University of Groningen

A tailor-made approach for causality assessment for ADR reports on drugs and vaccines

Oosterhuis, Ingrid; Zweers, Petra; Rümke, Hans; Muller-Hansma, Annemarie; van Puijenbroek, Eugène Paul

Published in:
Pharmacoepidemiology and Drug Safety

DOI:
[10.1002/pds.4637](https://doi.org/10.1002/pds.4637)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Oosterhuis, I., Zweers, P., Rümke, H., Muller-Hansma, A., & van Puijenbroek, E. P. (2019). A tailor-made approach for causality assessment for ADR reports on drugs and vaccines. *Pharmacoepidemiology and Drug Safety*, 28(4), 544-550. <https://doi.org/10.1002/pds.4637>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

ORIGINAL REPORT

A tailor-made approach for causality assessment for ADR reports on drugs and vaccines

Ingrid Oosterhuis¹ | Petra Zweers¹ | Hans Rümke¹ | Annemarie Muller-Hansma¹ | Eugène Paul van Puijenbroek^{1,2}

¹Netherlands Pharmacovigilance Centre
Lareb, 's-Hertogenbosch, the Netherlands

²Groningen Research Institute of Pharmacy,
Unit of Pharmacotherapy, Epidemiology &
Economics, University of Groningen,
Groningen, The Netherlands

Correspondence

I. Oosterhuis, Netherlands Pharmacovigilance
Centre Lareb, Goudsbloemvallei 7, 5237 MH,
's-Hertogenbosch, The Netherlands.
Email: i.oosterhuis@lareb.nl

Abstract

Purpose: To estimate causation of adverse drug reaction (ADR) reports, causality methods were developed from a theoretical perspective. In daily practice, not all information is relevant or available, decreasing the applicability. We developed a new causality documentation tool (CausDoc) where an algorithm is combined with expert judgement. The aim of this study is to test the validity and reliability of CausDoc for ADR reports on drugs and vaccines.

Methods: CausDoc provides 9 structured relevant questions. If information is available, an answer will be chosen. If not, the question is excluded. Causality outcome is based on the sum score of all answers divided by the included questions: $\leq 30\%$: unlikely, 31% to 70%: possible, 71% to 90%: probable, and $>90\%$: certain. Other relevant information is taken into account by expert judgement in the final step by adjusting the outcome to a limited extent. After testing face validity on 12 ADR reports, sensitivity and specificity were tested on 40 ADR reports, compared with the Naranjo algorithm and WHO AEFI criteria, using the expert panel's judgements as a standard. Inter-rater reliability was tested using weighted Cohen kappa coefficient.

Results: Average sensitivity and specificity with CausDoc were 47% and 83% for drugs (29% and 78% with Naranjo) and 72% and 89% for vaccines (65% and 87% with WHO AEFI criteria). Reliability between the 2 couples of assessors: κ 0.48 and 0.75.

Conclusions: CausDoc shows a better performance and allows for a better documentation of ADRs in clinical practice. This approach is useful in assessing the causality of adverse drug reactions.

KEYWORDS

adverse drug reactions, causality assessment, drug safety, drugs, pharmacoepidemiology, pharmacovigilance, vaccines

1 | INTRODUCTION

When a patient experiences a suspected adverse drug reaction (ADR), the question raises how likely it is that the drug actually causes the

reaction in this specific patient.¹ Causality assessment helps to answer this question and reflect the process of estimating the likelihood or causal relationship.² It has similarities with the differential diagnostic process, where other less likely probabilities are eliminated.^{2,3}

Prior postings and presentations

The abstract of this manuscript was presented as a poster at the International Society of Pharmacovigilance (ISoP): Oosterhuis, I, Zweers, PGMA, Rümke, HC, Hansma, AHG. A combined causality method for spontaneous reports of adverse reactions on drugs and vaccines. *Drug Safety* 2015; 38 (10): 1017.

Bradford Hill⁴ discussed 9 elements that should be considered in determining the strength of the likelihood of a causal relationship, but it depends on the specific case if an element is relevant. In pharmacovigilance, multiple causality methods were developed over time to estimate the likelihood of a causal relationship. Frequently used methods are based upon expert judgement (for example, the criteria for Adverse Events Following Immunization (AEFI) from WHO⁵), algorithms like the Naranjo algorithm,⁶ or upon probabilistic methods like the French method.⁷ Three main issues are important in performing a proper causality assessment of a reported ADR:

1.1 | Extrinsic and intrinsic causality

A distinction must be made between extrinsic factors in the assessment (can the drug cause the ADR?) and intrinsic factors (has the drug actually caused the event in this given patient?).¹ For extrinsic factors, information can be obtained from scientific literature. For intrinsic causality, information should be gathered from the report itself. Intrinsic factors are more prone to be judged differently by pharmacovigilance assessors because the available information is often limited and therefore more prone to subjective interpretation, which may lead to a different causality outcome and thus interobserver variability.^{3,8-10} All causality methods take extrinsic and intrinsic factors into account to determine the strength of the relationship. Items to be assessed should both be relevant and of sufficient quality. One should also determine if the information is actually available or not. The relationship between extrinsic and intrinsic factors and relevance and availability of the latter category is shown in Figure 1. We will elaborate on both relevance and availability because these aspects are key elements in the assessment of the causal relationship.

1.2 | Relevance

Mainly for the assessment of intrinsic factors, one must determine which items are of relevance for the assessment of the causal relationship, as drawn in Figure 1. Moreover, in clinical practice, some information has relatively more or less weight in contributing to the strength of the causal relationship. Information that is not relevant may distort causality assessment, if it is nevertheless taken into account when this is required by the specific method that is being used. If for example a reported reaction is by its nature irreversible (for example in case of an acute cardiac death following administration of methylphenidate), a dechallenge or rechallenge is not applicable. However, in most algorithms, information on dechallenges and rechallenges has to be taken into account. As a consequence, these factors are labelled as unknown. This will distort causality assessment and increase uncertainty, for instance leading to a "possible" causal relationship instead of a "certain" relationship. The fact that information should be taken into account that is not applicable or not relevant for a given situation is a major drawback of many algorithms, including for instance the Naranjo algorithm.¹⁰

1.3 | Documentation grade (availability)

For every relevant item, information can either be available or not available (see Figure 1). In the latter situation, this can be labelled as

KEY POINTS

- In causality assessment, 3 main issues are important: extrinsic and intrinsic causality, relevance, and documentation grade.
- Most existing causality methods are using nonrelevant or not available information, leading to a less precise causality outcome, or cannot be used to attribute weights to the various items.
- With a new developed causality approach, an algorithm is combined with expert judgement.
- In this approach, causality is calculated based only on relevant and available items and gives transparency in the weights of the items.
- This approach shows better performance as compared to the Naranjo algorithm and WHO AEFI criteria.

"unknown", whereas in case of nonrelevant information, this should not be taken into account in the final causality assessment. In an ideal situation, a report is of high quality, with objective clinical data and information about dechallenge or rechallenge.¹¹ In daily practice, most reports are not that unambiguous and (clinical) information is lacking.² Most causality methods assume for example that there is enough information about the exact latency time and presence or absence of a rechallenge. However, as an example, determining the exact date on which alopecia actually started is often not realistic. A high (clinical) documentation grading increases the accuracy of the causality outcome. As a result, it may not be possible to rate the causal relationship as "highly likely" or even certain when because of a poor documentation, essential information is lacking.

1.4 | The use of various types of information in existing models for causality assessment

Probabilistic methods, and also algorithms, are using also nonrelevant information and not available information in determining the causal outcome. It is not possible to exclude this information because of irrelevancy. When using expert judgement (Figure 1), the assessor determines which information is relevant and causality assessment is based on relevant information only, irrespective if these elements are reported or not reported. Irrelevant information is not taken into account. However, expert judgement cannot be used to attribute weights to the various items for a large variety of situations, but may be a useful adjunct to causality assessment.

Although there are differences in using expert judgement, algorithms, or probabilistic methods, none of the current approaches will be applicable for all different types of situations.⁸⁻¹⁰ Most causality methods are developed from a theoretical perspective, but in daily practice, not all information is relevant or available. Causality assessment therefore needs a more tailor-made approach, where assessment, documentation, and transparency are combined.

FIGURE 1 Relationship between extrinsic/intrinsic causality, relevance, and documentation grade and the use of causality. Whereas probabilistic methods and causality algorithms also take nonrelevant intrinsic factors into account, the information used in expert judgement and the newly developed causality documentation tool is limited to extrinsic factors and relevant intrinsic factors

Until recently, the Netherlands Pharmacovigilance Centre Lareb used the WHO AEFI criteria⁵ (for vaccine reports) and the Naranjo algorithm⁶ (for drug reports) for causality assessment. We combined these methods into a flexible causality method in which a standard set of relevant items are included with quantifiable information, but with the flexibility of expert judgement. Nonrelevant items are not taken into account. In this model (Figure 2), a set of most relevant questions is presented, based on practical experience. For information that is only relevant in specific cases, eg, toxic drug levels, expert judgement can be taken into account for the final causality outcome. In addition, for every available item, it can be decided if it has relatively more or less weight in contributing to the strength of the causal relationship, leading to transparency in the choices made during causality assessment. This approach is perceived as being a more closer reflection of the process of clinical decisions as made in clinical practice. As far as we know, this has never been applied in existing models. However, we did not study the performance of this new approach compared to the currently used approaches. The aim is to test the validity and reliability of the causality documentation tool (CausDoc) for ADR reports on drugs and vaccines.

2 | METHODS

The causality documentation tool provides a structured approach with 9 relevant fixed questions regarding intrinsic and extrinsic causality (Figure 2). The questions concerning extrinsic causality are mandatory. For every question regarding intrinsic causality, the assessor decides if this information is available in the reported drug-event association (column 1). If so, the answer best fitting the specific situation of the patient will be chosen (column 2) and the question contributes to the overall assessment. If no information is available, the question is excluded in determining the causality outcome. The overall causality outcome is based on the sum score of all answers given divided by the number of all questions on which information is available. This relates to corresponding causality outcome categories of the score to be achieved: $\leq 30\%$: unlikely, 31% to 70%: possible, 71% to 90%: probable, and $>90\%$: certain. The limits of these categories were determined in the development process of the model by the expert panel, based on their experiences with test cases.

If other relevant information which is not present in the 9 questions is nevertheless considered to be important, this information can be taken into account by manual adjustment in the final step of the assessment. When, based on expert judgement, a solid argument

for disagreement with the causality outcome exists, adjustment to 1 lower or higher outcome category is allowed. The reason for this adaptation has to be justified. For further instruction on how to use CausDoc, see Electronic Supplementary Material number 1: User instruction CausDoc.

We performed a validation study using the dataset of the Netherlands Pharmacovigilance Centre Lareb using ADR reports submitted in the period from January 1 until November 30 2014.

A report can refer to multiple drugs and multiple ADRs, which can be strongly related to each other. An example is reporting both headache and abdominal pain with an underlying diagnosis of influenza possibly associated with the use of adalimumab. This requires 2 possible drug-event combinations which are closely related. To avoid similar causality rating of these related drug-event combinations in this study, we only rated the first drug-event combination in every report.

2.1 | Validity and reliability testing

For validity and reliability testing, reports were selected at random from our database. They were selected by the expert panel and judged with a diversity of causality outcomes from “unlikely” to certain. The expert panel consists of 3 medical doctors (PZ, HR, AM) and 1 pharmacist (IO), all with broad experience in pharmacovigilance.

Face validity of this causality method (CausDoc) was tested by 2 couples of experienced pharmacovigilance assessors using 12 ADR reports in total (6 reports per couple).

After adjustment of the causality method, further testing was performed by assessing a total of 40 ADR reports (20 reports on drugs and 20 reports on vaccines). These reports, with all various causality outcomes, were selected from our database by the expert panel. Both couples scored 10 reports on drugs and 10 reports on vaccines using the new causality method, and both causality methods currently used in our organization. The degree of agreement between the assessors was calculated using weighted Cohen kappa coefficient (κ).¹² For the κ , the following standards for strength of agreement are used: slight (≤ 0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), and almost perfect (0.81–1.0).¹² Outcomes from both methods were compared, and agreement was reached using the Delphi method.¹³ Sensitivity and specificity of the final causality outcome with CausDoc and the currently used causality methods were calculated to test the validity, using the expert panel's judgements as a gold standard for the categories to be assigned (unlikely, possible, probable, and certain). The expert panel scored all ADR reports

Factor	Information available?	Question	Answer to question
Extrinsic	2 points (mandatory)	1. Is there previous information on this association in literature or SmPC ?	Well known: 2 Little/ not known: 1
Extrinsic	2 points (mandatory)	2. Is there a plausible mechanism / relationship of this reaction with this (similar) drug (incl the device)?	Yes: 2 No/ don't know: 1 Contradictory: 0
Extrinsic	1 point (mandatory)	3. Is there a high background incidence of this event in general in this group (age, gender)?	Yes: 0 No/ don't know: 1
Intrinsic	Yes, 3 points No, NA	4. Is there a plausible latency between the administration of the drug and the reaction?	Very plausible: 3 Plausible: 2 Unknown/ possible: 1 Pleads against: 0
Intrinsic	Yes, 2 points No, NA	5. Did the reversible adverse reaction improve after a (plausible) time interval after discontinuation of the drug or after single administration?	Plausible: 2 Possible: 1 No: 0
Intrinsic	Yes, 3 points No, NA	6. Did the adverse reaction reappear when the drug was re-administered or did a similar adverse reaction appear in the past after the drug was administered?	Yes: 3 (drugs) Yes: 2 (vaccines) No: 0
Intrinsic	Yes, 2 points No, NA	7. Was the reaction more severe when the dose was increased or less severe when decreased? (Vaccines: NA)	Yes: 2 No: 0
Intrinsic	Yes, 2 points No, NA	8. Did the patient have a similar reaction to a similar drug in any previous exposure?	Yes: 2
Intrinsic	Yes, 1 point No, NA	9. Are there alternative known causes in this patient that on their own could have caused the reaction (e.g. disease, concomitant medication)?	Yes: 0 No/ not sure: 1
	Total points column 1		Total points column 2
	Causality outcome = total points column 2/ total points column 1 * 100% ≤30%: unlikely 71-90%: probable 31-70%: possible >90%: certain		

If necessary, causality outcome can be adjusted based on expert judgement with one outcome level lower or higher, provide with argumentation.

NA = not applicable

SmPC = Summary of product characteristics

FIGURE 2 The new causality method, CausDoc, for ADR reports on drugs and vaccines

individual for causality based on expert judgement, and discussed different outcomes until agreement was reached.

3 | RESULTS

3.1 | Validity and reliability testing

Face validity testing showed that this approach is transparent, easy to use, and mostly reflects the impression of the assessor. Some of the

questions or answer options were not clear: Minor textual changes were made. Testing of 40 ADR reports with the causality documentation tool showed a higher average sensitivity for reports on drugs and vaccines compared to the conventional causality methods used in our centre. Specificity was almost the same for reports on vaccines and slightly higher for reports on drugs (Table 1). Reliability between the assessors demonstrated a weighted Cohen kappa (κ) of 0.48 and 0.75.

For 26 out of the 40 ADR reports (65%), the scores from CausDoc of the couples corresponded to the judgement of the expert panel. In

TABLE 1 Sensitivity and specificity testing compared with expert judgement with reports on drugs and vaccines with CausDoc

Causality Method	Sensitivity	Specificity
Drugs (n = 20)		
Naranjo	29%	78%
CausDoc	47%	83%
Vaccines (n = 20)		
AEFI WHO	65%	87%
CausDoc	72%	89%

the case of different scores, CausDoc scored in 7 ADR reports higher and in 7 ADR reports lower, compared to the judgement of the expert panel (Table 2). Only in 1 ADR report the expert panel and CausDoc differed 2 categories. This drug-event association concerned ear pain because of drug substitution of sildenafil. The assessor using CausDoc only scored the event ear pain, which was scored as "probable", while the expert panel scored the ear pain because of drug substitution and considered this as unlikely.

In general, the new approach was found easy to use. When difficulties with completing CausDoc were mentioned, these concern mostly specific ADR associations, like drug withdrawal reactions, which are rarely reported.

4 | DISCUSSION

4.1 | Validity and reliability testing

In this study, a new causality approach, combining an algorithm with expert judgement, was compared with the Naranjo algorithm (drug reports) and WHO AEFI criteria (vaccine reports). The new approach performed better than the Naranjo algorithm and showed a higher sensitivity and specificity. Although sensitivity and specificity are not remarkably higher when comparing with the AEFI criteria, CausDoc gives more insight in argumentation.

In a previous study, Theophile et al¹⁰ compared the Naranjo algorithm with expert judgement. For analysis purposes, dichotomization of the causality outcomes (varying from unlikely to certain) was made, with the outcome possible in either favour or in disfavour of drug causation. Sensitivity was respectively 100% or 42%, and specificity was respectively 11% and 89%. Because Theophile had 2 causality outcomes, instead of the 5 causality outcomes of CausDoc, these results are not comparable to our study.

When comparing the causality outcomes of CausDoc with the judgement of the expert panel, more pronounced causality outcomes like probable were obtained, where Naranjo tends to more possible causality outcome compared with the judgement of the expert panel.

TABLE 2 Causality outcomes of causality methods and compared with expert judgement

Drugs (n = 20)					
Naranjo Algorithm					
Expert judgement		Unlikely	Possible	Probable	Certain
	Unlikely		2		
	Possible		5	2	
	Probable		5	4	
	Certain		2		
CausDoc					
Expert judgement		Unlikely	Possible	Probable	Certain
	Unlikely	1		1	
	Possible		5	2	
	Probable		3	6	
	Certain			2	
Vaccines (n = 20)					
AEFI WHO					
Expert judgement		Unlikely	Possible	Probable	Certain
	Unlikely	5	1		
	Possible	4	3		
	Probable		2	1	
	Certain				4
CausDoc					
Expert judgement		Unlikely	Possible	Probable	Certain
	Unlikely	3	3		
	Possible	1	5	1	
	Probable		1	2	
	Certain				4

The better sensitivity and specificity of this new causality approach are considered high enough to be useful in assessing the causality of adverse drug reactions in both drugs and in vaccines.

4.2 | Strengths and limitations of this study

The major strength of this new causality tool is that we tested a causality approach that can be used for reports on both drugs and vaccines providing a close reflection of the decision making process of health care professionals. In contrast to many other causality methods, only information that is relevant for the causality assessment of a specific association in a specific report is taken into account, leading to a more precise causality outcome. Information, not reflected in the 9 questions of CausDoc, can nevertheless be relevant for the assessment of the causal relationship. This information can be taken into account in the adjustment of the temporary causality outcome, leading to the final causality outcome. A final advantage is the fact that this approach provides insight in the scores for each question, and thus increases the transparency of the causality assessment process. This is a clear advantage above expert judgement.

Testing with 40 ADR reports with 2 causality methods is highly time and recourse intensive. Nevertheless, it gives an impression of the performance of the new causality method. In similar studies, a comparable number of cases were used (eg, 30 cases in the French method⁷, 40 cases for the Liverpool causality assessment tool⁹, 59 drug-event pairs in comparing the Naranjo and Liverpool algorithms and probabilistic method¹⁰). We realize that this limited number of reports may nevertheless not be sufficient to see if this method performs well in all possible situations.

Because testing was performed by assessing a total of 40 ADR reports (20 reports on drugs and 20 reports on vaccines), these number of reports are too small for further stratification on (in) complete documented reports.

To test our new causality approach in daily circumstances, we also evaluated the causality documentation tool on 100 subsequent incoming ADR reports by 14 assessors who were hitherto unknown with this model until they were instructed how to use this approach. The real-life data were representative for the majority of reports coming into our pharmacovigilance centre, and the results were comparable with our previous findings. Adapting the calculated causality outcome was necessary in 5 out of hundred reports, of which 4 reports concerned well-documented reports. This supports our assumption that a better impression of causality can be made when there is more clinical information available.

Our study also has a few limitations. The test cases were manually, and not randomly, selected by the expert panel out of our database to test all various causality outcomes. However, sensitivity and specificity are characteristics of the test. The distribution of cases with varying degrees of causal relationship in the sample does not affect the calculation of sensitivity and specificity.¹³ Also, causality was assessed by 2 experienced assessors and 2 experts for the expert judgement with the Delphi method,¹⁴ by which we tried to minimize subjectivity. Because these models were not blinded, this may have resulted in information bias. However, because of the large differences between the models, blinding was not possible.

Another limitation is that we intended only to rate the first drug-event combination in every report to avoid multiple causality rating of several drug-event combinations. In at least 1 test case, showing ear pain because of drug substitution, this leads to different causality outcomes because the assessor only rated the ear pain (outcome probable) whereas the expert panel rated the ear pain because of drug substitution (outcome unlikely). Apparently, in daily practice, the whole spectrum of reported ADRs is assessed. CausDoc cannot overcome this completely, but if there is a solid argument for disagreement of the calculated causality outcome, the temporary causality outcome can be adjusted with 1 category, leading to the final causality outcome.

Although there is no objective gold standard, we used expert judgement as a gold standard instead. However, expert judgement is also prone to subjectivity. Theophile showed that in only 36% of the cases, agreement was found between the experts.¹⁰ When performing a causality assessment, subjectivity of judgements because of clinical interpretation of the often limited available information cannot be excluded, especially regarding the interpretation of possible other causes.⁸ It requires knowledge by the assessor of clinical aspects and context of health care and pharmacology.² This is probably the reason why our reliability testing with a weighted Cohen kappa of "moderate" and "substantial" is not quite high.

The expert judgement of the assessor in causality assessment is more important than a calculated percentage, or textual causality outcome. However, with this new causality approach, an algorithm and expert judgement is combined in assessing the causality of only relevant and available information of an ADR report. This will lead to more insight in the reasons behind allotted causality outcomes in a pharmacovigilance centre. Therefore, this model seems to fit best in causality assessment in daily practice. However, further studies are necessary to see if this new approach continues to work well with more diverse drug-event associations.

5 | CONCLUSION

A new causality approach takes only items which are relevant and for which information is available into account in determining causality. This approach shows a better performance and allows for a better documentation of ADRs in clinical practice compared with the Naranjo algorithm and WHO AEFI criteria.

ETHICS STATEMENT

The authors state that no ethical approval was needed.

ACKNOWLEDGEMENTS

The authors would like to thank the assessors who helped us testing the causality documentation tool.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

FUNDING

None.

ORCID

Ingrid Oosterhuis <http://orcid.org/0000-0002-8403-2418>

REFERENCES

1. Meyboom RH, Hekster YA, Egberts AC, Gribnau FW, Edwards IR. Causal or casual? The role of causality assessment in pharmacovigilance. *Drug Saf*. 1997;17(6):374-389.
2. Ralph E. I. Causality assessment in pharmacovigilance: still a challenge. *Drug Saf*. 2017;40(5):365-372.
3. Lanctot KL, Naranjo CA. Comparison of the Bayesian approach and a simple algorithm for assessment of adverse drug events. *Clin Pharmacol Ther*. 1995;58(6):692-698.
4. Hill AB. The environment and disease: association of causation? *Proc R Soc Med*. 1965;58:295-300.
5. WHO. Causality assessment of an adverse event following immunization (AEFI). (version date: 2013, access date: 10-11-2017) http://www.who.int/vaccine_safety/publications/gvs_aefi/en/
6. Naranjo CA, Busto U, Sellers EM, et al. A method for estimating the probability of adverse drug reactions. *Clin Pharmacol Ther*. 1981;30(2):239-245.
7. Arimone Y, Begaud B, Miremont-Salame G, et al. A new method for assessing drug causation provided agreement with experts' judgment. *J Clin Epidemiol*. 2006;59(3):308-314.
8. Agbabiaka TB, Savovic J, Ernst E. Methods for causality assessment of adverse drug reactions: a systematic review. *Drug Saf*. 2008;31(1):21-37.
9. Gallagher RM, Kirkham JJ, Mason JR, et al. Development and inter-rater reliability of the Liverpool adverse drug reaction causality assessment tool. *PLoS One*. 2011;6(12):e28096.
10. Theophile H, Andre M, Miremont-Salame G, Arimone Y, Begaud B. Comparison of three methods (an updated logistic probabilistic method, the Naranjo and Liverpool algorithms) for the evaluation of routine pharmacovigilance case reports using consensual expert judgement as reference. *Drug Saf*. 2013;36(10):1033-1044.
11. Hauben M, Aronson JK. Defining 'signal' and its subtypes in pharmacovigilance based on a systematic review of previous definitions. *Drug Saf*. 2009;32(2):99-110.
12. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. 2005;85(3):257-268.
13. Šimundić AM. Measures of diagnostic accuracy: basic definitions. *EJIFCC*. 2009;19(4):203-211.
14. Rowe G, Wright G. The Delphi technique as a forecasting tool: issues and analysis. *Int J Forecast*. 1999;15(4):353-375.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Oosterhuis I, Zweers P, Rümke H, Muller-Hansma A, van Puijenbroek EP. A tailor-made approach for causality assessment for ADR reports on drugs and vaccines. *Pharmacoepidemiol Drug Saf*. 2019;28:544-550. <https://doi.org/10.1002/pds.4637>