Citation: Clin Transl Sci (2021) 14, 756–763; doi:10.1111/cts.12944

ARTICLE

In Silico Approach to Predict Severe Cutaneous Adverse Reactions Using the Japanese Adverse Drug Event Report Database

Kaori Ambe[1], Kazuyuki Ohya[1], Waki Takada[1], Masaharu Suzuki[1] and Masahiro Tohkin[1,*]

Severe cutaneous adverse reactions (SCARs), such as Stevens–Johnson syndrome/toxic epidermal necrolysis and drug-induced hypersensitivity syndrome, are rare and occasionally fatal. However, it is difficult to detect SCARs at the drug development stage, necessitating a new approach for prediction. Therefore, in this study, using the chemical structure information of SCAR-causative drugs from the Japanese Adverse Drug Event Report (JADER) database, we tried to develop a predictive classification model of SCAR through deep learning. In the JADER database from 2004 to 2017, we defined 185 SCAR-positive drugs and 195 SCAR-negative drugs using proportional reporting ratios as the signal detection method, and the total number of reports. These SCAR-positive and SCAR-negative drugs were randomly divided into the training dataset for model construction and the test dataset for evaluation. The model performance was evaluated in the independent test dataset inside the applicability domain (AD), which is the chemical space for reliable prediction results. Using the deep learning model with molecular descriptors as the drug structure information, the area under the curve was 0.76 for the 148 drugs of the test dataset inside the AD. The method developed in the present study allows for utilizing the JADER database for SCAR classification, with potential to improve screening efficiency in the development of new drugs. This method may also help to noninvasively identify the causative drug, and help assess the causality between drugs and SCARs in postmarketing surveillance.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?
☑ Severe cutaneous adverse reactions (SCARs) have a low incidence, and the mechanism remains unknown. In addition, it is difficult to predict the relationship between drugs and SCARs at the development stage because a wide variety of drugs cause SCARs.

WHAT QUESTION DID THIS STUDY ADDRESS?
☑ Development of an efficient and reliable *in silico* method for predicting SCAR-causative drugs.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?
☑ The combination of the Japanese Adverse Drug Event Report database, a large-scale postmarketing drug adverse effect database, and deep learning was proven to be a new approach that enables evaluating the risk of SCAR-causative drugs from only chemical structure information.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCES?
☑ A rapid and noninvasive identification method of SCAR-causative drugs by deep learning is considered to be a useful technique to evaluate the causal relationship between the corresponding drug and SCAR. The results of this study will therefore help to reduce the risk of SCARs in clinical trials and during postmarketing surveillance, in addition to improving screening efficiency in new drug development.

Stevens–Johnson syndrome/toxic epidermal necrolysis and drug-induced hypersensitivity syndrome are representative severe cutaneous adverse reactions (SCARs), which are considered to be idiosyncratic adverse drug reactions. Although the incidence of SCAR is low, it remains a serious problem owing to its high mortality. Many antibiotics, anticonvulsants, and antipyretic analgesics have been reported to be related to SCARs, but all drugs can be potential causative agents.[1–6] Because SCAR has high species differences and a low occurrence, prediction of SCAR is difficult at the new drug development stage. In other words, SCAR is often found for the first time during postmarketing surveillance. In addition, it is extremely difficult to identify the causative agent of an SCAR from multiple candidate drugs. To overcome these problems, the development of an efficient approach to predict SCAR is very important.

[Correction added on 22nd January 2021, after first online publication: the term 'Repot' has been corrected to 'Report' in the article title.]

[1]Department of Regulatory Science, Graduate School of Pharmaceutical Sciences, Nagoya City University, Nagoya, Japan. *Correspondence: Masahiro Tohkin (tohkin@phar.nagoya-cu.ac.jp)

Received: September 8, 2020; accepted: November 8, 2020. doi:10.1111/cts.12944

Recently, an *in silico* method has drawn attention as a powerful approach for predicting idiosyncratic adverse drug reactions, the mechanism of which is unknown and complicated. By analyzing a large amount of existing data, machine learning can help learn certain rules and predict new data. Some studies have attempted to predict drug-induced liver injury by machine learning.[7–9] Therefore, it is desirable to develop an *in silico* prediction method for SCARs by machine learning using postmarketing data.

Postmarketing safety information is one of the main sources of SCARs because these reactions may only appear after a drug is already approved and used by a large number of patients. In recent years, the number of reports for spontaneous reporting systems (SRSs) has been increasing, offering an important source of information as a large-scale database for the detection of rare and severe adverse events.[10] The Pharmaceuticals and Medical Devices Agency (PMDA), the Japanese regulatory authority, has accumulated comprehensively adverse event reports related to pharmaceuticals and has been published in the Japanese Adverse Drug Event Report (JADER) database. Many studies have investigated the relationship between a drug and adverse drug events using the JADER database.[11–14] In the SRS database, information on the possibility of a causal relationship between adverse events and drugs that have not been known so far could be detected using signal detection approaches.[15–17]

In this study, we focused on developing an SCAR classification model using drug structure information from the JADER database. First, datasets from the JADER database were obtained using signal detection. Thereafter, we calculated molecular descriptors as the structure information of drugs and developed classification models using deep learning. Hence, our *in silico* prediction model could improve screening efficiency at the developmental stage of new drugs. This method may also support the identification of causative agents of SCARs from multiple medications both rapidly and noninvasively.

METHODS

Definition of SCARs

Adverse event reports from the first quarter of 2004 to the second quarter of 2017 were obtained from the JADER database, a drug adverse effect database constructed by the PMDA (https://www.pmda.go.jp/safety/info-services/drugs/adr-info/suspected-adr/0003.html, accessed on February 6, 2018). The JADER database contains information on adverse drug reactions and patients in Japan accumulated during postmarketing surveillance from April 2004, and the database structure is compiled based on the international safety reporting guidelines of the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) E2B (https://www.ich.org/products/guidelines/efficacy/article/efficacy-guidelines.html, accessed on July 6, 2019). The JADER database includes four data tables: (1) patient demographics information (DEMO), (2) drug information (DRUG), (3) adverse events (REAC), and (4) primary disease information (HIST). Data were combined using the ID numbers of the respective tables, and necessary data extraction was performed. Although the drugs were classified into three groups, "suspected drug," "interacting drug," and "concomitant drug," in the DRUG file, our analysis was conducted on only the "suspected drug" group. In addition, because SCARs are based on systemic adverse effects, drugs that are known to hardly migrate to the blood and unknown route drugs were eliminated considering the administration route.

Adverse event names are registered based on the Preferred Term of Medical Dictionary for Regulatory Activities (MedDRA) compiled by the ICH (https://www.meddra.org/, accessed on February 21, 2018). SCARs were defined according to Standardized MedDRA Queries severe cutaneous adverse reactions/20000020, described in MedDRA/J version 20.1, and 14 preferred terms included in a narrow area were used (**Table S1**).

To detect drugs that were likely to cause SCARs, the proportional reporting ratio (PRR) was used as a signal detection method. A signal was detected if the PRR was 2 or more, the χ^2 value was 4 or more, and the number of co-occurrences of interest was 3 or more.[15] PRR and χ^2 values were calculated from two-by-two contingency tables relating the presence of a particular drug and the presence of SCAR (**Table 1**). Furthermore, in the SRS database, the effects of concomitant drugs need to be considered.[18–20] In other words, the 10 most frequently reported SCAR-positive candidate drugs may increase the signal of other SCAR-positive drugs. Therefore, we selected the 10 most frequently reported SCAR-positive candidate drugs, which were considered to have a large impact. Among other SCAR-positive candidates that were not selected as the 10 most reported, we calculated the reporting odds ratio (ROR) and its 95% confidence interval (CI) from two-by-two contingency tables of the presence or absence of combined use with the 10 most reported SCAR-positive candidate drugs and the presence or absence of SCAR.[17] If the lower limit of the 95% CI was > 1, it was regarded as a drug affected by the combined use of the 10 most reported drugs. We excluded these drugs and created another new list considering the effects of concomitant drugs. SCAR-negative candidate drugs were extracted according to the following criteria: (1) no SCAR reports and (2) the total number of reports is ≥ 20.

Predictive model development

To develop a classification model for SCARs, we focused on deep learning.[21] Deep learning is an effective tool for

Table 1 A two-by-two contingency table

	With SCAR	Without SCAR
With a drug of interest	n_{11}	n_{12}
Without a drug of interest	n_{21}	n_{22}

n_{11}, the number of reports with a particular drug and SCAR; n_{12}, the number of reports with a particular drug, but without SCAR; n_{21}, the number of reports without a particular drug, but with SCAR; n_{22}, the number of reports with neither a particular drug nor SCAR; SCAR, severe cutaneous adverse reactions.

The proportional reporting ratio value was calculated as $n_{11}/(n_{11} + n_{12})$ divided by $n_{21}/(n_{21} + n_{22})$. The χ^2 value was calculated as $(n_{11} + n_{12} + n_{21} + n_{22})(|n_{11}n_{22} - n_{12}n_{21}| - (n_{11} + n_{12} + n_{21} + n_{22})/2)^2$ divided by $(n_{11} + n_{12})(n_{21} + n_{22})(n_{11} + n_{21})(n_{12} + n_{22})$.

modeling complex nonlinear relationships, such as the relationship between SCARs and various causative drugs.

To create a dataset for the model, any mixtures, large peptides, herbal products, bacterial preparations, inorganic compounds, organometallic compounds, and unspecified names or abbreviated drugs were removed. Furthermore, if the same generic name existed for SCAR-positive and SCAR-negative candidate drugs, both drugs were excluded. With the help of a PubChem search (https://pubchem.ncbi.nlm.nih.gov/, accessed on February 21, 2018) and the Kyoto Encyclopedia of Genes and Genomes DRUG database (https://www.genome.jp/kegg/drug/, accessed on February 21, 2018), the canonical Simplified Molecular Input Line Entry System was collected, and all salts were removed and then the main structures were neutralized whenever possible. In addition, the Anatomical Therapeutic Chemical (ATC) classification information of drugs was extracted using the Kyoto Encyclopedia of Genes and Genomes DRUG database. ATC is an international classification of drugs devised by the World Health Organization (https://www.whocc.no/atc/structure_and_principles/, accessed on July 6, 2019). We used the second-level therapeutic subgroup of the ATC classification.

Following the curation of drugs, the SCAR-positive and SCAR-negative drugs in the original dataset were represented by two-dimensional molecular descriptors calculated using Dragon 7 software (Talete srl., Milano, Italy). Molecular descriptors represent structured data that facilitate calculation and offer the additional advantage of versatility. Dragon 7 can calculate up to 5,270 molecular descriptors consisting of 30 blocks, such as atom type and functional group/fragment count. After preprocessing the descriptors, they were used in the model. The original dataset was randomly divided at a 1:1 ratio into a training dataset and a test dataset, comprising the same proportion of positive and negative drugs as the original. In addition, the balance of the ATC classification was also maintained to eliminate bias in drug efficacy. The training dataset was used for model construction and the independent test dataset was used for model evaluation.

Furthermore, the applicability domain (AD), which is the drug structure space that is reliably predicted, was defined in the test dataset for evaluating the model using a distance-based method.[22] First, the average distance of all data in the training dataset was calculated from the Euclidean distance, and then a threshold was set. If the distance between a drug in the test dataset and all training data was within the threshold, it was considered to be inside the AD.[23,24] Euclidean distance, as the distance between any two drugs, was calculated from the descriptors used for the model.[25]

The balanced accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and area under the receiver operating characteristic curve (AUC)[26] were calculated as the performance indices.

Deep learning was run through the package "h2o" (3.16.0.2) in R (version 3.4.4) software (h2o: R Interface for "H2O." https://CRAN.R-project.org/package=h2o, accessed on July 6, 2019), which is based on a multilayer feedforward artificial neural network. Our classification model consists of four layers: the input layer, two hidden layers, and the output layer. We fixed two layers as hidden layers because three or more did not improve the accuracy, and the model including more layers took a long time to run. Using the training dataset, we optimized the activation functions and regularization of l1 and l2. For the activation function, we examined the functions of Tanh, TanhWithDropout, Rectifier, and RectifierWithDropout, which can be selected in h2o for deep learning. We also examined an activation function that implements dropout to avoid overfitting. The Gedeon method was used to calculate the variable importance.[27] Details of the preprocessing of descriptors and evaluation methods are provided in **Supplementary Material S1**. The codes for setting the AD and SCAR prediction models for new data are provided in **Supplementary Material S2**. The training and test datasets are provided in **Supplementary Material S3**.

Patient characteristics

Patient characteristics associated with SCAR-positive and SCAR-negative drugs in the original dataset were examined using data from the JADER database from the first quarter of 2004 to the second quarter of 2017. The characteristics investigated were sex and age. Patient age was divided into four groups (0–9, 10–19, 20–59, and ≥ 60 years). To determine if the patient characteristics affected the onset of SCAR, we compared the characteristic factors of the SCAR-positive and SCAR-negative groups. The standardized difference was used to compare the mean of categorical baseline variables between the SCAR-positive and SCAR-negative treatment groups[28] using the following formula:

$$d = (P_{\text{positive}} - P_{\text{negative}}) \,/\, \sqrt{(P_{\text{positive}}(1 - P_{\text{positive}}) + P_{\text{negative}}(1 - P_{\text{negative}}))\,/\,2}$$

where P_{positive} and P_{negative} denote the proportion of the variable of interest in the SCAR-positive and SCAR-negative group, respectively. The standardized difference is the comparison of the mean difference in accumulated standard deviation units. Moreover, regardless of the sample size, the standardized difference can be compared with variables measured in different units. If the absolute value of the standardized difference is < 0.1, the difference between the treatment groups is considered to be negligible.[29]

Software

JADER database management and analyses were performed using SAS (version 9.4; SAS Institute, Cary, NC). Mathematical processing of the data and model construction were performed using R (version 3.4.4).

RESULTS

Curated dataset

The extraction process of SCAR-related reports and drugs is shown in **Figure 1**. The JADER database contained 3,908,356 combination unit reports from the first quarter of 2004 to the second quarter of 2017. The combination unit consists of the common name of the drug and an adverse

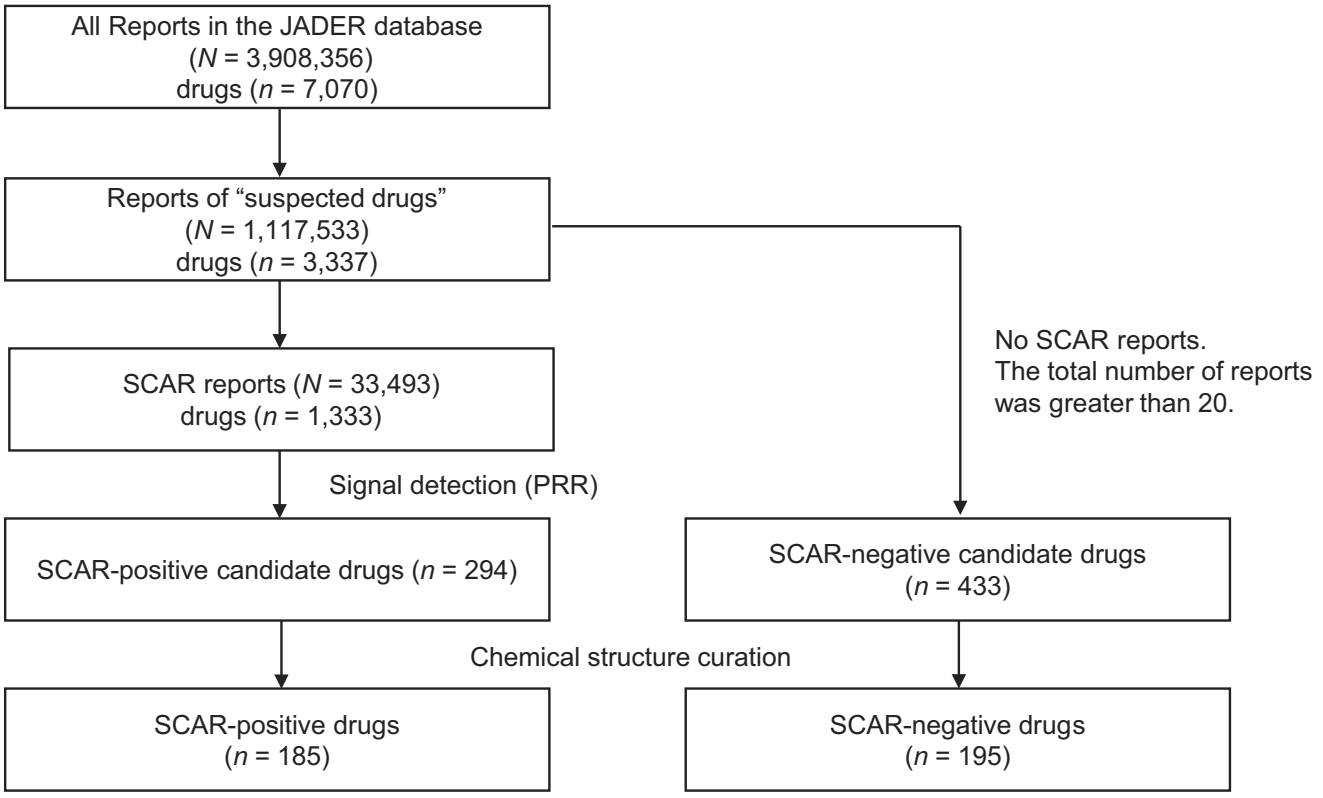

Figure 1 Extraction process of SCAR-related reports and drugs. Using the JADER database, 185 SCAR-positive drugs and 195 SCAR-negative drugs were extracted as the original dataset based on the PRR and the number of total reports, respectively. The number of reports reflects the combination unit reports, which include the common name of the drug and the adverse event at a one-to-one correspondence. The number of reports of "suspected drugs" is limited to routes of administration that are transferred to the blood. JADER, Japanese Adverse Drug Event Report; PRR, proportional reporting ratio; SCAR, severe cutaneous adverse reaction.

event in a one-to-one correspondence. The total number of drugs excluding common name duplication was 7,070. Furthermore, only "suspected drugs" were selected and limited only to the administration route that is likely to be transferred to the blood. Subsequently, 294 SCAR-positive candidate drugs were detected using PRR criteria,[15] and 433 SCAR-negative candidate drugs were detected using 2 criteria: no SCAR reports and ≥ 20 total reports. Other conditions were tested, including 3 total reports or more and 10 total reports or more, but setting the condition of a total of 20 reports or more was well balanced with the number of SCAR-positive drugs (data not shown).

Among SCAR-positive candidate drugs, the 10 most frequently reported drugs were selected (**Table 2**), which are all known to cause SCARs.[1–6,11–14,30] Furthermore, in the SRS database, the effects of concomitant drugs need to be considered.[18–20] To investigate the influence of the combined use of the 10 most frequently reported drugs, SCAR-positive candidate drugs affected by the combination were extracted using ROR as the risk ratio. For 284 unselected drugs, excluding the top 10 drugs, cross tables were created for the presence or absence of combined use with the 10 most frequently reported drugs and for the presence or absence of SCAR (**Figure 2**). The lower limit of the 95% CI for ROR exceeded 1 for 22 drugs. These 22 drugs were then excluded from the list of candidate SCAR-positive drugs as they were found to be affected by the combination with the 10 most frequently reported drugs (**Table S2**).

The chemical structures of the SCAR-positive and SCAR-negative candidate drugs were confirmed. The original dataset of the SCARs classification predictive model consisted of 380 unique drugs, comprising 185 SCAR-positive drugs and 195 SCAR-negative drugs, and the model dataset considering the effects of concomitant drugs consisted of

Table 2 The 10 most frequently reported drugs as SCAR-positive drugs

SCAR-positive drugs		n_{11}	PRR	χ^2 value
1	Carbamazepine	1,741	6.98	8,707.73
2	Lamotrigine	1,337	4.54	3,647.27
3	Allopurinol	1,309	7.88	7,776.16
4	Loxoprofen	832	3.68	1,626.61
5	Acetaminophen	785	7.81	4,683.53
6	Clarithromycin	597	5.18	2,037.47
7	Amoxicillin	569	8.44	3,773.43
8	Celecoxib	479	4.20	11,85.59
9	Cold agent	465	9.73	3,692.90
10	Lansoprazole	456	3.22	707.98

n_{11}, the number of reports with a particular drug and SCAR; PRR, proportional reporting ratio; SCAR, severe cutaneous adverse reactions.

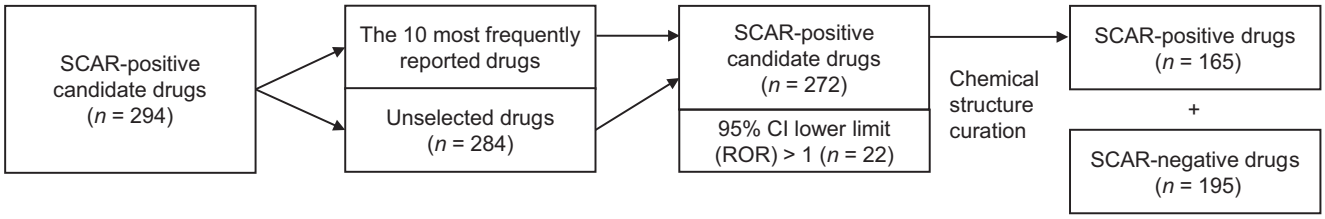

Figure 2 Construction process of the dataset considering the effects of concomitant drugs. A total of 22 SCAR-positive candidate drugs that were affected by the combined use of the 10 most frequently reported drugs were excluded from the original dataset, and a new SCAR-positive list was created considering the effects of concomitant drugs. CI, confidence interval; ROR, reporting odds ratio; SCAR, severe cutaneous adverse reaction.

360 unique drugs, comprising 165 SCAR-positive drugs and 195 SCAR-negative drugs. The generic names of SCAR-positive drugs and SCAR-negative drugs, ATC classification, calculated PRR, χ^2 values, and the number of reports are summarized in **Table S2**.

Patient characteristics

The characteristics (sex and age) of patients in the database that had used the 185 SCAR-positive and 195 SCAR-negative drugs among the 380 drugs classified as "suspected drug" from the original dataset were investigated (**Table 3**). A difference in patient characteristics between the SCAR-positive and SCAR-negative groups could be a potential factor affecting SCARs development. The total number of reports was 71,277 for SCAR-positive drugs and 13,650 for SCAR-negative drugs. When the standardized difference between SCAR-positive and SCAR-negative groups was determined for each variable, the value for sex was 0.02 and that for age was in the range of 0.03–0.07. The difference between groups was considered to be insignificant when the absolute difference was < 0.1; thus, there was no difference between SCAR-positive and SCAR-negative groups with respect to sex and age.

Model performance

The test dataset was evaluated using the drugs that were inside the AD, which is considered to be a reliable chemical space for prediction.[22] The coverage reflects the percentage of drugs inside the AD of the test dataset. Among the total 380 drugs in the original dataset of the SCARs classification model, there were 191 training drugs, comprising 93 positive and 98 negative drugs. The test dataset for the AD included 148 drugs, and the coverage was 0.78. The dataset of the model considering the effects of concomitant drugs included 360 drugs, comprising 181 drugs in the training set (83 positive and 98 negative), and 144 drugs in the test set for the AD, with a coverage of 0.80. As a preliminary analysis, we performed an exploratory search for the number of nodes and regularization of l1 and l2. We performed a random search in various ranges and evaluated the logloss function of cross-validation. We found a narrow range for the regularization of l1 and l2, where performance improvement was expected; however, the number of nodes did not show any performance improvement in a specific range (data not shown). Therefore, the number of nodes in the hidden layer was adopted as the default value (200). The number of descriptors with preprocessed and prediction categories was used as the number of nodes in the input layer and the output layer, respectively. Therefore, the final SCARs classification model consisted of 656 nodes for the input layer, 200 nodes for each of the 2 hidden layers, and 2 nodes of the output layer. The model considering the effects of concomitant drugs consisted of 647 nodes for the input layer, 200 nodes for each of the 2 hidden layers, and 2 nodes of the output layer. Finally, we optimized the activation functions (Tanh, TanhWithDropout, Rectifier, and RectifierWithDropout) and regularization of l1 (0–0.00001) and l2 (0–0.001) by 3-fold cross-validation using AUC as the evaluation criterion (**Table S3**). Other parameters were set to the default values. Different molecular descriptors were used in the SCARs classification model and in the model considering the effect of concomitant drugs (**Table S4**).

The prediction performances are summarized in **Table 4**. For the SCARs classification model, the test dataset inside

Table 3 Patient characteristics by groups

	SCAR-positive *N* = 71,277		SCAR-negative *N* = 13,650		
	n	%	*n*	%	Standardized difference
Men	35,890	50.4	6,992	51.2	0.02
Age					
≤ 9	3,639	5.2	492	3.6	0.07
10–19	2,730	3.8	432	3.2	0.04
20–59	25,238	35.4	4,660	34.1	0.03
≥ 60	39,670	55.7	8,066	59.1	0.07

Standardized difference shows absolute value.
SCAR, severe cutaneous adverse reactions.

Table 4 SCAR prediction performance

Model Number of drugs (test dataset inside AD)	Balanced accuracy	Sensitivity	Specificity	PPV	NPV	AUC
SCAR classification model *n* = 380 (148)	0.69	0.81	0.58	0.64	0.76	0.76
Model considering effects of concomitant drugs *n* = 360 (144)	0.65	0.63	0.66	0.62	0.67	0.73

AD, applicability domain; AUC, area under the receiver operating characteristic curve; NPV, negative predictive value; PPV, positive predictive value; SCAR, severe cutaneous adverse reactions.

the AD had an AUC of 0.76. In the deep learning SCARs classification model, the five most important descriptors extracted were "C-005: CH_3X," "VE1sign_Dz(p): coefficient sum of the last eigenvector from Barysz matrix weighted by polarizability," "nRCOOH: number of carboxylic acids (aliphatic)," "B05[O-O]: presence/absence of O-O at topological distance 5," and "SpMin2_Bh(m): smallest eigenvalue n. 2 of Burden matrix weighted by mass" (https://chm.kode-solutions.net/products_dragon_descriptors.php, accessed on June 10, 2020). There was no improvement in prediction performance using the model in which the effect of concomitant drugs was considered, with an AUC of 0.73.

The 380 drugs in the original dataset were aggregated by ATC classification. These were divided into 62 ATC classifications. The largest number was 57 drugs in the classification "J01 Antibacterials for systemic use" (**Table 5**). There were 199 drugs included in the 10 most frequently classified ATC classifications. The remaining 181 drugs were classified into 52 ATC classifications; therefore, various medicinal effects were included in the original dataset. The prediction results of the 148 test datasets inside the AD were then evaluated for each ATC classification (**Table 5**). "J01 Antibacterials for systemic use," "L01 Antineoplastic agents," "A02 Drugs for acid related disorders," "M01 Anti-inflammatory and antirheumatic products," "R05 Cough and cold preparations," and "R06 Antihistamines for systemic use" had high accuracy rates, ranging from 0.75 to 1.00. Except for "L01 Antineoplastic agents," these ATC classifications included more SCAR-positive drugs than SCAR-negative drugs. "C01 Cardiac therapy" and "N07 Other nervous system drugs" had an accuracy rate lower than 0.5. There was a tendency for the prediction of negative drugs to be incorrect.

DISCUSSION

In this study, we developed a deep learning prediction model to classify SCARs caused by a variety of drugs using only drug structure information. Prediction results with 148 test drugs in the test dataset inside the AD demonstrated an AUC of 0.76. In addition, the sensitivity was 0.81 and the negative predictive value was 0.76, indicating few false negatives (**Table 4**). This acceptable prediction performance indicates that this is a suitable model for reducing overlooked SCAR-positive drugs and excluding SCAR-negative drugs. With our model, it is possible to predict SCAR for drugs not only within the JADER database but also out of the JADER database if the structural information of the drugs is available. In addition, we set the AD as the chemical space in our model, which can provide highly reliable prediction results if the drug is inside the AD. Therefore, our model can be applied at the global scale even though we used the Japanese adverse effect database for the present model development and assessment.

Because it is difficult to obtain the exact SCAR risk from the JADER database, we used the presence or absence of a signal to calculate the risk of SCAR for a given drug. The PRR has been widely used for the signal detection of adverse drug reactions in adverse reaction reporting databases.[15] Therefore, we regarded drugs that showed an

Table 5 Prediction results and ATC classification

ATC classification of 380 drugs in the dataset		Results of SCAR classification of test dataset inside AD (*n* = 148)		
ATC classification	Number of drugs (positive/negative)	Positive	Negative	Accuracy
J01 Antibacterials for systemic use	57 (53/4)	19/20	0/1	0.90
L01 Antineoplastic agents	33 (3/30)	0/1	9/9	0.90
N05 Psycholeptics	20 (7/13)	2/2	3/7	0.56
A02 Drugs for acid related disorders	16 (15/1)	4/5	-	0.80
M01 Anti-inflammatory and antirheumatic products	14 (14/0)	6/6	-	1.00
R05 Cough and cold preparations	13 (12/1)	3/4	-	0.75
C01 Cardiac therapy	12 (2/10)	-	2/5	0.40
R06 Antihistamines for systemic use	12 (10/2)	4/5	0/1	0.80
N02 Analgesics	11 (4/7)	0/2	2/4	0.50
N07 Other nervous system drugs	11 (3/8)	1/1	0/2	0.33
Others	181 (62/119)	19/26	28/47	0.64

AD, applicability domain; ATC, Anatomical Therapeutic Chemical; SCAR, severe cutaneous adverse reactions.

adverse reaction signal based on the PRR as SCAR-positive drugs. There is a known risk of PRR value inflation, which may cause false signal detection if the number of reports is very small.[31] However, the maximum PRR value in this study was 33 (**Table S2**) because the number of reports of other drugs was high even if the number of SCAR reports for the drugs of interest was small. This result suggested that PRR value inflation was not occurring in our SCAR-positive dataset, and that the PRR-based criterion for SCAR-positive definition was appropriate in this case. We also considered the potential risk of SCARs in drugs that are specifically known to be related to the onset of SCARs, even if the total number of reports was small. Therefore, the total number of reports was not taken into consideration for SCAR-positive drug selection (**Figure 1**). By contrast, we defined SCAR-negative drugs as those with no SCAR-related reports and 20 or more of the total number of reports.

The 10 most frequently reported drugs among the SCAR-positive candidate drugs included carbamazepine and lamotrigine as anticonvulsants, allopurinol for hyperuricemia, loxoprofen and acetaminophen as antipyretic analgesics, clarithromycin and amoxicillin as antibacterial drugs, celecoxib as a nonsteroidal anti-inflammatory drug, cold agent, and lansoprazole as a proton pump inhibitor (**Table 2**). All 10 drugs have been previously reported to be related to SCARs, which is also included in the package inserts.[1–6,11–14,30] This concordance suggests that our criteria for SCAR-positive candidate drugs was appropriate for model construction.

Using the standardized difference, we compared the patients' general demographic characteristics in the original dataset (**Table 3**). Because the standardized difference was < 0.1, it indicated a negligible difference in the mean between two groups,[28,29] indicating that any bias of patient sex and patient age did not affect the construction of the model.

In the JADER database, it is important to consider whether the reported adverse event is due to a drug or only appears with the combination of other drugs. Although analyses of drug interactions in the SRS database have been performed, there is no established method.[18–20] Therefore, to examine whether the SCAR-positive drugs in the dataset used for the prediction model were appropriate, we built a model considering the effects of the concomitant drugs (**Figure 2**). It was hypothesized that if the predictive performance of this model was improved over the original model, the 22 drugs excluded would be less likely to cause SCARs. Interestingly, the prediction performance of the model in which the effects of combinations were considered (AUC = 0.73) did not improve compared with the performance of the original model (AUC = 0.76). Although the 22 drugs excluded may be affected by the combination of the 10 most frequently reported drugs, these drugs are highly likely to cause SCARs. In conclusion, our analysis suggests that SCAR-positive drugs in the original dataset are appropriate.

The efficacy of the 380 drugs in the original dataset was confirmed using the ATC classification (**Table 5**). The prediction result of "J01 Antibacterials for systemic use," which had the largest number of drugs, was almost perfectly in line with the candidate SCAR-positive drugs. This is considered to be due to the large number of cephem antibiotics having a β-lactam structure in the drugs included in the ATC classification "J01 Antibacterials for systemic use." Therefore, there is a possibility that the prediction of SCAR-negative drugs belonging to "J01 Antibacterials for systemic use" by our model has not been correctly evaluated. The ATC classification that classified the second highest number of drugs was "L01 Antineoplastic agents." Antineoplastic agents are associated with numerous adverse effects, and there are many such reports in the JADER database. As a result, antineoplastic agents tend to be extracted as SCAR-negative drugs because they easily meet the criteria.

In the SCARs classification model, "nRCOOH" was selected as one of the most highly important descriptors. This descriptor belongs to the functional group count and indicates the number of carboxylic acids (aliphatic). Of the 53 SCAR-positive drugs in the ATC classification "J01 Antibacterials for systemic use," 19 drugs contain aliphatic carboxylic acids, and 18 drugs contain aliphatic carboxylic acids and a β-lactam structure.[32] However, of the 53 SCAR-positive drugs in the ATC classification "J01 Antibacterials for systemic use," 28 drugs contain a β-lactam structures. Because the variable importance of "nBeta-Lactams: number of β-lactams" ranked 28th, "nRCOOH," which ranked third, may be a key functional group for discriminating SCAR-positive drugs and SCAR-negative drugs among drugs with various chemical structures.

In studying adverse effects based on the SRS database, it is necessary to interpret the results carefully owing to reporting biases, such as under-reporting and the impact of safety information.[33] In addition, the frequency of occurrence of adverse events cannot be determined using SRS databases, including the JADER database. Although SRS databases are subject to a large amount of noise, most of the information is available as real-world data, including a vast amount of accumulating information about the spontaneous reporting of adverse effects. Thus, an SRS database offers a good source of information for SCARs owing to its low incidence and the fact that it often is only detected during postmarketing surveillance. The combination of machine learning and an SRS database to construct a predictive model of SCARs will be an important approach for further research on adverse effects. Moreover, the JADER database contains well-validated information because it examines the reported adverse effects information to identify suspected drugs. Therefore, it is considered that the accuracy of the relationship between the drug and the adverse effect is relatively high. We considered that this advantage of the JADER database was important for training data of the model.

In conclusion, we developed a classification model of SCARs in humans based on the chemical structure information of SCAR-causative drugs in the JADER database. Although this model is not suitable for the quantitative risk assessment of SACRs, we consider that the model constructed in this study is useful for qualitative risk assessment. This prediction model is a noninvasive and highly efficient support method for clinical trials and postmarketing surveillance. Furthermore, utilizing the JADER database and the machine learning method is expected to promote the use of SRS databases in the research on adverse effects, such as drug-induced liver injury.

Supporting Information. Additional supporting information may be found in the online version of this article at the publisher's web site:

Acknowledgment. The authors thank Editage (www.editage.com) for English language editing.

Funding. This research is partially supported by a grant from the Research Foundation for Pharmaceutical Sciences, Grant-in-Aid for Research in Nagoya City University (Grant Number 1922008) and JSPS KAKENHI (Grant Number JP20K16050).

Conflict of Interest. All authors declared no competing interests for this work.

Author Contributions. K.A., K.O., and M.T. wrote the manuscript. K.A., K.O., and M.T. designed the research. K.A., K.O., W.T., and M.S. performed the research. K.A., K.O., and W. T. analyzed the data.

1. Gerull, R., Nelle, M. & Schaible, T. Toxic epidermal necrolysis and Stevens-Johnson syndrome: a review. *Crit. Care Med.* **39**, 1521–1532 (2011).
2. Lerch, M., Mainetti, C., Terziroli Beretta-Piccoli, B. & Harr, T. Current perspectives on Stevens-Johnson syndrome and toxic epidermal necrolysis. *Clin. Rev. Allergy Immunol.* **54**, 147–176 (2018).
3. Roujeau, J.C. *et al.* Medication use and the risk of Stevens-Johnson syndrome or toxic epidermal necrolysis. *N. Engl. J. Med.* **333**, 1600–1607 (1995).
4. Wang, Y.H. *et al.* The medication risk of Stevens-Johnson syndrome and toxic epidermal necrolysis in Asians: the major drug causality and comparison with the US FDA label. *Clin. Pharmacol. Ther.* **105**, 112–120 (2019).
5. Bocquet, H., Bagot, M. & Roujeau, J.C. Drug-induced pseudolymphoma and drug hypersensitivity syndrome (Drug Rash with Eosinophilia and Systemic Symptoms: DRESS). *Semin. Cutan. Med. Surg.* **15**, 250–257 (1996).
6. Shiohara, T. & Kano, Y. Drug reaction with eosinophilia and systemic symptoms (DRESS): incidence, pathogenesis and management. *Expert Opin. Drug Saf.* **16**, 139–147 (2017).
7. Hong, H., Thakkar, S., Chen, M. & Tong, W. Development of decision forest models for prediction of drug-induced liver injury in humans using a large set of FDA-approved drugs. *Sci. Rep.* **7**, 17311–17325 (2017).
8. Przybylak, K.R. & Cronin, M.T. In silico models for drug-induced liver injury–current status. *Expert Opin. Drug Metab. Toxicol.* **8**, 201–217 (2012).
9. Zhu, X.W. & Li, S.J. In silico prediction of drug-induced liver injury based on adverse drug reaction reports. *Toxicol. Sci.* **158**, 391–400 (2017).
10. Edwards, I.R., Lindquist, M., Wiholm, B.E. & Napke, E. Quality criteria for early signals of possible adverse drug reactions. *Lancet* **336**, 156–158 (1990).
11. Abe, J. *et al.* Analysis of Stevens-Johnson syndrome and toxic epidermal necrolysis using the Japanese Adverse Drug Event Report database. *J. Pharm. Health Care Sci.* **2**, 14–23 (2016).
12. Hosohata, K., Inada, A., Oyama, S., Niinomi, I., Wakabayashi, T. & Iwanaga, K. Adverse cutaneous drug reactions associated with old- and new- generation antiepileptic drugs using the Japanese pharmacovigilance database. *Clin. Drug Investig.* **39**, 363–368 (2019).
13. Nagai, J., Uesawa, Y., Shimamura, R. & Kagaya, H. Characterization of the adverse effects induced by acetaminophen and nonsteroidal anti-inflammatory drugs based on the analysis of the Japanese Adverse Drug Event Report database. *Clin. J. Pain* **33**, 667–675 (2017).
14. Shirakuni, Y., Okamoto, K., Kawashita, N., Yasunaga, T. & Takagi, T. Signal detection of drug complications applying association rule learning for Stevens-Johnson syndrome. *J. Comp. Aided Chem.* **10**, 118–127 (2009).
15. Evans, S.J., Waller, P.C. & Davis, S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol. Drug Saf.* **10**, 483–486 (2001).
16. Sakaeda, T., Tamon, A., Kadoyama, K. & Okuno, Y. Data mining of the public version of the FDA Adverse Event Reporting System. *Int. J. Med. Sci.* **10**, 796–803 (2013).
17. Van Puijenbroek, E.P., Bate, A., Leufkens, H.G., Lindquist, M., Orre, R. & Egberts, A.C. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. *Pharmacoepidemiol. Drug Saf.* **11**, 3–10 (2002).
18. Susuta, Y. & Takahashi, Y. Safety risk evaluation methodology in detecting the medicine concomitant use risk which might cause critical drug rash [in Japanese]. *Jpn. J. Pharmacoepidemiol.* **19**, 39–49 (2014).
19. Tatonetti, N.P. *et al.* Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels. *Clin. Pharmacol. Ther.* **90**, 133–142 (2011).
20. Van Puijenbroek, E.P., Egberts, A.C., Meyboom, R.H. & Leufkens, H.G. Signalling possible drug-drug interactions in a spontaneous reporting system: delay of withdrawal bleeding during concomitant use of oral contraceptives and itraconazole. *Br. J. Clin. Pharmacol.* **47**, 689–693 (1999).
21. Hinton, G.E., Osindero, S. & Teh, Y.W. A fast learning algorithm for deep belief nets. *Neural Comput.* **18**, 1527–1554 (2006).
22. OECD. Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models. OECD Series on Testing and Assessment 69 <https://doi.org/10.1787/9789264085442-en>. (2014). Accessed November 3, 2020.
23. Ambe, K. *et al.* In silico prediction of chemical-induced hepatocellular hypertrophy using molecular descriptors. *Toxicol. Sci.* **162**, 667–675 (2018).
24. Tetko, I.V. *et al.* Critical assessment of QSAR models of environmental toxicity against tetrahymena pyriformis: focusing on applicability domain and overfitting by variable selection. *J. Chem. Inf. Model* **48**, 1733–1746 (2008).
25. Mathea, M., Klingspohn, W. & Baumann, K. Chemoinformatic classification methods and their applicability domain. *Mol. Inform.* **35**, 160–180 (2016).
26. Lasko, T.A., Bhagwat, J.G., Zou, K.H. & Ohno-Machado, L. The use of receiver operating characteristic curves in biomedical informatics. *J. Biomed. Inform.* **38**, 404–415 (2005).
27. Gedeon, T.D. Data mining of inputs: analysing magnitude and functional measures. *Int. J. Neural Syst.* **8**, 209–218 (1997).
28. Cohen, J. Statistical Power Analysis for the Behavioral Sciences, 2nd ed (Routledge, New York, NY, 1988). https://doi.org/10.4324/9780203771587
29. Austin, P.C. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav. Res.* **46**, 399–424 (2011).
30. Lombardo, C. & Bonadonna, P. Hypersensitivity reactions to proton pump inhibitors. *Curr. Treat. Options Allergy* **2**, 110–123 (2015).
31. Almenoff, J.S., LaCroix, K.K., Yuen, N.A., Fram, D. & DuMouchel, W. Comparative performance of two quantitative safety signalling methods: implications for use in a pharmacovigilance department. *Drug Saf.* **29**, 875–887 (2006).
32. Lin, Y.F. *et al.* Severe cutaneous adverse reactions related to systemic antibiotics. *Clin. Infect. Dis.* **58**, 1377–1385 (2014).
33. Poluzzi, E., Raschi, E., Moretti, U. & De Ponti, F. Drug-induced torsades de pointes: data mining of the public version of the FDA Adverse Event Reporting System (AERS). *Pharmacoepidemiol. Drug Saf.* **18**, 512–518 (2009).