreadsheet programme on either paper or computer. It calculates and provides instant numerical and graphical feedback as soon as new pieces of evidence of the suspected ADR are evaluated [26]. Case reports are read and descriptions that fit reports from the literature are noted. The software consists of worksheets to impute case parameters, one for case findings and another for scoring. Although this method requires some expertise to operate, it can evaluate more than two possible causes at the same time. The spreadsheet allows rapid calculations and interaction during the process [29].

MacBARDI spreadsheet [14]

MacBARDI-Q&A is based on an earlier computerized spreadsheet that assessed cases of neutropenia suggested as being drug induced. This spreadsheet (MacBARDI) also has been used for cases of pulmonary fibrosis associated with antiarrhythmics, cutaneous reactions associated with sulfonamides, and anticonvulsants, fetal alcohol syndrome and benzodiazepine withdrawal. The MacBARDI computerized spreadsheet contains or requires five types of information: (i) pure information lines that describe the input needed; (ii) input lines that are the parameters used to calculate each of the six factors; (iii) assumption lines, which are built-in inputs used in the calculations; (iv) calculation lines that calculate and show the value of each term in the assessment; and (v) output lines, which show the value of each factor necessary to calculate the posterior odds and the posterior odds itself. MacBARDI facilitates updating case analyses as new information becomes available, has all the criteria necessary for a good causality assessment method (e.g. explicitness, flexibility), encourages learning and modelling, and substantially decreases the time required to assess cases. Using the spreadsheet required knowledge of BARDI, however, and it did not have automatic database lookup, meaning the assessor had to find relevant information on the database and enter it on the spreadsheet.

Other prototypes of BARDI developed for diagnosing ADRs include a model for the prediction of risk of pseudo-allergic reaction and histamine release in patients undergoing surgery [30] and a diagnostic aid for pseudomembranous colitis [29]. The merits of BARDI are reliability (the same input information generates the same output) explicitness and transparency (final results show clearly what information is considered and its contribution in the assessment) and aetiological balancing (all drug and non-drug possible causes are considered in the assessment). The significant amount of time, resources and complex calculations involved are obvious limitations of this approach.

CAUSALITY ASSESSMENT OF VACCINE RELATED ADVERSE EVENTS [31]

Vaccines are administered on large scale to healthy individuals for anticipated benefits. As infants and neonates are most common beneficiaries, vaccines must meet a high degree of safety. Hence causality assessment of adverse events associated with vaccines has to be given utmost importance. We describe here a method developed by the Advisory Committee on Causality Assessment (ACCA) in Canada. The most serious and unusual reactions requiring detailed review are submitted to ACCA; at each twice yearly meeting. ACCA is composed of specialist in paediatrics, epidemiology, infectious diseases, immunology, neurology, pathology, adverse event surveillance, and microbiology and has been reviewing individual cases in a systematic stepwise manner to categorize them on a specially designed causality assessment form.

This causality assessment form consists of seven sections. Section one relates to the reason for reporting and whether the committee agreed with both the diagnosis that was made and the statement of severity. Section two takes the evaluators through several important factors like frequency of occurrence of adverse events, similar events known to occur with other diseases, vaccine-event interval compatible with event, similar symptoms in past, concomitant drugs or other conditions; for assessment of causality. Section
three relates to causality assessment by using WHO-UMC criteria. Section four permits brief summary of case with important elements and discussion which contributed to the final assessment of causality. Section five permits recommendations for improving immunization or case reporting procedures to be written. Section six considers whether the case could be useful for educational purpose. Section seven considers whether the case could be useful for publication.

CONCLUSION

The numerous published methods for causality assessment in ADRs have various advantages and disadvantages. The idea of creating standardized causality assessment systems to provide reliable and reproducible measures of the relationship-likelihood in suspected cases of ADR seems unfeasible, since no single method has achieved this to date. The differences in ADR causality criteria and the unavoidable subjectivity of judgements may be responsible for the lack of reproducibility of most published methods. So far, no ADR causality assessment method has shown consistent and reproducible measurement of causality; therefore, no single method is universally accepted.

REFERENCES


