

# Validation of a novel causality assessment scale for adverse events in non-small cell lung carcinoma patients treated with platinum and pemetrexed doublet chemotherapy

Jyoti Bhagatram Sharma, Manjunath Nookala Krishnamurthy<sup>ID</sup>, Ankita Awase, Amit Joshi, Vijay Patil, Vanita Noronha, Kumar Prabhaskar and Vikram Gota

## Abstract

**Aim:** Accurate causality assessment (CA) of adverse events (AEs) is important in clinical research and routine clinical practice. The Naranjo scale (NS) used for CA lacks specificity, leading to a high rate of false positive causal associations. NS is a simple scale for CA; however, its limitations have reduced its popularity in favour of other scales. We therefore attempted to improvise the algorithm by addressing specific lacunae in NS.

**Methods:** We attempted to modify the existing NS by (a) changing the weightage given to certain responses, (b) achieving higher resolution to certain responses for delineating drug related and unrelated AEs and (c) modifying the slabs for classification of association as 'likely' and 'unlikely'. The new scale, named as the Sharma-Nookala-Gota (SNG) algorithm, was evaluated in a training set of 19 AEs in a tertiary care cancer hospital in western India, and further validated in a set of 104 AEs. Consensus of four physician opinion was taken as gold standard for comparison.

**Results:** Of the 19 AEs in the training set, 6 were described by the treating physician as 'not related' and 13 as related to the drug. The SNG algorithm had 100% concordance with physician opinion, whereas the NS had only 73.7% concordance. NS showed a tendency to misclassify AEs as 'related' when they were indeed 'not related'. In the validation set of 104 AEs, NS and SNG algorithms misclassified 30 and 2 AEs, respectively, leading to a concordance of 70.2% and 98.1%, respectively, with physician opinion.

**Conclusion:** Decisive modifications of the NS resulted in the SNG scale, with superior specificity while retaining sensitivity against the gold standard.

**Keywords:** SNG algorithm, adverse drug reactions, adverse events, drug safety, pharmacovigilance

Received: 26 August 2020; revised manuscript accepted: 7 January 2021.

## Plain Language Summary

### SNG algorithm – A novel tool for causality assessment of adverse drug reactions

Adverse events (AEs) can cause increased morbidity, hospitalisation, and even death. Hence it is essential to recognise AEs and to establish their correct causal relationship to a drug. Many causality assessment methods, scales and algorithms are available to assess the relationship between an AE and a drug. The Naranjo algorithm is most commonly employed

*Ther Adv Drug Saf*

2021, Vol. 12: 1–12

DOI: 10.1177/  
2042098621991280

© The Author(s), 2021.  
Article reuse guidelines:  
[sagepub.com/journals-](https://sagepub.com/journals-permissions)  
permissions

Correspondence to:

**Manjunath Nookala  
Krishnamurthy**

Department of Clinical  
Pharmacology, Advanced  
Centre for Treatment,  
Research and Education  
in Cancer, Tata Memorial  
Centre, Sector 22,  
Kharghar, Maharashtra,  
Navi Mumbai 410210,  
India; Homi Bhabha  
National Institute,  
Mumbai, India

[mnookala@actrec.gov.in](mailto:mnookala@actrec.gov.in)

**Jyoti Bhagatram Sharma  
Ankita Awase**

Department of Clinical  
Pharmacology, Advanced  
Centre for Treatment  
Research and Education  
in Cancer, Tata Memorial  
Centre, Navi Mumbai, India

**Amit Joshi**

Department of Medical  
Oncology, Tata Memorial  
Hospital, Tata Memorial  
Centre, Dr. E. Borge's  
Road, Parel, Mumbai,  
India; Homi Bhabha  
National Institute,  
Mumbai, India

**Vijay Patil**

Department of Medical  
Oncology, Tata Memorial  
Hospital, Tata Memorial  
Centre, Dr. E. Borge's  
Road, Parel, Mumbai,  
India; Homi Bhabha  
National Institute,  
Mumbai, India

**Vanita Noronha**

Department of Medical  
Oncology, Tata Memorial  
Hospital, Tata Memorial  
Centre, Dr. E. Borge's  
Road, Parel, Mumbai,  
India; Homi Bhabha  
National Institute,  
Mumbai, India

**Kumar Prabhaskar**

Department of Medical  
Oncology, Tata Memorial  
Hospital, Tata Memorial  
Centre, Dr. E. Borge's  
Road, Parel, Mumbai,  
India; Homi Bhabha  
National Institute,  
Mumbai, India

**Vikram Gota**

Department of Clinical Pharmacology, Advanced Centre for Treatment Research and Education in Cancer, Tata Memorial Centre, Navi Mumbai, India  
Homi Bhabha National Institute, Mumbai, India

in spite of its many drawbacks as it is simple to use. Concerns have been raised regarding the performance of the scale, and researchers have tried to answer them, but none of them could address all issues satisfactorily. We too experienced many problems while using it in our routine clinical practice and in clinical trials. For instance, the Naranjo scale is non-specific and shows a bias toward implicating the drug as the causal factor for AEs. This improper assessment has often led to drug discontinuation, thereby compromising the efficacy of treatment. Hence, we modified the existing Naranjo scale to a new one (the Sharma-Nookala-Gota – SNG algorithm) to address these shortcomings. We piloted the SNG causality assessment algorithm in patients suffering from AEs due to various drugs. The SNG algorithm was found to have good concordance with the physicians' assessment of causality. As a next step, we validated the SNG algorithm in patients receiving a standard drug combination of pemetrexed and carboplatin for lung cancer combination. Out of the 104 AEs observed in 65 patients, the SNG causality assessment algorithm showed good concordance (except in two cases) with the physicians' decision of causality assessment, while the Naranjo algorithm was not so successful. Hence, the SNG algorithm can be a better guide for causality assessment of AEs.

### Introduction

Adverse drug reactions (ADRs) are one of the major causes of mortality and morbidity, driving healthcare costs higher every year.<sup>1</sup> Against this backdrop, accurate assessment of causality of adverse events (AEs) is extremely important from the perspective of both day-to-day practice as well as drug development. Accurate assessment of causality avoids mislabelling of events as drug related, thus avoiding unnecessary dechallenge of drugs.<sup>2</sup> Causality assessment (CA) can also influence insurance coverage and resource allocation in hospitals.<sup>3–5</sup> From a legal perspective, accurate CA may aid in settlement of litigation for medical malpractice/negligence.<sup>6</sup> Similarly, proper CA is also important in drug development for establishing the safety profile of the drug and its labelling.<sup>1</sup> CA is undertaken extensively during pharmacovigilance and post-marketing surveillance, with the goal of better establishing the safety profiles of drugs.<sup>7,8</sup> Regulatory authorities routinely evaluate ADR reports, where CA can aid risk–benefit assessment of new medicines.<sup>9</sup>

There are several methods for CA of AEs including expert judgment, Bayesian methods, algorithms and scales. However, no single tool is universally accepted or considered a gold standard.<sup>10,11</sup> Among the CA scales, the Naranjo Scale (NS) is used widely because it is simple and quick to carry out. It was initially developed to serve as an ADR monitoring tool in clinical trials and registration trials of new drugs. However, it has some limitations that have been questioned

previously by a number of investigators. First of all, NS is biased in favour of a positive causal association between the drug and an AE.<sup>12</sup> This is a consequence of inappropriate weightage given to some questions,<sup>13</sup> and, in our opinion, setting a rather low threshold for calling an event as causally related to the drug. Second, the entire scale of 10 questions was constructed based on experience from only 63 cases, which has resulted in a less specific and less reliable assessment tool.<sup>14,15</sup> Therefore, not surprisingly, in phase I clinical trials conducted at our centre, these limitations of NS came to the fore, resulting in gross disagreements in CA as per NS and expert opinion.

There is an urgent need to develop accurate and simple tools for CA. Despite its limitations, major strength of the NS tool is its simplicity. Therefore, we decided to revise the NS through modifications in weightage assignment and redefining the threshold for relatedness, at the same time retaining its simplicity. The development and validation of the new scale, the Sharma-Nookala-Gota (SNG) algorithm, is reported here.

### Patients and methods

This study was conducted in a tertiary care cancer hospital in western India. The study was approved by our Institutional Ethics Committee (approval number: IEC/0416/1511/001) and was conducted in accordance with Good Clinical Practice guidelines and the Declaration of Helsinki. ADR data used in this study was collected from the respective

studies after obtaining informed consent from patients.

Many studies and researchers have suggested modification to NS.<sup>2,8,16–18</sup> Our experience with causality assessment from early phase trials using the NS prompted us to modify the scale. Five AEs observed in two clinical trials conducted at our centre are discussed (Supplemental material D). The AEs could not be assessed satisfactorily for causality using the NS. The NS was therefore modified (Table 1) as discussed below, which led to the development of the SNG algorithm.

### Steps involved in deriving the SNG scale from the NS

#### *Appropriateness of questions*

Question 3 was modified to also include situations where drugs are administered cyclically, such as cancer chemotherapy. By its very nature, cyclical administration of drugs mimics the scenario of dechallenge. Question 5 of NS was modified to reduce the probability of falsely concluding a causal association in situations where there is high level evidence available for an alternative cause (Table 1).

#### *Weightage*

The scoring of ‘Yes’ and ‘No’ responses of some of the questions were changed based on their strength of association with causality. The changes in scores and the questionnaire are shown in Table 1. Assigning appropriate weightage with adequate reasoning and justification is an important attribute of a good CA scale. Seger *et al.* challenged NS saying that a positive temporal relationship (question 2) would place an AE in the ‘possible’ category of association even before any other questions are answered.<sup>16</sup> We agree with this suggestion because several factors including disease progression, comorbidities and concomitant medications can lead to AEs that may coincide with the introduction of a drug. Therefore, temporal association cannot be solely relied upon for CA. Hence for the response ‘Yes’, the score was reduced to ‘+1’ from ‘+2’. No alterations were made to the ‘No’ and ‘Do not know’ responses.

Question 3 refers to improvement in AE after discontinuing the drug (dechallenge) or

administration of a specific antagonist. Response to dechallenge is a critical parameter capable of strongly implicating the drug with an AE or absolving it.<sup>19</sup> This is also an immediate reaction to an AE by any physician, especially if the AE is severe. As a result, the score assigned to ‘Yes’ was increased by one to ‘+2’ and that of ‘No’ response was changed to ‘–1’ from ‘0’ allowing greater resolution of causality assessment depending on the outcome of dechallenge. It is pertinent to note that if the AE does not improve as a response to dechallenge but decreases following initiation of specific AE-related treatment (symptomatic or otherwise), then the response to dechallenge should be considered negative and a score of ‘–1’ should be given.

In our SNG scale, no question was deleted from the existing NS scale. However, some authors have recommended removal of a few questions that are not applicable for a specific population of patients. For instance, Kane-Gill *et al.* raised doubts about the utility of question 4 (dechallenge) in Intensive Care Units (ICU).<sup>17</sup> The SNG algorithm is yet to be validated in the ICU setting, but the question on rechallenge is nevertheless relevant to the majority of clinical situations. However, the scoring of response to rechallenge was modified as discussed subsequently. An answer of ‘No’ applies if rechallenge was done, but the AE did not reappear. This is a strong indicator of the drug not being a causal factor for the event.<sup>20</sup> Hence, we increased the score from ‘–1’ to ‘–2’. This pattern of scoring will also offset any false positive responses there may be for other questions of the NS wrongly associating the drug with the event. It is important to allow a higher negative weightage for answers that are strongly predictive of the drug not being associated with the event in order to have a more balanced assessment of causality and minimise false positive outcomes. Although, for the response ‘Yes’ to rechallenge indicating the AE resolved or improved when the medication was stopped and subsequently reappeared unequivocally upon restarting the drug, the earlier NS scoring of ‘+2’ was retained. It is pertinent to note that the NS also allows for a ‘Yes’ if the causal association is well known and rechallenge cannot be done for clinical or ethical reasons.<sup>15</sup>

Kane-Gill also suggested that question 6 (administration of placebo) should be removed because it is highly unlikely to be answered.<sup>17</sup> However, our

**Table 1.** Suggested modifications in the Naranjo scale.

Naranjo question no.	Question	Response	Scoring as per NS	Scoring as per SNG algorithm
2	Did the AE appear after the suspected drug was administered?	Yes	+2	+1
		No	-1	-1
		Don't know	0	0
3	Did the AE improve when the drug was discontinued or a specific antagonist was administered? OR In case of chemotherapy administered cyclically, did the AE improve before the next dose (with or without causing a delay in administering the dose)	Yes	+1	+2
		No	0	-1
		Don't know	0	0
4	Did the AE reappear when the drug was re-administered?	Yes	+2	+2
		No	-1	-2
		Don't know	0	0
5	Are there alternative causes that could, on their own, have caused the reaction? If yes, your decision is based on (a) Case reports/Cohort study/Expertise RCT (large/small)	Yes	-1	(a) -1/(b) -2
		No	+2	+1
		Don't know	0	0
7	Was the drug detected in blood or other fluids in concentrations known to be toxic?	Yes	+1	+2
		No	0	-1
		Don't know	0	0
8	Was the reaction more severe when the dose was increased or less severe when the dose was decreased?	Yes	+1	+2
		No	0	0
		Don't know	0	0

AE, adverse event; NS, Naranjo scale; SNG, Sharma-Nookala-Gota; RCT, randomised controlled trial.

objective was to validate a scale, sufficiently generic, that it can be used in any situation. For example, many functional disorders mimic drug-induced AEs. Use of placebo would be important to confirm the causal association in such cases. Therefore, this question was also retained.

Question 5 deals with alternative causes that could, on their own, have caused the reaction. Since AEs are often nonspecific and can be manifestations of the disease being treated or an unrelated, concurrent disease or condition, other diagnoses need to be considered and excluded.<sup>10</sup> Also, concomitant

medications may be involved in causation of AEs. The strength of association between an alternative cause and the AE should determine the score for this question. We therefore believe that the level of evidence available for an alternative cause should be considered while scoring question 5. Hence, we split the scoring of a 'Yes' response to '-1' if the evidence in favour of an alternative causes are case-control studies and expert opinions, or '-2' if the evidence is coming from randomised controlled studies (RCTs) or meta-analysis/systematic reviews. In addition, Rehan *et al.* made a very pertinent observation that the source or cause of

disagreement in CA between raters was attributable to this question in most cases because of the difference in scores allowed for 'no' (score = +2) and 'do not know' (score = 0) responses.<sup>18</sup> Absence of an alternative cause is an indirect, and hence weak, pointer towards causal association between the drug and event. Hence, the scoring for absence of an alternative explanation was reduced from '+2' to '+1'. This will also ensure that any disagreement between raters will not impact the overall score significantly.

Question 7 refers to levels of drug detected in blood or other fluids in concentrations known to be toxic. Kane-Gill observed that this question should not be retained in NS since it is not practical to monitor all drugs by blood concentration.<sup>17</sup> This question applies specifically to exposure-dependent ADRs where drug levels above a certain threshold are known to be toxic. Drug level monitoring is fairly routine practice for anti-epileptics, anti-tubercular, some anti-cancer drugs, immunosuppressants and antifungals.<sup>21–25</sup> Question 7 is relevant in such situations. For instance, visual disturbance and hallucinations are known to occur at trough voriconazole levels above 6 mg/l. Thus, a patient presenting with hallucinations has trough voriconazole levels above this threshold at the same time, it will certainly aid in implicating the drug as a causal factor for the AE.<sup>26</sup> Similarly, non-toxic or sub-threshold levels can be good negative predictors of association. As a result, the score of response 'Yes' indicating objective evidence of toxic drug levels was changed from '+1' to '+2' and for 'No' from '0' to '-1'. No alterations were made to the 'Do not know' response.

Kane-Gill challenged the practicality of the 8th question, which deals with intensity of the AE when the dose was increased or decreased. We felt that, in many cases, physicians cannot discontinue medication as the drug may be critical for survival and all reactions are not dose dependent. In such cases, response to questions regarding dechallenge and rechallenge cannot be answered (questions 3 and 4, respectively). Question 8 establishes the dose–response relationship between the drug and the AE.<sup>17</sup> While a positive dose–response relationship strongly implicates the drug as the causal factor, absence of dose–response relationship does not necessarily absolve the causal association. For instance, the hypersensitivity reaction

to penicillin is not dose dependent and hence there will not be any change in intensity of reactions even if doses are reduced.<sup>27</sup> Therefore, the scoring was appropriately changed from '+1' to '+2' for 'Yes'. The score of '0' was retained for absence of dose–response relationship without resorting to negative scoring. Again, no alteration in scoring was made to the 'Do not know' response.

#### *Interpretation of scoring*

The NS categorises AEs into four categories: definite ( $\geq 9$ ), probable (8–5), possible (4–1) and doubtful ( $\leq 0$ ). In the SNG algorithm, slabs for each grade of causality was moved upwards by 2 points from 'doubtful' and above and a new category, 'not related', was introduced for scores  $\leq -1$  as shown in Table 2.

The SNG algorithm is shown in Table 2. It was initially tested in a training set of 19 AEs from the outpatient clinics of our hospital. These cases were assessed using the NS and SNG causality assessment algorithms and compared against experienced physicians' opinion as gold standard. Identified cases were verified for being reported or not in a drug database Micromedex® Medication, Disease and Toxicology Management Solutions (Truven Health Analytics, Grand Rapids, MI, USA). The cases were selected such that they covered the entire spectrum of causality from 'doubtful' to 'definite' as per NS.

The SNG algorithm was further validated with AEs noted in patients enrolled in the study entitled 'Monitoring adverse drug reactions of pemetrexed and platinum doublet chemotherapy in patients with non-squamous non-small cell lung cancer or mesothelioma'. AEs were captured prospectively from February 2016 to August 2017. Patients of either gender aged 18 years and above with histologically proven non-squamous non-small cell lung cancer (NSCLC) and mesothelioma were included in the study. The treatment regimen comprised of four cycles of pemetrexed–carboplatin doublet chemotherapy followed by maintenance pemetrexed at 3-weekly intervals until disease progression.

AEs were identified primarily in the out-patient department (OPD) during regular follow up.

**Table 2.** SNG causality assessment algorithm.

Sr. no:	Questions	Yes	No	Don't know
1	Presence of previous conclusive report on adverse reaction.	+1	0	0
2	Did AE appear subsequent to administration of suspected drug?	+1	-1	0
3	(a) Did the AE improve when the drug was discontinued or a specific antagonist was administered? <b>OR</b> (b) In case of chemotherapy administered cyclically, did the AE improve before the next dose (with or without causing a delay in administering the dose)	+2	-1	0
4	Did the AE reappear when the drug was re-administered?	+2	-2	0
5	Are there any alternative causes other than the suspected drug that could have caused the reaction on their own? (a) Case reports/Cohort study/Expertise RCT (large/small)	-1 -2	+1	0
6	Did the AE reappear when a placebo was administered?	-1	+1	0
7	Was the incriminated drug detected in toxic concentrations in blood (fluids)?	+2	-1	0
8	Did the AE worsen on increasing the dose or decreased in severity with lower doses?	+2	-1	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 AE confirmed by objective evidence?	+1	0	0
Total score:				
Interpretation of score: $\leq -1$ , Not Related; 0–2, Doubtful; 3–6: Possible; 7–10, Probable; $\geq 11$ , Definite. AE, adverse event; RCT, randomised controlled trial; SNG, Sharma-Nookala-Gota.				

Three members of the study team determined causality using both NS and SNG algorithms, which were subsequently compared against physician's opinion as gold standard.

#### Statistical analysis plan

Kappa statistics was used to test concordance between the two CA scales and the gold standard – physician's opinion. Sensitivity, specificity, kappa values and accuracy of CA scales against the gold standard was carried out using  $2 \times 2$  contingency table function in SPSS® (version 21.0, IBM, Chicago, IL, USA). Categorical data was compared between groups using the chi-square test. Outcomes of the assessment results (related or unrelated) of SNG algorithm and NS were compared against the gold standard by McNemar's test.

#### Results

During the training set, assessment of the original 19 AEs by the authors using the NS categorised the AEs as 'doubtful' ( $n=1$ ), 'possible' ( $n=8$ ), 'probable' ( $n=5$ ) and 'definite' ( $n=5$ ). The same cases assessed using the SNG algorithm categorised the AEs as 'not related' ( $n=1$ ), 'doubtful' ( $n=5$ ), 'possible' ( $n=4$ ), 'probable' ( $n=7$ ) and 'definite' ( $n=2$ ). The number of questions answered to arrive at the final causality score in each case is shown in Table 3. The CA of all cases using NS and SNG algorithms arranged in decreasing order of the strength of association is also shown in Table 3.

The validation set consisted of 104 AEs observed in 65 patients. The cases were categorised as 'doubtful' ( $n=6$ , 6%), 'possible' ( $n=36$ , 35%), 'probable' ( $n=50$ , 48%) and 'definite' ( $n=12$ , 12%) in the NS, whereas the SNG algorithm

**Table 3.** Causality assessments of AEs using SNG causality assessment algorithm and Naranjo scales arranged in decreasing order of the strength of association.

Sr. no	Drug	AE	No of questions answered	Naranjo scale		SNG causality assessment algorithm	
				Score	Category	Score	Category
1	Sunitinib	Leukopenia and neutropenia	8	11	Definite	13	Definite
2	Gemcitabine	Transaminitis	7	10	Definite	11	Definite
3	Telmisartan	Hyperkalemia	7	9	Definite	10	Probable
4	Carboplatin	Hyponatremia	6	9	Definite	9	Probable
5	Carboplatin	Febrile neutropenia and Leukopenia	6	9	Definite	9	Probable
6	Linezolid	Black hairy tongue	6	8	Probable	8	Probable
7	Irinotecan	Febrile neutropenia	7	7	Probable	8	Probable
8	Pemetrexed	Thrombocytopenia	6	7	Probable	7	Probable
9	Telmisartan	Increased serum creatinine	7	6	Probable	7	Probable
10	Gemcitabine	Thrombocytosis	4	6	Probable	5	Possible
11	Gefitinib	Transaminitis	4	5	Probable	5	Possible
12	Paclitaxel+ carboplatin	Febrile neutropenia and Leukopenia	4	5	Probable	5	Possible
13	TKI (drug under phase I study)	Thromboembolism	4	4	Possible	3	Possible
14	Gefitinib	Hyponatremia	2	3	Possible	2	Doubtful
15	Pegaspargase	Febrile neutropenia	3	2	Possible	1	Doubtful
16	Tyrosine Kinase Inhibitor	Dyspnoea	3	2	Possible	1	Doubtful
17	Tyrosine Kinase Inhibitor	Thrombocytopenia	3	2	Possible	0	Doubtful
18	Tyrosine Kinase Inhibitor	Diarrhoea	4	1	Possible	0	Doubtful
19	Vinorelbine	Anemia	2	-1	Doubtful	-2	Not related

AE, adverse event; SNG, Sharma-Nookala-Gota; TKI, tyrosine kinase inhibitor.

classified the causality as ‘not related’ ( $n=18$ , 17%), ‘doubtful’ ( $n=19$ , 18%), ‘possible’ ( $n=30$ , 29%), ‘probable’ ( $n=35$ , 34%) and ‘definite’ ( $n=02$ , 2%). The NS called 98 out of 104 (94%) events as ‘likely’ related to the drug. On the other hand, the SNG algorithm was more conservative with 37 out of 104 (35.6%) being called ‘unlikely’ to be related to the drug (Table 4). The performance indicators of NS and SNG algorithm in

training and validation sets respectively against the gold standard is shown in Table 5. The distribution of scores for each of the 104 cases on the NS and SNG algorithm is shown in Figure 1

McNemar’s test for comparing the outcomes of assessment showed that the causality assessments by SNG and gold standard (physicians’ consensus) were not statistically different ( $p=1$ )

**Table 4.** Validation set results.

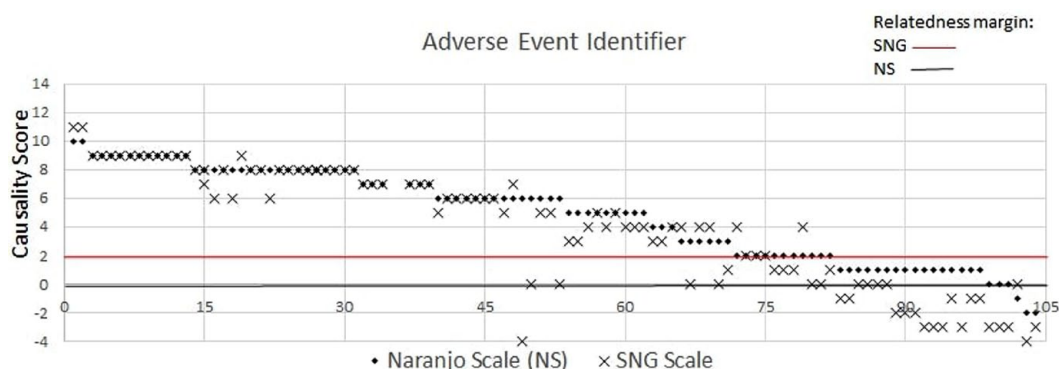
Category (n = 104)	Causality	NS		SNG causality assessment algorithm	
		No. of cases	Percentage (%)	No. of cases	Percentage (%)
Not Related	Unlikely	–	–	18	17.31
Doubtful		06	5.77	19	18.27
Possible	Likely	36	34.62	30	28.85
Probable		50	48.08	35	33.65
Definite		12	11.54	2	1.92

NS, Naranjo scale; SNG, Sharma-Nookala-Gota.

**Table 5.** Comparison of NS and SNG causality algorithm performance against the gold standard in training and validation sets.

Set and analysis	Training set (n = 19)				Validation set (n = 104)			
	NS compared with gold standard		SNG algorithm compared with gold standard		NS compared with gold standard		SNG algorithm compared with gold standard	
	Values	95% CI	Values	95% CI	Values	95% CI	Values	95% CI
Kappa	0.215	-0.082–0.215	1	0.451–1	0.20	0.058–0.341	0.958	0.9–1
Concordance	73.70%	0.637–0.737	100%	0.763–1	70.19%	61.4–79%	98.08%	95.4–100.7%
Sensitivity	1	0.927–1	1	0.827–1	100%	94.64–100%	98.51%	91.96–99.96%
Specificity	0.167	0.009–0.167	1	0.624–1	16.62%	6.19–32.01%	97.3	85.84–99.93%
Positive predictive value	0.722	0.670–0.722	1	0.827–1	68.37%	65.23–71.35%	98.51%	90.52–99.78%
Negative predictive value	1	0.055–1.0	1	0.624–1	100%	0.530–1.000	97.3%	83.72–99.6%

CI, confidence interval; NS, Naranjo scale; SNG, Sharma-Nookala-Gota.



**Figure 1.** Graphical representation of the distribution of scores by the NS (◆) and SNG causality algorithm (×) scales for each of the 104 AEs. Causality scores above the horizontal lines indicate relatedness according to NS (black line) and modified NS (red line), respectively. AE, adverse event; NS, Naranjo scale; SNG, Sharma-Nookala-Gota.



(Table 6), while the outcomes of assessments by NS were significantly different from gold standard. ( $p=0.001$ ) (Table 7).

## Discussion

Causality assessment of AEs is ‘a quest for the holy grail’ in drug development and pharmacovigilance, since an accurate and reliable tool does not exist for routine use in clinical trials and practice. The existing causality scales suffer from lack of clarity, over-estimation of causal association and lack of consensus with expert judgement,<sup>12,28,29</sup> and attempts to improve the scales have been unsuccessful in overcoming these limitations satisfactorily. A recent model developed by Roche, an individual case safety report (ICSR) causality decision support tool called MONARCSi, is a logistic transformation of the NS with complex mathematical equations.<sup>30</sup> In spite of being mathematically intensive, the model exhibited moderate sensitivity (65%) with Roche safety professionals’ assessment of causality.

In our study, we observed that only 68 out of 104 (64.42%) AEs documented in the validation set were listed in Micromedex<sup>®</sup> for pemetrexed and carboplatin combined. The NS called 98 events as causally related, of which 30 were not listed in Micromedex<sup>®</sup>. This is a significant overestimation by NS because it is rather unusual for so many unreported ADRs to occur in a small cohort of 65 patients involving 104 AEs with a drug regime that has been in use since 2004.<sup>31</sup> On the other hand, the SNG algorithm called 67 AEs as causally related, all of which are listed on Micromedex<sup>®</sup>. However, the high specificity of the SNG algorithm came at the cost of slight loss in sensitivity, with one AE being ruled as ‘not related’ when it was related to the drug regimen.

Many researchers and clinicians have suggested modification to NS. However, none of the modified scales could completely answer the majority of the researchers’ issues. However, the SNG causality algorithm is able to answer queries relating to most of the researcher’s difficulties. We did also consider the recommendations of other researchers in finalising the questionnaire. In this manuscript we discuss about the issues faced by the various clinicians/researchers and how they are answered by the SNG causality algorithm.

**Table 6.** 2 × 2 Contingency table between the results of causality assessment by physicians’ consensus and SNG causality assessment algorithm.

Physician's consensus \ SNG algorithm	Related	Unrelated	Total
Related	66	1	67
Unrelated	1	36	37
Total	67	37	104

SNG, Sharma-Nookala-Gota.

**Table 7.** 2 × 2 Contingency table between the results of causality assessment by physicians’ consensus and NS causality assessment.

Physician's consensus \ Naranjo scale	Related	Unrelated	Total
Related	67	31	98
Unrelated	0	6	6
Total	67	37	104

NS, Naranjo scale.

Seeger *et al.* opined that positive response to the second question in NS placing the drug to be in the ‘possible’ category (of causing an ADR) even before any other questions are answered.<sup>16</sup> We agree with their opinion and hence the weight of the ‘yes’ response was reduced in the SNG. This reduction in weight allows other questions to have an impact on overall causality assessment.

Kane-Gill *et al.* opined that rechallenge is not commonly used in ICUs and hence suggested that this rechallenge question should be removed from NS.<sup>17</sup> We are not in agreement with Kane-Gill *et al.*, as the causality assessment questionnaire is not only for the ICU setting as rechallenge is relevant in all other scenarios and hence should be retained. The SNG causality algorithm is yet to be validated in the ICU setting.

Rehan *et al.* observed a high inter-rater disagreement observed among users of the NS; they attributed this to question #5 of NS in most cases because of the difference in scores assigned for ‘No’ and ‘Do not know’ responses (2 and 0,

respectively) in the NS.<sup>18</sup> We agree with the authors, because absence of an alternative cause is an indirect and hence a weak pointer towards causal association between the drug and an event. Hence, the scoring for absence of an alternate explanation was reduced from '+2' to '+1' in the SNG. Often, the existence of an alternative cause is not obvious and hence raters opinion may be divided between 'No alternative cause' and 'Do not know', leading to high inter-rater disagreements. Reducing the score from +2 to +1 in the SNG for the absence of an alternative explanation will also ensure that any disagreement between raters highlighted by Rehan *et al.* will not impact the overall score significantly.

Kane-Gill *et al.* suggested for removal of sixth question of NS because they felt administration of a placebo is not common in any patient population.<sup>17</sup> However, we feel that this question should be retained because there are functional disorders that may mimic drug-induced AEs. Use of placebo would be important to confirm the causal association in such cases. Furthermore, the purpose is to develop a general scale that can be used in any situation. The same authors felt that it is not practical to monitor all drugs by blood concentrations, and that question #7 of NS should be removed. However, we feel that drug level monitoring is a routine practice for anti-epileptics, anti-tubercular, some anti-cancer drugs, immunosuppressants and so on. If the scale must be relevant for all situations, this question should be retained. This will ensure the importance of measuring the drug levels.

We concur with the suggestion of Kane-Gill *et al.* regarding question 8 of NS that the question needs more clarification because all reactions are not dose dependent. Furthermore, wherever a reaction shows dose dependency, it strongly implicates the drug as the causal factor. Hence, the score has been changed from +1 to +2 in the SNG scale. However, the absence of dose-response relationship for an AE does not necessarily absolve the drug. Hence, the score of '0' is retained in the SNG scale, without resorting to negative scoring. These authors have also felt that question 10 of NS overlaps with question 7; however, we feel that there is no overlap because question 7 refers to detection of toxic levels of drug in blood or body fluids, whereas question 10 refers to objective evidence confirming the AE.

Researchers that grades of causality (e.g. 'possible', 'probable', 'definite') offer little practical advantage and only 'related' *versus* 'unrelated' is needed for regulatory reporting requirements.<sup>8</sup> However, we feel that the causality algorithm is intended not only in regulatory reporting but also in routine clinical practice. The strength of association becomes important if the scale must be used in clinical practice. Hence the grading 'possible', 'probable', 'definite' becomes important. For regulatory reporting, any associations 'possible' and beyond can be considered 'related'. 'Doubtful' and 'Not related' may be considered 'unrelated'.

Avner *et al.* raised a pertinent point that the NS does not deal adequately with the time relationship between drug administration and the AE.<sup>2</sup> They felt that more information is needed in addition to question 2, for example whether the appearance of symptoms corresponds to the typical time of appearance described for the specific ADR in the literature? We too agree with the authors, but we feel that the algorithm should be sufficiently flexible for use in situations where not much is known about the drug, as in early clinical development stages. Therefore, we did not modify the wording of question number 2 in the SNG algorithm; instead, we reduced the score by 1 to minimise false positive associations. In any case, in routine clinical practice, physicians are likely to respond to temporal association based on their knowledge of the typical time of appearance of AEs, whether or not it is stated explicitly.

A major limitation of our study is the limited number of cases in which the SNG algorithm was validated. Indeed, it was also one of the criticisms made about the NS by many commentators. Besides, the generalisability of the SNG algorithm to niche therapeutic areas such as paediatric and ICUs cannot be ascertained. More studies involving larger numbers of cases are required in these therapeutic domains to establish the superiority of SNG algorithm over NS. Besides, some of the shortcomings of NS also apply to the SNG algorithm. For instance, a conclusion of definite association cannot be made with the SNG algorithm despite a positive dechallenge and rechallenge response, if information on drug levels and dose-response relationships are not available. This information may not be available in majority of cases. Thus, although the SNG algorithm correctly categorised such events as 'related', the

strength of association would be underestimated owing to the conservative nature of the tool.

We validated the SNG algorithm in a clinical practice setting, which is its strength, thereby overcoming a major drawback of the NS, which was validated in the setting of an RCT.<sup>14,32</sup> The chosen regimen of pemetrexed and carboplatin has been in existence for more than a decade, and, therefore, expert's judgement can serve as gold standard in this setting owing to accumulated evidence of ADRs from a large body of patients. The poor performance of NS against physician's opinion in our study underscores its limitations. On the other hand, superior concordance achieved with the SNG algorithm supports our strategy of reallocation of slabs for categorisation of causality and revised weight allocation for certain questions.

Validation of the SNG algorithm in different settings including paediatrics and ICUs using larger cohorts of patients is underway. Validation is being carried out using drug regimens whose ADR profile is fairly well characterised. Once this is done, the SNG algorithm will be piloted in an early phase clinical trial where not much would be known about the ADR profile of the investigational drug. A series of such validations would be required before the SNG algorithm can replace NS as the preferred tool for CA.

To conclude, the SNG algorithm allowed better causality assessment between the drug and an AE, and showed high concordance with physician opinion. Unlike NS, the SNG algorithm was validated in a clinical setting by a team of oncologists and clinical pharmacologists with several years of experience treating patients with the chosen drug regimen. The NS showed a bias towards classifying AEs as 'drug related'. Decisive alterations including revised scoring of some questions and rearranging the slabs that determined the strength of association resulted in an algorithm with superior specificity while retaining the sensitivity of the NS. The SNG algorithm can be used for routine pharmacovigilance and causality assessment activities in routine clinical practice.


### Conflict of interest statement

The authors declare that there is no conflict of interest.

### Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

### ORCID iD

Manjunath Nookala Krishnamurthy  <https://orcid.org/0000-0002-9187-842X>

### Supplemental material

Supplemental material for this article is available online.

### References

1. Impicciatore P, Choonara I, Clarkson A, *et al.* Incidence of adverse drug reactions in paediatric in/out-patients: a systematic review and meta-analysis of prospective studies. *Br J Clin Pharmacol* 2001; 52: 77–83.
2. Avner M, Finkelstein Y, Hackam D, *et al.* Establishing causality in pediatric adverse drug reactions: use of the Naranjo probability scale. *Paediatr Drugs* 2007; 9: 267–270.
3. Powe NR, Eggers PW and Johnson CB. Early adoption of cyclosporine and recombinant human erythropoietin: clinical, economic, and policy issues with emergence of high-cost drugs. *Am J Kidney Dis* 1994; 24: 33–41.
4. de Vries EN, Ramrattan MA, Smorenburg SM, *et al.* The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care* 2008; 17: 216–223.
5. Pirmohamed M, James S, Meakin S, *et al.* Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *BMJ* 2004; 329: 15–19.
6. Kessler DP. Evaluating the medical malpractice system and options for reform. *J Econ Perspect* 2011; 25: 93–110.
7. Rehan HS, Chopra D and Kakkar AK. Physician's guide to pharmacovigilance: terminology and causality assessment. *Eur J Intern Med* 2009; 20: 3–8.
8. Naidu RP. Causality assessment: a brief insight into practices in pharmaceutical industry. *Perspect Clin Res* 2013; 4: 233–236.
9. Venulet J. Role and place of causality assessment. *Pharmacoepidemiol Drug Saf* 1992; 1: 225–234.
10. Khan LM, Al-Harthi SE, Osman A-MM, *et al.* Dilemmas of the causality assessment tools in the diagnosis of adverse drug reactions. *Saudi Pharm J* 2016; 24: 485–493.

11. Agbabiaka TB, Savović J and Ernst E. Methods for causality assessment of adverse drug reactions: a systematic review. *Drug Saf* 2008; 31: 21–37.
12. Mittal N and Gupta MC. Comparison of agreement and rational uses of the WHO and Naranjo adverse event causality assessment tools. *J Pharmacol Pharmacother* 2015; 6: 91–93.
13. García-Cortés M, Lucena MI, Pachkoria K, *et al.* Evaluation of Naranjo adverse drug reactions probability scale in causality assessment of drug-induced liver injury. *Aliment Pharmacol Ther* 2008; 27: 780–789.
14. Naranjo CA, Busto U, Sellers EM, *et al.* A method for estimating the probability of adverse drug reactions. *Clin Pharmacol Ther* 1981; 30: 239–245.
15. 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.
16. Seger D, Barker K and McNaughton C. Misuse of the Naranjo adverse drug reaction probability scale in toxicology. *Clin Toxicol (Phila)* 2013; 51: 461–466.
17. Kane-Gill SL, Kirisci L and Pathak DS. Are the Naranjo criteria reliable and valid for determination of adverse drug reactions in the intensive care unit? *Ann Pharmacother* 2005; 39: 1823–1827.
18. Rehan H, Chopra D and Kakkar A. Causality assessment of spontaneously reported adverse drug events: comparison of WHO-UMC criteria and Naranjo probability scale *Int J Risk Saf Med* 2007; 19: 223–227.
19. Banu AB, Alias Balamurugan SA and Thirumalaikolundusubramanian P. Detection of dechallenge in spontaneous reporting systems: a comparison of Bayes methods. *Indian J Pharmacol* 2014; 46: 277–280.
20. Doherty MJ. Algorithms for assessing the probability of an adverse drug reaction. *Respir Med CME* 2009; 2: 63–67.
21. Jacob S and Nair AB. An updated overview on therapeutic drug monitoring of recent antiepileptic drugs. *Drugs R D* 2016; 16: 303–316.
22. Herviou P, Thivat E, Richard D, *et al.* Therapeutic drug monitoring and tyrosine kinase inhibitors. *Oncol Lett* 2016; 12: 1223–1232.
23. Mohammadpour N, Elyasi S, Vahdati N, *et al.* A review on therapeutic drug monitoring of immunosuppressant drugs. *Iran J Basic Med Sci* 2011; 14: 485–498.
24. Ashbee HR, Barnes RA, Johnson EM, *et al.* Therapeutic Drug Monitoring (TDM) of antifungal agents: guidelines from the British Society for Medical Mycology. *J Antimicrob Chemother* 2014; 69: 1162–1176.
25. Joerger M, Schellens JHM and Beijnen JH. Therapeutic drug monitoring of non-anticancer drugs in cancer patients. *Methods Find Exp Clin Pharmacol* 2004; 26: 531–545.
26. Mikulska M, Novelli A, Aversa F, *et al.* Voriconazole in clinical practice. *J Chemother* 2012; 24: 311–327.
27. Legendre DP, Muzny CA, Marshall GD, *et al.* Antibiotic hypersensitivity reactions and approaches to desensitization. *Clin Infect Dis* 2014; 58: 1140–1148.
28. Théophile H, Arimone Y, Miremont-Salamé G, *et al.* Comparison of three methods (consensual expert judgement, algorithmic and probabilistic approaches) of causality assessment of adverse drug reactions: an assessment using reports made to a French pharmacovigilance centre. *Drug Saf* 2010; 33: 1045–1054.
29. Arimone Y, Miremont-Salamé G, Haramburu F, *et al.* Inter-expert agreement of seven criteria in causality assessment of adverse drug reactions. *Br J Clin Pharmacol* 2007; 64: 482–488.
30. Comfort S, Dorrell D, Meireis S, *et al.* MODified NARAnjo causality scale for ICSRs (MONARCSi): a decision support tool for safety scientists. *Drug Saf* 2018; 41: 1073–1085.
31. Zinner RG and Herbst RS. Pemetrexed in the treatment of advanced non-small-cell lung cancer: a review of the clinical data. *Clin Lung Cancer* 2004; 5(Suppl. 2): S67–S74.
32. Davies EC, Rowe PH, James S, *et al.* An investigation of disagreement in causality assessment of adverse drug reactions. *Pharm Med* 2011; 25: 17–24.