Drug Saf (2015) 38:383–393
DOI 10.1007/s40264-014-0257-5

ORIGINAL RESEARCH ARTICLE

Assessment of a New Instrument for Detecting Preventable Adverse Drug Reactions

Raja Benkirane · Rachida Soulaymani-Bencheikh · Asmae Khattabi · Ghita Benabdallah · Loubna Alj · Houda Sefiani · Khedidja Hedna · Lahcen Ouammi · Sten Olsson · Shanti N. Pal

Published online: 24 December 2014

Abstract

Background Pharmacovigilance centres (PVCs) in the World Health Organization (WHO) Programme for International Drug Monitoring have demonstrated their ability to detect preventable adverse drug reactions (ADRs) in their databases. In this field, there is no gold-standard method for detecting medication errors and evaluating ADR preventability. Therefore, we developed, from existing tools, a preventability assessment method: the 'P Method' (PM).

Objective To present the PM and to evaluate its inter-rater reliability.

Methods The PM includes 20 explicit criteria for assessing ADR preventability. This approach is based on identification of any potentially preventable risk factor that increases the likelihood of ADR occurrence. The outcome of the preventability assessment results in one of three possible scores: 'preventable', 'non-preventable' or 'not

Key Points

The new adverse drug reaction (ADR) preventability assessment method developed by the World Health Organization (the 'P Method') has a level of agreement between reviewers that is similar to those of existing tools.

The P Method provides a substantial basis for further development and for signalling possible preventability.

A documented case report with a case narrative facilitates assessment of preventability.

Assessment of ADR preventability should explore patient and drug-related characteristics, healthcare professionals' practices and also environmental factors.

Electronic supplementary material The online version of this article (doi:10.1007/s40264-014-0257-5) contains supplementary material, which is available to authorized users.

R. Benkirane · G. Benabdallah · L. Alj · H. Sefiani · L. Ouammi
Moroccan Pharmacovigilance Centre, Rabat, Morocco

R. Benkirane (✉)
Centre AntiPoison et de Pharmacovigilance, Rue Lamfedel Cherkaoui, BP 6671, Madinate Al Irfane, 10100 Rabat, Morocco
e-mail: raja.benkirane@gmail.com

R. Soulaymani-Bencheikh
Faculty of Medicine, Mohamed V, Rabat, Morocco

A. Khattabi
Institut National Administration Sanitaire, Rabat, Morocco

K. Hedna
Department of Drug Research/Clinical Pharmacology, Linköping University, Linköping, Sweden

S. Olsson
World Health Organization Collaborating Centre for International Drug Monitoring, Uppsala Monitoring Centre, Uppsala, Sweden

S. N. Pal
Quality Assurance and Safety of Medicines, World Health Organization, Geneva, Switzerland

assessable'. The PM was tested in a multicentre study involving nine national PVCs. Two experienced reviewers at each participating PVC independently analysed the preventability of 183 ADRs, applying the PM.

Results The overall agreement between all reviewers for assessment of ADR preventability was 'fair', with a kappa value of 0.27 [95 % confidence interval (CI) 0.21–0.40]. The level of agreement between reviewer pairs ranged from 'slight', with a kappa value of 0.12 (95 % CI −0.03 to 0.27), to 'substantial', with a kappa value of 0.69 (95 % CI 0.48–0.89).

Conclusion The analysis of the agreements and disagreements between reviewers highlighted where improvements might be made. Given that no standard assessment tool exists in the WHO Programme, the transparency of the assessment process in this method provides a substantial basis for further development and for support in signalling possible preventability.

1 Background

A new challenge for pharmacovigilance centres (PVCs) in the World Health Organization (WHO) Programme for International Drug Monitoring is to extend their role to include collection of information on adverse incidents related to medication errors (MEs). For 40 years, PVCs have been concerned with preventing adverse drug reactions (ADRs), and it is well established that MEs are the most preventable cause of ADRs [1]. Some PVCs have been exploring their capacity to detect preventable ADRs [1–4]. In the principal conducted studies, the estimated rate of preventable ADRs as a proportion of all ADRs varied between 18.7 and 80 %, depending on the setting and the method used to assess their 'avoidability' [5]. Recent literature has indicated that the drug classes mostly involved in MEs are analgesics, anxiolytics, antidiabetic agents, anticoagulants and anticonvulsants [6, 7]. Currently, there is no gold-standard method for evaluating ADR preventability. A systematic review has highlighted that existing approaches to assess ADR preventability are not satisfactory [8]. These approaches rely either on the judgment of experts, which is less reproducible, or on use of predefined criteria, which cannot always be applied to every individual case. Hence, the authors propose a theoretical approach based on analysis of ADR mechanisms [9].

In this field, the Monitoring Medicines (MM) Project—set up by the WHO, coordinated by the Uppsala Monitoring Centre (UMC) and funded by the European Commission (FP7)—was initiated to expand the role and scope of national PVCs (NPVCs) to prevent medicine-related adverse events [10]. One of the most important steps forward in this project was to develop a method to assess ADR preventability.

For that purpose, the Moroccan PVC, which is a partner in this project, has developed a method based on previously existing preventability tools published in the literature: the 'P Method' (PM), which relies on explicit criteria for assessing ADR preventability [8]. The aim of the current study was to present the PM and to evaluate its inter-rater agreement.

2 Methods

2.1 Development of the P Method

The purpose of the PM is to explore the whole medication use process from prescription to drug use monitoring, aiming to identify preventable risk factors that increase the likelihood of ADRs. These risk factors constitute the 20 criteria that are used to assess the avoidability of ADRs. It explores risk factors in relation to healthcare professionals' practices (criteria 1–16), patient behaviour (criteria 19 and 20) and drug quality (criteria 5, 6, 17 and 18) (Table 1). The reference documents that were used to assess ADR preventability were the summary of product characteristics (SmPC) and the updated international or standard guideline in relation to drug use.

The PM requires a 'yes', 'no', 'not applicable' or 'unknown' response to each of the 20 criteria for each ADR (Table 1). Answering 'yes' to any one of the criteria involved in the ADR's occurrence classifies the event as preventable. This implies that the cause of the ADR is known, which facilitates identification of the critical criteria that are potentially involved in the occurrence of the ADR. These critical criteria vary according to the cause of the ADR. For example, if the cause of the ADR is linked to the dose, the critical criteria to be explored are criteria 1, 2, 3, 4, 9, 10, 12, 13 and 16. If the ADR is time related, the critical criteria are criteria 3, 4, 7 and 15. Criteria 9, 10 and 11 are the critical criteria for an ADR that is related to patient susceptibility [17]. Patient behaviour, the indication for the drug and drug quality should be explored systematically; they could increase the likelihood of any ADR (criteria 5, 6, 8, 17, 18, 19 and 20).

A criterion is considered 'not applicable' when it is not critical (e.g. prescription of two medicines with similar ingredients does not influence the occurrence of an allergy). More than one criterion could be detected. The outcome of preventability assessment will result in one of three possible scores: 'preventable', 'non-preventable' or 'not assessable'. When one or more critical criteria are met, it implies that a risk factor or factors contributed to the likelihood of the ADR, therefore the ADR is

Table 1 Preventability criteria in the P Method (PM)

Factors related to:	Preventability criteria	Yes	No	Unknown[a]	Not applicable[b]
Healthcare professionals' practices ('Pr')	1. Incorrect dose?				
	2. Incorrect drug administration route?				
	3. Incorrect drug administration duration?				
	4. Incorrect drug dosage formulation administered?				
	5. Expired drug administered?				
	6. Incorrect storage of drug?				
	7. Drug administration error (timing, rate, frequency, technique, preparation, manipulation, mixing)?				
	8. Wrong indication?				
	9. Inappropriate prescription according to the characteristics of the patient (age, sex, pregnancy, other)?				
	10. Inappropriate prescription for patient's underlying medical condition (renal failure, hepatic failure, etc.) or underlying pathology?				
	11. Documented hypersensitivity to administered drug or drug class?				
	12. Labelled drug–drug interaction?				
	13. Therapeutic duplication (prescription of two or more medicines with similar ingredients)?				
	14. Necessary medication not given?				
	15. Withdrawal syndrome (due to abrupt discontinuation of treatment)?				
	16. Incorrect laboratory or clinical monitoring of medicine?				
Product/drug ('Pd')	17. Poor-quality drug administered?				
	18. Counterfeit drug administered?				
Patient ('Pa')	19. Non-compliance?				
	20. Self-medication with non-OTC drug?				

ADR adverse drug reaction, *OTC* over-the-counter, *SmPC* summary of product characteristics

[a] A criterion that is not documented in the report form is considered unknown

[b] A criterion is deemed 'not applicable' if it is not involved in ADR occurrence according to the reviewer evaluation. The indication is evaluated according to the SmPC and the updated international or standard guideline in relation to drug use. The drug–drug interaction is evaluated according to the SmPC

categorized as 'preventable'. The ADR is deemed 'non-preventable' if none of the critical criteria are identified in the individual case safety report (ICSR). The case is categorized as 'not assessable' if there are no data or insufficient data for assessment (e.g. an anaphylactic reaction due to penicillin is deemed 'not assessable' if the patient's previous history of drug allergy is not documented) or if the situation is controversial (e.g. use of a drug that does not have a paediatric indication but is commonly used in children) (Fig. 1). An application of the PM is presented in Table 2.

2.2 Evaluation of the P Method

This was a multicentre study involving nine NPVCs (in Brazil, Iran, Morocco, New Zealand, Nigeria, Spain, Switzerland, Thailand and Tunisia), with the Moroccan PVC acting as the leader/coordination centre, executed within the framework of the MM Project [10]. The PVCs selected to participate in this study were already involved in patient safety activities (ME detection and prevention) [11]. For confidentiality reasons, a number was allocated to each NPVC (PVC1–PVC9).

2.2.1 *Data Collection*

Each PVC was required to send the Moroccan PVC a set of 30 randomly selected ICSRs: 10 ICSRs with ADRs to paracetamol in children, 10 with ADRs to anticoagulants and 10 with ADRs to non-steroidal anti-inflammatory drugs (NSAIDs). In total, 242 ICSRs were received from the nine PVCs, instead of the 270 that were requested. Overall, 67 ICSRs related to anticoagulants, 86 ICSRs related to NSAIDs and 89 ICSRs related to use of paracetamol in children. These 242 reports were sent back to all nine NPVCs for blinded preventability assessment applying the PM (Fig. 2). For that purpose, an assessment summary table (AST) was sent to the reviewers in order for

Table 2 Application of the P Method (PM), using the example of a patient who developed facial oedema in response to ibuprofen, after a history of four prior episodes of allergy with non-steroidal anti-inflammatory drugs (NSAIDs). First step: case causality assessment: the relationship is scored 'probable'; second step: the adverse drug reaction (ADR) is related to patient susceptibility; third step: it is a preventable ADR because the previous episodes of allergy with NSAIDs were documented (one critical criterion is identified)

Factors related to:	Preventability criteria	Yes	No	Unknown	Not applicable
Healthcare professionals' practices ('Pr')	1. Incorrect dose?				X
	2. Incorrect drug administration route?				X
	3. Incorrect drug administration duration?				X
	4. Incorrect drug dosage formulation administered?				X
	5. Expired drug administered?			X	
	6. Incorrect storage of drug?			X	
	7. Drug administration error (timing, rate, frequency, technique, preparation, manipulation, mixing)?				X
	8. Wrong indication?		X		
	9. Inappropriate prescription according to the characteristics of the patient (age, sex, pregnancy, other)?		X		
	10. Inappropriate prescription for patient's underlying medical condition (renal failure, hepatic failure, etc.) or underlying pathology?		X		
	11. Documented hypersensitivity to administered drug or drug class?	X			
	12. Labelled drug–drug interaction?				X
	13. Therapeutic duplication (prescription of two or more medicines with similar ingredients)?		X		
	14. Necessary medication not given?		X		
	15. Withdrawal syndrome (due to abrupt discontinuation of treatment)?		X		
	16. Incorrect laboratory or clinical monitoring of medicine?				X
Product/drug ('Pd')	17. Poor-quality drug administered?			X	
	18. Counterfeit drug administered?			X	
Patient ('Pa')	19. Non-compliance?			X	
	20. Self-medication with non-OTC drug?		X		

OTC over-the-counter

them to report, for each ICSR, their judgment process based on their use of the preventability criteria (see the electronic supplementary material).

2.2.2 *Quality of the Individual Case Safety Reports*

Qualitative information provided in the ICSR is essential for assessing both case causality and the preventability of the ADR [12]. Thus, on the basis of the quality of the provided information, the ICSRs were classified into two categories: 'documented' and 'non-documented'. The 'documented' category encompassed the ICSRs that provided relevant information to identify potential risk factors that may have influenced the likelihood of ADR occurrence. In addition to the previous items, we considered that information about a previous history of drug allergy or susceptibility was useful in evaluating the avoidability of hypersensitivity reactions. When present, the 'case narrative' section may have provided pertinent information about the circumstances of the ADR, allowing identification of a possible ME [13].

An ICSR was considered 'non-documented' when the available data were insufficient for assessing the known risk factors of the ADR.

2.2.3 *Causality Assessment*

The ADR causality assessment was performed by two experienced reviewers (physicians or pharmacists) at each participating PVC, according to the WHO-UMC causality assessment system [14].

2.2.4 *Assessment of Adverse Drug Reaction Preventability*

The same reviewers independently analysed the 242 ICSRs. These reviewers were trained to use the PM, during a meeting organized by the Moroccan PVC as part of the MM Project. The meeting was focused on identifying, analysing and preventing MEs. The definitions and approach were explained, and the preventability criteria were discussed. Each criterion in the scale was reviewed by the participants at a consensus session to assess whether it

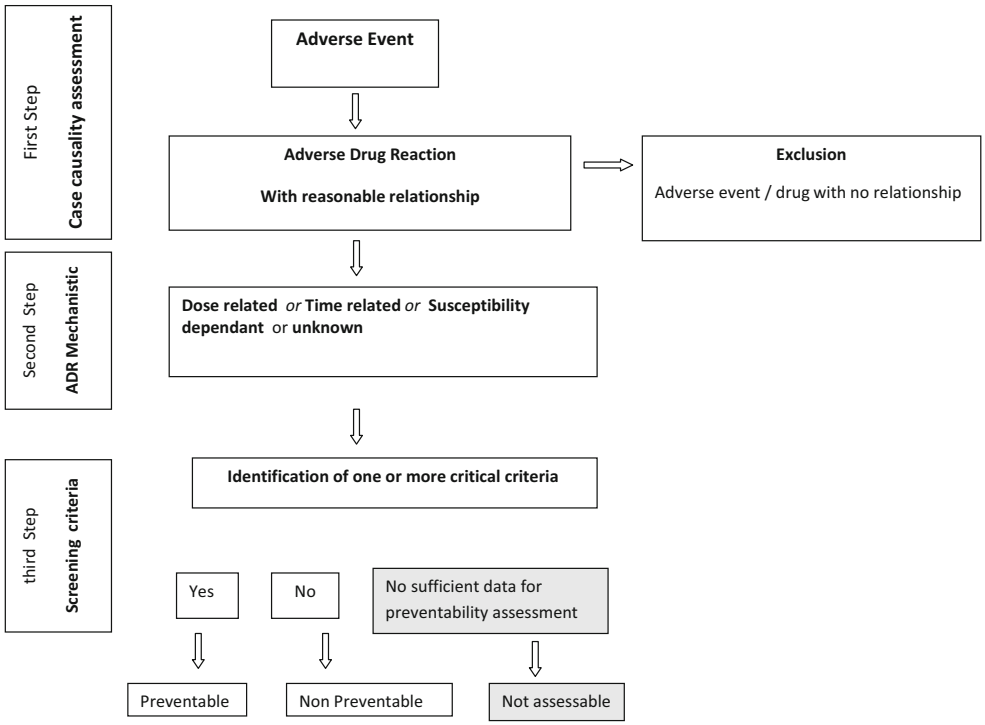

Fig. 1 Description of the process for assessing the adverse drug event, using the P Method (PM). *ADR* adverse drug reaction

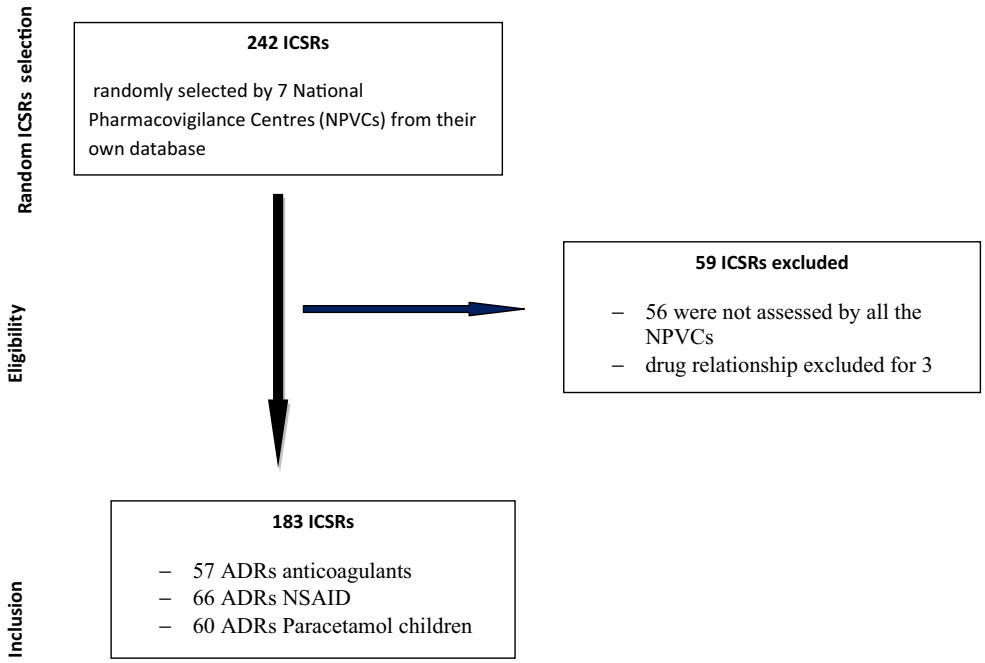

Fig. 2 Diagram flow for data collection and analysis. *ADR* adverse drug reaction, *ICSR* individual case safety report, *NPVC* national pharmacovigilance centre, *NSAID* non-steroidal anti-inflammatory drug

was appropriate to incorporate it or to discard it. Harmonization of the preventability criteria formulation with the WHO Adverse Reaction Terminology (WHO-ART) [15] and the Medical Dictionary for Regulatory Activities (MedDRA) terminology [16] was achieved in order to facilitate recording of the ICSR in VigiFlow (a web-based ICSR management system designed for sending reports from national centres to the WHO Programme for International Drug Monitoring) and to classify the types of MEs that are detected.

2.2.5 Data Analysis

For testing the reproducibility of the PM, we evaluated the inter-reviewer agreement. This evaluation relies on the percentage of agreement between the reviewers with regard to the outcome of assessment of ADR preventability. The inter-reviewer agreement was expressed as a kappa statistic with a 95 % confidence interval (CI) and as the percentage of ADRs for which there was agreement. The level of agreement between two reviewer pairs was analysed using the Cohen kappa test. The Fleiss–Cohen kappa test was used to study the level of agreement between the seven reviewer pairs. A kappa value between 0.00 and 0.20 was classified as 'slight', between 0.21 and 0.40 as 'fair', between 0.41 and 0.60 as 'moderate', between 0.61 and 0.80 as 'substantial' and between 0.81 and 1.00 as 'almost perfect' [17]. The chi-squared test was used to compare the kappa values (significant $p < 0.05$).

3 Results

Among the nine NPVCs enrolled in the study, seven sent back the preventability assessment and two were unable to meet the deadline. Of the 242 ICSRs that were initially received from the PVCs, only 183 were included in the inter-reviewer agreement analysis. Overall, 57 ICSRs were related to anticoagulants, 66 ICSRs concerned NSAIDs and 60 ICSRs were related to paracetamol. Among the 59 excluded ICSRs, 56 could not be assessed by all reviewers as they were written in French and three were categorized as 'unlikely' according to the WHO-UMC causality assessment gradation (Fig. 1). Therefore, only those ICSRs with a causality assessment of 'certain' (8.1 %), 'probable' (33.3 %) or 'possible' (58.5 %) were retained. The underlying mechanism of the ADRs was dose related (49.1 %) in approximatively half of the cases, susceptibility dependent in 28.4 %, time related in 0.5 % and unknown in 22 %. The documentation grading of the ICSRs revealed that 89 (48 %) were 'documented' and 94 (52 %) were 'non-documented'. In total, 32 (18 %) of the documented ICSRs contained a 'case narrative'. For 41 ICSRs (22.4 %), one risk factor was identified, and for 19 ICSRs (10.4 %), more than one risk factor was found (Fig. 3).

The mean number for each category of preventability was 74 (95 % CI 59–89) for 'preventable' ADRs, 72 (95 % CI 34.2–111) for 'non-preventable' ADRs and 37 (95 % CI −6 to 79.4) for 'not assessable' ADRs (Table 3).

All reviewers were in consensus for 17 ICSRs (9.3 %) that were scored as 'preventable' (Fig. 3). Six ICSRs were related to paracetamol, six to NSAIDs and five to anticoagulant drugs. Among them, the underlying mechanism was dose related in 15 cases, related to patient susceptibility in one case and time related in one case.

3.1 Level of Agreement Between All Reviewers

The overall agreement between all reviewers for assessment of ADR preventability was 'fair', with a kappa value of 0.27 (95 % CI 0.21–0.40). Regarding the three categories of preventability, we found that the overall levels of agreement were 'fair' for the 'preventable' category and 'slight' for the other categories (Table 4).

3.2 Level of Agreement Between All Reviewers for Each Medication

The overall agreement between all reviewers for assessment of ADR preventability related to anticoagulants and NSAIDs was 'fair', with kappa values of 0.28 (95 % CI 0.21–0.40) and 0.25 (95 % CI 0.21–0.40), respectively. However, the agreement was 'slight' for ADRs related to paracetamol, with a kappa value of 0.18 (95 % CI 0.01–0.20).

3.3 Level of Agreement Between Pairs of Reviewers

We compared the results obtained by each NPVC with the Moroccan PVC assessment. The level of agreement between each reviewer pair and the Moroccan PVC ranged from 'slight', with a kappa value of 0.12 (95 % CI −0.03 to 0.27), to 'substantial', with a kappa value of 0.69 (95 % CI 0.48–0.89) (Table 5).

Regarding the three categories of preventability, for the 'preventable' category, the level of agreement between the pairs of reviewers and the Moroccan PVC ranged from 'fair', with a kappa value of 0.28 (95 % CI 0.21–0.35), to 'perfect', with a kappa value of 0.81 (95 % CI 0.76–0.85). The estimated kappa values for the 'non-preventable' category demonstrated 'slight' agreement, with a kappa value of 0.002 (95 % CI −0.001 to 0.019), to 'substantial' agreement, with a kappa value of 0.61 (95 % CI 0.55–0.66). For the 'not assessable' category, the level of agreement was main ly 'slight' (Table 5).

3.4 Level of Agreement According to the Quality of Individual Case Safety Reports

We found that the 30 ICSRs provided by PVC3 were categorized as 'documented'. The level of agreement between all reviewers when assessing those ICSRs was 'moderate', with a kappa value of 0.41 (95 % CI 0.40–0.60).

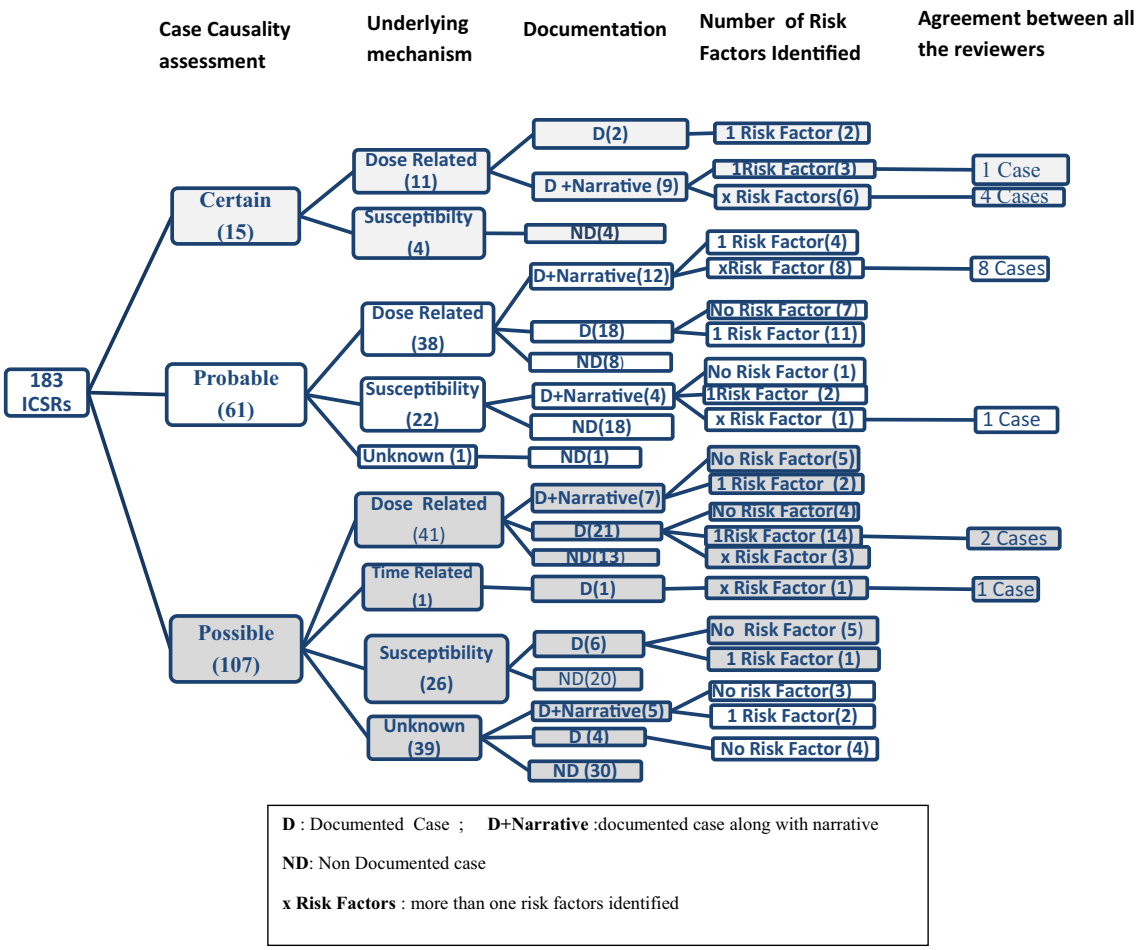

Fig. 3 Analysis of individual case safety report (ICSR) characteristics. *D* documented case, *D + Narrative* documented case along with narrative, *ND* non-documented case, *x Risk Factor* more than one risk factor identified

Table 3 Assessment of preventability of adverse drug reactions (ADRs)

NPVC	ADRs [*n*]		
	Preventable	Non-preventable	Not assessable
NPVC1	71	85	27
NPVC2	57	120	8
NPVC3	68	91	24
NPVC4	68	41	74
NPVC5	98	85	0
NPVC6	92	83	8
NPVC7	64	3	116

NPVC national pharmacovigilance centre

4 Discussion

The main outcome of the present study was to assess the reliability of a new method to detect preventable ADRs. We found that the level of agreement between all reviewers was 'fair' when applying the PM. This is in agreement with similar studies that used explicit criteria consisting of statements indicating that an ME was present. Indeed, Kunac et al. [18] reported 'fair' agreement between three reviewers when using the scale of Schumock and Thornton [19], with a kappa value of 0.37 (95 % CI 0.33–0.41). In the Schumock and Thornton algorithm, which is one of the most widely used tools, preventability is classified dichotomously—'preventable' versus 'not preventable'—on the basis of seven statements concerning possible contraindications, inappropriate dosage, inappropriate therapeutic drug monitoring, not considering previous allergic reactions to the drug, drug interactions, toxic serum drug concentrations and compliance. If at least one statement was valid, the case was deemed preventable. Moreover, Olivier et al. [5] reported 'fair' agreement between four reviewers when applying the standardized French preventability scale proposed by Imbs et al. [20], reporting a kappa value from 0.14 to 0.58. The scale of Imbs et al. contains numerous items concerning knowledge about ADRs to a drug and communication of this knowledge, the risk factors of the patient, drug management, the conditions of prescription and management of the ADR. Each item is

Table 4 Inter-rater agreement for all reviewers

	Frequency reported by each PVC's reviewers [*n* (%)]								
	RC1	RC2	RC3	RC4	RC5	RC6	RC7	Kappa value (95 % CI)	*p* value
Preventability category									
Preventable	71 (39.0)	57 (31.0)	68 (37.2)	68 (37.2)	98 (53.5)	92 (50.3)	64 (35.0)	0.38 (0.356–0.403)	$<10^{-5}$
Non-preventable	85 (46.4)	120 (65.6)	91 (49.7)	41 (22.4)	85 (46.4)	83 (45.2)	3 (1.6)	0.19 (0.167–0.212)	$<10^{-4}$
Not assessable	27 (14.8)	8 (4.4)	24 (13.1)	74 (40.4)	0	8 (4.4)	116 (63.4)	0.002 (−0.006 to 0.010)	$<10^{-5}$
All categories	0.27 (0.21–0.40)								$<10^{-5}$

CI confidence interval, *PVC* pharmacovigilance centre, *RC* reviewers from the participating NPVC, *RC1* reviewers from the NPVC1

scored, and the range of the global score would be −11 to +18. A positive score suggests that an ADR was potentially preventable, and a negative score suggests that it was potentially not preventable.

A recent systematic review, evaluating the methods assessing the preventability of adverse drug events, underlined that their reliability is not often tested [21]. While comparing the reliability of those tools through published articles, the authors of that review found that implicit methods resulted in excellent reliability, in contrast with explicit algorithms. These results were unexpected, since explicit methods are based on predefined criteria, which are easily reproducible [22], whereas the measurement procedure in implicit review requires reviewers to form their criteria and apply them, and thus a source of variability is included in the measurement of reliability.

Considering that the external reliability of the PM is a crucial element for full assessment of the utility of this new tool for international use in the WHO Programme or elsewhere, we would like to underline that in comparison with other studies carried out in one specific place [5, 18], our study yielded substantial results with respect to the judgment of ADR preventability, since it was a multicentre study involving nine NPVCs with different backgrounds and expertise. This method was developed in the framework of the MM Project and represented one component of the guideline *Reporting and Learning Systems for Medication Errors: The Role of Pharmacovigilance Centres*, issued by the WHO in November 2014 [23].

The PM's external reliability was not as high as we expected. In order to investigate specific points of agreement and disagreement between reviewers, we undertook a descriptive analysis of the cases regarding their causality assessment, underlying mechanism, documentation and presence/absence of any risk factor that might have contributed to their occurrence (Fig. 3). We also analysed the AST, where the reviewers reported their judgment process for each ICSR, when assessing the 20 criteria (see the electronic supplementary material).

Therefore, we found that overall agreement between all reviewers was more likely to be reached when the case causality assessment was 'certain' or 'probable' (14 versus 3; $p < 0.001$), the ICSRs were documented along with a case narrative (14 versus 3; $p < 0.001$) or the underlying mechanism of the ADR was known and more than one risk factor was identified (13 versus 4; $p < 0.001$). These findings confirmed that a known ADR with a documented ICSR facilitated the preventability assessment.

On the other hand, three main sources of disagreement were identified. The first one was related to misinterpretation of the scores for 'non-preventable' versus 'not assessable'. From the 88 documented cases, no risk factors were identified by the reviewers for 16 cases, suggesting that they should be scored as 'non-preventable', whereas some reviewers categorized them as 'not assessable'. In the same way, among the 95 non-documented cases, no risk factors were identified by the reviewers for 22 cases. These cases were categorized as 'not assessable' by some reviewers (in concordance with the PM approach) and as 'non-preventable' by the others (Table 3).This was confirmed by the high standard deviation of the average number of the 'non-preventable' and 'not assessable' categories. In fact, for both categories, the proportions found by the seven PVCs were spread out over a large range of values. In this field, some authors suggest that having fewer categories than reviewers' capability to discriminate may result in loss of information, which may lead to reduced reliability [22]. After reduction of the three preventability categories to two categories ('preventable' versus 'non-preventable'), which was achieved by merging the two categories 'non-preventable' and 'not assessable', agreement between the reviewers improved to 'moderate', with a kappa value of 0.46 (95 % CI 0.40–0.60).

'Substantial' agreement between PVC2 and the Moroccan reviewers was observed, presumably because one member of the PVC2 team had had an opportunity to train on the PM for 1 week. This implies that reviewers need more practice to become familiar with the definitions

Table 5 Level of agreement between reviewer pairs

	Reliability between reviewer pairs: kappa value (95 % CI) and *p* value											
	RC1 vs RC2	*p* value	RC1 vs RC3	*p* value	RC1 vs RC4	*p* value	RC1 vs RC5	*p* value	RC1 vs RC6	*p* value	RC1 vs RC7	*p* value
Preventability category												
Preventable	0.81 (0.76–0.85)	10^{-3}	0.28 (0.21–0.35)	10^{-3}	0.35 (0.28–0.42)	10^{-3}	0.35 (0.29–0.42)	10^{-3}	0.31 (0.24–0.38)	10^{-3}	0.46 (0.40–0.53)	10^{-3}
Non-preventable	0.61 (0.58–0.66)	10^{-3}	0.30 (0.23–0.37)	10^{-3}	0.10 (0.48–0.89)	10^{-3}	0.25 (0.16–0.30)	10^{-3}	0.23 (0.16–0.30)	0.002	0.002 (0.0–0.01)	NS
Not assessable	0.24 (0.14–0.33)	10^{-3}	0.00	NS	0.07 (0–0.13)	NS	0.00	NS	0.12 (0–0.13)	0.03	0.09 (0.05–0.13)	0.02
All categories	0.69 (0.48–0.89)	10^{-3}	0.37 (0.17–0.57)	NS	0.23 (0.10–0.36)	0.002	0.22 (0.08–0.36)	NS	0.16 (0.05–0.38)	10^{-3}	0.12 (0.0–0.27)	10^{-3}

CI confidence interval, *NS* non-significant, *RC* reviewers from the participating NPVC, *RC1* reviewers from the NPVC1

and better definition of what should be included in the term 'non-assessable'.

The second source of disagreement related to the reviewers' knowledge of standards of medication use. Assessment of ADR preventability requires up-to-date knowledge of standards of medication use. Of the 36 dose-related ADRs where a unique risk factor was identified, drug–drug interactions were implicated in 23 cases. In this field, it is important to underline that national attitudes to various classes of drugs could be different and have evolved over the years. An example of this evolution is the prescription of vitamin K anticoagulants in atrial fibrillation patients at risk of stroke, who are already receiving antiplatelet drugs despite the increased risk of bleeding [24]. Opinions about some drugs are controversial. In some countries, metamizol was withdrawn in the 1970s because of the risk of agranulocytosis, whereas it is still widely used in other countries [25, 26].

Individual judgment represented a significant factor of disagreement between reviewers. Two examples illustrating this point are as follows:

- Case 1: An infant presented with excessive drowsiness 2 h after a single dose of paracetamol (50 mg/kg) administered by her mother because "she was crying too much" and so her mother thought the infant might be in pain. This case was assessed by some reviewers as 'preventable' because they considered that the indication was wrong. Others considered the dose of paracetamol incorrect, whereas some reviewers deemed it 'not assessable' because some criteria were not recorded.
- Case 2: An infant presented with a skin rash 3 h after administration of ketoprofen for a cold. Some reviewers deemed this case 'preventable' because it was the wrong indication. Others categorized it as 'not assessable' because the previous history of drug allergy was not documented.

The first case concerned a non-described ADR with an unknown underlying mechanism, which led to different interpretations by the assessors, whereas the second case was related to an allergic ADR for which a previous history of drug allergy would represent a well-known risk factor. These cases outline that assessment of ADR preventability should be extended beyond the strict confines of exploring known risk factors. Therefore, the exploration should include patient and drug-related characteristics, healthcare professionals' practices and also environmental factors.

4.1 Limitations

Our study had some limitations. Because it was a retrospective ADR record study, the reporters could not provide

some of the relevant information, as some incidents had happened a long time ago. Therefore, we could not complete the case history and information that may have been lacking in the report form. Moreover, although this was not clearly stated for all participating PVCs, it seems likely that for some PVCs, the personnel who submitted the reports were those who subsequently analysed them, using the PM. This might have constituted bias due to familiarity of the assessors with their own reports.

Also, the reviewers did not have an opportunity to discuss and reconsider their review to obtain a consensus. The validity of the method was not evaluated, since there is no gold-standard method with which the current results could be compared. Another limitation was that there is a range of experience between PVCs regarding assessment of ADR preventability. Some PVCs, such as those in Brazil, Iran, Morocco, New Zealand and Switzerland, were already involved in this field and had developed their own database on MEs [11].

5 Conclusion

The analysis of the agreements and disagreements between reviewers using the PM highlighted where improvements might be made. It was shown that assessment of ADR preventability is optimum in case reports where there is reasonable link documented between the reaction and the suspected drug, including the case narrative and sufficient information to more completely explore patient and drug-related characteristics. The results were also affected by the assessors' value judgments in matters such as individual interpretation, up-to date knowledge of standards of medication use, healthcare professionals' practices and also environmental factors. It is essential to refine the preventability subcategories in order to avoid reviewer misinterpretation, suggesting that more training should be scheduled for reviewers in the future.

The results of this study showed that a fair to moderate consensus is possible for a tool that is practical to use and explicit in its outcomes. It is important to emphasize that none of the previous methods developed for detecting MEs have demonstrated a high level of inter-rater agreement. Some of these methods have been widely used, and that has led to their improvement. We believe that, similarly, the dissemination of this method could lead to its improvement. It therefore seems reasonable to propose use of the PM more widely among national centres within and outside the WHO Programme.

Acknowledgments The authors would like to thank their colleagues from the national pharmacovigilance centres for their valuable contributions to this work: Luciana Soubihe and Eliane Gil Rodrigues de Castro from Brazil; Gloria Shalviri and Naghmeh Javidnikou from Iran; Uchenna Elemuwa, Cassandra A. Elagbaje, Desirée Kunac and Michael Tatley from New Zealand; Mariano Madurga and Cristina Fernandez from Spain; Alexandra Raetz Bravo and Wolfgang Renftl from Switzerland; Pattreya Pokhagul and Watcharee Rungapiromnan from Nigeria; Mohamed lakhal and Sihem El Aïdli from Tunisia; and Fatima Abadi, Driss Soussi Tanani and Souad Skalli from Morocco. The authors would also like to thank Wendy Shepherd for proof-reading the text.

Funding This project was funded by the EU-FP7 Monitoring Medicines Project (grant number 223566). The funding organization (the European Commission) had no role either in the conduct of the present study or in the writing of the present manuscript.

Conflicts of interest Raja Benkirane, Rachida Soulaymani-Bencheikh, Asmae Khattabi, Ghita Benabdallah, Loubna Alj, Houda Sefiani, Khedidja Hedna, Lahcen Ouammi, Sten Olsson and Shanti N. Pal have no conflicts of interest that are directly relevant to the content of this study.

Theme issue This article is part of a theme issue co-edited by Elliot G. Brown, Shanthi Pal and Sten Olsson. No external funding was used to support the publication of this theme issue.

References

1. Alj L, Touzani MDW, Benkirane R, Edwards IR, Soulaymani R. Detecting medication errors in pharmacovigilance database: capacities and limits. Int J Risk Saf Med. 2007;19(4):187–94.
2. Kunac DL, Tatley MV. Detecting medication errors in the New Zealand pharmacovigilance database: a retrospective analysis. Drug Saf. 2011;34(1):59–71.
3. Caduff-Janosa P. Medication errors in Switzerland: an analysis of the spontaneous adverse reactions reports received at the Swissmedic Pharmacovigilance Centre. Drug Saf. 2004;27:831–977.
4. Bencheikh RS, Benabdallah G. Medication errors: pharmacovigilance centres in detection and prevention. Br J Clin Pharmacol. 2009;67(6):687–90.
5. Olivier P, Boulbés O, Tubery M, Lauque D, Montastruc JL, Lapeyre-Mestre M. Assessing the feasibility of using an adverse drug reaction preventability scale in clinical practice: a study in a French emergency department. Drug Saf. 2002;25(14):1035–44.
6. Desai RJ, Williams CE, Greene SB, Pierson S, Caprio AJ, Hansen RA. Analgesic medication errors in North Carolina nursing homes. J Pain Palliat Care Pharmacother. 2013;27(2):125–31.
7. Rajanayagam J, Bishop JR, Lewindon PJ, Evans HM. Paracetamol-associated acute liver failure in Australian and New Zealand children: high rate of medication errors. Arch Dis Child. doi:10.1136/archdischild-2013-304902. (Epub 2014 Sep 16).
8. Ferner RE, Aronson JK. Preventability of drug-related harms—part I: a systematic review. Drug Saf. 2010;33(11):985–94.
9. Aronson JK, Ferner RE. Preventability of drug-related harms—part II: proposed criteria, based on frameworks that classify adverse drug reactions. Drug Saf. 2010;33(11):995–1002.
10. Uppsala Monitoring Centre. Monitoring medicines. 2014. www.monitoringmedicines.org. Accessed 16 Mar 2014.
11. Benabdallah G, Benkirane R, Khattabi A, Edwards IR, Soulaymani Bencheikh R. The involvement of pharmacovigilance centres in medication errors detection: a questionnaire-based analysis. Int J Risk Saf Med. 2011;23(1):17–29.

12. UMC presents 'ICSRs and VigiBase—the vital importance of quality'. 9 March 2012. http://www.who-umc.org/DynPage.aspx?id=108464&mn1=7347&mn2=7489&mn3=7248&newsid=10762. Accessed 30 Nov 2012.
13. Morimoto T, Gandhi TK, Seger AC, Hsieh TC, Bates DW. Adverse drug events and medication errors: detection and classification methods. Qual Saf Health Care. 2004;13:306–14.
14. Uppsala Monitoring Centre. The use of the WHO-UMC system for standardised case causality assessment. http://www.who-umc.org/Graphics/24734.pdf. Accessed 20 Jan 2013.
15. Uppsala Monitoring Centre. WHO-ART—WHO adverse reaction terminology. http://www.umc-products.com/DynPage.aspx?id=73589&mn1=1107&mn2=1664. Accessed 10 Jan 2012.
16. MedDRA. Medical dictionary for regulatory activities. http://www.meddra.org. Accessed 10 Jan 2012.
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–74.
18. Kunac DL, Reith DM, Kennedy J, Austin NC, Williams SM. Inter- and intra-rater reliability for classification of medication related events in paediatric inpatients. Qual Saf Health Care. 2006;15:196–201.
19. Schumock GT, Thornton JP. Focusing on the preventability of adverse drug reactions. Hosp Pharm. 1992;27:538.
20. Imbs JL, Pletan Y, Spriet A, et al. Evaluation de la iatrogénèse médicamenteuse évitable: méthodologie. Therapie. 1998;53:365–70.
21. Hakkarainen KM, Andersson Sundell K, Petzold M, Hägg S. Methods for assessing the preventability of adverse drug events: a systematic review. Drug Saf. 2012;35(2):105–26.
22. Zegers M, de Bruijne MC, Wagner C, et al. The inter-rater agreement of retrospective assessments of adverse events does not improve with two reviewers per patient record. J Clin Epidemiol. 2010;63(1):94–102.
23. World Health Organization. Reporting and learning systems for medication errors: the role of pharmacovigilance centres. 2014. http://apps.who.int/iris/bitstream/10665/137036/1/9789241507943_eng.pdf?ua=1. Accessed 28 Nov 2014.
24. Lip GY, et al. Bleeding risk assessment and management in atrial fibrillation patients: a position document from the European Heart Rhythm Association, endorsed by the European Society of Cardiology Working Group on Thrombosis. Europace. 2011;13:723–46.
25. Ibanez L, Vidal X, Ballarin E, Laporte JR. Population-based drug-induced agranulocytosis. Arch Intern Med. 2005;165(8):869–74.
26. Andrade SE, Martinez C, Walker AM. Comparative safety evaluation of non-narcotic analgesics. J Clin Epidemiol. 1998;51(12):1357–65.