Algorithm and hyperparameter optimizations for hetero-device classification by near-infrared spectra of falsified and substandard amoxicillin capsules

Yusuke Hattori1 · Yuka Hoshi1 · Naofumi Hashimoto2 · Yasunori Ichimura2 · Yasuo Sugiura2 · Makoto Otsuka3

Received: 13 January 2022 / Accepted: 2 June 2022 / Published online: 8 August 2022
© The Author(s), under exclusive licence to The Japan Society for Analytical Chemistry 2022

Abstract
In this work, we optimized classification algorithms and the hyperparameters for screening falsified and substandard amoxicillin capsules. The distribution of low-quality medical products is a serious problem, especially in low- and middle-income countries. Near-infrared (NIR) spectroscopy has been proposed as the first choice for a screening device. However, preparation of the reference library for the classification training is a highly difficult process. We herein propose a hetero-device classification between training and test devices. In this proposal, Fourier-transform NIR spectrometer and portable wavelength dispersive NIR spectrometer were used as training and test devices, respectively. As the classifier candidates, we examined 13 algorithms and selected 8. We then optimized the hyperparameters for these classifiers by the grid search and cross validation methods. In the final analysis, few classifiers were found to give acceptable prediction results by the hetero-device classification. When using these methods, it is crucial to examine the results by the classification probability, due to the trade-off between sensitivity and specificity. Finally, we suggest that k-nearest neighbors, extra trees, and gradient boosting classifiers are the optimal algorithms with high classification probability for the substandard and falsified amoxicillin capsules.

Keywords Near-infrared spectroscopy · Machine learning · Substandard medicine · Falsified medicine · Hetero-device classification

Introduction
Substandard medicine is defined as a product containing less of the active pharmaceutical ingredient (API) than the stated content due to poor manufacturing processes and/or degradation during inappropriate storage or transportation. Falsified medicine is a product containing no API or an incorrect API. These low-quality products are found anywhere in the world, and falsified antibiotics and antimalarial drugs will often be found in low- and middle-income countries (LMICs) [1, 2].

These low-quality medicines in the marketplace have a significant negative impact on public health. The World Health Organization (WHO) estimated that the observed failure rate of falsified medical products including medicines and medical devices in LMICs was approximately 10.5% [1]. One of the factors contributing to the distribution of falsified agents is a dysfunctional process of periodic inspection by national regulatory authorities. Particularly in LMICs, the supply chain of medical products is often insufficiently controlled.

The WHO [1] reported that the prevalence of low-quality medicines was 11.8% for antimalarials and 7.2% for antibiotics. Bate et al. [3] reported that the discovery rates of substandard and falsified ciprofloxacin were 4.1% and 5.8% in LMICs, respectively. Tabernero et al. [4] reported that the prevalence of low-quality amoxicillin capsules was 16.0% in Laos. Substandard antibiotic medicines are especially pernicious, since they not only provide an inadequate therapeutic
effect but can also contribute to the emergence of drug-resistant strains.

The Asia Development Bank (ADB) [5] and many other researchers [6] have reported the use of the spectroscopic method to screen for these low-quality products. Such spectroscopic screening is becoming dramatically more prevalent due to breakthrough technologies that allow miniaturized spectrometry and low-cost mass production of spectrometers. Raman spectroscopy [7–10] and near-infrared spectroscopy (NIRS) [11–14] are now widely used for screening. In terms of the cost–benefit performance of these two methods, NIRS maintains the best cost efficiency while matching the performance of Raman spectroscopy [5]. In addition to these spectroscopic methods, liquid chromatography [15, 16] and microfluidic “lab-on-a-chip” technology [17] are highly accurate methods for classifying both falsified and substandard products; however, both these methods require sample destruction and are difficult to perform. The portable Minilab analytical device (Global Pharma Health Fund (GPHF), Giessen, Germany) is a thin-layer chromatography-based system that requires destruction of samples and has low sensitivity to substandard products [18–22]. Paper chromatography is a simpler and less expensive destructive method [23, 24].

By contrast to the destructive methods, NIRS has advantages as a non-destructive tool for the screening of product quality. The only drawback of NIRS is the requirement of a reference library as the training data for classification. Usually, the reference library is prepared using the same device as used for the testing, or at least a same model. In our previous report [25], we suggested the use of hetero-devices that perform both training and testing of classification. The hetero-device classification is expected to encourage the use of low-cost NIRS devices for screening. We believe that this method would be effective to protect the health of people in LMICs by eradicating low-quality medicines.

In this proposed method, under the assumption of using a portable device for testing and using a Fourier-transform (FT) NIR spectrometer for training. Because the FT-NIR spectra has a superior resolution and other advantageous qualities compared to the wavelength dispersive (WD) spectra via a portable device, and it is possible to transform and interpolate the FT spectra to wavelength spacing spectra. Thus, the training and testing of classification were performed using wavelength spacing spectra with the same intervals.

The hetero-device classification method successfully screened out falsified amoxicillin capsules by principal component analysis (PCA) and a support vector machine (SVM) classifier [25]. However, the accuracy of this method for classifying the substandard products was quite low, and it most cases it failed to screen such products out. An improvement of the classification accuracy is thus strongly required before NIRS can be successfully applied for the field testing of medicines. In this study, we examined the accuracy of 13 different classification algorithms for screening amoxicillin capsules. Subsequently, optimization of hyperparameters and spectral preprocessing methods were examined and the classification performances discussed.

**Experimental**

**Materials**

Fifteen marketed brands of amoxicillin capsules were obtained in eight Asian and three African countries, as shown in Table 1. These commercial products were stated to contain 250 mg or 500 mg of amoxicillin in each colored capsule. These capsules were packaged in blister packs of colorless transparent plastic and aluminum sheets in small quantity. The bulk powder of amoxicillin trihydrate used as a standard was purchased from Tokyo Chemical Industry (Tokyo, Japan).

**Apparatus**

Near-infrared (NIR) spectra of amoxicillin capsules were acquired using an FT NIR spectrometer (MPA; Bruker Optics, Ettlingen, Germany) and WD NIR spectrometer (NIRONE S2.0 model; Spectral Engines, Helsinki, Finland). The content of amoxicillin in each capsule was determined using a high-performance liquid chromatography (HPLC) system composed of an auto-sampler (SIL-20AC; Shimadzu, Kyoto, Japan), a pump (LC-20AB; Shimadzu), a UV–Vis detector (SPD-M20A; Shimadzu) and a C18 column (ϕ4.6 mm, L250 mm; InertSustain, GL Sciences, Tokyo).

**Methods**

For the validation of non-destructive analysis, diffuse reflectance NIR spectra of capsules were collected at a depth beyond the underside of the plastic sheet of the blister pack. As the training data, 30 spectra of each brand product were collected using the MPA in the wavenumber range of 4000–9500 cm−1 with 16 cm−1 resolution. Fifteen WD NIR spectra were collected for each product using NIRONE in the wavenumber range of 1550–1950 nm at an interval of 10 nm. In total, 450 FT NIR spectra and 225 WD NIR spectra were obtained for the training and test data sets, respectively.

After the spectral collection, amoxicillin capsules were removed from the blister pack and the content was assayed using an HPLC system. A sample capsule was dissolved in 1 L of water, and the solution was centrifuged.
The supernatant liquid was filtered using a membrane filter (0.45 μm pore size). Standard amoxicillin solution of 0.500 g/L was prepared by dissolving in water. Five percent acidic methanol solution (pH 4.5) was used as the mobile phase. The mobile phase and sample solution were flowed at 0.9 mL/min at 25 °C. The separated amoxicillin was detected by UV absorbance at 230 nm. Ten capsules were used for the assay of each brand product.

### Classification methods

**Preprocessing of spectra** Data processing and classifications were performed using Python (version 3.8.8) [26] via Jupyter notebook (version 6.3.0) [27] as the computational environment. The figures shown in this paper were prepared using Matplotlib (version 3.3.4) [28] in Python. The array data of spectra were processed using the modules of Pandas (version 1.2.4) [29], NumPy (version 1.20.1) [30], and SciPy (version 1.6.2) [31] in Python. The FT-NIR spectra collected using MPA were converted from wavenumber spectra to wavelength spectra. To unify the wavelength intervals of MPA data and NIRONE data, the non-constant wavelength intervals of the converted MPA spectra were standardized to 10 nm by interpolation using a cubic curve as the approximate curve. The interpolation was performed using the SciPy module. All of the equally spaced wavelength spectra were processed using the standard normal variate (SNV), and then first or second derivation was conducted using a Savitzky-Golay filter with a window size of 5 and a cubic curve as the approximation to cancel baseline variation. The SciPy module was also applied for the Savitzky-Golay filter in Python.

**Machine learning classification** The preprocessed data sets of MPA and NIRONE in the wavelength range of 1550–1770 nm were concatenated into a single array data set of array. The spectra in the range 1770–1950 nm includes absorbance due to water; thus, the range was removed from the data set. Principal component analysis (PCA) was then perform using the scikit-learn library (version 0.24.1) [32] in Python. The obtained principal components (PCs) were used for the following analysis of machine learning classification.

**Optimization of classifier algorithms and hyperparameters** As the machine learning classifier candidates, the 13 algorithms listed in Table 2 were used for classification with the PCs of both MPA and NIRONE by the scikit-learn library in Python. The classification accuracy was evaluated using the k-fold cross-validation with a fold number (k) of 5, default parameters of scikit-learn, and data shuffling prior to splitting into the folds.

The hyperparameters were then optimized for the potential classifiers with high average of accuracy score on the former k-fold cross-validation. The optimization was performed using both data of MPA and NIRONE by a grid search of the scikit-learn library. The outputs of the grid search were evaluated based on the accuracy score of the k-fold cross-validation with five folds and shuffling of the data.

### Table 1 Sample information of amoxicillin capsules

| Market | Lot No | Stated vendor                  | Stated content (mg) | Determined content (%) | SD (%) |
|--------|--------|--------------------------------|---------------------|------------------------|--------|
| A Vietnam | 1,041,017 | DOMESCO                  | 500                 | 99.5                   | 1.57   |
| B China     | 50,817  | Accord Healthcare           | 500                 | 99.9                   | 0.94   |
| C India     | DBC028  | Marksans Pharma             | 500                 | 97.8                   | 1.54   |
| D Thailand  | X25T    | GlaxoSmithKline             | 500                 | 97.9                   | 1.22   |
| E Mongolia  | V26998  | KRKA                       | 500                 | 99.0                   | 0.87   |
| F Cameroon  | H02605  | Maxxheal Pharmaceuticals    | 500                 | 98.0                   | –      |
| G India     | 684     | Maxx Pharmaceuticals        | 500                 | 94.9                   | 1.34   |
| H Zambia    | 180,631 | North china pharmaceutical | 250                 | 99.0                   | 0      |
| I Japan     | 2790    | Astellas Pharma            | 250                 | 95.9                   | 2.84   |
| J Zambia    | BP8CO2M03 | Shalina Healthcare      | 500                 | 96.8                   | 1.25   |
| K Senegal   | 1,831,311,083 | Ubithera          | 500                 | 83.7b                  | 3.58   |
| L Laos      | 0329/98 | Codupha                    | 500                 | 99.1                   | 0.94   |
| M Vietnam   | 3418    | Sandoz                     | 500                 | 111                    | 1.26   |
| N Myanmar   | 1,706,201 | Kunming Baker Norton Pharmaceutical | 250 | 111 | 2.19 |
| O Congo     | 17-L8   | Phatkin                    | 500                 | 111                    | 2.19   |

*a Falsified drug  
b Substandard drug  
c Standard deviation
Results and discussion

HPLC assay of the amoxicillin capsules showed that 13 products included more than 95% of the stated content as standard products. The content of product “L” was 83.7 ± 3.58% on average, and no amoxicillin was detected in product “F” by HPLC. Thus, the products “L” and “F” were classified as substandard and falsified medicines, respectively. Referencing the classes of standard, substandard, and falsified, we then applied the machine learning classifiers for the capsules.

First, the obtained MPA spectra were interpolated to be even wavelength intervals of 10 nm, and then the concatenated array data set of MPA and NIRONE data was processed by SNV and derivation using a Savitzky-Golay filter. Figure 1 shows the processed spectra of amoxicillin capsules and amoxicillin trihydrate. In the range of 1550–1770 nm, some characteristic peaks of amoxicillin were found in both first and second derivative spectra. All found peaks at 1640 nm, 1680 nm, and 1720 nm could be attributed to the first overtone of C–H stretching mode [33]. Intensive peaks at 1720 nm and 1680 nm were clearly found in each spectrum of both MPA and NIRONE, and the weak peak at 1640 nm on the sample spectra was unclear. Comparing between MPA and NIRONE data, the peaks on the NIRONE data were broader than the peaks on the MPA data. The first derivative spectrum of sample “F”, which was falsified, showed a slight decrease around the peak at 1640 nm. The substandard product “L” showed spectra similar to those of the other standard products, such that it would be extremely difficult to discriminate these products based on the spectra.

In a previous study [25], we suggested that amoxicillin capsules could be discriminated based on the PC scores derived from PCA of these spectra. Our rationale was that these spectra involve discarded information such as device characteristics, capsule materials and the plastic sheet of the blister pack. PCA functions to remove these unnecessary contributions from the spectra, yielding classification results superior to those obtained by the direct use of spectra for machine learning classifiers. Other researchers reported advantages of the combined use of PCA and machine

| Classifier type         | Classifier                           | 1st derived | 2nd derived |
|-------------------------|--------------------------------------|-------------|-------------|
| Nearest neighbor-based  | Radius neighbors classifier           | 82.98       | 82.98       |
|                         | K-nearest neighbors classifier (KNN) | 96.03       | 95.74       |
| Tree                    | Decision tree classifier              | 93.62       | 93.33       |
|                         | Extra tree classifier                 | 91.06       | 92.91       |
| Ensemble                | AdaBoost classifier                   | 84.82       | 91.35       |
|                         | Bagging classifier                    | 94.18       | 94.61       |
|                         | Gradient boosting classifier (GBC)    | 95.60       | 95.74       |
|                         | Extra trees classifier (ETC)          | 95.60       | 95.89       |
|                         | Random forest classifier (RFC)        | 95.32       | 95.04       |
| Support vector machine  | Linear support vector classifier      | 92.77       | 83.40       |
|                         | Support vector classifier (SVC)       | 96.31       | 95.18       |
| Discriminant analysis   | Linear discriminant analysis (LDA)    | 95.60       | 94.61       |
|                         | Quadratic discriminant analysis (QDA) | 96.31       | 94.18       |

![Fig. 1 Preprocessed spectra of MPA (solid lines) and NIRONE (dashed lines) data by first (a) and second (b) derivations](image)
learning algorithms such as SVM [34, 35] and linear discriminant analysis (LDA) [35]. The use of SNV-processed spectra for PCA did not provide clear separation by the PC scores (data not shown). Because the peak separation in the spectra was not enough for the analysis, the first and second derivative spectra were used for PCA.

The scatter plots of PC scores were shown in Fig. S1 (supporting information), which were resulting from PCA of the first derivative spectra. With respect to the first PC (PC1), the scores were clearly divided into negative and positive clusters, which were attributed to the MPA data and NIRONE data, respectively. The PC1 is caused by the device characteristics on the spectra and should be removed for further analysis of classification.

The distribution of PC2 scores also has two clusters consisting of a positive minor cluster in red and major cluster at around zero. The scores of the minor cluster are due to the falsified products “F”; thus, the PC2 score will be the critical parameter for its classification. Although the other higher PC scores make a smaller contribution to the classification, the scatter plots of PC2 versus PC3 and PC2 versus PC4 seem to reflect score variations depending on the content of amoxicillin. In these plots, not only the scores of falsified products, but also the scores of the substandard product are plotted on the near side of the major cluster to the falsified cluster. The results of the PCA of the second derivative spectra showed similar patterns of score plots, as shown in Fig. S2 (supporting information), and the third PC score reflected a clear separation between falsified and other products.

Figure 2 shows the principal axes (PAs) of each PC from the first derivative spectra. These profiles represent the directions of maximum variance in the spectra. Actually, the PA of the PC2 denoted a pattern similar to the amoxicillin spectrum, which means that the PC2 scores make a critical contribution to classification of the products based on the spectral features of amoxicillin. The PA of PC1 also makes large contributions around the 1650–1750 nm range. In this wavelength range, peaks due to the first overtone of the C–H stretching mode are observed, as mentioned above. The capsule film is usually made of gelatin, and the plastic film of the blister pack may be composed of a synthetic resin such as polypropylene. The methyl and/or ethyl groups in these materials contribute to the PC1. These contributions on the spectra were about the same among the products; thus, the differences between devices may have dominantly appeared on the PC1.

The PAs of PC3 and PC5 represented characteristic peaks around 1650 nm and 1700 nm. As shown in Fig. 1, since the absorbance around the wavelength denoted the differences between falsified and other products, the PC3 and PC5 had a high probability of impacting the classification. The contributing ratios of variance explained by each PC were 72.6% (PC1), 15.0% (PC2), 8.3% (PC3), 2.4% (PC4), and 0.5% (PC5). The PCs higher than PC5 may have had less impact on the classification.

The PC scores excluding the first PC score were then used for the machine learning of classification. First, 15 classifier candidates were tested and screened by the cross-validation (CV) scores of classifications based on the PC scores of MPA and NIRONE data sets. As shown in Table 2, seven classifiers with higher accuracy than 95% of CV scores were used for further evaluation and the hyperparameter optimization.

Similar results were given by the second derivatives of the spectral preprocessing with the first derivative data. Although, when using discriminant analysis, the CV scores by the second derivative resulted in lower accuracy than the CV scores of the first derivative, both linear and quadratic discriminant analyses were employed for the next hyperparameter optimization.

The hyperparameter optimization was performed by a grid search using several sets of parameters, and the optimized hyperparameters are listed in Table S1 and S2 (supporting information). Based on these hyperparameters, classification models were prepared using the PCs of only MPA data. Subsequently, the quality classification was predicted using the models and PCs of NIRONE data. The hetero-device predictions were performed by changing the number of PCs from the second PC.

To determine the optimum number of PCs and the preprocess method of spectra, the accuracy scores were obtained by the hetero-device prediction and examined as shown in Fig. 2. When second derivative spectra were used for PCA, the predictions via LDA, SVC, and KNN resulted...
in high accuracy comparing to other classifiers. LDA and SVC respectively provided best accuracies of 94.2% and 93.3% with a low number of PCs, and the best accuracy of KNN was 91.1% with 5 PCs.

The first derivative gave higher accuracy scores compared to the results from the second derivative spectra for most of the employed classifiers. In particular, discriminant analysis of QDA and LDA gave best accuracies of 96.4% and 94.7%, respectively, with 3 PCs (PC2, PC3, and PC4). RFC also gave high accuracy of 96.0% with 3 PCs, and other classifiers provide the highest accuracy with 3 PCs. The impact of these PCs on the classification was already anticipated by the results in Fig. S1 (supporting information), which supported their high contribution. For further discussion, we will focus on the classification results using first derivative spectra due to their better classification scores.

In Table 3, the classification performances were evaluated with the sensitivity, specificity, and precision for the classification of substandard and falsified amoxicillin capsules by the hetero-device method. The sensitivity and specificity were defined as the true positive rate, $R_{tp}$, and true negative rate, $R_{tn}$, respectively, which were represented as follows:

\[
R_{tp} = \frac{N_{tp}}{N_p} = 1 - R_{fp},
\]

\[
R_{tn} = \frac{N_{tn}}{N_n} = 1 - R_{fn},
\]

where $N_{tp}$ and $N_{tn}$ are the predicted numbers of samples as true positive and true negative cases, respectively. $N_p$ and $N_n$ are the actual numbers of positive and negative samples, respectively. $R_{fp}$ and $R_{fn}$ are the false negative rate and false positive rate, which are related to $R_{tp}$ and $R_{tn}$ as represented in these equations, respectively. In our prediction experiments, 15 samples were used and a total of 225 spectra were collected via NIRONE. Hence, the positive and negative numbers of substandard and falsified products were $N_p = 15$ and $N_n = 210$, respectively. In addition, the precision of the prediction was calculated as a positive predictive value, $V_{pp}$, by the following equation:

\[
V_{pp} = \frac{N_{tp}}{N_{tp} + N_{fp}},
\]

where $N_{fp}$ is the predicted number of false positive cases.

The predictive results of falsified classification were excellent for all of the classifiers, with high sensitivity, specificity, and precision. As previously reported, falsified amoxicillin capsules were detectable using hetero-devices regardless of the prediction methods used [25].

As shown in Fig. 3, the total accuracy of RFC and QDA was relatively high among the classifiers; however, the sensitivities for the substandard product were only 0.60 and

|       | Substandard |   | Falsified |   |
|-------|-------------|--|--|----------|--|
|       | $R_{tp}$    | $R_{tn}$ | $V_{pp}$ | $R_{tp}$ | $R_{tn}$ | $V_{pp}$ |
| KNN   | 0.80        | 0.94     | 0.50      | 1.0      | 1.0      | 1.0      |
| GBC   | 0.73        | 0.95     | 0.52      | 1.0      | 1.0      | 1.0      |
| ETC   | 0.60        | 0.95     | 0.47      | 1.0      | 1.0      | 1.0      |
| RFC   | 0.60        | 0.99     | 0.60      | 1.0      | 1.0      | 1.0      |
| SVC   | 0.40        | 0.99     | 0.67      | 1.0      | 1.0      | 1.0      |
| LDA   | 0.47        | 1.0      | 1.0       | 1.0      | 1.0      | 1.0      |

![Fig. 3](image-url) Variations in the accuracy scores of hetero-device predictions via eight classifiers using first derivative (a) and second derivative (b) spectra.
0.47 for RFC and QDA, respectively. On the other hand, KNN and GBC gave higher sensitivity than the discriminant analysis. In particular, KNN gave a sensitivity of 0.80 but specificity of 0.94, and the latter value was lower than that of the other classifiers. These classifiers logically trade off specificity against sensitivity. Comparing the precision of the substandard classification, KNN and GBC gave a precision of around 0.5. SVC gave neither true positive nor false positive results for this substandard classification. The sensitivity of SVC was too low to detect substandard products.

Consequently, there was a trade-off between sensitivity and specificity for the substandard classification; thus, we suggest using several classifiers and carefully examining the probability of classification. Figure S3 (supporting information) shows the classification probability charts for each product. QDA and KNN denoted that the falsified product “F” could be 100% detectable, and the probability decreased to around 80% using SVC. The substandard probabilities of product “L” were around 44% and 66.7% on average by LDA and KNN, respectively. The lowest probability was given by SVC. Despite the high specificity of LDA for the substandard classification, some products have around 10% probability for false classification as substandard. KNN yields a substandard probability of 44% and 29% for products “C” and “J”, respectively. The other probabilities of the false classification are comparable with the results of LDA.

Machine learning can provide predictive classification; however, the classification is somewhat less than perfect. Figure 4 shows the classification probability of each product by KNN classifier. Visualizing the probability chart would be helpful to approach the correct classification.

As a result, KNN demonstrated that the product “L” is predicted to be substandard with high probability. ETC and GBC also gave high probability for “L” as substandard as shown in Fig. S3 (supporting information). The product “C” is a standard product; however, the classification results were over 40% probability by the classifiers of KNN, ETC, and GBC. This suggests the need for rigorous analysis to confirm the result. Therefore, it should be noted that the machine learning classification does not make a decision but rather provides information for making a decision.

Finally, the hetero-device method provides two crucial advantages. One is reducing the cost for the device, and another is unnecessity of preparing a reference library. The universal library is the biggest benefit of this method using an NIR device.

Conclusions

We proposed a classification method using hetero-spectroscopic devices between training data collection and test data collection. The hetero-device method is expected to encourage the use of portable spectroscopic devices for the field tests of medicines and any other products. In this study, we used screening machine learning classifiers of hetero-device classification for detecting falsified and substandard amoxicillin capsules. Following this process, hyperparameter optimization and elaboration of the predicted results would be needed to obtain more accurate classification.

Consequently, 7 classifiers were selected out of 13 algorithms, and the hyperparameters were optimized and predicted the quality of amoxicillin capsules by the hetero-device method. Additionally, the spectral preprocess critically operates the classification; thus, first and second derivations were conducted and tested as the preprocess. The first derivation gave higher accuracy in the classification than the second derivation. Subsequently, in the evaluation of the prediction performance of sensitivity, specificity, and precision, it was clearly shown that the classification trades off sensitivity against specificity. Finally, KNN, GBC, and ETC were found to give superior classification results for amoxicillin capsules. We emphasize that a close examination based on the prediction probability is necessary to obtain significant and accurate results.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s44211-022-00142-2.

Acknowledgements This study was supported by the Casio Science Promotion Foundation, by a Gakuin-tokubetsu kenkyu-hi Grant (Musashino University), and by the National Center for Global Health.
and Medicine, Japan (19A03). We thank the staff of the National Center for Global Health and Medicine (Tokyo, Japan) for collecting the samples in Asian and African countries.

Data availability The datasets generated during and analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

References

1. WHO, A study on the public health and socioeconomic impact of substandard and falsified medical products, https://www.who.int/publications/i/item/9789241513432
2. S. Ozawa, D.R. Evans, S. Bessias, D.G. Haynie, T.T. Yemeke, J. Pharm. Biomed. Anal. 15, 160 (2016)
3. A. Petersen, N. Held, L. Heide, on behalf of the D.-E.M.S. Group, Anal. Methods 7, 130, 122 (2014)
4. P. Tabernero, I. Swamidoss, M. Mayxay, M. Phonlavong, C. Vilayhong, S. Yeuchaixiong, C. Sichanh, S. Senling, M.P. Grobusch, Malar. J. 18, 403 (2019)
5. A. Guillemain, K. Dégardin, Y. Roggo, J. Pharm. Biomed. Anal. 74, 2417 (2019)
6. T. Kluyver, B. Ragan-Kelley, D. Avila, S. Abdalla, C. Willing, Jupyter development team, Jupyter Notebooks—a publishing format for reproducible computational workflows, ed. Fernando Loizides and Birgit Scmidt (IOS Press, 2016)
7. J.D. Hunter, Comput. Sci. Eng. 9, 90 (2007)
8. W. McKinney, Proc. 9th Python Sci. Conf. 56 (2010)
9. C.R. Harris, K.J. Millman, S.N.A. Amartey, G. Nti, M.P. Grobusch, Malar. J. 15, 160 (2016)
10. K. Dégardin, A. Guillemain, N.V. Guerreiro, Y. Roggo, J. Pharm. Biomed. Anal. 128, 89 (2016)
11. A. Guillemain, K. Dégardin, Y. Roggo, Talanta 165, 632 (2017)
12. H. Chen, Z. Lin, C. Tan, Anal. Chem. 90, 113514 (2020)
13. B.J. Visser, J. Meerveld-Gerrits, D. Kroon, J. Mougoula, R. Vingerling, E. Bache, J. Boersma, M. van Vugt, S.T. Agnandji, H. Kaur, M.P. Grobusch, Malar. J. 14, 273 (2015)
14. A. Petersen, N. Held, L. Heide, on behalf of the D.-E.M.S. Group, PLoS ONE 12, e0184165 (2017)
15. K.F.-M. Opuni, H. Nettey, M.A. Larbi, S.N.A. Amartey, G. Nti, A. Dzidoum, P. Owusu-Danso, N.A. Owusu, A.K. Nyarko, Malar. J. 18, 403 (2019)
16. S. Schäfermann, C. Hauk, E. Wemakor, R. Nci, G. Mutombo, E. Ngah Ndze, T. Cletus, F. Nyaah, M. Pattinora, D. Wistuba, I. Helmle, C. Häfele-Abah, H. Gross, L. Heide, Am. J. Trop. Med. Hyg. 103, 894 (2020)
17. M.T. Koesdjojo, Y. Wu, A. Boonloed, E.M. Dunfield, V.T. Remcho, Talanta 130, 122 (2014)
18. R.W.O. Jähnke, G. Küsters, K. Fleischer, Drug Inf. J. 35, 941 (2001)
19. B.J. Visser, J. Meerveld-Gerrits, D. Kroon, J. Mougoula, R. Vingerling, E. Bache, J. Boersma, M. van Vugt, S.T. Agnandji, H. Kaur, M.P. Grobusch, Malar. J. 14, 273 (2015)
20. A. Petersen, N. Held, L. Heide, on behalf of the D.-E.M.S. Group, PLoS ONE 12, e0184165 (2017)
21. K.F.-M. Opuni, H. Nettey, M.A. Larbi, S.N.A. Amartey, G. Nti, A. Dzidoum, P. Owusu-Danso, N.A. Owusu, A.K. Nyarko, Malar. J. 18, 403 (2019)
22. S. Schäfermann, C. Hauk, E. Wemakor, R. Nci, G. Mutombo, E. Ngah Ndze, T. Cletus, F. Nyaah, M. Pattinora, D. Wistuba, I. Helmle, C. Häfele-Abah, H. Gross, L. Heide, Am. J. Trop. Med. Hyg. 103, 894 (2020)
23. A.A. Weaver, H. Reiser, T. Barstis, M. Benvenuti, D. Ghosh, M. Hunckler, B. Joy, L. Koenig, K. Raddell, M. Lieberman, Anal. Chem. 85, 6453 (2013)
24. H.-H. Chen, C. Higgins, S.K. Laing, S.L. Biese, M. Lieberman, S. Ozawa, Med. Access Point Care 5, 2399202620980303 (2021)
25. Y. Hattori, Y. Hoshi, Y. Ichimura, Y. Sugiura, M. Otsuka, Appl. Spectrosc. 75, 1251 (2021)
26. G. Van Rossum, F.L. Drake, Python 3 Reference Manual (CreateSpace, Scotts Valley, CA, 2009)
27. T. Kluiver, B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. Hamrick, J. Grout, S. Corlay, P. Ivanov, D. Avila, S. Abdalla, C. Willing, Jupyter development team, Jupyter Notebooks—a publishing format for reproducible computational workflows, ed. Fernando Loizides and Birgit Scmidt (IOS Press, 2016)
28. J.D. Hunter, Comput. Sci. Eng. 9, 90 (2007)
29. W. McKinney, Proc. 9th Python Sci. Conf. 56 (2010)
30. C.R. Harris, K.J. Millman, S.N.A. Amartey, G. Nti, M.P. Grobusch, Malar. J. 15, 160 (2016)
31. P. Virtanen, R. Gommers, T. Oliphant, C.R. Harris, K.J. Millman, N. Mayorov, S.J. del Río, M. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, İ Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, Nat. Methods 17, 261 (2020)