Adverse Drug Reaction Causality Assessment Tools for Drug-Induced Stevens-Johnson Syndrome and Toxic Epidermal Necrolysis: Room for Improvement

Jennifer L. Goldman[a], **Wen-Hung Chung**[b], **Brian R. Lee**[a], **Chun-Bing Chen**[b,c], **Chun-Wei Lu**[b], **Wolfram Hoetzenecker**[d,e], **Robert Micheletti**[h], **Sally Usdin Yasuda**[g], **David J. Margolis**[f,h], **Neil H. Shear**[i], **Jeffery P Struewing**[j], **Munir Pirmohamed**[k]

[a]Department of Pediatrics, Children's Mercy Hospitals & Clinic, Kansas City, MO 64108 USA [b]Department of Dermatology, Drug Hypersensitivity Clinical and Research Center, Chang Gung Memorial Hospital, Taipei, Linko and Keelung [c]College of Medicine, Chang Gung University, Taoyuan, Taiwan [d]Department of Dermatology, University of Zurich, Gloriastrasse 31, 8091 Zurich, Switzerland [e]Department of Dermatology, University Hospital Linz, Krankenhausstrasse 9, 4020 Linz, Austria [f]Department Dermatology and Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, PA 19104 USA [g]Food and Drug Administration, Silver Spring, MD 20993, USA [h]Department of Dermatology, University of Pennsylvania, Philadelphia, PA 19104 USA [i]Division of Dermatology, Department of Medicine, Sunnybrook Health Sciences Center and University of Toronto, Toronto, ON, Canada [j]Division of Genomic Medicine, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892 USA [k]Department of Molecular and Clinical Pharmacology, University of Liverpool, Liverpool, UK

Corresponding Author: Jennifer Goldman MD, MS, Assistant Professor, Department of Pediatrics, Divisions of Pediatric Infectious Diseases & Clinical Pharmacology, Children's Mercy, 2401 Gillham Rd., Kansas City, MO 64108. Phone: 816-234-3061, Fax: 816-346-1328, jlgoldman@cmh.edu.

Contribution of Authors statement

Dr. Goldman contributed to the concept and design, analysis and interpretation of data, drafting and revising the manuscript.

Dr. Chung contributed to the concept and design, acquisition of data, analysis and interpretation of data, drafting and revising the manuscript.

Dr. Lee contributed to the concept and design, analysis and interpretation of data, drafting and revising the manuscript

Dr. Chen contributed acquisition of data, drafting and revising the manuscript.

Dr. Lu contributed acquisition of data, drafting and revising the manuscript.t

Dr. Hoetzenecker contributed to the concept and design, analysis and interpretation of data, drafting and revising the manuscript.

Dr. Micheletti contributed to the concept and design, analysis and interpretation of data, drafting and revising the manuscript.

Dr. Yasuda contributed to the concept and design, analysis and interpretation of data, drafting and revising the manuscript.

Dr. Margolis contributed to the concept and design, analysis and interpretation of data, drafting and revising the manuscript.

Dr. Shear contributed to the concept and design, analysis and interpretation of data, drafting and revising the manuscript.

Dr. Struewing contributed to the concept and design, analysis and interpretation of data, drafting and revising the manuscript.

Dr. Pirmohamed contributed to the concept and design, analysis and interpretation of data, drafting and revising the manuscript.

Ethical statement

Since the data received by the SJS/TEN Standardizing Case Definitions Working Group was de-identified and no potentially identifiable variables were provided, no ethical approval was needed.

Conflict of interest

The authors declare they have no conflict of interest.

Abstract

Purpose: Establishment of causality between drug exposure and adverse drug reactions (ADR) is challenging even for serious ADRs such as Stevens-Johnson syndrome/toxic epidermal necrolysis (SJS/TEN). Several causality assessment tools (CAT) exist, but the reliability and validity of such tools is variable. The objective of this study was to compare the reliability and validity of existing ADR CAT on SJS/TEN cases.

Methods: Seven investigators completed three CAT (ALDEN, Naranjo, Liverpool) for 10 SJS/TEN cases. Each CAT categorized the causality of 30 potential drugs as definite/very probable, probable, possible, or doubtful/unlikely. An additional reviewer provided expert opinion by designating the implicated drug(s) for each case. A Kappa score was generated to compare CAT responses both by method (reliability of all 7 reviewers, by CAT) and by reviewer (reliability of the 3 CAT, by reviewer). A c-statistic was calculated to assess validity.

Results: Inter-rater reliability by CAT was poor to fair: ALDEN 0.22, Naranjo 0.11, and Liverpool 0.12. Reliability was highest when causality classification was definite/very probable (0.16-0.41). Similarly, intra-rater reliability by reviewer was poor. When comparing the validity of the overall CAT to expert reviewer, area under the curve was highest for ALDEN (c-statistic 0.65) as compared to Liverpool (0.55) or Naranjo (0.54).

Conclusion: Available CAT have poor reliability and validity for drug-induced SJS/TEN. Due to the importance of determining ADR causality for research, industry and regulatory purposes, development of an enhanced tool that can incorporate data from immunological testing and pharmacogenetic results may strengthen CAT usefulness and applicability for drug-induced SJS/TEN.

Keywords

INTRODUCTION

The establishment of causality between drug exposure and adverse drug reactions (ADRs) is challenging. To date there are few diagnostic tools available to confirm or refute an implicated drug. Thus obtaining a comprehensive history, including timing of exposure, onset of ADR symptoms, previous reactions to similar medication, and other associated risk factors, and having an understanding of the pharmacological profile of the implicated drug as related to an ADR is critical in assessing causality.

A standardized approach for drug causality assessment is recommended [1]. Causality assessment tools (CAT) provide guidance for gathering critical information related to an ADR event. Several CAT exist which categorize the relationship between drug exposure and ADR as unlikely, possible, probable, or definite. Some CAT have been developed for specific types of ADRs such as the Roussel Uclaf Causality Assessment Method for drug induced liver injury or the algorithm of drug causality for epidermal necrolysis (ALDEN) specific for cases of Stevens-Johnson Syndrome (SJS) and toxic epidermal necrolysis (TEN)

[2,3]. Other CAT such as the Naranjo or Liverpool are not specific to a clinical presentation and can thus be used for a variety of ADRs [4,5].

Though CAT were developed to assist in determining a link between drug exposure and ADR, agreement between causality tools is poor [6]. To date, no studies exist comparing the ALDEN, Liverpool, and Naranjo in the assessment of SJS/TEN ADR cases. The objective of this study was to compare the reliability of these three CAT in assessing SJS/TEN cases, and quantify the validity by comparing the results to expert judgement.

METHODS

Causality assessment tools

Seven reviewers independently completed three CAT (ALDEN, Liverpool, Naranjo) for 11 Stevens-Johnson syndrome/Toxic epidermal necrolysis (SJS/TEN) cases. Briefly, the Naranjo consists of 10 questions with yes/no/do not know options for each response [5]. A score is provided for each question based on the response, and the sum of the scores determines the causality classification of doubtful, possible, probable, or definite. Similarly, the ALDEN consists of 6 criteria with an associated question and score based on response, with the total score determining the causality as very unlikely, unlikely, possible, probable, or very probable [3]. The Liverpool Tool is a visual algorithm consisting of yes/no questions that determine the path to the next question and final causality classification of unlikely, possible, probable, or definite [4]. Each CAT was applied to categorize all potential drugs as definite/very probable, probable, possible, or doubtful/unlikely as causing SJS/TEN. For analysis, the ALDEN results of very unlikely and unlikely were grouped together and classified as doubtful/unlikely. An additional reviewer (NS) provided expert opinion by designating the most likely implicated drug(s) for each case using clinical judgement without use of a CAT.

Clinical cases

Eleven real and randomly chosen historical clinical cases were provided by the authors CC and CL, who did not take part in the causality assessments. All cases were diagnosed on the basis of RegiSCAR criteria and had undergone rigorous evaluation by a dermatologist and a SJS/TEN review committee at the time of clinical presentation [7,8]. For this study, information for each SJS/TEN clinical case was de-identified and then provided to the 7 reviewers. The information included a brief medical history, detailed clinical presentation and relationship of the cutaneous ADR onset to any recent drug exposure. Laboratory evaluations and data regarding skin biopsy results were provided when available. The timing of the initiation through discontinuation was provided for each drug when available.

Statistical analysis

Reliability was evaluated using Cohen's Kappa. Agreement was measured 1) by method when comparing all reviewers within each CAT ["inter-rater reliability"], 2) by reviewer when comparing reliability across the 3 CAT for each reviewer ["intra-rater reliability"], and 3) by case when comparing all reviewers within each method. Kappa results were interpreted as follows: values < 0 as indicating no agreement and 0.01–0.20 as none to

slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement.[9] Somers' D and a c-statistic were calculated to assess validity of the 7 reviewers' results when compared to expert opinion (NS). A c-statistic with a value of 0.5 indicates that the CAT is no better at identifying the implicated drug than random chance when compared to expert opinion, a value over 0.7 indicates a good model, and a value of 1 means that the CAT perfectly predicted agreement with expert opinion.[10] We report confidence intervals for the Kappa scores based on bias-corrected and accelerated (BCa) bootstrapped intervals, using 1000 replications. All analyses were completed using the *'irr' 'pROC'* and '*Hmisc*' packages in R (version 3.3.2).

RESULTS

Eleven SJS/TEN cases were initially examined. We excluded one case from the final analysis as reviewers determined the cutaneous reaction occurred prior to drug exposure. The final analysis included 10 cases involving 30 drugs (Table 1).

Overall inter-rater reliability by CAT was poor to fair. The Kappa for ALDEN was 0.223 (95% confidence interval (Cl) 0.141, 0.355), Naranjo 0.112 (95% Cl 0.019, 0.266), and Liverpool 0.124 (95% Cl 0.034, 0.273). In general, the Kappa increased with increasing perceived likelihood of the drug causing the reaction (Table 2). Moderate agreement occurred when the ALDEN response classified a drug as definite/very probable and this was the highest level of agreement achieved across all CAT.

Similarly, intra-rater reliability by reviewer was generally poor when comparing across the 3 CAT (Supplemental Table 1). Only a single reviewer (reviewer #1) achieved overall moderate agreement (Kappa: 0.466) when evaluating the same drugs using the different CAT. Similar to the inter-rater reliability, the Kappa was highest with a definite/very probable response. Examination of all reviewer results stratified by both case and method failed to improve agreement (Fig. 1).

When comparing the validity of an individual CAT to the expert reviewer, the area under the curve was highest for the ALDEN (c-statistic; 0.65) as compared to the Naranjo (0.52) or Liverpool (0.54). Agreement between CAT and expert review occurred most frequently when the reviewers' response for a given drug was definite/very probable. A definite result by the Naranjo aligned with the expert reviewer 100% of the time, though only 2 responses fell into this category by Naranjo scoring. Using Liverpool CAT, 8 responses were deemed definite with 88% agreement with expert opinion as compared to the ALDEN with 36 responses deemed definite with 86% agreement. Agreement was lowest (56%) when comparing responses determined by ALDEN as unlikely as compared to expert reviewer.

DISCUSSION

Determining the likelihood of a drug exposure resulting in an ADR is important, yet the effectiveness of available CAT is insufficient [11,12]. The ability to discern whether or not a drug resulted in an adverse reaction has several implications. First, a medical provider must make future prescribing decisions for a patient following an adverse reaction. Establishment of causality helps guide which drug classes should and should not be used in the future.

Second, pharmacovigilance programs used by healthcare systems, the pharmaceutical industry, and regulatory agencies are reliant on determining causality between drug exposure and adverse reactions for the detection of both existing and new ADR signals. Third, research focused on identifying ADR predictors requires detailed phenotyping of ADR patients including drug exposure and causality. The findings from this study demonstrate overall poor performance of the three CAT based on inter-rater reliability, reliability by reviewer when comparing across the 3 CAT, and validity when compared to expert opinion.

Our study demonstrates CAT results have low overall agreement even when used by specialists in the field of drug safety. Even when reviewers were provided the same data, interpretation was highly variable. CAT agreement appeared highest when the drug was deemed the definitive culprit. Inter-rater agreement has been previously shown to be highest when results are more conclusive such as a 'definite' classification [13]. Inter-rater reliability was poor when the drug was determined as unlikely to have caused SJS/TEN. Additionally, the CAT used for this study were only slightly better than chance of predicting the implicated drug when compared to an expert reviewer.

This study is unique as the seven reviewers provide geographical representation across the globe making the findings more generalizable. Our ADR cases were limited to severe cutaneous reactions. The selection of CAT included for this study had not been previously compared. These specific CAT were selected due to their unique characteristics: the Naranjo tool which can be applied to all ADRs regardless of phenotype, the ALDEN is specific to SJS/TEN cases, and the Liverpool is non-ADR specific but presented in a flow diagram as compared to Naranjo and ALDEN table scoring systems. Despite these differences, our findings align with previous studies demonstrating the overall poor reliability of causality assessment tools [11,13,12,14].

ADRs are under-recognized, underreported, and CAT are not consistently utilized in the medical setting [15,16]. To date, no universally accepted CAT has been identified as providing highly reliable and valid results, thus clinicians often rely on clinical judgment alone which can be highly subjective [17]. A CAT that is easy to use while providing useful and reliable results to help guide future prescribing is needed. As the field of drug safety evolves, more information becomes available regarding potential predictors associated with the development of ADRs. Efforts continue in the identification of genetic markers associated with ADRs, including serious skin reactions [18]. New information on drug metabolism and the immune system continues to advance our current understanding of ADR development and risk [19,20]. An enhanced tool for drug-induced SJS/TEN that can incorporate data from immunological testing (e.g. lymphocyte transformation test), pharmacogenetic results (e.g. human leukocyte antigen, drug metabolizing enzyme genotype), and pharmacokinetic data may strengthen the usefulness and applicability of CAT.

Our study has limitations. This study was retrospective and application of CAT was based on the case documentation provided. We reviewed 11 cases only - however, we simulated kappa calculations by randomly sampling cases with replacement from our existing data. Based on the simulation, 250 cases would be needed to see non-overlapping confidence intervals

between ALDEN and the other methods and the kappa scores did not vary, even when including 1000 simulated cases, supporting our findings that reliability is poor for the 3 CAT assessed. However the 250 simulated cases was only a calculated estimate. Larger studies in this population, such as the one performed by Sassolas et al. [3], should be performed in the future to further validate our findings. No validated testing to serve as the gold standard for SJS/TEN cases is available and thus we relied on expert opinion. A single expert with vast experience in making the clinical diagnosis of SJS/TEN served as the "gold standard" in this study to best represent what occurs in the clinical setting when causality was assessed. We recognize that there is potential variability in expert opinion, however evaluating variability in expert opinion was beyond the scope of this study. Not every reviewer completed CAT for every drug, resulting in sporadic missing data points. Only 30 drugs were included in this study, which may have impacted our ability to determine the true reliability of these assessment tools. For this study, our working group focused only on drug associated SJS/TEN. Future studies should include non-drug induced reactions.

In conclusion, the currently available CAT have poor reliability and validity for drug-induced SJS/TEN. Due to the importance of determining ADR causality for patient care, research, pharmaceutical industry and regulatory purposes, development of an enhanced tool for drug-induced SJS/TEN that can incorporate data from immunological testing and pharmacogenetic results may strengthen CAT usefulness and applicability.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

REFERENCES

1. Wheatley LM, Plaut M, Schwaninger JM, Banerji A, Castells M, Finkelman FD, Gleich GJ, Guttman-Yassky E, Mallal SA, Naisbitt DJ, Ostrov DA, Phillips EJ, Pichler WJ, Platts-Mills TA, Roujeau JC, Schwartz LB, Trepanier LA (2015) Report from the National Institute of Allergy and Infectious Diseases workshop on drug allergy. J Allergy Clin Immunol 136 (2):262–271 e262. doi: 10.1016/j.jaci.2015.05.027 [PubMed: 26254053]

2. Danan G, Benichou C (1993) Causality assessment of adverse reactions to drugs--I. A novel method based on the conclusions of international consensus meetings: application to drug-induced liver injuries. Journal of clinical epidemiology 46 (11):1323–1330 [PubMed: 8229110]

3. Sassolas B, Haddad C, Mockenhaupt M, Dunant A, Liss Y, Bork K, Haustein UF, Vieluf D, Roujeau JC, Le Louet H (2010) ALDEN, an algorithm for assessment of drug causality in Stevens-Johnson Syndrome and toxic epidermal necrolysis: comparison with case-control analysis. Clin Pharmacol Ther 88 (1):60–68. doi:10.1038/clpt.2009.252 [PubMed: 20375998]

4. Gallagher RM, Kirkham JJ, Mason JR, Bird KA, Williamson PR, Nunn AJ, Turner MA, Smyth RL, Pirmohamed M (2011) Development and inter-rater reliability of the Liverpool adverse drug reaction causality assessment tool. PLoS One 6 (12):e28096. doi:10.1371/journal.pone.0028096 [PubMed: 22194808]

5. Naranjo CA, Busto U, Sellers EM, Sandor P, Ruiz I, Roberts EA, Janecek E, Domecq C, Greenblatt DJ (1981) A method for estimating the probability of adverse drug reactions. Clin Pharmacol Ther 30 (2):239–245. [PubMed: 7249508]

6. Macedo AF, Marques FB, Ribeiro CF, Teixeira F (2005) Causality assessment of adverse drug reactions: comparison of the results obtained from published decisional algorithms and from the evaluations of an expert panel. Pharmacoepidemiology and drug safety 14 (12):885–890. doi: 10.1002/pds.1138 [PubMed: 16059869]

7. The RegiScar Project. http://www.regiscar.org/cht/index.html. Last accessed May 21, 2018

8. Wang YH, Chen CB, Tassaneeyakul W, Saito Y, Aihara M, Choon SE, Lee HY, Chang MM, Roa FD, Wu CW, Zhang J, Nakkam N, Konyoung P, Okamoto-Uchida Y, Man-Tung Cheung C, Huang JW, Ji C, Cheng B, Chung-Yee Hui R, Chu CY, Chen YJ, Wu CY, Hsu CK, Chiu TM, Ho HC, Lin JY, Yang CH, Chang YC, Su SC, Wang CW, Chung WH, Asian Severe Cutaneous Adverse Reaction C (2018) The Medication Risk of Stevens-Johnson Syndrome and Toxic Epidermal Necrolysis in Asians: The Major Drug Causality and Comparison to the USA FDA Label. Clin Pharmacol Ther 105(1):112–120. doi:10.1002/cpt.1071 [PubMed: 29569740]

9. McHugh ML (2012) Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 22 (3):276–282 [PubMed: 23092060]

10. Hosmer DW, Lemeshow S. Applied Logistic Regression (2nd Edition). New York, NY: John Wiley & Sons; 2000.

11. Agbabiaka TB, Savovic J, Ernst E (2008) Methods for causality assessment of adverse drug reactions: a systematic review. Drug safety : an international journal of medical toxicology and drug experience 31 (1):21–37

12. Theophile H, Arimone Y, Miremont-Salame G, Moore N, Fourrier-Reglat A, Haramburu F, Begaud B (2010) 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 safety : an international journal of medical toxicology and drug experience 33 (11):1045–1054. doi: 10.2165/11537780-000000000-00000

13. Arimone Y, Begaud B, Miremont-Salame G, Fourrier-Reglat A, Moore N, Molimard M, Haramburu F (2005) Agreement of expert judgment in causality assessment of adverse drug reactions. Eur J Clin Pharmacol 61 (3):169–173. doi:10.1007/s00228-004-0869-2 [PubMed: 15827761]

14. Varallo FR, Planeta CS, Herdeiro MT, Mastroianni PC (2017) Imputation of adverse drug reactions: Causality assessment in hospitals. PLoS One 12 (2):e0171470. doi:10.1371/journal.pone.0171470 [PubMed: 28166274]

15. Hazell L, Shakir SA (2006) Under-reporting of adverse drug reactions : a systematic review. Drug safety : an international journal of medical toxicology and drug experience 29 (5):385–396

16. Smyth RM, Gargon E, Kirkham J, Cresswell L, Golder S, Smyth R, Williamson P (2012) Adverse drug reactions in children--a systematic review. PLoS One 7 (3):e24061. doi: 10.1371/journal.pone.0024061 [PubMed: 22403604]

17. Pande S (2018) Causality or Relatedness Assessment in Adverse Drug Reaction and Its Relevance in Dermatology. Indian J Dermatol 63 (1):18–21. doi:10.4103/ijd.IJD_579_17 [PubMed: 29527021]

18. Collins SL, Carr DF, Pirmohamed M (2016) Advances in the Pharmacogenomics of Adverse Drug Reactions. Drug safety : an international journal of medical toxicology and drug experience 39 (1): 15–27. doi:10.1007/s40264-015-0367-8

19. Cho T, Uetrecht J (2017) How Reactive Metabolites Induce an Immune Response That Sometimes Leads to an Idiosyncratic Drug Reaction. Chem Res Toxicol 30 (1):295–314. doi:10.1021/acs.chemrestox.6b00357 [PubMed: 27775332]

20. Ogese MO, Ahmed S, Alferivic A, Betts CJ, Dickinson A, Faulkner L, French N, Gibson A, Hirschfield GM, Kammuller M, Meng X, Martin SF, Musette P, Norris A, Pirmohamed M, Park BK, Purcell AW, Spraggs CF, Whritenour J, Naisbitt DJ (2017) New Approaches to Investigate Drug-Induced Hypersensitivity. Chem Res Toxicol 30 (1):239–259. doi:10.1021/acs.chemrestox. 6b00333 [PubMed: 27806199]

KEY POINTS

- Establishment of causality between drug exposure and serious cutaneous adverse drug reactions is challenging.

- Three existing causality assessment tools performed poorly when assessing causality for drug-induced Stevens-Johnson syndrome and toxic epidermal necrolysis.

- Future work on developing an enhanced tool that can incorporate important clinical and genetic data may improve the usefulness and applicability of causality assessment tools for serious adverse drug reactions.
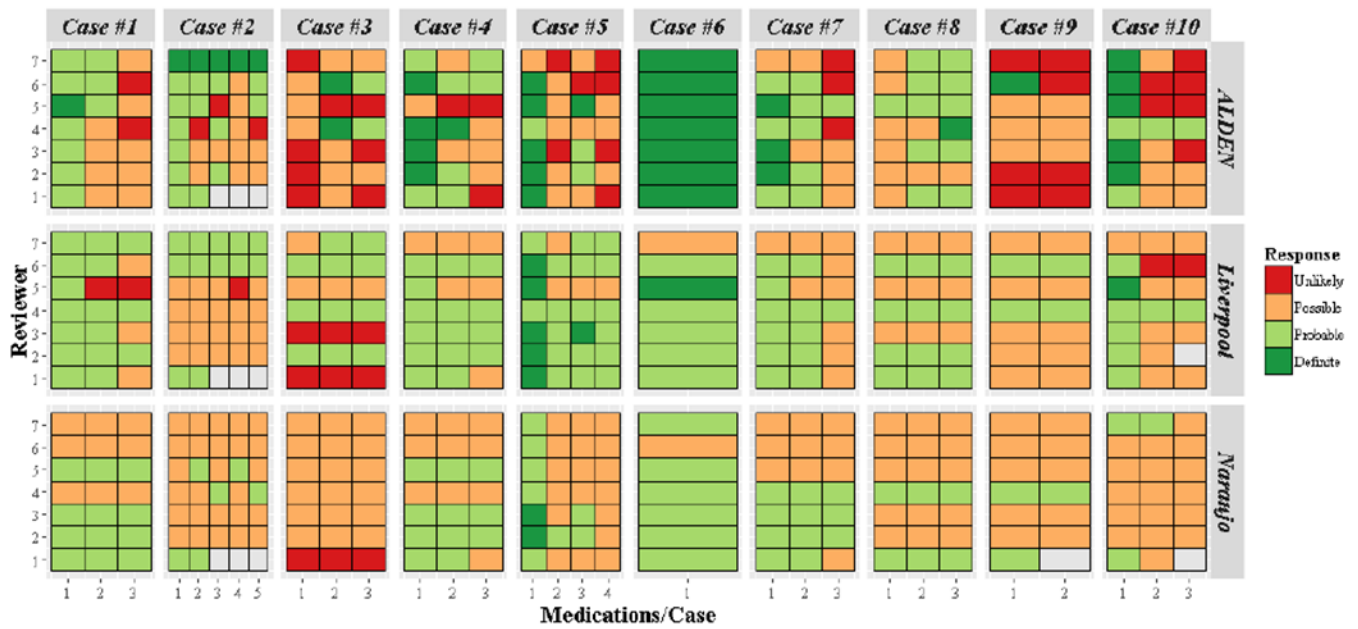
Fig. 1. Variability of Causality Assessment Tool (CAT) Results for 10 SJS/TEN Cases
Results of three CAT involving 10 cases of severe cutaneous drug reactions performed by 7 reviewers is displayed. Causality classification of unlikely, possible, probable, or definite is depicted by color/shading. The figure demonstrates the observed inter and intra-rater reviewer variability by case and CAT.

Table 1.

Cutaneous Skin Reaction Cases

Case	Underlying medical conditions	Drug Exposure	Skin Findings (% Total skin involvement)	Mucosa involved	Skin Biopsy performed	Diagnosed by Dermatologist	Cutaneous Reaction Classification [a]
1	pulmonary adenocarcinoma	dexamethasone, phenytoin, sulindac	erythema (60); blisters/ erosions (1)	lips, oral	Y	Y	Atypical SJS
2	Type II Diabetes Mellitus, Hepatitis B carrier, cirrhosis	amoxicillin, amoxicillin-clavulanate, ibuprofen, mefenamic acid, ofloxacin	erythema (50); blisters/ erosions (2)	conjunctivitis, lips, oral, genital	Y	Y	Typical SJS
3	herpetic gingivostomatitis, duodenal ulcer	amoxicillin, clarithromycin, esomeprazole	erythema (35); blisters/ erosions (28)	lips, genital	Y	Y	EEMM/SJS
4	Rheumatologic condition	acetaminophen, celecoxib, sulfasalazine	erythema (60); blisters/ erosions (35)	conjunctivitis, lips, oral, genital	N	Y	SJS/TEN overlap
5	Tuberculosis meningitis	ethambutol, isoniazid, pyrazinamide, rifampin	erythema (19); blisters/ erosions (1)	lips	Y	Y	Typical SJS
6	Not available	Allopurinol	erythema (55-60); blisters/ erosions (27)	conjunctivitis, lips, oral, genital	N	Y	SJS/TEN overlap
7	None	lorazepam, sulfasalazine, sulindac	erythema (30); blisters/ erosions (1)	conjunctivitis, lips, oral	Y	Y	Not available
8	Not available	aceclofenac, cephalexin, esomeprazole	erythema (80); blisters/ erosions (30)	conjunctivitis, lips, oral, genital	Y	Y	SJS/TEN overlap
9	None	acetylsalicylic acid, ibuprofen	erythema (35); blisters/ erosions (28)	conjunctivitis, lips, oral, genital	Y	Y	Typical SJS
10	Nasopharyngeal carcinoma, hypertension, stroke	lercanidipine, meloxicam, telmisartan	erythema (55); blisters/ erosions (1)	lips, oral	Y	Y	Typical SJS

[a] Atypical Stevens-Johnson syndrome (SJS) = blisters/erosions only involving oral mucosa; Typical SJS = blisters on macules, flat atypical targets, <10% blisters/detachment; erythema exudativum multiforme major (EEMM)/SJS = widespread typical target lesions or acral flat atypical lesions; SJS/toxic epidermal necrolysis (TEN) overlap = blisters on macules, flat atypical targets,10-29% blisters/detachment

Table 2.

Inter-rater reliability by Causality Assessment Tool (CAT)

CAT	Score	N	Kappa (95% Confidence Interval)	p-value
ALDEN		**27**	**0.223 (0.141, 0.355)**	**0.000**
	Unlikely		0.171	0.000
	Possible		0.101	0.016
	Probable		0.265	0.000
	Definite		0.413	0.000
Naranjo		**25**	**0.112 (0.019, 0.266)**	**0.005**
	Unlikely		-0.017	0.689
	Possible		0.113	0.009
	Probable		0.117	0.007
	Definite		0.157	0.000
Liverpool		**26**	**0.124 (0.034, 0.273)**	**0.000**
	Unlikely		0.065	0.130
	Possible		0.138	0.001
	Probable		0.077	0.072
	Definite		0.390	0.000