



# Causality Assessment Between Drugs and Fatal Cerebral Haemorrhage Using Electronic Medical Records: Comparative Evaluation of Disease-Specific and Conventional Methods

Miki Ohta<sup>1</sup> · Satoru Miyawaki<sup>2</sup> · Shinichiroh Yokota<sup>3</sup> · Makoto Yoshimoto<sup>1</sup> · Tatsuya Maruyama<sup>1</sup> · Daisuke Koide<sup>1</sup> · Takashi Moritoyo<sup>1</sup> · Nobuhito Saito<sup>2</sup>

Accepted: 13 December 2023  
© The Author(s) 2024

## Abstract

**Introduction** A new algorithm for causality assessment of drugs and fatal cerebral haemorrhage (ACAD-FCH) was published in 2021. However, its use in clinical practice has not been verified.

**Objectives** This study aimed to explore the practical value of the ACAD-FCH when applying information available in clinical practice.

**Methods** The medical records of patients who died at the University of Tokyo Hospital in 2020 were reviewed, and cases with intracranial haemorrhage were selected. Two evaluators independently assessed these cases using three methods (the ACAD-FCH, Naranjo algorithm, and WHO-UMC scale). The number of ‘Yes’, ‘No’, and ‘No information/Do not know’ responses to each question by both evaluators were summed and compared. Inter-rater reliability was evaluated for each method using agreement rates and kappa coefficients with 95% confidence intervals (CI).

**Results** Among 316 deaths, 24 cases with intracranial haemorrhage were evaluated. The proportion of ‘No information/Do not know’ responses for each question was 35.6% (95% CI 31.4–40.6%) for the ACAD-FCH and 66.9% (95% CI 62.5–71.1%) for the Naranjo algorithm. The respective agreement rates and kappa coefficients were 0.917 (0.798–1.00) and 0.867 (0.675–1.00) for the ACAD-FCH, 0.708 (0.512–0.904) and 0.139 (–0.236 to 0.513) for the Naranjo algorithm, and 0.50 (0.284–0.716) and 0.326 (0.110–0.541) for the WHO-UMC scale, respectively.

**Conclusion** Our findings suggest the utility of the ACAD-FCH when assessing death cases with intracranial haemorrhage. However, larger studies including intra-rater assessments are warranted for further validation of this algorithm.

## Key Points

For effective use of adverse drug reaction reports in pharmacovigilance, reports should contain sufficient information and apply feasible causality assessment tools.

The results of this study suggest that the new algorithm for specifically assessing fatal cerebral haemorrhage is more useful for assessing medicine causality using the information available in clinical practice (medical records) compared with existing methods.

✉ Miki Ohta  
mjohta@g.ecc.u-tokyo.ac.jp

<sup>1</sup> Clinical Research Promotion Centre, The University of Tokyo Hospital, Tokyo, Japan

<sup>2</sup> Department of Neurosurgery, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

<sup>3</sup> Department of Healthcare Information Management, The University of Tokyo Hospital, Tokyo, Japan

## 1 Introduction

In pharmacovigilance, adverse drug reaction reports are an essential source of information. Causality assessment of individual cases forms the basis of safety measures for medicinal products [1, 2]. Since clinical trials involve only a limited number and characteristics of patients, information regarding suspected adverse drug reactions obtained from post-marketing clinical practice is crucial to ensure medicinal product safety [2–4]. Specifically, in adverse drug reaction reports with death as the outcome, the extent to which the drug affects the course of events leading to death should be carefully assessed when considering the need for prompt safety measures.

Reports of suspected drug reactions to regulatory authorities should contain sufficient information to be used effectively [2, 5]. Additionally, proper methods for assessing the causal relationship between drugs and adverse events should be established [6, 7].

Sufficient information on suspected drug reactions is a prerequisite for assessing the causal relationships between medicines and adverse events. Annually, the Japanese regulatory authority Pharmaceuticals and Medical Devices Agency (PMDA) receives approximately 60,000 reports on adverse drug reactions from medical practices and pharmaceutical companies [8]. However, most of these reports contain insufficient information [9]. Thus, clarifying the information to be included in a suspected adverse drug reaction report is necessary to enable causal relationship assessment and establish practical tools for this assessment [10, 11]. Such a method would allow for consensus among regulatory authorities, pharmaceutical companies, and medical institutions and thus facilitate consistent and efficient causality assessment.

Various methods have been developed for causality assessment [7], including the widely used Naranjo algorithm [12] and the World Health Organization-Uppsala Monitoring Centre (WHO-UMC) scale [13]. However, both methods lack specificity with respect to the content of the items since they are not event specific; instead, they are designed to generally deal with adverse effects. Furthermore, when analysing the relationship between medicinal products and adverse reactions, these methods include items that are challenging to observe in fatal cases, including adverse reactions disappearing after withdrawal (dechallenge) or reappearing after re-administration (rechallenge). Therefore, event-specific methods that help evaluate fatal cases are warranted.

Cerebral haemorrhage is an important and potentially life-threatening adverse reaction that is most frequently reported to the PMDA [9]. Cerebral haemorrhage is one of the most challenging events to assess for a causal

relationship with a medicinal product owing to the diverse possible causes of its onset or exacerbation. In addition, 84% of fatal cerebral haemorrhage cases reported as suspected adverse drug reactions are published as ‘not assessable’ owing to a lack of information [9]. To overcome this, the first author previously proposed a novel algorithm (referred to here as ‘the ACAD-FCH: the causality assessment algorithm for fatal cerebral haemorrhage’ for convenience) that uses required information to assess the causal relationship between medicines and fatal cerebral haemorrhage [9].

The ACAD-FCH consists of two steps. The first step is ‘appropriate use of the drug’, ‘temporal association’, and ‘severity of brain haemorrhage (fatal or not)’; the second step is ‘assessment of the effect of the drug’ and ‘assessment of other factors’. The first step identifies the preconditions for assessing the causal relationship between the medicinal product and fatal brain haemorrhage, whereas the second step evaluates the causality balance between the suspected drug and other factors in fatal cerebral haemorrhage.

The use of the ACAD-FCH in clinical practice has not been verified. Thus, this study aims to explore the feasibility of implementing the ACAD-FCH in drug safety evaluation. Additionally, the assessment results from this algorithm were compared with those from the Naranjo algorithm and the WHO-UMC scale.

## 2 Methods

This exploratory study evaluated cerebral haemorrhage cases using information obtained from hospital electronic medical records to verify the feasibility of implementing the ACAD-FCH in drug safety evaluation. This was to determine whether the information required in the causality assessment using this algorithm was available in clinical practice. We also compared the assessment results from this algorithm with those from the Naranjo algorithm and the WHO-UMC scale.

### 2.1 Data Resource

The medical records of patients who died at the University of Tokyo Hospital between 1 January and 31 December 2020 and for whom a death certificate was created in the documentation system or death was registered as an outcome in the electronic medical records were reviewed. Cases with a head imaging scan [computed tomography (CT) or magnetic resonance imaging (MRI)], performed during the course leading to death and a documented intracranial haemorrhagic lesion, including a subarachnoid haemorrhage, subdural haematoma, and cerebral haemorrhage, were selected. Cases with unavailable medical records, without

administered medication, or with unknown administration status before the haemorrhagic lesion were excluded.

## 2.2 Causality Assessment

Two board-certified neurosurgeons (MO and SM) independently reviewed the electronic medical records of all the included cases. They examined the association between the medicines used and the incidence of intracranial haemorrhage and death. Three tools were used to assess causality: the ACAD-FCH (Table 1), Naranjo algorithm (Table 2), and WHO-UMC scale (Table 3). The ACAD-FCH was used to assess the causal relationship between death due to intracranial haemorrhage and drug use. Contrastingly, the Naranjo algorithm and WHO-UMC scale were used to assess the causal relationship between drugs and intracranial haemorrhage since these tools lack items related to death. More information on these methods can be found in previous studies [9, 12, 13].

Regarding the ACAD-FCH and Naranjo algorithms, the numbers of 'Yes', 'No', and 'No information/Do not know' responses to each question in each algorithm by both evaluators were summed. The frequency of the 'No information/Do not know' response was compared using Fisher's exact test. Statistical significance was set at  $P < 0.001$ .

## 2.3 Inter-Rater Reliability

Inter-rater reliability for each method was evaluated using percentages of agreement and kappa statistics with 95% confidence intervals (CIs). The kappa result was interpreted as follows: values  $< 0$ , none; 0.00–0.20, slight agreement; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; and 0.81–1.00, almost perfect agreement [14]. Computations were performed using STATA software (version 15.1 for Windows, STATA Corporation, College Station, USA).

## 3 Results

A total of 316 patients died at the University of Tokyo Hospital during the study period. Among them, 190 patients underwent head CT or MRI for diverse reasons, and 33 had intracranial haemorrhage. Of these, 24 patients who had received any medication before the haemorrhage were assessed. The Supplementary File contains details of the characteristics of the 24 cases, and the results of the assessments by the two evaluators.

Figure 1 illustrates the proportion of 'Yes', 'No', and 'No information/Do not know' responses selected for the nine and ten items of the ACAD-FCH and Naranjo algorithm, respectively, in the 24 case assessments recorded by both evaluators (i.e., 48 times for each question). The proportion

of 'No information/Do not know' responses for each question, indicating insufficient information for a decision, was 35.6% (95% CI 31.4–40.6) for the overall ACAD-FCH and 66.9% (95% CI 62.5–71.1) for the overall Naranjo algorithm, with a significant difference ( $P < 0.0001$ ). In the ACAD-FCH, the question with the least number of 'No information' selections was 3(i): 'It was a serious cerebral haemorrhage' (one response). Contrastingly, the question with the most frequent 'No information' selections was 2(ii): 'Blood levels of the suspected drug were sufficiently high at the onset of cerebral haemorrhage' (39 responses). In the Naranjo algorithm, questions 3, 4, 6, 8, and 9 were challenging or impossible to answer in assessing death cases, with the 'Do not know' response being selected in most cases.

Table 4 presents the distribution of results (A–C) and inter-rater agreement (D) of the 24 cases assessed by both evaluators using the three methods. There were notable differences in the inter-rater agreement using the three evaluation methods. The ACAD-FCH had a high Cohen's kappa coefficient of 0.867, indicating substantial agreement beyond chance. Additionally, the agreement rate was 91.7%, with a narrow CI (95% CI 79.8–100). These findings suggest a strong consensus between both evaluators when using the ACAD-FCH.

In contrast, the Naranjo algorithm yielded a considerably lower Cohen's kappa coefficient of 0.139, implying a slight agreement beyond chance, with a relatively high agreement rate (70.8%). This discrepancy suggests that the true agreement level is limited when the Naranjo algorithm is used. The WHO-UMC scale yielded intermediate results with a Cohen's kappa coefficient of 0.326, indicating a fair agreement. The agreement rate was 50.0%, with a relatively wide CI (95% CI 28.4–71.6). These outcomes suggest substantial variability in agreement between the observers.

## 4 Discussion

Our findings suggest that the event-specific ACAD-FCH is more useful for assessing medicine causality using the information available in clinical practice (medical records) compared with existing methods. The high frequency with which the evaluator was able to select 'Yes' or 'No' rather than 'No information' implies that the information required for evaluation using the ACAD-FCH is sufficiently available from the medical records, and thus it is feasible. In addition, the extremely high inter-rater agreement indicates that the ACAD-FCH is reliable.

These results may have been obtained because the ACAD-FCH excludes items, such as rechallenge, which cannot be easily determined when assessing death cases. The Naranjo algorithm and WHO-UMC scale include conclusive categories, such as 'Definite' and 'Certain', which require

**Table 1** The causality assessment algorithm for fatal cerebral haemorrhage (ACAD-FCH) [9]

STEP I				Category	
1. Proper use?	Yes		⇒ OK		
	No		⇒	Unassessable	
	No information		⇒	Unassessable	
2. Temporal relationship?	(i) The suspected drug was administered before the occurrence of fatal cerebral haemorrhage.	(ii) Blood levels of the suspected drug were sufficiently high at the onset of cerebral haemorrhage.	⇒ OK		
			Yes		⇒ OK
			No		⇒ OK
	No	No information	Yes		⇒ Unlikely
			No		⇒ Unlikely
			No information		⇒ Unlikely
	No information	No information	Yes		⇒ OK
			No		⇒ Unassessable
			No information		⇒ Unassessable
3. Direct cause of death?	(i) It was a serious cerebral haemorrhage.	(ii) Mortality due to complications can be excluded.	⇒ OK		
			Yes		⇒ Unlikely
			No		⇒ OK
	No	No information	Yes		⇒ Unlikely
			No		⇒ Unlikely
			No information		⇒ Unlikely
	No information	No information	Yes		⇒ Unassessable
			No		⇒ Unlikely
			No information		⇒ Unassessable

If all items 1 to 3 are 'OK' → Go to STEP II.

If either item 2 or item 3 is 'Unlikely' → The assessment is completed, and the assessment result for the case is 'Unlikely'.

Otherwise → the assessment is completed, and the assessment result for the case is 'Unassessable'.

**STEP II**

4. Can the pathological condition be explained by the action of the suspected drug?	(i) The suspected drug has the potential to cause the rupture of cerebral blood vessels.  or (ii) The suspected drug has the potential to cause haematoma extension.	(iii) There are findings suggesting (i) or (ii) in the report.	Yes	2
			No	0
			No information	1
	No or Unknown	No information	Yes	1
			No	0
			No information	C
5. Were alternative aetiologies other than the suspected drug, such as patient's background taken into consideration?	(i) There was no or well-controlled medical history of increasing the risk of developing fatal cerebral haemorrhage.	(ii) No abnormalities were found in imaging and autopsy findings that could cause cerebral haemorrhage.	Yes	3
			No	1
			No information	2
	No	No information	Yes	0
			No	0
			No information	0
No information	No information	Yes	1	
		No	0	
		No information	C	

**Decision of Categories**

item 4	item 5	
2	3	Probably
	2	Probably
	1	Probably
	0	Possibly
	C	Possibly
1	3	Probably
	2	Probably
	1	Possibly
	0	Unlikely
	C	Unassessable
0	3	Possibly
	2	Possibly
	1	Unlikely
	0	Unlikely
	C	Unlikely
C	3	Possibly
	2	Possibly
	1	Unassessable
	0	Unlikely
	C	Unassessable

extensive solid evidence. However, removing the items with inaccessible data in death cases in both methods would not yield a conclusive result. The score range for the Naranjo algorithm is  $-4$  to  $13$ . However, after excluding items that are challenging to assess in critical cases, including death (question 3, 4, 6, 8, and 9), the score range becomes  $-2$  to  $7$  points, and the assessment result cannot be ‘Definite’ (need  $\geq 8$  points). In the WHO-UMC scale, in cases in which it is challenging to determine the ‘response to withdrawal’, few cases are ‘Certain’ or ‘Probable’. Proving that the drug is 100% responsible is challenging, and this is not limited to fatal cases [6, 15]. Causality assessment can only be performed by determining whether the drug or other factors are more influential (causality balance) using the available information. The ACAD-FCH, which is structured mainly using items for which data can be easily obtained in clinical practice, such as temporal relationships, imaging findings, and patient background, allows for efficient assessment of the causality balance.

Another reason is that the ACAD-FCH has a two-step structure. Step 1 confirms the five items that form the premise for causal assessment. Step 2 assesses the balance of causal relationships with four items, which facilitates the correction of the assessment trajectory compared with those of the Naranjo algorithm or WHO-UMC scale, in which these items are assessed together. For example, in the ACAD-FCH, events without a temporal causal relationship are determined as ‘Unlikely’ in Step 1 and are not affected by the results of Step 2. In contrast, in the Naranjo algorithm, the score for no temporal relationship ( $-1$  point) is cancelled out by the scores for the other items. Accordingly, the assessment result can be ‘Possible’ or indicate a stronger causal relationship [16]. In the WHO-UMC scale, the effect of the medication, other factors, and the temporal association

are assessed on the basis of the evaluator’s impression. The assessment result is based on the one that fits all the criteria, leading to high variability between evaluators [17]. In the ACAD-FCH, nine items are rated in two steps; therefore, we could obtain inter-rater agreement beyond chance.

Previous studies using the Naranjo algorithm and WHO-UMC scale revealed agreement rates of 43–93% [18–20] and 24–73% [17, 20, 21], respectively, with kappa coefficients of 0.14–0.86 [18–20, 22] and 0.11–0.71 [17, 20, 21], respectively. The results vary on the basis of study conditions, including the subject of the evaluation. These values are consistent with our findings. Furthermore, the Naranjo algorithm is likely to have a ‘Probable’ or ‘Possible’ result [19]; and the WHO-UMC scale is not reproducible since it is based on individual judgement [17, 23]. Similar trends were observed in this study.

No single method is universally accepted [7]; nonetheless, the previously developed assessment methods have their advantages [20]. Moreover, using and comparing several assessment methods simultaneously or using different methods, depending on the situation, is beneficial. For example, a generic method is required when evaluating cases without specifying the event in which a specific drug was used. Contrastingly, when focussing on a specific event, an assessment method involving items appropriate to that event is more efficient. The ACAD-FCH is the only causality assessment method specific to drugs and fatal cerebral haemorrhage. Consequently, it could be useful when the analysis focusses on the occurrence of severe intracranial haemorrhage without identifying the suspected drug. This method is expected to contribute to compensating for the shortcomings of existing methods when assessing fatal cases with cerebral haemorrhage. Furthermore, while the ACAD-FCH is currently specific to cases of fatal cerebral

**Table 2** Naranjo algorithm [12]

Question	Yes	No	Do not know
1. Are there previous conclusive reports on this reaction?	+1	0	0
2. Did the adverse event appear after the suspected drug was administered?	+2	-1	0
3. Did the adverse event improve when the drug was discontinued or a specific antagonist was administered?	+1	0	0
4. Did the adverse event reappear when the drug was readministered?	+2	-1	0
5. Are there alternative causes that could on their own have caused the reaction?	-1	+2	0
6. Did the reaction reappear when a placebo was given?	-1	+1	0
7. Was the drug detected in blood or other fluids in concentrations known to be toxic?	+1	0	0
8. Was the reaction more severe when the dose was increased or less severe when the dose was decreased?	+1	0	0
9. Did the patient have a similar reaction to the same or similar drugs in any previous exposure?	+1	0	0
10. Was the adverse event confirmed by any objective evidence?	+1	0	0

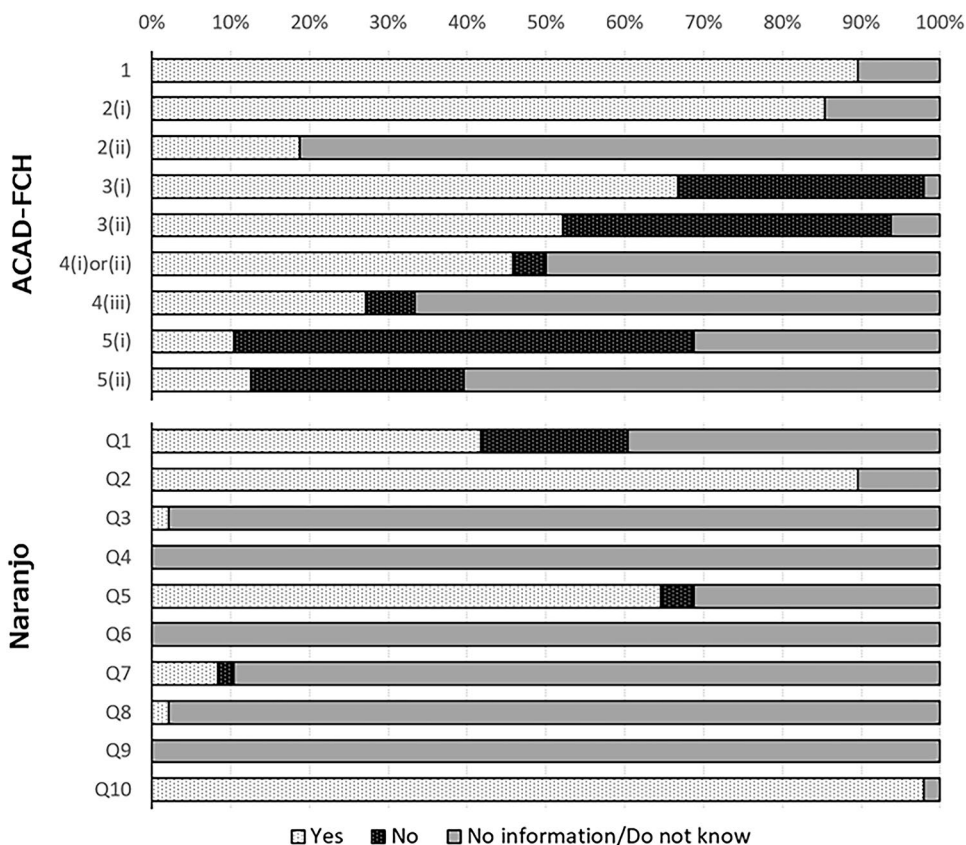
> 8: Definite, 5–8: Probable, 1–4: Possible, < 1: Doubtful

**Table 3** WHO-UMC scale [13]

Causality term	Assessment criteria (all points should be reasonably complied)
Certain	Event or laboratory test abnormality, with plausible time relationship to drug intake Cannot be explained by disease or other drugs Response to withdrawal plausible (pharmacologically, pathologically) Event definitive pharmacologically or phenomenologically (i.e. an objective and specific medical disorder or a recognized pharmacologic phenomenon) Rechallenge satisfactory, if necessary
Probable/likely	Event or laboratory test abnormality, with reasonable time relationship to drug intake Unlikely to be attributed to disease or other drugs Response to withdrawal clinically reasonable Rechallenge not required
Possible	Event or laboratory test abnormality, with reasonable time relationship to drug intake Could also be explained by disease or other drugs Information on drug withdrawal may be lacking or unclear
Unlikely	Event or laboratory test abnormality, with a time to drug intake that makes a relationship improbable (but not impossible) Disease or other drugs provide plausible explanation
Conditional/unclassified	Event or laboratory test abnormality More data for proper assessment needed, or Additional data under examination
Unassessable/unclassifiable	Report suggesting an adverse reaction Cannot be judged because information is insufficient or contradictory Data cannot be supplemented or verified

WHO-UMC World Health Organization-Uppsala Monitoring Centre

**Fig. 1** The proportion of ‘Yes’, ‘No’, and ‘No information/Do not know’ responses selected by both evaluators for each question of the ACAD-FCH and Naranjo algorithm



**Table 4** Causality assessment by two raters

A: ACAD-FCH						
Rater #1	Rater #2					
	Probably	Possibly	Unlikely	Unassessable		
Probably	0	0	0	0	0	
Possibly	0	5	1	0	0	
Unlikely	0	0	11	1	0	
Unassessable	0	0	0	6	0	

B: Naranjo Algorithm					
Rater #1	Rater #2				
	Definite	Probable	Possible	Doubtful	
Definite	0	0	0	0	0
Probable	0	0	3	0	0
Possible	0	1	16	0	0
Doubtful	0	0	3	1	0

C: WHO-UMC scale						
Rater #1	Rater #2					
	Certain	Probable	Possible	Unlikely	Conditional	Unassessable
Certain	0	0	0	0	0	0
Probable	0	0	1	0	0	0
Possible	0	0	5	5	2	0
Unlikely	0	0	0	6	0	0
Conditional	0	0	0	0	0	0
Unassessable	0	0	0	2	2	1

D: Inter-rater agreement			
	ACAD-FCH	Naranjo	WHO-UMC
Kappa (95% CI)	0.867 (0.675–1.00)	0.139 (–0.236 to 0.513)	0.326 (0.110–0.541)
Agreement rate (95% CI)	0.917 (0.798–1.00)	0.708 (0.512–0.904)	0.50 (0.284–0.716)

ACAD-FCH algorithm for causality assessment of drugs and fatal cerebral haemorrhage; WHO-UMC World Health Organization-Uppsala Monitoring Centre

haemorrhage, its applicability could extend to other side effects. This adaptation could be achieved by adjusting the details of the assessment items, or in cases where the outcome is other than death, by removing the item 3 (assessing whether the event was fatal). In cases where the involvement of multiple drugs is suspected, including instances where drug interactions may be the cause, the causality assessment can be conducted by replacing item 4 with the following: 'Can the pathological condition be explained by the drug interactions?'

While the ACAD-FCH proves useful for extracting and organising the necessary information for causality assessment from large amounts of clinical information, its usefulness may diminish in instances where there is insufficient information, even in clinical practice. For instance, scenarios lacking a history of hospital visits and providing minimal

background information could limit the algorithm's utility. Additionally, if, for any reason, a head imaging scan cannot be performed, the usefulness of this algorithm might not be demonstrated, even in cases where cerebral haemorrhage is strongly suspected on the basis of clinical symptoms. From our perspective, information derived from head imaging studies plays a crucial role in assessing causality concerning drugs.

The ACAD-FCH was developed with the purpose of assessing individual cases, aiming to improve the quality of individual case reports and streamline the assessment process. However, to determine the true causal relationship between drugs and adverse events, rigorous large-scale controlled trials, controlling for confounding factors, become imperative. Numerous confounding elements, such as age and underlying medical conditions, can be associated with

head imaging tests and drug use. This complexity makes it challenging to assess the causal relationship between drugs and cerebral haemorrhage. The ACAD-FCH has the potential to facilitate early signal detection, potentially prompting controlled trials and creating a virtuous circle wherein trial outcomes contribute to refining the assessment of the causality balance.

This study has several limitations. First, both evaluators are neurosurgeons; therefore, their judgement may be similar. Therefore, future studies should consider assessments by doctors from other medical specialities or pharmacists. Second, since this study was based on data from a single centre, it is difficult to extrapolate the results, and further multi-centre studies are warranted. In addition, this study did not analyse intra-rater agreement, which should be performed after a certain period in the future. Furthermore, given the insufficient number of cases for comparing drug- and non-drug-related cerebral haemorrhages and the potential bias introduced by the COVID-19 pandemic in 2020, an assessment of cases from the pre-pandemic period becomes imperative—a challenge to be addressed in future. Overcoming these challenges and obtaining additional data from diverse real-world sources will enable the validation of the usefulness of ACAD-FCH by analysing its sensitivity, specificity and positive predictive value in comparison with existing databases. Once the algorithm's robustness is confirmed, there may be an opportunity to develop it into a compelling study comparing artificial intelligence (AI) and human assessments.

## 5 Conclusion

This study explored the usefulness of the ACAD-FCH, a method developed to assess the causal relationship between drugs and fatal cerebral haemorrhage, using medical information obtained from cases of intracranial haemorrhage where death occurred. Compared with the Naranjo algorithm and WHO-UMC scale, which are widely used, the results suggest that the disease-specific ACAD-FCH may be more effective in assessing intracranial haemorrhage cases with death. However, larger studies, including intra-rater assessments, are warranted to validate this algorithm.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s40801-023-00413-y>.

**Acknowledgements** The authors acknowledge Dr Ruri Kaneda (Clinical Research Promotion Centre, The University of Tokyo Hospital) for her critical review of the manuscript.

## Declarations

**Conflict of Interest** MO, SM, SY, MY, TMa, DK, TMo, and NS declare that they have no relevant financial or non-financial interests to disclose that might be relevant to the contents of this manuscript.

**Funding** This study was supported by MEXT KAKENHI (grant no. JP22K15335) and MHLW Health and Labour Sciences Research Grants (grant no. JPMH23KC1005).

**Ethical Approval** This study was approved by the Ethical Review Committee of the University of Tokyo Hospital (approval date: 2 May 2022; approval number: 2021294NI).

**Consent to Participate** The Ethical Review Committee determined that individual-level consent was not required, as the subjects of this study have already died. The research methodology was published on the website of the Clinical Research Promotion Centre of the University of Tokyo Hospital, and the patient's family could opt out at any time.

**Consent for Publication** Not applicable.

**Data Availability** Supplementary Tables are available online with this publication.

**Author Contributions** MO developed the concept for this study, coordinated the data collection, performed the causality assessment of cases and the data analysis, and drafted the manuscript. SM conducted the causality assessment of cases and the data analysis. SY determined a method of extracting cases for analysis from the electronic medical record and performed the extraction. MY and TMo contributed to the interpretation of the analysis results and development of the discussion. DK contributed to the study design, determination of the analysis method, and interpretation of the analysis results. TMa and NS contributed to the study conception and design. All authors contributed to the article and approved the final manuscript.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

## References

1. Bergvall T, Norén GN, Lindquist M. *vigiGrade*: a tool to identify well-documented individual case reports and highlight systematic data quality issues. *Drug Saf.* 2014;37:65–77. <https://doi.org/10.1007/s40264-013-0131-x>.
2. Brajovic S, Piazza-Hepp T, Swartz L, Dal Pan G. Quality assessment of spontaneous triggered adverse event reports received by the Food and Drug Administration. *Pharmacoepidemiol Drug Saf.* 2012;21:565–70. <https://doi.org/10.1002/pds.3223>.
3. Lindquist M. Data quality management in pharmacovigilance. *Drug Saf.* 2004;27:857–70. <https://doi.org/10.2165/00002018-200427120-00003>.



4. Marwitz K, Jones SC, Kortepeter CM, Dal Pan GJ, Muñoz MA. An evaluation of postmarketing reports with an outcome of death in the US FDA adverse event reporting system. *Drug Saf.* 2020;43:457–65. <https://doi.org/10.1007/s40264-020-00908-5>.
5. Plessis L, Gómez A, García N, Cereza G, Figueras A. Lack of essential information in spontaneous reports of adverse drug reactions in Catalonia—a restraint to the potentiality for signal detection. *Eur J Clin Pharmacol.* 2017;73:751–8. <https://doi.org/10.1007/s00228-017-2223-5>.
6. Hutchinson TA, Lane DA. Assessing methods for causality assessment of suspected adverse drug reactions. *J Clin Epidemiol.* 1989;42:5–16. [https://doi.org/10.1016/0895-4356\(89\)90020-6](https://doi.org/10.1016/0895-4356(89)90020-6).
7. Agbabiaka TB, Savović J, Ernst E. Methods for causality assessment of adverse drug reactions: a systematic review. *Drug Saf.* 2008;31:21–37. <https://doi.org/10.2165/00002018-200831010-00003>.
8. Japanese Adverse Drug Event Report database (JADER). <https://www.pmda.go.jp/safety/info-services/drugs/adr-info/suspected-adr/0004.html>. Accessed 21 Sept 2023.
9. Ohta M. Causality assessment between reported fatal cerebral haemorrhage and suspected drugs: developing a new algorithm based on the analysis of the Japanese Adverse Event Report (JADER) database and literature review. *Eur J Clin Pharmacol.* 2021;77:1443–52. <https://doi.org/10.1007/s00228-021-03131-y>.
10. Bandekar M, Anwikar S, Kshirsagar N. Quality check of spontaneous adverse drug reaction reporting forms of different countries. *Pharmacoepidemiol Drug Saf.* 2010;19:1181–5. <https://doi.org/10.1002/pds.2004>.
11. Oosterhuis I, Rolfes L, Ekhardt C, Muller-Hansma A, Härmark L. First experiences with a tool to measure the level of clinical information present in adverse drug reaction reports. *Expert Opin Drug Saf.* 2018;17:111–5. <https://doi.org/10.1080/14740338.2018.1400008>.
12. Naranjo CA, Busto U, Sellers EM, et al. A method for estimating the probability of adverse drug reactions. *Clin Pharmacol Ther.* 1981;30:239–45. <https://doi.org/10.1038/clpt.1981.154>.
13. Centre UM. The use of the WHO-UMC system for standardised case causality assessment. Uppsala: Uppsala Monitoring Centre; 2018.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–74. <https://doi.org/10.2307/2529310>.
15. Meyboom RH, Hekster YA, Egberts AC, Gribnau FW, Edwards IR. Causal or casual? The role of causality assessment in pharmacovigilance. *Drug Saf.* 1997;17:374–89. <https://doi.org/10.2165/00002018-199717060-00004>.
16. Théophile H, André M, Miremont-Salamé G, Arimone Y, Bégaud 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;36:1033–44.
17. de Menezes RR, Silva MDG, Ribeiro ALP, et al. Causality assessment of adverse drug reactions by applying a global introspection method in a high complexity hospital. *Explor Res Clin Soc Pharm.* 2021;3:100064. <https://doi.org/10.1016/j.rcsop.2021.100064>.
18. Kane-Gill SL, Kirisci L, Pathak DS. Are the Naranjo criteria reliable and valid for determination of adverse drug reactions in the intensive care unit? *Ann Pharmacother.* 2005;39:1823–7. <https://doi.org/10.1345/aph.1G177>.
19. 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: e28096. <https://doi.org/10.1371/journal.pone.0028096>.
20. Thaker SJ, Sinha RS, Gogtay NJ, Thatte UM. Evaluation of inter-rater agreement between three causality assessment methods used in pharmacovigilance. *J Pharmacol Pharmacother.* 2016;7:31–3. <https://doi.org/10.4103/0976-500x.179361>.
21. Mouton JP, Mehta U, Rossiter DP, Maartens G, Cohen K. Inter-rater agreement of two adverse drug reaction causality assessment methods: a randomised comparison of the Liverpool Adverse Drug Reaction Causality Assessment Tool and the World Health Organization-Uppsala Monitoring Centre system. *PLoS ONE.* 2017;12: e0172830. <https://doi.org/10.1371/journal.pone.0172830>.
22. Leopoldino RWD, de Oliveira LVS, Fernandes FEM, et al. Causality assessment of adverse drug reactions in neonates: A comparative study between Naranjo's algorithm and Du's tool. *Int J Clin Pharm.* 2023;45(4):1–7. <https://doi.org/10.1007/s11096-023-01595-9>.
23. Behera SK, Das S, Xavier AS, Velupula S, Sandhiya S. Comparison of different methods for causality assessment of adverse drug reactions. *Int J Clin Pharm.* 2018;40:903–10. <https://doi.org/10.1007/s11096-018-0694-9>.