BPB Reports

Regular Article

Development of a Visual Inspection Method for Identifying Falsified Medicines Obtained by Personal Import via the Internet

Naoko Yoshida,[*,a,b,c] and Shoei Maeda[c]

[a]*AI Hospital/Macro Signal Dynamics Research and Development Center, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University; Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan;* [b]*Society for Medicines Security Research, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan;* [c]*Clinical Pharmacy and Healthcare Sciences, Faculty of Pharmacy, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan*

Received January 17, 2025; Accepted February 26, 2025

To avoid harm caused by falsified medicines, we aimed to devise a method to identify falsified medicines in Japan using visual observation among medicines obtained from personal import agency websites, which are the main conduits through which falsified medicines are obtained. We recorded details regarding the information provided on personal import agency websites used to purchase medicines, the outer package received, the customs declaration description, and the product appearance for 212 samples of medicines obtained through personal import via the Internet. We investigated the relationship between each observed item and the rate of falsified medicines. We developed a classification and prediction model to identify falsified medicines using items that could be visually observed. The results showed that the rate of falsified medicines was significantly higher for websites that did not contain identifying information such as the name and address of the contact or import agency, as well as for products that did not contain the name and address of the manufacturer, indicating that these items may be useful in the identification of falsified medicines. In the prediction model constructed, we extracted features such as the country of dispatch and address of the import agency, and a prediction model was created to identify falsified medicines and websites selling these medicines. Careful observation of the identified features and use of our prediction model will help to prevent harm owing to the use of falsified medicines.

Key words falsified medicines, visual inspection, prediction, personal import agency

INTRODUCTION

The problem of substandard and falsified medicines is a global issue that must be addressed. Falsified medicines were previously referred to as counterfeit.[1)] The term "counterfeit" in reference to falsified medicinal products was changed to spurious/falsely labelled/falsified/counterfeit (SFFC). The World Health Organization (WHO) currently defines "falsified" medicines as a medical product that is deliberately or fraudulently mislabeled with regard to identity, composition, or origin.[2)] In this paper, "falsified" is used according to the WHO definition.

There are many personal import agency websites on the Internet, and falsified medicines have been confirmed to enter Japan as medicines imported by individuals via these websites.[3-7)] The Ministry of Health, Labour and Welfare (MHLW) has alerted consumers to the dangers of personal import of medicines and has taken measures such as closing down illegal websites.[8)] Pharmaceutical companies have also taken steps to devise ways to prevent falsification and ensure the safety of their supply chains, but falsification of medicines is increasing.[9)] Although websites may sell genuine products, other risks associated with the personal import of medicines, such as insufficient or improper indications, cannot be avoided. Thus, obtaining medicinal products via personal import should be discouraged.

As a method of identifying falsified medicines, we have conducted authenticity investigations of manufacturers. Authenticity investigations involve sending questionnaires, photographs and, if necessary, the actual product to the manufacturers of medicines to determine whether the product has been authorized and licensed by the manufacturing company. However, authenticity investigations require the cooperation of the manufacturer and, depending on the investigation, can be labor-intensive and time-consuming. We have conducted quality tests such as qualitative and quantitative analyses, content uniformity tests, and dissolution tests, in accordance with the Pharmacopoeia as a method for detecting falsified medicines. However, in addition to acquiring equipment and reagents, it is necessary to be familiar with the experimen-

*To whom correspondence should be addressed. e-mail: naoko@p.kanazawa-u.ac.jp

tal techniques used. Additionally, investigations into the actual distribution of medicines require testing a large number of samples; as the number of samples increases, this requires a great deal of labor, cost, and time. Therefore, we sought to determine a method for efficiently detecting falsified medicines by extracting the characteristics of websites selling falsified medicines, as well as the characteristics of the falsified medicines obtained.

In the present study, we investigated the information obtained visually from personally imported medicines and the websites of personal import agencies from which such medicines were obtained. We also examined the relationship between the above information and the rate of falsified medicines personally imported to Japan. Additionally, with the aim of extracting the characteristics of falsified medicines and websites selling falsified medicines with greater precision, we applied machine learning to construct a model for predicting and classifying falsified medicines using product information that could be visually obtained to develop a method for identifying falsified medicines and the websites selling these medicines.

MATERIALS AND METHODS

Target Websites and Products

The present study was conducted using a total of 212 samples imported personally via the Internet, which were obtained in our previous surveys; the target medicines were 43 Cialis tablets, 28 Levitra tablets, 22 Viagra tablets, 11 Diflucan tablets, 12 amoxicillin-clavulanic acid combination tablets/capsules, 23 dexamethasone tablets, 19 orlistat capsules, 40 metformin tablets, and 13 ivermectin tablets.[2,5-7,10-13] Of these, 34 samples of Cialis tablets, 17 samples of Levitra tablets, 18 samples of Viagra tablets, 2 samples of Diflucan tablets, and 3 samples of Orlistat capsules were falsified medicines. These falsified medicines contained insufficient amounts of active pharmaceutical ingredients or different ingredients, and they were confirmed as being falsified in the authenticity investigation conducted in previous studies.[3,5-7] In cases where the product authenticity was unknown, we excluded the sample from the study to avoid incorrect feature extraction and derivation of predictive models. If several products with different packaging were purchased from the same website or arrived from the same website in multiple deliveries, these were considered different samples.

Observation Items

Personal Import Agency Websites We recorded the following 27 items from personal import agency websites: contact details (telephone number); contact details (fax number); contact address (e-mail); enquiry mail form; import agency name or personal name; import agency address (head office); name of import agency representative (responsible person); Japan branch; medicine price; shipping costs; payment time; method of payment; delivery time; special conditions for returns; product photographs; product name; dosage and administration; indications; side effects; recommendation to consult a doctor or pharmacist; information about personal imports; limitations on the quantity of personal import purchases; reference to the Act on Specified Commercial Transactions;[14] reference to the Act on Quality, Efficacy and Safety Assurance of Medical Devices (Pharmaceutical and Medical Device Act);[15] information about consultation; use of SSL data encryption; and privacy policy.

For each of these items, we determined the proportion of medicines that were falsified among our samples.

Obtained Products For each of the sample products obtained, the following 12 items were recorded: country of dispatch, name of the shipping company or shipper, shipper's company address, customs declaration description, package insert, lot/batch number, expiry date, country of manufacture, manufacturer's address, name of manufacturer, date of manufacture, and instruction leaflet.

The proportion of medicines that were falsified from among all obtained products was determined.

Analysis

Statistical analysis was performed using IBM SPSS version 27 (IBM Corporation, Tokyo, Japan), with cross-tabulation for each item, as well as Pearson's chi-square test or Fisher's exact probability test (two-tailed) with a significance level of 5%.

IBM SPSS Modeler version 18.3 (IBM Corporation, Tokyo, Japan) was used to conduct machine learning, and classification and prediction models were created based on information regarding the authenticity of medicines and information obtained visually from personal import agency websites and from the products themselves. For the purpose of conducting hold-out validation, the data were randomly classified as training:testing (70:30). A prediction and classification model was created using chi-square automatic interaction detection (CHAID), which is a form of decision tree analysis used to build tree models based on chi-square and F-tests. In feature extraction, a feature was considered to be present when the probability of it being genuine or falsified was 90% or more. Discrimination performance was assessed by comparing the predictions made using the created model with the actual classification. We used the area under the receiver operating characteristic curve (AUC) and the Gini impurity, which indicates the purity of the classification with the decision tree (before classification as 1, the closer to 0, the higher the purity) as indicators of generalization performance.

RESULTS

Incidence of Falsified Medicines

Personal Import Agency Websites Using the 27 items recorded from websites selling medicines for personal import, we determined the presence or absence of website information and the authenticity of the medicines sold online, and we calculated the rate of falsified medicines (Table 1).

The rate of falsified medicines sold on websites with a telephone number was 33.3%, compared with 69.0% on websites without a telephone number, which was significantly higher ($p < 0.01$). The rate of falsified medicines sold on websites with an email address was 50.4%, compared with 38.9% for websites without an email address, with no significant difference. The rate of falsified medicines sold on websites with a stated enquiry mail form was 31.4%, compared with 58.9% for websites without a stated enquiry email form, which was significantly higher ($p < 0.01$).

The rate of falsified medicines on websites with an import agency name or personal name was 29.3%, compared with 76.4% on websites without this information, which was significantly higher ($p < 0.01$). The rate of falsified medicines

Table 1. Items Listed on Personal Import Agency Websites and Falsified Medicines

No.	Observation items	Status	G	F	Rate of F (%)	p-value
1	Contact details (telephone number)	Stated	94	47	33.3	<0.01
		Not stated	22	49	69.0	
2	Contact details (fax number)	Stated	36	3	7.7	<0.01
		Not stated	80	93	53.8	
3	Contact address (E-mail)	Stated	58	59	50.4	0.099
		Not stated	58	37	38.9	
4	Enquiry mail form	Stated	72	33	31.4	<0.01
		Not stated	44	63	58.9	
5	Import agency name or personal name	Stated	99	41	29.3	<0.01
		Not stated	17	55	76.4	
6	Import agency address (head office)	Stated	104	39	27.3	<0.01
		Not stated	12	57	82.6	
7	Name of representative of import agency (responsible person)	Stated	71	27	27.6	<0.01
		Not stated	45	69	60.5	
8	Japan branch	Stated	6	7	53.8	0.574
		Not stated	110	89	44.7	
9	Price	Stated	116	92	44.2	<0.05
		Not stated	0	4	100.0	
10	Shipping costs	Stated	103	76	42.5	0.060
		Not stated	13	20	60.6	
11	Payment time	Stated	88	60	40.5	<0.05
		Not stated	28	36	56.3	
12	Method of payment	Stated	112	90	44.6	0.354
		Not stated	4	6	60.0	
13	Time of delivery	Stated	111	85	43.4	0.067
		Not stated	5	11	68.8	
14	Special conditions for returns	Stated	110	86	43.9	0.193
		Not stated	6	10	62.5	
15	Product photograph	Stated	99	69	41.1	<0.05
		Not stated	17	27	61.4	
16	Product name	Stated	115	86	42.8	<0.01
		Not stated	1	10	90.9	
17	Dosage and administration	Stated	56	70	55.6	<0.01
		Not stated	60	26	30.2	
18	Indications	Stated	72	87	54.7	<0.01
		Not stated	44	9	17.0	
19	Side effects	Stated	54	58	51.8	0.053
		Not stated	62	38	38.0	
20	Recommendation to consult a doctor or pharmacist	Stated	73	34	31.8	<0.01
		Not stated	43	62	59.0	
21	Information on personal imports	Stated	95	29	23.4	<0.01
		Not stated	21	67	76.1	
22	Limitations on the quantity of personal import purchases	Stated	86	28	24.6	<0.01
		Not stated	30	68	69.4	
23	Reference to the Act on Specified Commercial Transactions	Stated	58	22	27.5	<0.01
		Not stated	58	74	56.1	
24	References to the Act on Quality, Efficacy and Safety Assurance of Medical Devices	Stated	16	1	5.9	<0.01
		Not stated	100	95	48.7	
25	Consultation	Stated	23	6	20.7	<0.01
		Not stated	93	90	49.2	
26	Use of SSL encryption	Stated	41	16	28.1	<0.01
		Not stated	75	80	51.6	
27	Privacy policy	Stated	50	20	28.6	<0.01
		Not stated	66	76	53.5	

G, genuine; F, falsified.

was 27.3% on websites with an import agency address and 82.6% on sites without an address, which was also significant ($p < 0.01$). The rate of falsified medicines sold on websites with the name of the import agency representative was 27.6%, compared with 60.5% on websites without this information, which was significantly higher ($p < 0.01$). The rate of falsified medicines on websites with a Japanese branch listed was 53.8%, compared with 44.7% on websites without a branch in Japan, with no significant difference.

The rate of falsified medicines sold on these websites with a stated price was 44.2%, compared with 100% on websites without a stated price, which was significantly higher ($p < 0.05$). The rate of falsified medicines on websites where shipping costs were stated was 42.5%, compared with 60.6% for sites where these costs were not stated, with no significant difference. The rate of falsified medicines on websites with a payment time statement was 40.5%, compared with 56.3% for those without a payment time statement, which was significantly higher ($p < 0.05$). The rate of falsified medicines sold on websites with a stated payment method was 44.6%, compared with 60.0% on websites without a stated payment method, with no significant difference. The rate of falsified medicines on websites with a description of the delivery time was 43.4%, compared with 68.8% for sites without such a description, which was not significant. The rate of falsified goods on websites with a return policy was 43.9%, compared with 62.5% on websites without a return policy, with no significant difference. The rate of falsified medicines sold on websites that provided a photograph of the product was 41.1%, compared with 61.4% on websites without a photograph of the product, which was significantly higher ($p < 0.05$). Some websites provided a photograph of the product, but the packaging of the actual product differed from the photograph.

The rate of falsified medicines available on websites with a product name was 42.8%, compared with the rate on sites without a product name (90.9%), which was significantly higher ($p < 0.01$). The rate of falsified medicines on websites with a stated dosage was 55.6%, which was significantly higher ($p < 0.01$) than the 30.2% on websites without a stated dosage. The rate of falsified medicines on websites with indications listed was 54.7% which was significantly higher ($p < 0.01$) than the 17.0% on sites without listed indications. The rate of falsified medicines on websites where side effects were mentioned was 51.8%, compared with 38.0% on websites where these were not mentioned, with no significant difference. The rate of falsified medicines available on websites with a recommendation to consult a doctor or pharmacist was 31.8%, compared with 59.0% on sites without such a recommendation, which was significantly higher ($p < 0.01$). Consultation with a doctor or pharmacist was classified as not stated if the conditions were limited or stated, for example, that consultation was recommended in case of side effects or with pre-existing medical conditions or concomitant medications. The rate of falsified medicines on websites with information about personal import was 23.4%, compared with 76.1% on websites with no information about personal import, which was significantly higher ($p < 0.01$). The rate of falsified medicines on websites with a stated limit on the quantity that could be purchased for personal import was 24.6%, compared with 69.4% on websites without this information, which was significantly higher ($p < 0.01$). The presence of a statement similar to "importing quantities not exceeding those for personal use" was classified as having an appropriate description; any statement similar to "in principle, no restrictions" was classified as not having an appropriate description. The rate of falsified medicines on websites with reference to the Act on Specified Commercial Transactions was 27.5%, compared with 56.1% on websites without such a reference, which was significantly higher ($p < 0.01$). The rate of falsified medicines on websites with a reference to the Pharmaceutical and Medical Device Act was 5.9%, compared with 48.7% on sites without such a reference, which was significantly higher ($p < 0.01$). The rate of falsified medicines on websites with information regarding where to obtain a consultation was 20.7%, compared with 49.2% on websites without this information, which was significantly higher ($p < 0.01$). The rate of falsified medicines on websites that used SSL data encryption was 28.1%, compared with 51.6% in those without the use of SSL, which was significantly higher ($p < 0.01$). The rate of falsified medicines available on websites with a stated privacy policy was 28.6%, compared with 53.5% for websites without a stated privacy policy, which was significantly higher ($p < 0.01$).

Personally Imported Medicines We examined 12 items that could be visually observed on all sample products for their relevance to authenticity; the rates of falsified medicines for these items are given in Table 2.

The following information was recorded from the shipping label attached to the packages containing medicine samples: country of origin, name of the shipping company or shipper, address of the shipping company, and customs declaration description. The medicines in this study were shipped from 10 countries. All products shipped from Malaysia, Taiwan, and Thailand were genuine whereas all products shipped from Puerto Rico and Japan were falsified.Although Japan was not stated as the country of dispatch for products shipped from Japan; we determined that these products had indeed been shipped from Japan because they arrived with Japanese postage stamps on them. Sixty-two of the 64 products shipped from China were falsified, representing a proportion of falsified medicines of 96.9%. The rate of falsified medicines for products with the name of the shipping company or shipper was 50.5%, compared with 40.0% for websites without this information, with no significant difference. The rate of falsified medicines with a shipping company address was 48.0%, and the rate of falsified medicines without a shipping company address was 41.2%, with no significant difference. The rate of falsified medicines with a customs declaration was 71.1%, which was significantly higher ($p < 0.01$) than the rate of falsified medicines without a customs declaration (26.2%).

The obtained products were checked for the package insert, lot/batch number, expiry date, country of manufacture, manufacturer, manufacturer's address, date of manufacture, and instruction leaflet. The rate of falsified medicines with a package insert was 48.4%, and the rate without a package insert was 41.1%, with no significant difference. The rate of falsified medicines with a lot/batch number was 45.8%, compared with 43.2% for those without a lot/batch number, with no significant difference. The rate of falsified medicines with an expiration date was 47.7%, compared with 33.3% for those without an expiration date, with no significant difference. Products labeled as manufactured in Australia, Turkey, France, the United Kingdom, Canada, Thailand, Taiwan, and Japan were all genuine, and those labeled as manufactured in Sweden and China were all falsified. The rate of falsified medicines was

Table 2. Items Listed on Product and Falsified Medicines

No.	Product visual information	Status	G	F	Rate of F (%)	p-value
1	Country of dispatch	Puerto Rico	0	2	100.0	n.a.
		United States	21	2	8.7	
		India	7	2	22.2	
		Singapore	37	4	9.8	
		Thailand	10	0	0.0	
		Malaysia	3	0	0.0	
		Hong Kong	7	20	74.1	
		Taiwan	29	0	0.0	
		China	2	62	96.9	
		Japan	0	4	100.0	
2	Name of the shipping company or shipper	Stated	53	54	50.5	0.132
		Not stated	63	42	40.0	
3	Shipper's company address	Stated	66	61	48.0	0.398
		Not stated	50	35	41.2	
4	Customs declaration description	Stated	26	64	71.1	<0.01
		Not stated	90	32	26.2	
5	Package insert	Stated	63	59	48.4	0.33
		Not stated	53	37	41.1	
6	Lot/batch number	Stated	91	77	45.8	0.865
		Not stated	25	19	43.2	
7	Expiry date	Stated	92	84	47.7	0.142
		Not stated	24	12	33.3	
8	Country of manufacture	United States	10	9	47.4	n.a.
		United Kingdom	6	0	0.0	
		India	15	7	31.8	
		Australia	1	0	0.0	
		Netherlands	7	0	0.0	
		Canada	1	0	0.0	
		Sweden	0	2	100.0	
		Thailand	2	0	0.0	
		Germany	8	3	27.3	
		Turkey	1	0	0.0	
		France	5	0	0.0	
		Taiwan	1	0	0.0	
		China	0	16	100.0	
		Japan	1	0	0.0	
		Not listed	58	59	50.4	
9	Manufacturer's address	Stated	93	38	29.0	<0.01
		Not stated	23	58	71.6	
10	Name of manufacturer	Stated	84	55	39.6	<0.05
		Not stated	32	41	56.2	
11	Manufacture date	Stated	37	27	42.2	0.652
		Not stated	79	69	46.6	
12	Instruction leaflet	Stated	2	2	50.0	1.000
		Not stated	114	94	45.2	

G, genuine; F, falsified; n.a., not applicable.

Table 3. Performance of the Model in Predicting Personal Import Agency Websites Selling Falsified Medicines

Data set	Matching matrix			Accuracy	AUC	Gini
		Genuine	Falsified			
Training	Genuine	62	3	89.58	0.906	0.812
	Falsified	12	67			
Test	Genuine	31	0	82.35	0.889	0.779
	Falsified	12	25			

AUC, area under the receiver operating characteristic curve.

29.0% for products with the manufacturer's name on the outer package and 71.6% for products without the manufacturer's name, which was significantly higher ($p < 0.01$). The rate of falsified medicines was 39.6% for products with a manufacturer's address and 56.2% for products without a manufacturer's address, with a significantly higher rate ($p < 0.05$). The rate of falsified medicines was 42.2% for products with a manufacturing date, and 46.6% for products without a manufacturing date, with no significant difference. The rate of falsified medicines was 50.0% for products with an instruction leaflet, and 45.2% for products without an instruction leaflet, with no significant difference.

Creating Classification and Prediction Models Using Machine Learning

Personal Import Agency Websites A classification and prediction model was created for the 27 items observed on each website, based on the presence or absence of each item on each website and the authenticity of medicines obtained through that website. We extracted the characteristics of websites selling falsified medicines (Fig. 1). The results of performance assessment for the classification and prediction models for the items listed on personal import agency websites are

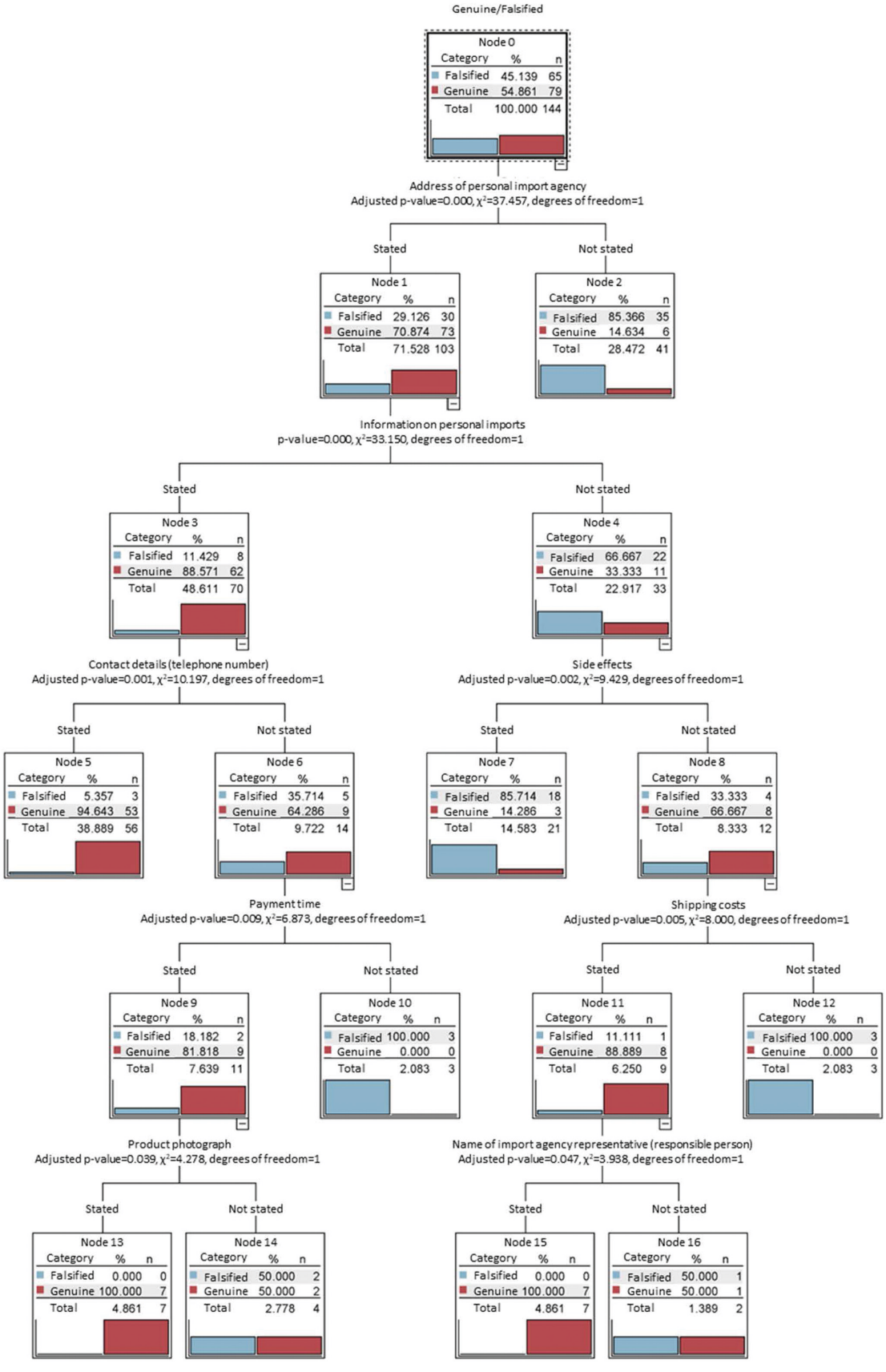

Fig. 1. Tree Diagrams Created from the Information Listed on Personal Import Agency Websites

The ratio of genuine to falsified is shown in each node branched by observation item.

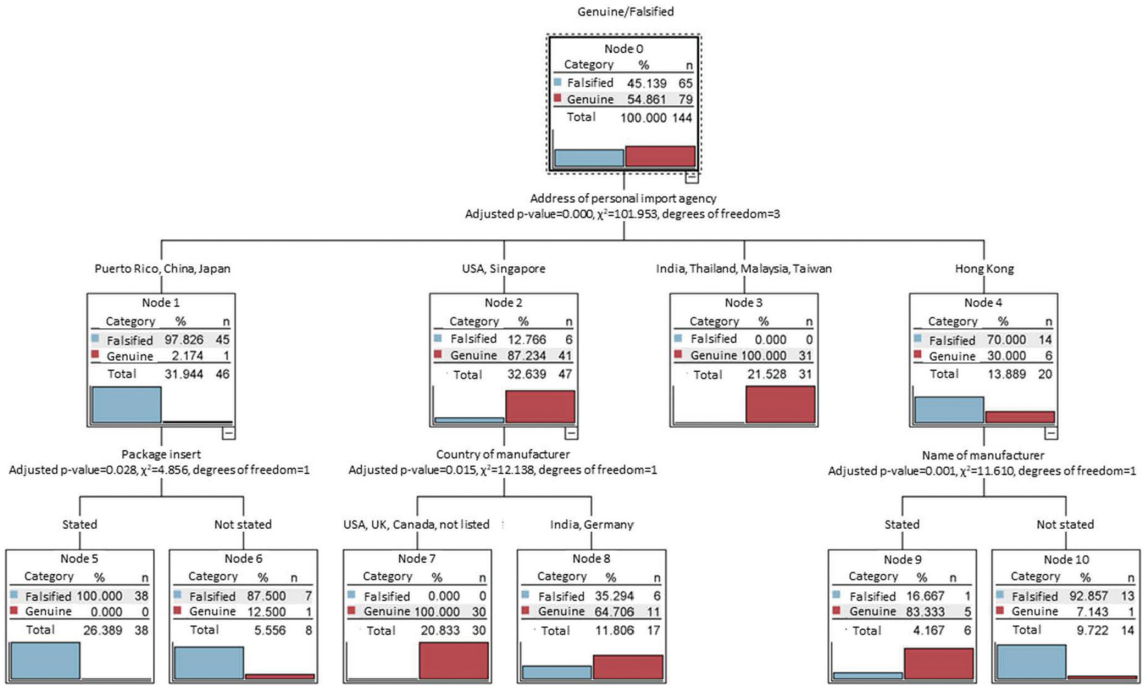

Fig. 2. Tree Diagrams Created from the Information on Each Product Obtained

The ratio of genuine to falsified is shown in each node branched by observation item. USA, United States; UK, United Kingdom.

shown in Table 3.

The conditions for a 90% or higher probability of being a genuine product were the presence of a personal import agency address and a statement pertaining to personal import. The conditions for a 90% or higher probability of being a falsified medicine were the absence of a personal import agency address and the presence of a statement regarding the payment time.

As for the accuracy of the constructed model, the correct discrimination rates in the training and test data were 89.58% and 82.35%, respectively. As an assessment of the generalization performance of the constructed model in predicting websites that sold falsified medicines, the AUC of the training and test data was 0.906 and 0.889, respectively, and the Gini impurity for the training and test data was 0.812 and 0.779, respectively.

Personally Imported Medicines A classification and prediction model was created using the 12 items observed for the obtained products based on the information described in each item and the product's authenticity. Items that were characteristically observed in falsified medicines were extracted (Fig. 2). The results of performance assessment for the classification and prediction models in using visually observable product characteristics to predict falsified medicines are shown in Table 4.

The condition for a 90% or higher probability of being a genuine medicine was that it was shipped from the United States, Singapore, Thailand, Malaysia, or Taiwan. The condition for a 90% or higher probability of being a falsified medicine was that the product was shipped from China or Japan, or that the product was shipped from India or Hong Kong, and the name of the manufacturing company was not indicated.

As for the accuracy of the constructed model, the correct discrimination rates in the training and test data were 93.75% and 94.12%, respectively. As an assessment of the generalization performance of the constructed model in predicting falsified medicines, the AUC of the training and test data was 0.985 and 0.925, respectively, and the Gini impurity of the training and test data was 0.970 and 0.851, respectively.

Items Listed on Websites and Products A classification and prediction model was created on the basis of the presence or absence of the 27 items for websites and 12 items for products in combination, as well as the authenticity of the obtained products, to identify falsified medicines (Fig. 3). The results of performance evaluation for the classification and prediction models are shown in Table 5.

The conditions for a 90% or higher probability of being a genuine medicine were that the product was shipped from the United States, Singapore, Thailand, Malaysia, or Taiwan, or that the product was shipped from India or Hong Kong and the address of the personal import agency was listed. The conditions for a 90% or higher probability of being a falsified medicine were that the products were shipped from China or Japan, or that the products were shipped from India or Hong Kong and the address of the personal import agency was not listed.

As for the accuracy of the constructed model, the correct discrimination rates in the training and test data were 93.75% and 94.12%, respectively. As an assessment of the generalization performance of the constructed model to predict falsified medicines, the AUC of the training and test data was 0.987 and 0.925, respectively, and the Gini impurity of the training and test data was 0.973 and 0.851, respectively.

DISCUSSION

Relationship Between the Frequency of Falsified Medicines and Visual Observations

Personal Import Agency Websites As a result of our

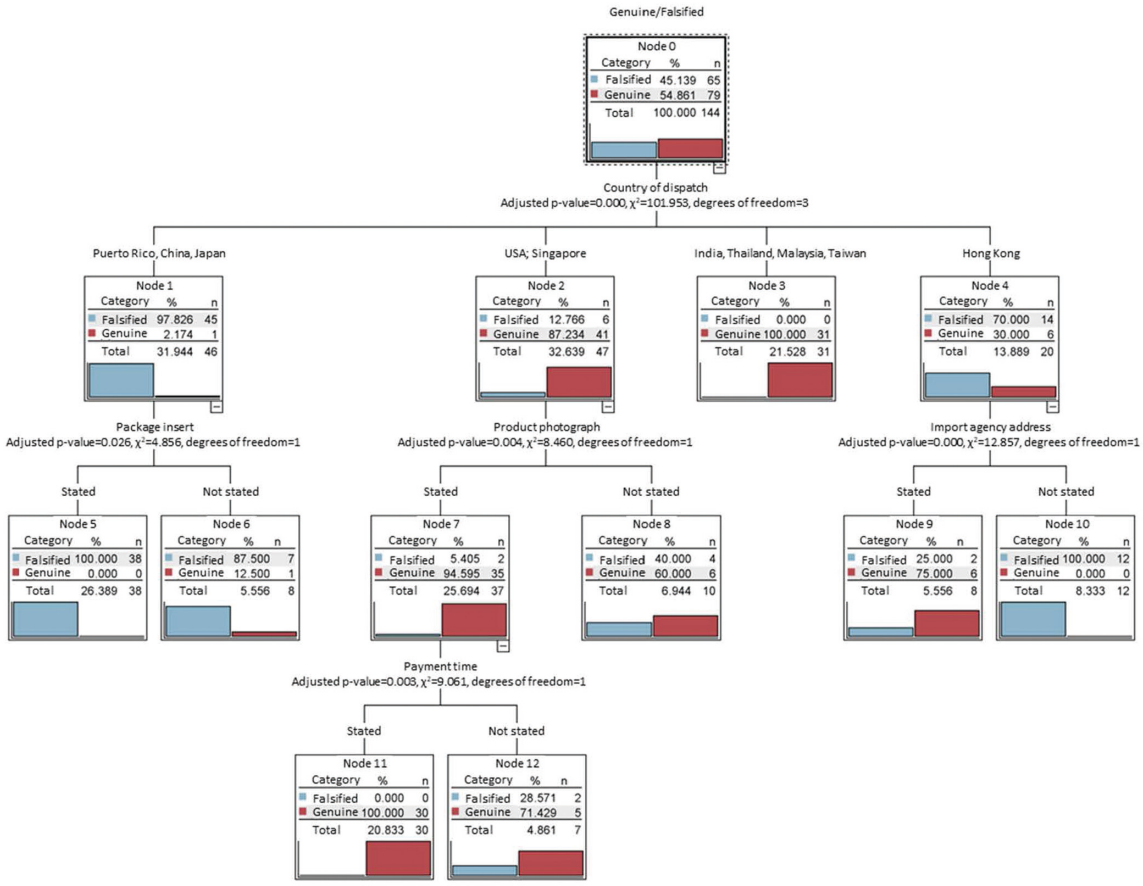

Fig. 3. Tree Diagrams Created from the Information on Both the Import Agency Website and the Product Obtained

The ratio of genuine to falsified is shown in each node branched by observation item. USA, United States.

Table 4. Performance of the Model in Predicting Falsified Medicines Using Product Visual Information

Data set	Matching matrix			Accuracy	AUC	Gini
		Genuine	Falsified			
Training	Genuine	58	7	93.75	0.985	0.97
	Falsified	2	77			
Test	Genuine	29	2	94.12	0.925	0.821
	Falsified	2	35			

AUC, area under the receiver operating characteristic curve.

Table 5. Performance of the Model in Predicting Falsified Medicines Using Both Website and Product Visual Information

Data set	Matching matrix			Accuracy	AUC	Gini
		Genuine	Falsified			
Training	Genuine	57	8	93.75	0.987	0.973
	Falsified	1	78			
Test	Genuine	29	2	94.12	0.925	0.851
	Falsified	2	35			

AUC, area under the receiver operating characteristic curve.

investigation regarding the incidence of falsified products based on website information, significant differences were found in the following 20 out of 27 items: contact details (telephone number), contact details (fax number), contact address (e-mail), enquiry mail form, import agency name or personal name, import agency address, name of import agency representative, product price, photograph of product, product name, dosage and administration, indications, recommendation to consult a doctor or pharmacist, information about personal import, limitations on the quantity of personal import purchases, reference to the Act on Specified Commercial Transactions, reference to the Pharmaceutical and Medical Device Act, information about consultation, use of SSL encryption, and privacy policy. For most items, the rate of falsified products was significantly higher when these items were not present on the website. In contrast, the rate of falsified medicines was significantly higher on websites where dosage, administration, and indications were stated. However, determining the authenticity of medicines based on the presence or absence of these items may lead to mistaken judgment. It was shown that not

only websites selling falsified medicines but also some selling genuine products provide insufficient or improper information (Table 1).

The rate of falsified medicines was significantly higher when the address of the personal import agency and the name of the representative were not stated, indicating that many websites selling falsified medicines do not provide information about the company behind the personal import agency. The rate of falsified products was also higher when no contact details (telephone or fax number) were provided, suggesting that websites without an address or contact details are more likely to be selling falsified products. The World Health Organization has issued a warning about websites that do not state a physical address or telephone number to avoid the risk of purchasing falsified medicines; this advice is consistent with the characteristics identified in our study.[16)]

The product price and product name were listed on nearly all websites; these items were not considered important as checkpoints in detecting falsified medicines. Some websites provided a photograph, but the product was delivered in a different packaging from that in the photograph. Not obtaining the desired product is a potential risk when importing medicines personally because the website may show a photograph of a product that differs from the actual product sent.

The rate of falsified medicines was significantly higher when there was no information about personal import or reference to the Act on Specified Commercial Transactions, suggesting that websites selling falsified products are more likely to omit mention of regulatory-related items.

Statistical analysis showed that the incidence of falsified medicines was significantly higher for medicines obtained from personal import agencies that did not state identifying information such as contact details or the name and address of the import agency, as well as from agencies that did not include a reference to regulations regarding topics such as personal import or specified commercial transactions. Refraining from purchasing products from such websites may help prevent the acquisition of falsified medicines.

Personally Imported Medicines As a result of investigating the rate of falsified products based on items listed on the actual product samples, significant differences were found for 3 of 12 items: customs declaration information on the outer package, manufacturing company name, and address listed on the product package (Table 2). We observed that if a customs declaration was listed, the product was likely to be falsified. Additionally, if the product had no manufacturer name or address, it was likely to be a falsified product requiring caution. Although tools for the visual inspection of medicines have been published for identifying suspected falsified medicines through product observation,[17)] this study is the first to investigate suspected falsification, including information about the product's transport.

Among our samples, all products shipped from Puerto Rico and Japan were falsified, and 96.9% of products shipped from China were falsified. The rate of falsified medicines in personally imported medicines shipped from these countries was found to be particularly high, with a bias in the rate of falsified medicines according to the country of dispatch. Among the observed visually information, information on the country of dispatch was considered to be an item that should be noted as a characteristic of falsified medicines entering Japan through personal import via the Internet. However, the total number of international mail shipments from foreign countries is enormous, so it would be unrealistic to rely solely on this information to identify falsified goods. Additionally, products shipped from Japan that arrive by post with Japanese postage stamps cannot be considered personal import products, which would be in conflict with the law. However, tracking the shipper of falsified goods received by post with Japanese postage is difficult owing to insufficient information about the shipper.

All products with no mention of the shipping company name or name of the shipper were described as undecipherable; there were no products with no mention of either. There was no significant difference in the rate of falsified products depending on whether the shipper's information was specified or not, suggesting that there may be other items that can be used to better identify falsified products.

The shipper address was listed for all products, except for one product sent from Japan. This is thought to be owing to the fact that international mailing labels have fields for the shipper's name and customs declaration details. A detailed analysis of the information provided might reveal certain characteristics. However, the authenticity and appropriateness of the information can only be ascertained by checking whether it matches the package contents. Therefore, it would be quite difficult to determine whether an unopened package contains falsified products by relying solely on the shipper information.

After the outer package was opened, the information on the product was observed. Together with information on the outer label, such as the country of shipment and customs declaration, it was possible to use this information to determine whether the product might be falsified. Our results indicate that products shipped from China that includes a customs declaration and do not include the name or the address of the manufacturing company on the product packageare more likely to be falsified. These items can serve as checkpoints in the distribution process of medicinal products in settings outside of a research environment, as well as for individuals who have personally acquired imported medicinal products, to ascertain whether the products purchased might be falsified.

Identification Using Machine Learning

Personal Import Agency Websites Of the 27 items observed on websites of personal import agencies, the import agency address, payment time, and information on personal import were among the key features that could predict websites dealing in falsified medicine. This model might be used to predict falsified medicines, although the accuracy was approximately 80% (Fig. 1 and Table 3). Nevertheless, careful observation of the presence or absence of these items was shown to be useful in detecting possible falsified products, according to our statistical analysis. In this study, we found that, in particular, purchasing products from websites that do not list their address should be avoided to prevent the purchase of falsified products.

As for the payment time, although our statistical analysis showed that the rate of falsified medicines was significantly higher when payment time was not stated on the website, CHAID analysis predicted that websites stating this item were more likely to sell falsified medicines. This was based on the assumption that the personal import agency address was not provided, suggesting that if information about the payment time was present but the import agency address was absent, the website was considered more likely to be selling falsified medicines.

Personally Imported Medicines Of the 12 items that could be observed on the products themselves, the country of dispatch (United States, Singapore, Thailand, Malaysia, Taiwan, China, Japan, India, and Hong Kong) and information about the manufacturer were identified as key points in identifying falsified products. The model predicted falsified medicines with an accuracy of approximately 94% (Fig. 2 and Table 4). We showed that prediction using visually observable product information was more accurate than identifying falsified products using website information.

Manufacturer information was also identified as having a significantly different rate of being associated with falsified medicines (Table 1). Among the 12 observable items on the product, this item is considered particularly useful for identifying falsified products.

As for the prediction of falsified medicines using the country of dispatch, the results were mostly consistent with the trend in the rate of falsified products for each country of dispatch, as well as the predictions using CHAID. For products shipped from India and Taiwan, falsified products could be inferred in combination with information about the manufacturer.

Features of falsified products extracted using information available on the outer package and the products themselves indicated that products whose country of dispatch is China or Japan, or those whose country of dispatch is India or Hong Kong and that lack manufacturer information, are more likely to be falsified.

Items Listed on Websites and Products Of a total of 39 items (27 observable items on websites and 12 observable items on products), the country of dispatch (United States, Singapore, Thailand, Malaysia, Taiwan, China, Japan, India, or Hong Kong) and address of the personal import agency were extracted as important checkpoints for identifying falsified products. The country of dispatch was the most impactful. and the model including this item predicted falsified medicines with an accuracy of approximately 94% (Fig. 3 and Table 5).

Predictive modeling using website information showed little difference in terms of accuracy in predicting falsified products compared with the model using visual product information. However, the features observable on websites should not be ignored. Our findings suggest that the appearance of the product is important in the identification of falsified medicines. However, medicines for personal import are most commonly ordered from a personal import agency website. Even if the accuracy of falsified product identification might be slightly lower, taking note of information on the website is important to avoid the risk of purchasing falsified products.

Our results suggest that the country of dispatch can be used in combination with the presence/absence of the import agency address to predict falsified products. The prediction results can vary depending on whether these items are used separately or in combination, which should be taken into account when using this information to identify falsified medicines.

Feature Extraction and Identification Processes

In this study, we used statistical analysis and machine learning to extract and identify the features of falsified medicines. Statistical analysis can be carried out with relatively small amounts of data owing to its relatively simple structure, and this method can be used to draw appropriate inferences and conclusions if hypotheses about the distribution of data are met. In this study, 23 of 39 included items of observation were significantly associated with the proportion of falsified medicines, suggesting that a checklist for identifying falsified medicines could be developed. However, we considered that those 23 items should be weighted in developing a prediction method. We therefore used machine learning as an alternative approach.

Machine learning uses large amounts of data owing to the complexity of its algorithmic structure. This method has a high potential for more accurate prediction and analysis than statistical analysis if sufficient data can be obtained. In our study, a falsified medicines prediction model was created using CHAID; however, with the exception of lack of personal import agency address, the observation items found to have a statistically significant association with the rate of falsified medicines differed from those extracted using machine learning to predict falsified medicines. Although each feature was derived from medicines obtained through personal import, our study results suggest that prediction of falsified medicines using machine learning can be used to make effective inferences from multiple perspectives, and it might be possible to develop a method for predicting product falsification using fewer features than those needed for statistical analysis.

Research results have been reported using the latest analytical technologies to identify falsified medicines with high accuracy.[18,19] Additionally, non-destructive methods using the latest miniaturized devices is also being developed.[20,21] Ours is the first report of an approach to predict falsification solely on the basis of appearance. This approach is innovative in that it does not rely on comparisons with genuine medicines to predict falsified ones. The results of this study, derived with maximum consideration given to the possibility of implementation in the actual pharmaceutical distribution process, are expected to contribute to the optimization of pharmaceutical distribution in the future.

The features of falsified medicines were extracted in this study; these findings indicate that products with none of these features are more likely to be genuine. However, the results of this study cannot be used to ensure the authenticity of products obtained online by consumers. Even with a low probability of being a falsified product, the potential risks associated with personally imported medicines remain. If consumers purchase medicines from a personal import agency, information on the website, as well as the potential risks associated with imported medicines, should be carefully considered before making a final decision on whether to use the medicines obtained.

LIMITATIONS OF THIS STUDY

The scope of this study is limited to medicines imported personally to Japan and personal import agency websites through which these medicines can be purchased, which we have addressed in our previous research. We were unable to extract the characteristics of all medicines that can be purchased online for personal import. However, it might be possible to develop a visual inspection method to identify falsified medicines in circulation via a similar process from features that can be observed visually, according to the regulations and actual conditions of access to medicines in each country. Also, there are various approaches to the construction of prediction models. For the purpose of establishing an accurate method to identify falsified medicines using visual inspection, it is neces-

sary to collect, compare, and validate new data in line with the current trend in the future.

CONCLUSION

In this study, we identified the features of websites most likely to sell falsified medicines for personal import into Japan, as well as features of falsified medicines that could be observed visually. Furthermore, we constructed a prediction and classification model to predict falsified medicines, offering the possibility of identifying these falsified products without the need to conduct chemical analysis or other types of systematic investigation. The present results will contribute to improving consumer awareness and strengthening the fight against falsified medicines entering the country through personal import.

Acknowledgments The identification of websites dealing in falsified medicines was supported by MHLW Pharmaceuticals and Medical Devices Regulatory Science Policy Research Program (grant no. JPMHLW23KC2007). The development of a method for detecting falsified medicines using visual inspection was supported by JSPS KAKENHI (grant number JP22K04597). We thank Analisa Avila, MPH, ELS, of Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript.

Conflict of interest The authors declare no conflict of interest.

REFERENCES

1) World Health Organisation. WHO Member State mechanism on substandard/spurious/falselylabelled/falsified/counterfeit (SSFFC) medical products: Report by the Director-General (A70/23). Geneva: World Health Organisation; 2017 March. ‹http://apps.who.int/gb/ebwha/pdf_files/WHA70/A70_23-en.pdf›, accessed 18 February, 2025.
2) World Health Organization. WHO Global Surveillance and Monitoring System for substandard and falsified medical products, 2017.: ‹http://who.int/entity/medicines/regulation/ssffc/publications/GSMS_Report.pdf›, accessed 18 February, 2025.
3) Khan MH, Tanimoto T, Nakanishi Y, Yoshida N, Tsuboi H, Kimura K. Public health concerns for anti-obesity medicines imported for personal use through the internet: a cross-sectional study. *BMJ Open*, **2**, e000854 (2012)
4) Yoshida N, Numano M, Nagasaka Y, Ueda K, Tsuboi H, Tanimoto T, Kimura K. Study on health hazards through medicines purchased on the Internet: a cross-sectional investigation of the quality of anti-obesity medicines containing crude drugs as active ingredients. *BMC Complement. Altern. Med.*, **15**, 430 (2015).
5) Zhu S, Yoshida N, Kimura K, Matsushita R, Tsuboi H. Falsified vardenafil tablets available online. *J. Pharm. Biomed. Anal.*, **177**, 112872 (2020).
6) Sanada T, Yoshida N, Matsushita R, Kimura K, Tsuboi H. Falsified tadalafil tablets distributed in Japan via the internet. *Forensic Sci. Int.*, **307**, 110143 (2020).
7) Sanada T, Ohnishi M, Yoshida N, Kimura K, Tsuboi H. Quality assessment of Diflucan® tablets distributed online. *Med. Access Point Care*, **5**, 1–8 (2021).
8) Ministry of Health. Labour and Welfare. Information for those who are bringing medicines for personal use into Japan. ‹https://www.mhlw.go.jp/english/policy/health-medical/pharmaceuticals/01.html›, accessed 4 December 2024.
9) Shofuda K, Aragane K, Igari Y, Matsumoto K, Ito K. Anti-counterfeit activities of pharmaceutical companies in Japan: for patient safety. *Yakugaku Zasshi*, **134**, 203–211 (2014).
10) Yoshida N, Matsushita R, Takashima S, Rhahman MS, Yamashita H. Health and hygiene survey of personal imported amoxicillin/clavulanic acid combination products. Research on investigating the actual situation and devising countermeasures against falsified medicines distributed globally, *Research Report*, pp. 95–122 (2020). (In Japanese) ‹https://mhlw-grants.niph.go.jp/system/files/2019/193041/201925012B_upload/201925012B0011.pdf›, accessed 27 November, 2024.
11) Sanada T, Yoshida N, Kimura K, Tsuboi H. Discrimination of falsified erectile dysfunction medicines by use of an ultra-compact Raman scattering spectrometer. *Pharmacy (Basel)*, **9**, 3 (2020).
12) Zhu S, Yoshida N, Tsuboi H, Matsushita R, Kimura K. Quality and authenticity of metformin tablets circulating on Japanese websites. *Ther. Innov. Regul. Sci.*, **55**, 656–666 (2021).
13) Zhu S, Yoshida N, Matsushita R, Rahman MS, Kimura K. Circulation of COVID-19-related medicines on Japanese websites during the COVID-19 pandemic and their quality and authenticity. *Am. J. Trop. Med. Hyg.*, **111**, 1097–1106 (2024).
14) Act on Specified Commercial Transactions. ‹https://laws.e-gov.go.jp/law/351AC0000000057›, accessed 4 December 2024.
15) Act on Quality, Efficacy and Safety Assurance of Medical Devices. ‹https://laws.e-gov.go.jp/law/335AC0000000145›, accessed 4 December 2024.
16) World Health Organization. “Substandard and falsified medical products: Advice to patients and consumers. 9 December 2019.”: ‹https://www.who.int/news-room/questions-and-answers/item/substandard-and-falsified-medical-products-advice-to-patients-and-consumers›, accessed 18 February, 2025.
17) International Pharmaceutical Federation. Tool for Visual Inspection of Medicines. ‹https://www.fip.org/files/fip/counterfeit/VisualInspection/A%20tool%20for%20visual%20inspection%20of%20medicines%20EN.pdf›, accessed 4 December 2024.
18) Ciza PH, Sacre PY, Waffo C, *et al*. Comparing the qualitative performances of handheld NIR and Raman spectrophotometers for the detection of falsified pharmaceutical products. *Talanta*, **202**, 469–478 (2019).
19) Young JM, Liddicoat C, van Dijk KJ, *et al*. Environmental DNA as an innovative technique to identify the origins of falsified antimalarial tablets-a pilot study of the pharmabiome. *Sci. Rep.*, **12**, 21997 (2022).
20) Sanada T, Yoshida N, Kimura K, Tsuboi H. Discrimination of falsified erectile dysfunction medicines by use of an ultra-compact Raman scattering spectrometer. *Pharmacy (Basel)*, **9**, 3 (2020).
21) Sanada T, Yoshida N, Kimura K, Tsuboi H. Detection method of falsified medicines by using a low-cost Raman scattering spectrometer combined with soft independent modeling of class analogy and partial least squares discriminant analysis. *Biol. Pharm. Bull.*, **44**, 691–700 (2021).