

CASE REPORT

Open Access



# Case report: therapeutic monitoring of vancomycin in an acute liver failure patient with anuria under high-flow continuous hemodiafiltration

Yuriko Ito<sup>1\*</sup>, Junya Nakade<sup>1,2</sup>, Akihiro Seki<sup>3</sup>, Ryosuke Gabata<sup>4</sup>, Mitsuyoshi Okazaki<sup>4</sup>, Shinichi Nakanuma<sup>4</sup>, Arimi Fujita<sup>1,5</sup>, Tsutomu Shimada<sup>1,5</sup>, Taro Yamashita<sup>3</sup>, Shintaro Yagi<sup>4</sup>, Takumi Taniguchi<sup>6</sup> and Yoshimichi Sai<sup>1,5,7</sup>

## Abstract

**Background** High-flow continuous hemodiafiltration (HF-CHDF) combines diffusive and convective solute removal and is employed for artificial liver adjuvant therapy. However, there is no report on dosage planning of vancomycin (VCM) in patients with acute liver failure under HF-CHDF.

**Case presentation** A 20-year-old woman (154 cm tall, weighing 50 kg) was transferred to the intensive care unit (ICU) with acute liver failure associated with autoimmune liver disease. On the following day, HF-CHDF was started due to elevated plasma ammonia concentration. On ICU day 8, VCM was started for suspected pneumonia and meningitis (30 mg/kg loading dose, then 20 mg/kg every 12 hrs). However, on ICU day 10, VCM blood concentration was under the limit of detection ( $< 3.0 \mu\text{g/mL}$ ) and the patient developed anuria. The VCM dose was increased to 20 mg/kg every 6 hrs. Calculation with a one-compartment model using the HF-CHDF blood flow rate as a surrogate for VCM clearance, together with hematocrit and protein binding ratio, predicted a trough VCM blood concentration of  $15 \mu\text{g/mL}$ . The observed concentration was about  $12 \mu\text{g/mL}$ . The difference may represent non-HF-CHDF clearance. Finally, living donor liver transplantation was performed.

**Conclusion** We report an acute liver failure patient with anuria under HF-CHDF in whom VCM administration failed to produce an effective blood concentration, likely due to HF-CHDF-enhanced clearance. VCM dosage adjustment proved successful, and was confirmed by calculation using a one-compartment model.

**Keywords** High flow continuous hemodiafiltration, Anuric, Therapeutic drug monitoring, Vancomycin

\*Correspondence:

Yuriko Ito

yoda-yuriko818@staff.kanazawa-u.ac.jp

<sup>1</sup> Department of Hospital Pharmacy, University Hospital, Kanazawa University, 13-1 Takara-machi, Kanazawa, Ishikawa 920-8641, Japan

<sup>2</sup> Department of Infection Control and Prevention, University Hospital, Kanazawa University, 13-1 Takara-machi, Kanazawa, Ishikawa 920-8641, Japan

<sup>3</sup> Department of Gastroenterology, Graduate School of Medicine, Kanazawa University, 13-1 Takara-machi, Kanazawa, Ishikawa 920-8641, Japan

<sup>4</sup> Department of Hepato-Biliary-Pancreatic Surgery and Transplantation, Graduate School of Medical Science, Kanazawa University, 13-1 Takara-machi, Kanazawa, Ishikawa 920-8641, Japan

<sup>5</sup> Department of Clinical Pharmacokinetics, Graduate School of Medical Sciences, Kanazawa University, 13-1 Takara-machi, Kanazawa, Ishikawa 920-8641, Japan

<sup>6</sup> Intensive Care Unit, Kanazawa University Hospital, 13-1 Takara-machi, Kanazawa, Ishikawa 920-8641, Japan

<sup>7</sup> AI Hospital/Macro Signal Dynamics Research and Development Center, Kanazawa University, 13-1 Takara-machi, Kanazawa, Ishikawa 920-8641, Japan



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

High-flow continuous hemodiafiltration (HF-CHDF) is used as artificial liver adjuvant therapy for blood purification in acute liver failure [1–3], since it efficiently removes small-molecular compounds such as ammonia ( $\text{NH}_3$ ) and pathogenic cytokines, and promotes emergence from hepatic coma [1, 3]. However, coadministered therapeutic drugs with low molecular weight and low protein binding rate are also easily removed [4, 5], so it is necessary to ensure an appropriate administration dosage and schedule for patients receiving drug treatments.

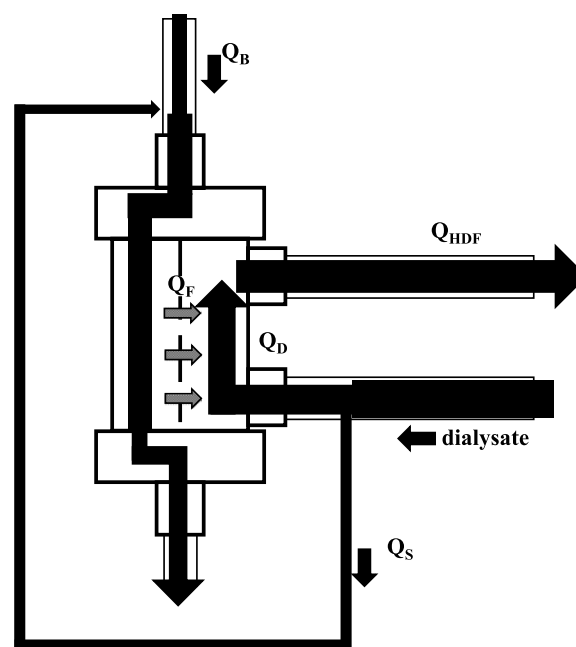
Vancomycin (VCM) is a glycopeptide antibiotic with activity against methicillin-resistant *Staphylococcus aureus* (MRSA). To ensure efficacy and to avoid adverse effects of VCM, therapeutic drug monitoring is critically important [6]. Furthermore, since VCM is removed by hemodiafiltration [7], individual dosage design is essential [8–10].

CHDF, used for renal replacement therapy, is a continuous dialysis method that reduces the blood and dialysate flow rates (generally, blood flow rate: about 80–100 mL/min > dialysis flow rate + filtration flow rate: about 10–25 mL/min) compared with normal dialysis. However, we experienced a case of vancomycin administration under HF-CHDF, involving continuous high-flow on-line hemodiafiltration (on-line HDF) (blood flow rate: 200 mL/min < dialysis flow rate + filtration flow rate: 600 mL/min) for 24 hrs. Although VCM clearance during 4 hrs of on-line HDF has been examined [11], there is no report on the dosage design of VCM during HF-CHDF.

Here, we report a one-compartment model developed to aid dosage planning of VCM in an acute liver failure patient with anuria who was treated with VCM while receiving HF-CHDF. The results of VCM monitoring are also presented.

## Case presentation

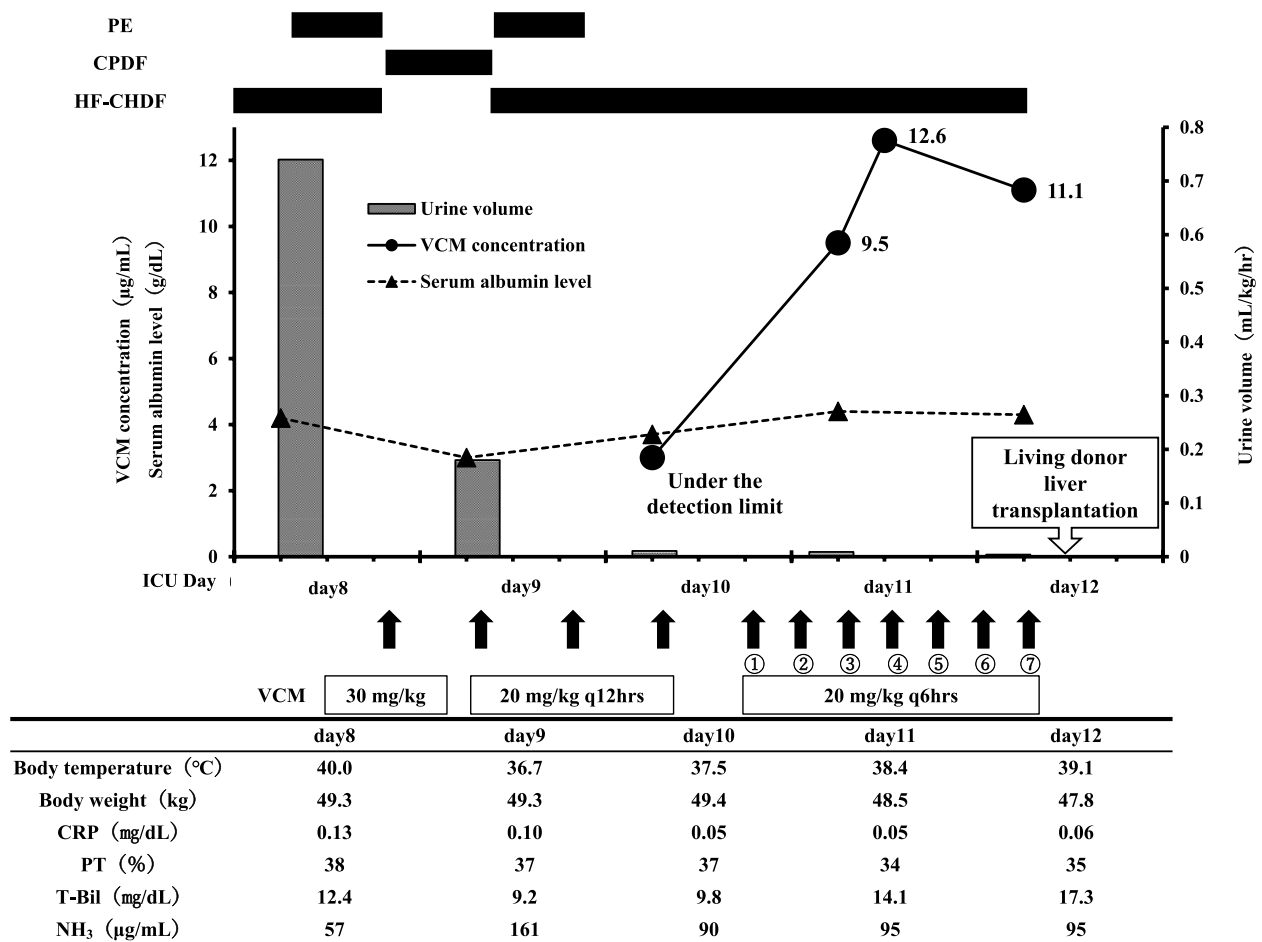
A 20-year-old woman (154 cm tall, weighing 50 kg), who had been under long-term prednisolone treatment for dermatomyositis and autoimmune hepatitis, was hospitalized with acute liver damage. On the 11th day of hospitalization, her prothrombin time (PT) activity was 25% and  $\text{NH}_3$  level was 113  $\mu\text{g/mL}$ . She was diagnosed with acute liver failure and transferred to the intensive care unit (ICU). Plasma exchange (PE) was conducted and steroid pulse therapy was started. On the following day (ICU day 2), HF-CHDF (Fig. 1: The system employs on-line HDF in a predilution mode) was started because the  $\text{NH}_3$  plasma concentration was elevated and she was diagnosed with coma II hepatic encephalopathy. Thereafter, HF-CHDF was mainly used in combination with PE and continuous plasma filtration with dialysis (CPDF; a



**Fig. 1** The HF-CHDF system used in this case. On-line hemodiafiltration in a predilution mode was performed continuously. Dialyzer: ABH-22PA (Asahi Kasei Medical Co., Ltd., Tokyo, Japan). Material: polysulfone membrane. Dialysate: Carbostar®-L (Yoshindo Inc., Toyama, Japan).  $Q_S$  = substitute fluid flow rate,  $Q_B$  = blood flow rate,  $Q_D$  = dialysis flow rate,  $Q_F$  = filtration flow rate,  $Q_{HDF}$  = dialysis outflow rate

combination of slow, continuous PE and hemodiafiltration [12]) to replenish coagulation factors and to control the  $\text{NH}_3$  level.

On ICU day 6, the patient developed fever at night, and piperacillin/tazobactam (PIPC/TAZ) treatment (4.5 g, every 6 hrs) was started for suspected ventilator-associated pneumonia (Culture result: [Supplemental Table](#)). On ICU day 7, cervical rigidity was observed, and PIPC/TAZ was changed to cefepime (CFPM) (2.0 g, every 12 hrs) for suspected meningitis. On ICU day 8, PE with HF-CHDF was changed to CPDF, and the antibiotic therapy was switched to VCM with meropenem (MEPM) (2.0 g, every 8 hrs) to achieve distribution to the cerebrospinal fluid (Fig. 2). At this time, accurate assessment of renal function was difficult because serum creatinine was removed by the dialysis, and the estimated glomerular filtration rate (eGFR) was more than 150 mL/min/1.73 m<sup>2</sup>. Since urine output was maintained (>0.5 mL/kg/hr), and renal function before admittance to ICU was preserved (eGFR > 100 mL/min/1.73 m<sup>2</sup>), a loading dose of 30 mg/kg and maintenance dose of 20 mg/kg VCM every 12 hrs were adopted, according to the TDM Guidelines for Antibiotics 2016 [6] (Japan, The Japanese Society of Chemotherapy and the Japanese Society of Therapeutic Drug Monitoring).



**Fig. 2** Clinical course and VCM concentrations from ICU day 8 to day 12. VCM: Vancomycin, PE: Plasma exchange, CPDF: Continuous plasma filtration with dialysis, HF-CHDF: High-flow continuous hemodiafiltration, CRP: C-reactive protein, PT: Prothrombin time, T-Bil: Total bilirubin, NH<sub>3</sub>: Ammonia

On ICU day 9, due to elevated NH<sub>3</sub> and prolonged disturbance of consciousness, dialysis was switched from CPDF to HF-CHDF again (flow rate: see Table 1), and PE was performed simultaneously for approximately 2hrs. Meningitis was ruled out by spinal fluid examination, but fever and high inflammatory response persisted. VCM administration was continued because we could not rule out ventilator-associated pneumonia or catheter-related bloodstream infection associated with gram-positive cocci. On ICU day 10, the initial TDM for VCM was performed, and the blood level was under the detection limit (<3.0 μg/mL), presumably due to the effects of HF-CHDF and PE. At the same time, the onset of anuria due to hepatorenal syndrome was noted. It was decided that a living donor liver transplantation (LDLT) would be performed 2 days later (ICU day 12), and HF-CHDF would be continuously performed until the day of LDLT. To maintain the VCM concentration in the therapeutic range prior to LDLT, the VCM dose was increased

**Table 1** HF-CHDF flow rate, VCM parameters, and patient factors

#### HF-CHDF flow rate

Blood flow rate( $Q_b$ )	200 mL/min
Dialysis flow rate( $Q_D$ )	433 mL/min
Filtration flow rate( $Q_F$ )	167 mL/min
Substitute fluid flow rate ( $Q_S$ )	167 mL/min
$Q_{HDF} = Q_D + Q_F$	600 mL/min

#### VCM parameters[13, 14]

Volume of distribution( $V_d$ )	0.7 L/kg
Protein binding rate( $f_p$ )	0.35
Protein unbound form( $f_u$ )	0.65

#### Patient factors

Body weight	50 kg
Hematocrit(Ht)	25%

to 20 mg/kg every 6 hrs from the evening of ICU day 10 (MEPM was changed to 1.5g every 6 hrs). The trough blood concentrations before the 3rd, and 4th (ICU day

11), and 7th (ICU day 12) VCM administrations were 9.5, 12.6, and 11.1 µg/mL, respectively. On ICU day 12, after the 7th administration of VCM, LDLT was conducted as scheduled and VCM administration was discontinued after the LDLT. No infection was apparent during VCM administration.

## Discussion

In this case, administration of VCM according to the TDM guideline resulted in VCM concentrations below the therapeutic range during HF-CHDF. There is no report on the dosage design of VCM under conditions of enhanced clearance, such as HF-CHDF. We therefore adjusted the maintenance dosing schedule of VCM, and subsequent calculation of VCM blood concentration using a one-compartment model with the parameters shown in Table 1 supported the appropriateness of the adjusted dosage.

## Clearance calculation

The total body clearance ( $CL_{tot}$ ) under dialysis is expressed as follows,

$$CL_{tot} = CL_R + CL_{NR} + CL_{HDF} \quad (1)$$

$CL_R$ :renal clearance,  $CL_{NR}$ :non-renal clearance,  $CL_{HDF}$ :HF-CHDF clearance.

Since the patient was anuric on ICU day 10 and VCM is excreted almost exclusively from the kidneys,  $CL_{tot}$  could be approximated by  $CL_{HDF}$  (Eq. 2).

$$CL_{tot} \approx CL_{HDF} \quad (2)$$

Clearance never exceeds the liquid inflow rate to dialyzer ( $Q_{B_{in}}$ ) or the dialysis outflow rate ( $Q_{HDF} = Q_F + Q_D$ ), and the slower flow is rate-determining in blood purification therapy [4, 15–17]. If  $Q_{B_{in}} < Q_D + Q_F$  and assuming complete removal of the drug from the blood, the maximum clearance is defined by  $Q_{B_{in}}$ .

$$CL = Q_{B_{in}} \quad (3)$$

Although  $Q_{B_{in}}$  appears to be the blood flow rate ( $Q_B$ ) + substitute fluid flow rate ( $Q_S$ ) in the pre-dilution mode, the drug concentration is decreased by dilution before dialysis. Therefore, corrected clearance should be considered, since dilution reduce the efficacy of solute removal. In a previous report, corrected clearance ( $CL'$ ) was defined [18] as follows:

$$CL' = CL \cdot Q_{BI}' / Q_{BI} \quad (4)$$

where  $Q_{BI}'$  is the blood flow rate before dilution, i.e.,  $Q_B$ , and  $Q_{BI}$  is blood flow rate after dilution, i.e.,  $Q_B + Q_S$ . Substituting Eq. (3) into Eq. (4) gives the following result for  $CL'$ :

$$CL' = Q_B \quad (5)$$

If the dialysis outflow rate ( $Q_{HDF} = Q_D + Q_F$ ) is greater than the blood flow rate ( $Q_B$ ), clearance is limited by  $Q_B$ . In this case,  $Q_{HDF}$  was greater than  $Q_B$ , so clearance is defined by  $Q_B$ . In fact, only unbound drug in plasma is eliminated, so actual  $CL_{HDF}$  can be expressed as follows [4].

$$CL_{HDF} = Q_B \cdot (1 - Ht/100) \cdot fu \cdot 60/1000 \text{ (L/hr)} \quad (6)$$

Substitution of  $Q_B = 200 \text{ mL/min}$  under this condition,  $Ht = 25\%$  in this patient and  $fu = 0.65$  from the literature value of VCM (Table 1) gave a calculated  $CL_{HDF}$  value of 5.9 L/hr. The rate of disappearance ( $k_e$ ) of VCM in this patient was calculated as  $0.17 \text{ hr}^{-1}$  based on  $V_d = 0.70 \text{ L/kg}$  from the literature (Table 1) according to the following equation.

$$k_e = CL_{tot}/V_d \approx CL_{HDF}/V_d \text{ (/hr)} \quad (7)$$

The  $t_{1/2}$  was calculated as 4.1 hrs using the formula  $t_{1/2} = \ln 2/k_e$ .

## Calculation of blood concentration

Since VCM is homogeneously distributed under steady-state conditions, a one-compartment model was applied. The steady-state blood concentration of the drug during intermittent infusion was approximated as follows,

$$R_0 = D/t_0 \quad (8)$$

$$C_{ss,max} = \frac{R_0}{k_e \cdot V_d} (1 - e^{-k_e \cdot t_0}) \left( \frac{1}{1 - e^{-k_e \cdot \tau}} \right) (\mu\text{g/mL}) \quad (9)$$

$$C_{ss,min} = C_{ss,max} \cdot e^{-k_e \cdot \tau} (\mu\text{g/mL}) \quad (10)$$

$R_0$ : dosing rate,  $D$ : dose,  $t_0$ : infusion time,  $\tau$ : dosing interval,  $C_{ss,max}$ : maximum blood concentration,  $C_{ss,min}$ : minimum blood concentration.

To achieve a sufficiently high trough concentration while avoiding adverse effects, it is necessary to increase the frequency of dosing rather than the dosage amount, considering the short half-life of VCM. When a dose of 20 mg/kg was administered every 6 hrs, the values of  $C_{ss,max}$  and  $C_{ss,min}$  were calculated as 41.6 µg/mL and 15.2 µg/mL, respectively, from eq. 9 and 10. These values are suitable for VCM treatment, and based on the half-life of 4.1 hr, a steady state would be reached after the 3rd to 4th administration.

## Comparison of measured and calculated values

The measured blood concentration was around 12 µg/mL at the 4th (ICU day 11) and 7th (ICU day 12) VCM



administrations after dose escalation to 20 mg/kg every 6 hrs at ICU day 10. This lies within the effective range for preventing infection before LDLT.

Nevertheless, this measured concentration (around 12 µg/mL) was 20% lower than the calculated steady-state concentration of around 15 µg/mL, presumably due to factors such as non-renal clearance and changes in protein binding rate.

The systemic clearance of VCM in healthy adults is about 100 mL/min, and the urinary excretion rate of unchanged drug is more than 90% [13]. However, VCM is slowly excreted via an unknown route in patients without renal function [19]. Indeed, non-renal clearance of vancomycin was suggested to amount to 1.05 L/hr/70 kg [20, 21]. Moreover, the protein binding rate influences VCM clearance [22]. We applied a protein binding rate of 0.35 based on literature values in our clearance calculation, but it remains possible that the protein binding fraction was different because of inadequate albumin synthesis due to liver failure [23], hyperbilirubinemia, and the effect of pre-dilution of blood flowing into the dialyzer [24], which may have increased the measured  $CL_{HDF}$ . All these factors might have contributed to a blood concentration lower than the calculated value. Thus, there is scope to increase the accuracy of the calculation of blood concentration by taking account of these factors.

In recent years, several academic societies in America (e.g., the Infectious Diseases Society of America) have recommended the use of AUC/MIC as an important biomarker for efficacy and safety evaluation [25], not just the trough concentration of VCM. In 2022, the Japanese TDM guideline was similarly revised, so the calculation of AUC should be employed in future work.

## Conclusion

We report an acute liver failure patient with anuria under HF-CHDF who was treated with VCM. Dosage adjustment was required, and success was confirmed by calculation of VCM blood concentration using a one-compartment model. This calculation employs the HF-CHDF flow rate as a surrogate for clearance. Nevertheless, the measured VCM concentration was 20% lower than the calculated value (15 µg/mL), suggesting that other factors, such as non-renal clearance and protein binding rate, will need to be taken into account to improve the prediction of VCM concentration in patients under HF-CHDF.

## Abbreviations

CFPM	Cefepime
CL'	Corrected clearance
CL <sub>tot</sub>	Total body clearance
CL <sub>NR</sub>	Non-renal clearance
CL <sub>R</sub>	Renal clearance

CPDF	Continuous plasma filtration with dialysis
CRP	C-reactive protein
C <sub>ss,max</sub>	Maximum blood concentration
C <sub>ss,min</sub>	Minimum blood concentration
D	Dose
eGFR	Estimated glomerular filtration rate
f <sub>p</sub>	Protein binding rate
f <sub>u</sub>	Protein unbound form
HF-CHDF	High-flow continuous hemodiafiltration
Ht	Hematocrit
ICU	Intensive care unit
k <sub>e</sub>	Rate of disappearance
LDLT	Living donor liver transplantation
MEPM	Meropenem
MRSA	Methicillin-resistant <i>Staphylococcus aureus</i>
NH <sub>3</sub>	Ammonia
on-line HDF	On-line hemodiafiltration
PE	Plasma exchange
PIPC/TAZ	Piperacillin/tazobactam
PT	prothrombin time
Q <sub>B</sub>	Blood flow rate
Q <sub>BI</sub>	Blood flow rate after dilution
Q <sub>BI</sub> '	Blood flow rate before dilution
Q <sub>B in</sub>	Liquid inflow rate to dialyzer
Q <sub>D</sub>	Dialysis flow rate
Q <sub>F</sub>	Filtration flow rate
Q <sub>HDF</sub>	Dialysis outflow rate
Q <sub>S</sub>	Substitute fluid flow rate
R <sub>0</sub>	Dosing rate
t <sub>0</sub>	Infusion time
τ	Dosing interval
T-Bil	Total bilirubin
VCM	Vancomycin
Vd	Volume of distribution

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40780-023-00283-0>.

### Additional file 1.

## Acknowledgements

Not applicable.

## Authors' contributions

YI designed the study, collected and analyzed data, and wrote the first draft of the manuscript. JN, AS, RG, MO, SN, AF, TY, SY and TT designed the study, collected and analyzed data, and contributed to the writing of the manuscript. TS and YS supervised the project and contributed to the writing of the manuscript. All authors reviewed the results and approved the final version of the manuscript.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Availability of data and materials

Not applicable.

## Declarations

### Ethics approval and consent to participate

Since this is a case report, ethical approval was deemed unnecessary by the Medical Ethics Committee of Kanazawa University. Verbal consent was obtained from the patient.

### Consent for publication

Not applicable.

# Competing interests

The authors have no conflicts of interest to declare.

Received: 15 December 2022 Accepted: 24 March 2023

Published online: 01 May 2023

# References

- Shinozaki H, Oda S, Abe R, et al. Blood purification in fulminant hepatic failure. *Contrib Nephrol*. 2010;166:64–72. <https://doi.org/10.1159/000314854>. Epub 2010 May 7
- Yoshida M, Inoue K, Sekiyama K, et al. Favorable effect of new liver support on survival of patients with fulminant hepatic failure. *Artif Organs*. 1996;20(11):1169–72. <https://doi.org/10.1111/j.1525-1594.1996.tb00657.x>
- Arata S, Tanaka K, Takayama K, et al. Treatment of hepatic encephalopathy by on-line hemodiafiltration: a case series study. *BMC Emerg Med*. 2010;10:10. <https://doi.org/10.1186/1471-227X-10-10>
- Pistolesi V, Morabito S, Mario F, et al. A guide to understanding antimicrobial drug dosing in critically ill patients on renal replacement therapy. *Antimicrob Agents Chemother*. 2019;63(8):e00583–19. <https://doi.org/10.1128/AAC.00583-19>. Print 2019 Aug
- Urata M, Narita Y, Fukunaga M, et al. Simple formula for predicting drug removal rates during hemodialysis. *Ther Apher Dial*. 2018;22(5):485–93. <https://doi.org/10.1111/1744-9987.12675>. Epub 2018 Jul 10
- Takesue Y, Omagari T, Okada K, et al. Vancomycin: TDM Guideline for Antibiotics, the Japanese Society of Chemotherapy and the Japanese Society of Therapeutic Drug Monitoring. Revised version. Tokyo: KYORINSHA; 2016. p. 35–58.
- Petejova N, Martinek A, Zahalkova J, et al. Vancomycin removal during low-flux and high-flux extended daily hemodialysis in critically ill septic patients. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*. 2012;156(4):342–7. <https://doi.org/10.5507/bp.2012.002>. Epub 2012 Jan 30
- Zelenitsky SA, Ariano RE, McCrae ML, et al. Initial vancomycin dosing protocol to achieve therapeutic serum concentrations in patients undergoing hemodialysis. *Clin Infect Dis*. 2012;55(4):527–33. <https://doi.org/10.1093/cid/cis458>. Epub 2012 May 9
- Frazee EN, Kuper PJ, Schramm GE, Larson SL, Kashani KB, Osmon DR, et al. Effect of continuous venovenous hemofiltration dose on achievement of adequate vancomycin trough concentrations. *Antimicrob Agents Chemother*. 2012;56(12):6181–5. <https://doi.org/10.1128/AAC.00459-12>. Epub 2012 Sep 17
- Tsuruyama M, Yamashina T, Tsuruta M, et al. Vancomycin pharmacokinetics in critically ill patients receiving continuous haemodiafiltration with a polyethyleneimine-coated polyacrylonitrile membrane. *J Clin Pharm Ther*. 2020;45(5):1143–8. <https://doi.org/10.1111/jcpt.13197>. Epub 2020 Jun 4
- Sombolos KI, Fragidis SK, Bamichas GI, et al. Subtherapeutic serum vancomycin concentration during on-line hemodiafiltration. *ASAIO J*. 2011;57(6):507–10. <https://doi.org/10.1097/MAT.0b013e3182306196>
- Komura T, Taniguchi T, Sakai Y, et al. Efficacy of continuous plasma dialysis therapy in critical patients with acute liver failure. *J Gastroenterol Hepatol*. 2014;29(4):782–6. <https://doi.org/10.1111/jgh.12440>
- Nakashima M, Katagiri K, Oguma T. Phase I studies on vancomycin hydrochloride for injection. *Chemotherapy*. 1992;40:210–24.
- Lewis P. Vancomycin area under the curve simplified. *Ther Drug Monit*. 2018;40(3):377–80. <https://doi.org/10.1097/FTD.0000000000000500>
- Yamamoto T, Yasuno N, Katada S, et al. Proposal of pharmacokinetically optimized dosage regimen of antibiotics in patients receiving continuous hemodiafiltration. *Antimicrob Agents Chemother*. 2011;55(12):5804–12. <https://doi.org/10.1128/AAC.01758-10>. Epub 2011 Sep 12
- Yamamoto T, Hisaka A, Suzuki H. Principle of dosage adjustment for patients receiving continuous renal replacement therapy (CRRT) based on the quantitative estimation of clearance by CRRT. *Jpn J Nephrol Pharmacother*. 2014;3(1):3–19.
- Akashita G, Hosaka Y, Noda T, et al. PK/PD analysis of biapenem in patients undergoing continuous hemodiafiltration. *J Pharm Health Care Sci*. 2015;1:31. <https://doi.org/10.1186/s40780-015-0031-6>. eCollection 2015
- Mineshima M, Hoshino T, Teraoka S, et al. Effect of dilution modality and filtration flow rate on solute removal characteristics in HF and HDF with high ultrafiltration. *Jpn J Artif Organs*. 1995;24(3):670–5.
- Thomson Micromedex. United States Pharmacopeial Convention: Drug Information for the Health Care Professional 27th edition, vol. I; 2007. p. 2868.
- Rotschafer JC, Crossley K, Zaske DE, et al. Pharmacokinetics of vancomycin: observation in 28 patients and dosage recommendations. *Antimicrob Agents Chemother*. 1982;22(3):391–4. <https://doi.org/10.1128/AAC.22.3.391>
- Emoto C, Johnson TN, McPhail BM, et al. Using a vancomycin PBPK model in special populations to elucidate case-based clinical PK observations. *CPT Pharmacometrics Syst Pharmacol*. 2018;7(4):237–50. <https://doi.org/10.1002/psp4.12279>. Epub 2018 Feb 15
- Harada H, Miyagawa S, Kawasaki S, et al. Study of the pharmacokinetics of vancomycin in patients with impaired liver function. *J Infect Chemother*. 1999;5(2):104–7. <https://doi.org/10.1007/s101560050018>
- Oetli K, Stauber RE. Physiological and pathological changes in the redox state of human serum albumin critically influence its binding properties. *Br J Pharmacol*. 2007;151(5):580–90. <https://doi.org/10.1038/sj.bjp.0707251>. Epub 2007 Apr 30
- Maheshwari V, Thijssen S, Tao X, et al. In silico comparison of protein-bound uremic toxin removal by hemodialysis, hemodiafiltration, membrane adsorption, and binding competition. *Sci Rep*. 2019;9(1):909. <https://doi.org/10.1038/s41598-018-37195-1>
- Rybak MJ, Le J, Lodise TP, et al. Therapeutic monitoring of vancomycin for serious methicillin-resistant *Staphylococcus aureus* infections: a revised consensus guideline and review by the American Society of Health-System Pharmacists, the Infectious Diseases Society of America, the Pediatric Infectious Diseases Society, and the Society of Infectious Diseases Pharmacists. *Am J Health Syst Pharm*. 2020;77(11):835–64. <https://doi.org/10.1093/ajhp/zxaa036>

# Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)



## MEDICAL IMAGING

# Preoperative assessment of pleural adhesions in patients with lung cancer based on quantitative motion analysis with dynamic chest radiography: A retrospective study

Rie Tanaka<sup>1</sup> | Isao Matsumoto<sup>2</sup> | Tetsuya Takayama<sup>3</sup> | Noriyuki Ohkura<sup>4</sup> | Dai Inoue<sup>5</sup>

<sup>1</sup>College of Medical, Pharmaceutical & Health Sciences, Kanazawa University, Kanazawa, Japan

<sup>2</sup>Department of Thoracic Surgery, Kanazawa University, Kanazawa, Japan

<sup>3</sup>Department of Thoracic Surgery, Kanazawa University, Kanazawa, Japan

<sup>4</sup>Department of Respiratory Medicine, Kanazawa University Hospital, Kanazawa, Japan

<sup>5</sup>Department of Radiology, Kanazawa University Hospital, Kanazawa, Japan

## Correspondence

Rie Tanaka, College of Medical, Pharmaceutical & Health Sciences, Kanazawa University, 5-11-80 Kodatsuno, Ishikawa, Kanazawa 920-0942 Japan.  
Email: rie44@mhs.mp.kanazawa-u.ac.jp

Isao Matsumoto, Department of Thoracic Surgery, Kanazawa University, Kanazawa 920-8641, Japan.  
Email: isa-mat@med.kanazawa-u.ac.jp

## Abstract

**Purpose:** Preoperative assessment of pleural adhesion is crucial for appropriate surgical planning. This study aimed to quantitatively evaluate the usefulness of motion analysis using dynamic chest radiography (DCR) for assessing pleural adhesions.

**Methods:** Sequential chest radiographs of 146 lung cancer patients with or without pleural adhesions ( $n = 25/121$ ) were obtained using a DCR system during respiration (registration number: 1729). The local motion vector was measured, and the percentage of poor motion area to the maximum expiration lung area (%lung area with poor motion) was calculated. Subsequently, percentage values  $\geq 49.0\%$  were considered to indicate pleural adhesions. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated to assess the prediction performance. The percentage of lung area with poor motion was compared between patients with and without pleural adhesions ( $p < 0.05$ ).

**Results:** DCR-based motion analysis correctly predicted pleural adhesions in 21 out of 25 patients, with 47 false-positive results (sensitivity, 84.0%; specificity, 61.2%; PPV, 30.9%; NPV, 94.9%). The lung with pleural adhesions showed a significantly greater %lung area with poor motion than the opposite lung in the same patient, similar to the cancerous lung in patients without pleural adhesions.

**Conclusion:** On DCR-based motion analysis, pleural adhesions could be indicated by an increase in the percentage of lung area with poor motion. Although the proposed method cannot identify the exact location of pleural adhesions, information regarding the presence or absence of pleural adhesions provided by DCR would help surgeons prepare for challenging surgeries and obtain informed consent from patients.

## KEYWORDS

digital radiography, lung cancer, motion analysis, pleural adhesions

## 1 | INTRODUCTION

Video-assisted thoracoscopic surgery, a less invasive and well-established approach for the surgical resection

of lung cancer,<sup>1–3</sup> is now commonly used as the first choice for many cases of lung cancer resection. However, the presence of pleural adhesions may increase the risk of lung injury from a trocar or scope, prolong the

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Journal of Applied Clinical Medical Physics* published by Wiley Periodicals, LLC on behalf of The American Association of Physicists in Medicine.

operation time, and lead to conversion to thoracotomy in severe cases.<sup>4</sup> Therefore, the preoperative assessment of pleural adhesions is crucial for planning appropriate operative approaches.

Recent advances in medical imaging technologies have allowed the preoperative prediction of pleural adhesions in patients with lung cancer based on findings from four-dimensional (4D) computed tomography (CT),<sup>5–8</sup> cine magnetic resonance imaging (MRI),<sup>9–11</sup> and ultrasonography.<sup>12–14</sup> These approaches assess pleural invasion and adhesion based on tumor motion during respiration relative to the adjacent lung structures. A lung tumor moving independently from the parietal or mediastinal pleura suggests that it does not invade or adhere strongly to the chest wall or pleura. Yamashiro et al.<sup>7</sup> demonstrated that 4DCT had perfect diagnostic accuracy for pleural invasion/adhesion (Sensitivity: 100%, Specificity: 100%) compared to conventional chest CT (Sensitivity: 60%, Specificity: 77%). Similarly, Sakai et al.<sup>9</sup> reported that the sensitivity, specificity, and accuracy of breathing dynamic cine MR in detecting chest wall invasion were 100%, 70%, and 76%, while those of conventional CT and MRI were 80%, 65%, and 68%, respectively. Regarding US approaches, a pooled analysis of 10 articles showed sensitivity and specificity of 71% and 96%, respectively.<sup>12</sup> Despite the effectiveness of these approaches for preoperative prediction of pleural invasion and adhesion of lung tumors, they have some issues, such as the low availability of MRI and 4DCT and a limited field of view for ultrasonography. To address some of these limitations, dynamic chest radiography (DCR) has been proposed as a potential approach to assess pleural adhesions.

DCR is a flat-panel detector (FPD)-based functional X-ray imaging technique that provides 15 sequential chest radiographs per second.<sup>15</sup> Improvements in FPD technology have enabled quasi-real-time observation of both lungs with extremely low-dose imaging; the total entrance surface dose can be less than the limit of two projections (1.9 mGy) recommended by the International Atomic Energy Agency.<sup>16,17</sup> Recent preliminary studies have demonstrated that DCR can detect pleural invasion or adhesion based on image findings,<sup>18,19</sup> such as localized reduction and distortion of lung motion, with a sensitivity of 88.0% and specificity of 83.5% at an average reading time of 20.4 min per case.<sup>20</sup> However, the subjective assessment of moving lung structures on DCR requires knowledge of lung respiratory dynamics and extensive reading experience, as well as a long reading time. Thus, to allow the utilization of the DCR as a routine adhesion assessment, a qualitative analysis method for lung motion is essential. Although computer-based approaches, such as optical flow and motion tracking techniques,<sup>21,22</sup> have been proposed, their diagnostic performance remains unevaluated in patients with cancer and pleural adhesions. This retrospective study aimed to quantitatively evaluate the

usefulness of motion analysis for the assessment of pleural adhesions in patients with lung cancer with DCR. We conducted a study to evaluate the feasibility of identifying pleural adhesions by comparing quantified lung motion between affected and unaffected lungs.

## 2 | METHODS

### 2.1 | Participants

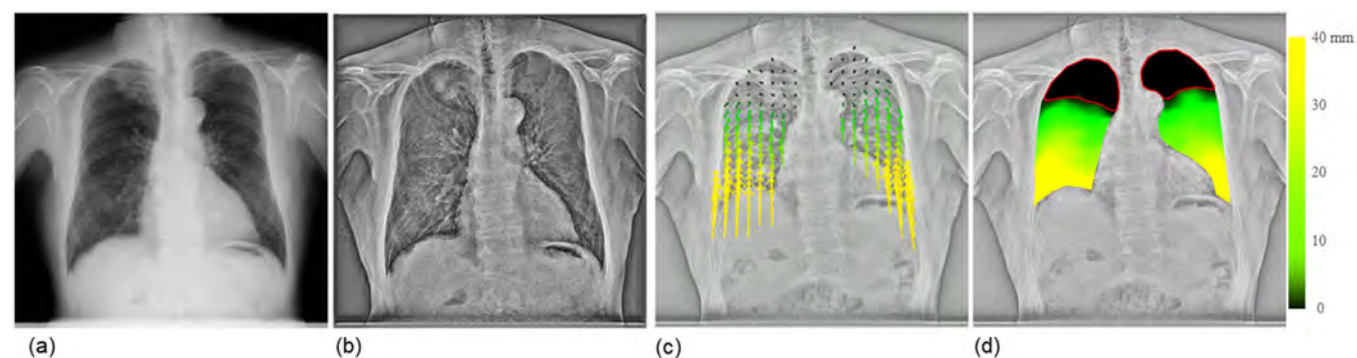
This study was a secondary analysis of the same clinical trial data from another study using a completely different analytical approach<sup>20</sup> and had the same trial registration number as the primary trial (registration number: 1729). While the previous study was conducted using a subjective approach, this study used an objective approach. Therefore, the analytical methods and conclusions presented herein do not duplicate those of the previous report. All study participants provided informed consent, and the study design was approved by the institutional review board.

A total of 148 consecutive patients with lung cancer who visited Kanazawa University hospital between February 2016 and July 2019 were initially identified for this retrospective study. The study included patients who underwent both DCR and conventional chest CT as part of their routine clinical care before lung cancer surgery and those who underwent surgical removal of lung cancer. We excluded patients whose lung lesions were confirmed to be mesothelial tumors ( $n = 1$ ) and those with an interval of more than 3 months between DCR and surgery ( $n = 1$ ). Eventually, 146 preoperative patients (age, 40–88 years; mean age, 70 years; male-to-female ratio, 99:47) were assessed in this study. The presence or absence of pleural adhesions and invasion was retrospectively determined based on surgical videos, which were used as the ground truth. The findings were further classified into five grades according to the extent of invasion or adhesions on the lung surface: grade 0, no adhesions; grade 1, less than one-third of the lung surface, no surgical disturbance; grade 2, one-third to two-third of the lung surface; grade 3, two-third to  $< 1$ ; and grade 4, adhesions on the entire lung surface. Pleural adhesions covering one-third or more of the lung surface area may require longer operation times because of the need to separate adhesions. Therefore, 146 patients were further classified into invasion or adhesion-negative (grades 0–1:  $n = 121$ ) and invasion or adhesion-positive (grades 2–4:  $n = 25$ ) groups. The patient characteristics are summarized in Table 1. Twenty-nine patients had tumors in the peripheral lung and/or adjacent to the chest wall (Grades 0–1:  $n = 26$ , Grades 2–4:  $n = 3$ ), and none of the three patients with tumors and pleural adhesions had findings suggestive of pleural adhesions on conventional CT.



**TABLE 1** Patient characteristics.

			Adhesion negative	Adhesion positive
Number			121	25
Age (years)			70.0 ± 9.5	71.0 ± 7.9
Sex (female/male)			44/77	3/22
Affected side	Right lung/left lung		81/39	17/8
	Others		1	0
Adhesion grade	Adhesion negative	Grade 0	92	
		Grade 1	29	
	Adhesion positive	Grade 2		13
		Grade 3		8
		Grade 4		4



**FIGURE 1** An example of a true-negative case correctly predicted as not showing pleural adhesion by dynamic chest radiography (DCR)-based quantitative motion analysis. Frame of the original DCR image (a), bone-suppressed DCR with enhanced high spatial frequency (b), vector-map images (c), and displacement-map image (d) of a 72-year-old man with right lung cancer (grade 0). The patient showed poor motion in 27.5% and 28.5% of the right and left lung areas, respectively, a predicted VC of 129.6%, predicted FEV1 of 90.7%, FEV1/FVC ratio of 52.4, and predicted DLco of 52.4%. DLco, diffusing capacity of the lung for carbon monoxide; FEV1, forced expiratory volume in 1 s; FEV1/FVC, forced expiratory volume in 1 s/forced vital capacity; VC, vital capacity.

## 2.2 | Imaging procedures

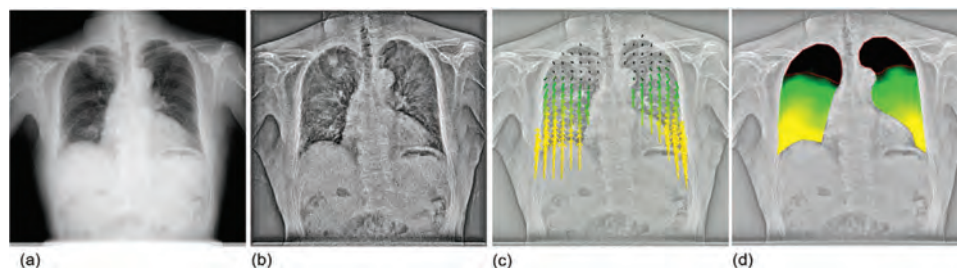
Sequential chest radiographs were obtained using a dynamic digital radiography system (Test Model; Konica Minolta, Tokyo, Japan) consisting of an indirect conversion FPD (PaxScan, 4343CB; Varex Imaging Corp., Utah, USA) and an X-ray generator and tube (DHF-155H II/UH-6QC-07E; Hitachi Healthcare, Ltd., Tokyo, Japan). Imaging was performed during forced breathing in the standing position and a posteroanterior direction, using pulsed irradiation of 15 frames/s (100 kV, 0.2 mAs/pulse, source-to-detector distance = 2.0 m).<sup>15</sup> An automatic voice system was used for patient instruction to include one full respiratory cycle of forced breathing in approximately 14 s of imaging time (Figure 1a, Video 1a). The matrix and pixel sizes were 1024 × 1024 pixels and 417 × 417 μm<sup>2</sup>, respectively, with a 16-bit grayscale. After selectively suppressing the shadows of bone structures in the lung region,<sup>23,24</sup> the high spatial frequency image components were enhanced to achieve selective motion tracking of lung structures, such as bronchi and blood vessels in the lungs (Figure 1b, Video 1b).

The enhancement of high-spatial frequency components was conducted based on the multiple resolution decomposition techniques; multiple images with different non-sharpness were subtracted from the original image to decompose into multiple images with different frequency components.<sup>25</sup> The images were then selectively combined so that the vascular shadows were emphasized.

## 2.3 | Motion analysis of lung structures

Sequential images with the diaphragm moving upward were selected to measure the motion vector of lung structures in the expiratory phase. The Dense Optical Flow by Gunnar FarneBack technique<sup>26</sup> was used to estimate pixel-by-pixel motion (optical flow) between two consecutive frames in each local lung area on bone-suppressed images. Optical flow was calculated as an array of flow vectors in terms of magnitude and direction. The total amount of movement was calculated by connecting the movement between two





**VIDEO 1** An example of a true-negative case (72-year-old man with grade 0 right lung cancer). The video shows the original dynamic chest radiography (DCR) images (a), bone-suppressed DCR with enhanced high spatial frequency (b), vector-map images (c), where the arrow length represents the movement from the first frame, and a displacement-map image (d), where larger motion is indicated in yellow, gradually changing to green, while the lung area with poor motion is indicated in black.

consecutive frames. An open-source computer vision library (OpenCV 3.4.3) was used. The resulting output is a two-dimensional vector field, where each vector is a displacement vector that shows the movement of points from a reference frame to the current frame. In this study, the reference frame was set to the maximum inspiratory frame, that is, the first frame in the sequential images of the expiratory phase.

## 2.4 | Visualization of the motion vector

To facilitate visual evaluation, the motion vector was visualized with arrows, indicating the orientation and displacement from the maximum inspiratory frame, as shown in Figure 1c and Video 1c (hereafter, vector-map images). Pleural adhesions were observed as locally restricted and/or distorted motion findings in the lung structures.<sup>20</sup> Therefore, we defined the local lung area with a displacement less than the threshold value as local lung areas with poor motion. In this study, the threshold value was set to 1.5 mm to account for errors in the motion analysis. The threshold value was determined on the basis of the receiver operating characteristic (ROC) analysis, where the area under the curve (AUC) was maximized with stable diagnostic performance. In addition, to account for the displacement attributable to the upward movement of the diaphragm during expiration, the horizontal component of the motion vector was ignored, and the ascending component was extracted. The downward component was treated as a negative value. The arrows were also expressed on a color scale, where the lung area with an upward movement was indicated by yellow (>40 mm), gradually changing to green (>10 mm), and the lung area with a displacement less than the threshold value was indicated by black (<1.5 mm). In addition, the local lung displacement was mapped onto the maximum expiratory frame according to the same color scaling, as shown in Figure 1d and Video 1d (hereafter, a displacement-map image). The local area of the lung with poor motion was surrounded by a red line on a displacement-map image.

## 2.5 | Calculation of the lung area with poor motion

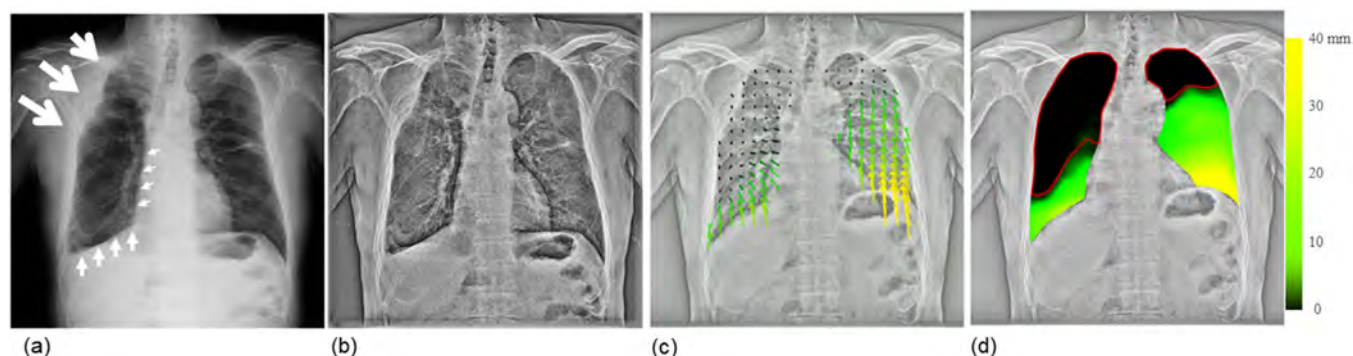
The percentage of the local lung area with a poor motion to the maximum expiration lung area (%lung area with poor motion) was used as an index to determine the presence or absence of pleural adhesions. Although the shape and volume of the left and right lungs are different, evaluation based on the percentage of the lung area with a poor motion can eliminate the difference. The %lung area with poor motion was calculated using the following formula:

$$\begin{aligned} &\% \text{lung area with poor motion} \\ &= \frac{\text{The local lung area with poor motion}}{\text{The maximum expiration lung area}} \end{aligned}$$

If the %lung area with poor motion was greater than the cut-off value, the patient was considered to have pleural adhesions. The sensitivity should be rather high to avoid false negatives in patients with severe adhesions. Thus, the cut-off value was determined using the Youden index<sup>27,28</sup> to have practical sensitivity and specificity values for the detection of pleural adhesions. In this study, we identified pleural adhesions, which are problematic in surgery, as a reduction in the “% area of lung in motion” compared to the contralateral lung of the same patient or the cancerous lung of a patient without pleural adhesions.

## 2.6 | Statistical analysis

Statistical analysis of the groups with or without pleural adhesions was performed using R software (Ver 4.1.3<sup>29</sup>) with the Wilcoxon signed-rank sum test to determine whether the %lung area with poor motion differed significantly between the lung with cancer and the opposite lung. In contrast, the Mann–Whitney *U* test was used to test whether the differences between patients with and without pleural adhesions were significant. Statistical significance was set at  $p < 0.05$ . The



**FIGURE 2** An example of a true-positive case correctly evaluated as showing pleural adhesion by dynamic chest radiography (DCR)-based quantitative motion analysis. Frame of the original DCR image (a), bone-suppressed DCR with enhanced high spatial frequency (b), vector-map images (c), and displacement-map image (d) of a 69-year-old man with right lung cancer (grade 4 pleural adhesion over the right lung). The patient showed poor motion in 69.9% and 27.3% of the right and left lung areas, respectively, with a predicted VC of 73.9%, predicted FEV1 of 84.9%, FEV1/FVC ratio of 84.2, and predicted DLco of 55.3%. White arrows indicate areas with pleural invasion or adhesions; large, middle, and small arrows indicate invasion and strong and moderate adhesions, respectively. DLco, diffusing capacity of the lung for carbon monoxide; FEV1, forced expiratory volume in 1 s; FEV1/FVC, forced expiratory volume in 1 s/forced vital capacity; VC, vital capacity.

diagnostic performance of DCR-based motion analysis was evaluated in terms of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).

### 3 | RESULTS

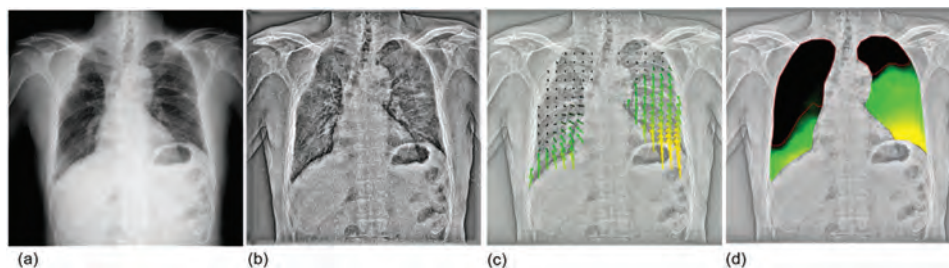
Vector analysis of DCR images quantified the respiratory motion of the lung structures. Figure 1 shows the results in a patient without pleural adhesions (a 72-year-old man with grade 0 right lung cancer). The resulting vector map quantitatively clarified the orientation and magnitude of the movement (Figure 1c, Video 1c). In general, patients without pleural adhesions showed right-left symmetry and a top-to-down gradually decreasing distribution of local vectors, resulting in a displacement-map image, as shown in Figure 1d. This patient had 27.5% and 28.5% lung areas with poor motion in the right and left lungs, respectively, and was correctly predicted to have no pleural adhesions. Figure 2 shows the results for a patient with pleural adhesions (a 69-year-old man with grade 4 right lung cancer). In this case, the lung cancer invaded the second to fifth ribs. In addition, the middle and lower lobes strongly adhered to the mediastinal and diaphragmatic surfaces. The resulting vector map indicated a right-left asymmetric and widely distorted distribution of motion vectors, as shown in Figure 2c and Video 2c. The right lung with pleural adhesions indicated a larger poor motion area than did the left lung in a displacement-map image, as shown in Figure 2d. This patient had 69.9% and 27.3% of lung areas with poor motion in the right and left lungs, respectively, and was correctly predicted to have pleural adhesions in the right lung.

**TABLE 2** Diagnostic performance of dynamic chest radiography-based quantitative motion analysis in each adhesion grade.

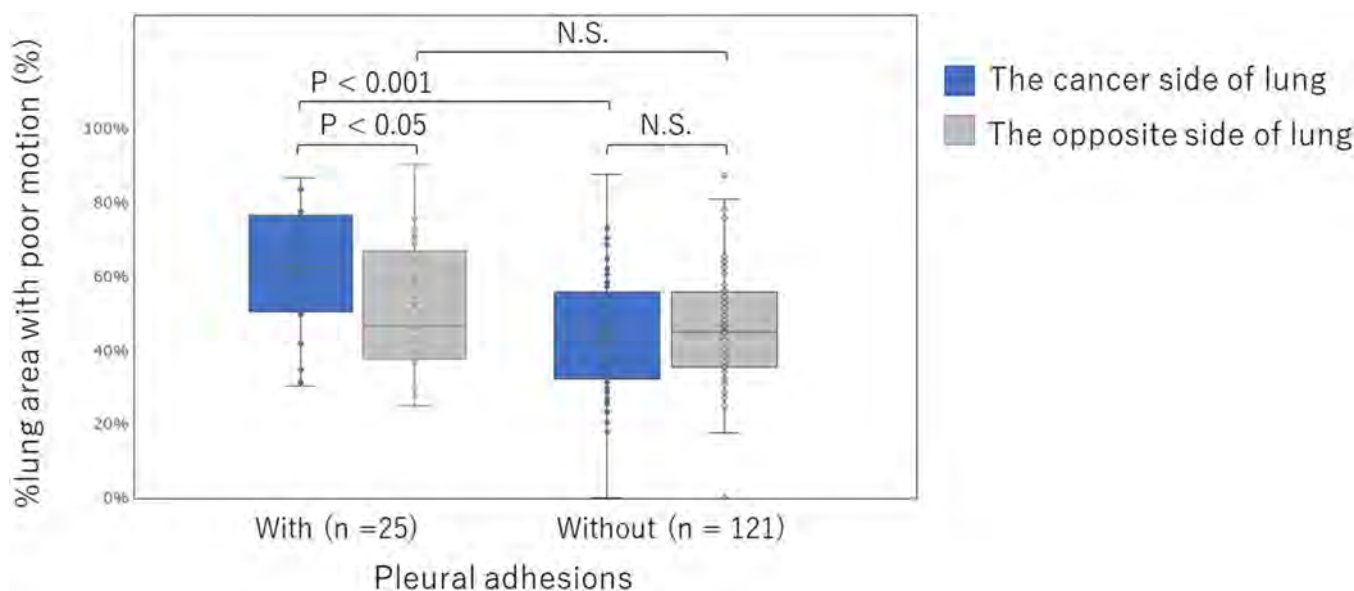
Adhesion grade	Performance	Results of motion analysis	
		Positive	Negative
Adhesion negative	Specificity		
Total ( $n = 121$ )	61.2% (74/121)	47	74
Grade 0	59.8% (55/92)	37	55
Grade 1	65.5% (19/29)	10	19
Adhesion positive	Sensitivity		
Total ( $n = 25$ )	84.0% (21/25)	21	4
Grade 2	92.3% (12/13)	12	1
Grade 3	62.5% (5/8)	5	3
Grade 4	100.0% (4/4)	4	0

In patients with pleural adhesions ( $n = 25$ ), the %lung area with a poor motion in the lung with cancer was significantly greater than that in the opposite lung ( $p < 0.05$ ) (Figure 3). In contrast, in patients without pleural adhesions ( $n = 121$ ), the %lung area with a poor motion did not differ significantly between the lungs with cancer and the opposite lung. In addition, the %lung area with poor motion on the cancerous side of the lung differed significantly between patients with and without pleural adhesions ( $p < 0.001$ ) (Figure 3). However, the findings for the opposite lung did not show significant differences between the patients with and without pleural adhesions.

Figure 4 shows the ROC curve for predicting pleural adhesions using the DCR-based quantitative motion analysis. The AUC was 0.793 and the cut-off value was 49.0% based on the Youden index. Table 2 shows the performance of DCR-based quantitative motion



**VIDEO 2** An example of a true-positive case (a 69-year-old man with right lung cancer and grade 4 pleural adhesion). The video shows the original dynamic chest radiography (DCR) images (a), bone-suppressed DCR with enhanced high spatial frequency (b), vector-map images (c), where the arrow length represents the movement from the first frame, and a displacement-map image (d), where larger motion is indicated in yellow, gradually changing to green, while the lung area with poor motion is indicated in black.



**FIGURE 3** Box-and-whisker plots showing the percentage of lung area with poor motion to the maximum expiration lung area (%lung area with poor motion) on the lung with cancer and the opposite lung in patients with/without pleural adhesions. The error bar indicates the maximum value in the upper part and the minimum value in the lower part.

analysis in pleural adhesion assessment for each adhesion grade. The present method correctly identified 21 out of 25 patients with pleural adhesions, including three patients with peripheral lung cancer who had no findings suggestive of pleural adhesions on conventional CT. The performance of DCR in detecting pleural adhesions had a sensitivity of 84.0%, specificity of 61.2%, PPV of 30.9%, and NPV of 94.9%.

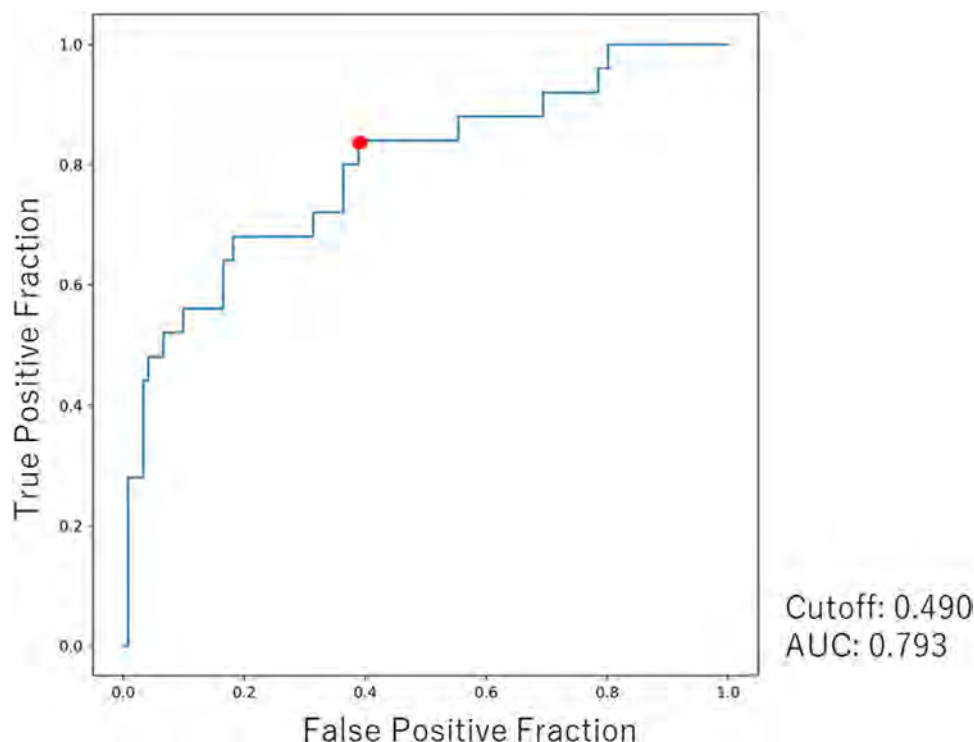
Figure 5 and Video 3 show a false-negative example of a 71-year-old man incorrectly predicted to have no pleural adhesions. In this case, the upper lobe (S2) slightly adhered to the dorsal and lateral thoracic surfaces. The middle and lower lobes strongly adhered to the diaphragmatic plane. Otherwise, the lateral thoracic region was free, resulting in only 34.4% of the area on the lung with cancer showing poor motion. Figure 6 and Video 4 show a false-positive example of a 78-year-old woman with incorrectly predicted pleural adhesions. This patient had emphysema with thickening of the

bronchial wall and insufficient respiration during imaging, resulting in 64.3% of the lung area with cancer showing poor motion.

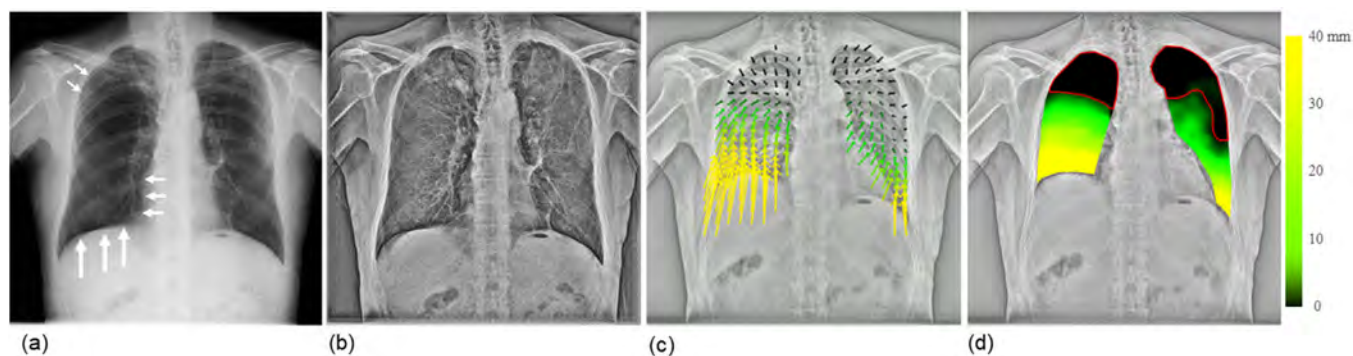
## 4 | DISCUSSION

We analyzed the DCR findings of 146 patients with or without pleural adhesions ( $n = 25/121$ ) and confirmed that poor lung motion due to pleural adhesions could be quantified by DCR-based motion analysis; lungs with pleural adhesions showed a significant increase in the %lung area with a poor motion than did the opposite lung in the same patient, similar to the findings for the lung with cancer in patients without pleural adhesions. DCR-based motion analysis correctly identified 21 out of 25 patients with pleural adhesions based on an increase in %lung area with poor motion and yielded a specificity of 61.2% (74/121).





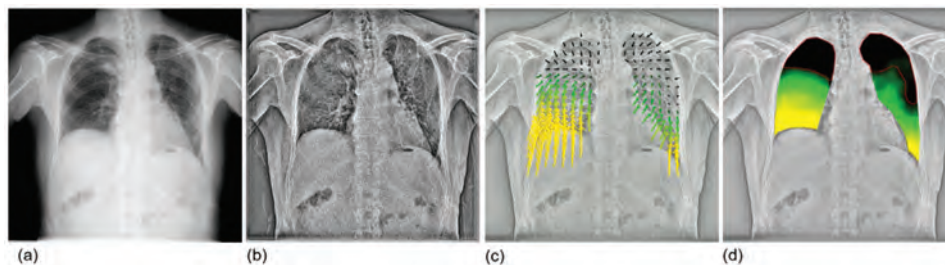
**FIGURE 4** Receiver operating characteristic (ROC) curve for predicting pleural adhesions by DCR-based quantitative motion analysis. The area under curve and cutoff value are 0.793 and 0.490, respectively.



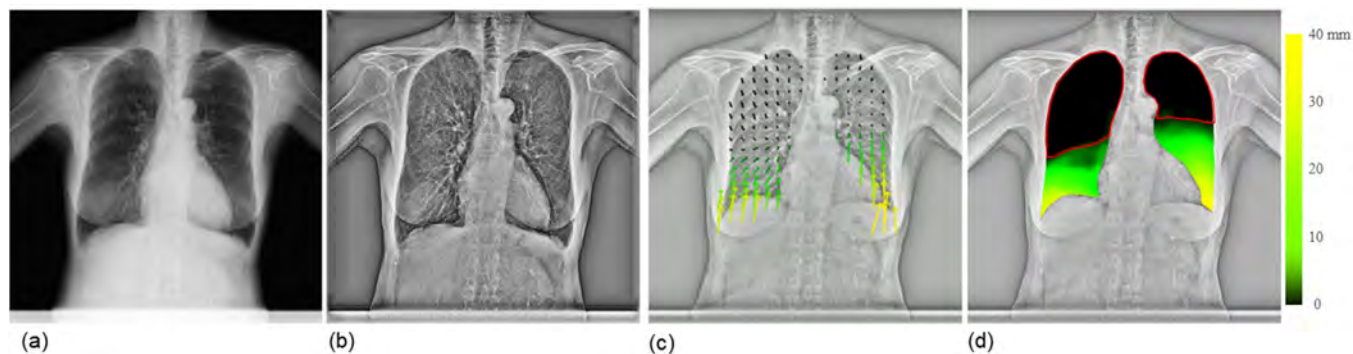
**FIGURE 5** An example of a false-negative case. Frame of the original dynamic chest radiography (DCR) images (a), bone-suppressed DCR with enhanced high spatial frequency (b), vector-map images (c), and displacement-map image (d) of a 71-year-old man with right lung cancer (grade 2 pleural adhesion of the entire diaphragmatic surface). The patient showed poor motion in 34.4% and 45.9% of the right and left lung areas, respectively, a predicted VC of 90.5%, predicted FEV1 of 78.1%, FEV1/FVC ratio of 66.6, and predicted DLco of 41.6%. White arrows indicate areas with pleural adhesions; large, middle, and small arrows indicate strong, moderate, and mild adhesions, respectively. DLco, diffusing capacity of the lung for carbon monoxide; FEV1, forced expiratory volume in 1 s; FEV1/FVC, forced expiratory volume in 1 s/forced vital capacity; VC, vital capacity.

Previous studies have reported that lung structures show restricted and/or distorted motion in DCR due to restricted lung sliding in patients with pleural adhesions and smooth and left-right symmetrical motions according to respiration in patients without adhesions.<sup>20</sup> This study showed no significant difference in the %lung area with poor motion between the left and right sides in patients without pleural adhesions. In contrast, in patients with pleural adhesions, the %lung

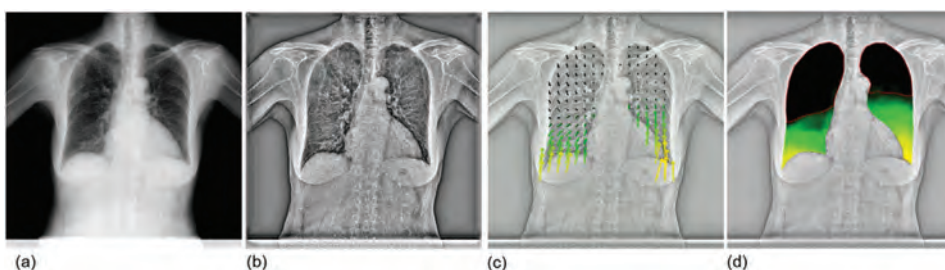
area with a poor motion was significantly greater in the lungs with pleural adhesions than in those without ( $p < 0.05$ ). In a comparison of lungs with cancer, the %lung area with a poor motion in patients with pleural adhesions was significantly greater than in those without ( $p < 0.001$ ). These results suggest that pleural adhesions could be predicted as an increase in the %lung area with poor motion on the DCR motion analysis.



**VIDEO 3** An example of a false-negative case (a 71-year-old man with right lung cancer and grade 2 pleural adhesion of the entire diaphragmatic surface). The video shows the original dynamic chest radiography (DCR) images (a), bone-suppressed DCR with enhanced high spatial frequency (b), vector-map images (c), where the arrow length represents the movement from the first frame, and a displacement-map image (d), where larger motion is indicated in yellow, gradually changing to green, while the lung area with poor motion is indicated in black.



**FIGURE 6** An example of a false-positive case. Frame of the original dynamic chest radiography (DCR) images (a), bone-suppressed DCR with enhanced high spatial frequency (b), vector-map images (c), and displacement-map image (d) of a 78-year-old female with right lung cancer (grade 0). The patient showed poor motion in 64.3% and 49.1% of the area of the right and left lungs, respectively; the predicted VC was 107.2%, predicted FEV1 was 124.0%, FEV1/FVC ratio was 69.2, and predicted DLco was 54.4%. DLco, diffusing capacity of the lung for carbon monoxide; FEV1, forced expiratory volume in 1 s; FEV1/FVC, forced expiratory volume in 1 s/forced vital capacity; VC, vital capacity.



**VIDEO 4** An example of a false-positive case (a 78-year-old woman with grade 0 right lung cancer). The video shows the original dynamic chest radiography (DCR) images (a), bone-suppressed DCR with enhanced high spatial frequency (b), vector-map images (c), where the arrow length represents the movement from the first frame, and a displacement-map image (d), where greater motion is indicated in yellow, gradually changing to green, while the lung area with poor motion is indicated in black.

Nevertheless, despite the use of optimal parameters in the DCR motion analyses, 4 and 47 cases showed false-negative and false-positive results, respectively. False-negative diagnoses may occur because the %lung area with a poor motion did not increase when pleural adhesions were localized just above the diaphragm or at the lung apex. In cases with pleural adhesions over the entire diaphragm, as in Figure 5d, pleural adhesions may appear to be absent because the two structures move together. In contrast, insuffi-

cient respiration and/or underlying pulmonary disease can make pleural adhesion assessments based on lung motion challenging, which was also observed in radiologists' readings.<sup>20</sup> As shown in Figure 6, patients with chronic obstructive pulmonary disease and interstitial pneumonitis associated with decreased pulmonary function are likely to show the restricted motion of the entire lung. DCR-based motion analysis showed one additional false-negative case compared to radiologists' readings, indicating that the performance of



DCR-based motion analysis in detecting pleural adhesions (sensitivity, 84.0%; specificity, 61.2%; PPV, 30.9%; NPV, 94.9%) was lower than that of radiologists' readings (sensitivity, 88.0%; specificity, 83.5%; PPV, 52.4%; NPV, 97.12%).<sup>20</sup> This performance degradation may have occurred because our computational algorithm detected pleural adhesions as an increase in %lung area with a poor motion, making it challenging to identify localized abnormal motion.

In relation to these findings, our study had several limitations. First, the %lung area with poor motion does not directly represent the location of adhesions. Thus, lung motion is reduced around adhesions, resulting in an increase in the %lung area with a poor motion in the lung. Therefore, the development of adhesion detection algorithms that focus on motion analysis of each local lung area is essential. The sensitivity of the proposed method to detect pleural adhesions could be improved by calculating the %lung area with poor motion for each lung area. In addition, comprehensive image interpretation combining image findings and quantified lung motion, such as displacement-map images and the %lung area with a poor motion, is recommended to predict the location of pleural adhesions by DCR. Combining DCR with conventional CT or MRI may provide another potential solution for more accurate assessment of pleural adhesions in patients with lung cancer. Second, an insufficient inspired volume or underlying disease can cause poor motion of lung structures, even in the absence of adhesions. Thus, detection sensitivity should be increased by repeatedly practicing breathing techniques before imaging. Third, cases with the same degree of adhesions on both sides might show no difference in the %lung area with a poor motion in both lungs. If the %lung area with a poor motion is greater than that of the normal lung, it might be picked up aggressively. This is because, false positives, along with false negatives caused by moderate adhesions, are rarely surgically problematic. Thus, the sensitivity should be rather high to avoid false negatives in patients with severe adhesions, such as in patients undergoing lower lobectomies with severe diaphragmatic adhesions that might lead to an unexpected prolongation of the operative time. For clinical implementation, however, additional algorithms must be developed to reduce the increased false positives, while maintaining performance levels similar to radiologists' readings.

Computer-based motion analysis offers several advantages over a radiologist's reading. A previous study based on subjective evaluations<sup>20</sup> reported three discordant cases in image interpretation of 146 cases between the two radiologists. With computerized approaches based on the %lung area with a poor motion, the evaluation is reproducible with no discordant cases. DCR can also detect local pulmonary impairments based on reduced respiration-induced changes in lung density.<sup>15</sup> Thus, computerized analysis

of lung motion combined with changes in lung density has the potential to provide new evaluation metrics for pleural adhesions. Furthermore, analyzing the motion of the left and right lung based on imaging has the potential to improve the accuracy of pulmonary assessment. In addition, DCR-based motion analysis was able to correctly identify pleural adhesions in peripheral lung cancers, whereas conventional CT did not show findings suggestive of pleural adhesions. The DCR system became commercially available in 2018 and was approved by the U.S. Food and Drug Administration in 2019. DCR can be performed as an additional examination in conventional radiography protocols. Although the patients were exposed to radiation during the examination, the total irradiation was approximately twice that of a conventional chest radiograph. Thus, DCR is acceptable because of its high information yield and rapid and simple imaging procedures in daily clinical settings. The present study demonstrated that DCR-based motion analysis correctly identified 21 out of 25 patients with pleural adhesions based on an increase in the %lung area with poor motion and yielded a specificity of 61.2% (74/121). Although some issues need to be addressed in its clinical implementation, the feasibility of DCR-based quantitative motion analysis for the preoperative evaluation of pleural adhesions was ascertained in this study.

## 5 | CONCLUSION

This retrospective study revealed that pleural adhesions could be predicted by DCR-based quantitative motion analysis as an increase in the %lung area with a poor motion. Although the proposed method cannot identify the exact location of pleural adhesions, it may be useful to determine the presence or absence of surgically problematic pleural adhesions by comparing them between lungs. It may also help surgeons prepare for difficult surgeries and obtain informed consent from the patients. Further development of computerized methods to locally assess lung motion will pave the way for the routine preoperative evaluation of pleural adhesions using DCR.

## AUTHOR CONTRIBUTIONS

All authors made substantial contributions to all of the following: (1) the conception and design of the study, or acquisition of data, or analysis and interpretation of data, (2) drafting the article or revising it critically for important intellectual content, (3) final approval of the version to be submitted.

## ACKNOWLEDGMENTS

The authors sincerely thank Noritsugu Matsutani from Konica Minolta, Inc., Staffs in Department of Thoracic Surgery and Radiology at Kanazawa University Hospital for the data acquisitions and technical support during these studies.

## CONFLICT OF INTEREST STATEMENT

This research was partially supported by Konica Minolta, Inc. (Tokyo, Japan). Konica Minolta had no role in the design, analyses, or reporting of this study.

## REFERENCES

- McKenna RH, Houck W, Fuller CB. Video-assisted thoracic surgery lobectomy: experience with 1,100 cases. *Ann Thorac Surg*. 2006;81:421-425.
- Committee for Scientific Affairs, The Japanese Association for Thoracic Surgery, Shimizu H, Okada M, Toh Y, et al, Committee for Scientific Affairs, The Japanese Association for Thoracic Surgery. Thoracic and cardiovascular surgeries in Japan during 2018: annual report by the Japanese Association for Thoracic Surgery. *Gen Thorac Cardiovasc Surg*. 2021;69:179-212.
- Yan TD, Black D, Bannon PG, McCaughan BC. Systematic review and meta-analysis of randomized and nonrandomized trials on safety and efficacy of video-assisted thoracic surgery lobectomy for early-stage non-small-cell lung cancer. *J Clin Oncol*. 2009;27:2553-2562.
- Marty-Ane CH, Canaud L, Solovei L, Alric P, Berthet JP. Video-assisted thoracoscopic lobectomy: an unavoidable trend? A retrospective single-institution series of 410 cases. *Interact Cardiovasc Thorac*. 2013;17:36-43.
- Sakuma K, Yamashiro T, Moriya H, Murayama S, Ito H. Parietal pleural invasion/adhesion of subpleural lung cancer: quantitative 4-dimensional CT analysis using dynamic-ventilatory scanning. *Eur J Radiol*. 2017;87:36-44.
- Yamashiro T, Moriya H, Tsubakimoto M, Nagatani Y, Kimoto T, Murayama S, investigators of ACTIve study group. Preoperative assessment of parietal pleural invasion/adhesion of subpleural lung cancer: advantage of software-assisted analysis of 4-dimensional dynamic-ventilation computed tomography. *Eur Radiol*. 2019;29:5247-5252.
- Nagatani Y, Hashimoto M, Oshio Y, et al, investigators of ACTIve study group. Preoperative assessment of localized pleural adhesion: utility of software-assisted analysis on dynamic-ventilation computed tomography. *Eur J Radiol*. 2020;133:109347.
- Hashimoto M, Nagatani Y, Oshio Y, et al, investigators of ACTIve study group. Preoperative assessment of pleural adhesion by four-dimensional ultra-low-dose computed tomography (4D-ULDCT) with adaptive iterative dose reduction using three-dimensional processing (AIDR-3D). *Eur J Radiol*. 2018;98:179-186.
- Sakai S, Murayama S, Murakami J, Hashiguchi N, Masuda K. Bronchogenic carcinoma invasion of the chest wall: evaluation with dynamic cine MRI during breathing. *J Comput Assist Tomogr*. 1997;21:595-600.
- Shiotani S, Sugimura K, Sugihara M, et al. Diagnosis of chest wall invasion by lung cancer: useful criteria for exclusion of the possibility of chest wall invasion with MR imaging. *Radiat Med*. 2000;18:283-290.
- Akata S, Kajiwara N, Park J, et al. Evaluation of chest wall invasion by lung cancer using respiratory dynamic MRI. *J Med Imaging Radiat Oncol*. 2008;52:36-39.
- Cassabelli N, Caroli G, Dolci G, et al. Accuracy of transthoracic ultrasound for the dictation of pleural adhesions. *Eur J Cardiothorac Surg*. 2012;42:813-818.
- Wei B, Wang T, Jiang F, Wang H. Use of transthoracic ultrasound to predict pleural adhesions: a prospective blinded study. *Thorac Cardiovasc Surg*. 2012;60:101-814.
- Bandi V, Lunn W, Ernst A, Eberhardt R, Hoffmann H, Herth F. Ultrasound vs. CT in detecting chest wall invasion by tumor: a prospective study. *Chest*. 2007;133:881886.
- Tanaka R. Dynamic chest radiography: flat-panel detector (FPD) based functional X-ray imaging. *Radiol Phys Technol*. 2016;9:139-153.
- International Basic Safety Standards for Protection Against Ionizing Radiation and for the Safety of Radiation Sources. *IAEA Safety Series No. 115*. Vienna: International Atomic Energy Agency (IAEA); 1996;279.
- Annex IV, Schedule III. Guidance levels of dose, dose rate and activity for medical exposure. *Radiological Protection for Medical Exposure to Ionizing Radiation. Safety Guide. IAEA Safety Standards Series No. SSG-46*. Vienna: International Atomic Energy Agency (IAEA); 2018.
- Tamura M, Matsumoto I, Saito D, et al. Dynamic chest radiography: novel and less-invasive imaging approach for preoperative assessments of pleural invasion and adhesion. *Radiol Case Rep*. 2020;15:702-704.
- Tamura M, Matsumoto I, Saito D, Yoshida S, Takata M, Takemura H. Case report: uniportal video-assisted thoracoscopic resection of a solitary fibrous tumor preoperatively predicted visceral pleura origin using dynamic chest radiography. *J Cardiothorac Surg*. 2020;15:166.
- Tanaka R, Inoue D, Izumozaki A, Takata M, Tamura M, Matsumoto I. Preoperative evaluation of pleural adhesions with dynamic chest radiography: a retrospective study of 146 patients with lung cancer. *Clin Radiol*. 2022;77:e689-e696.
- Tanaka R, Sanada S, Suzuki M, Kobayashi T, Matsui T, Inoue H. Breathing chest radiography using a dynamic flat-panel detector combined with computer analysis. *Med Phys*. 2004;31:2254-2262.
- Tanaka R, Samei E, Segars WP, et al. Assessment of pleural invasion and adhesion of lung tumors with dynamic chest radiography: a virtual clinical imaging study. *Med Phys*. 2021;48:1616-1623.
- Miyoshi T, Yoshida J, Aramaki N, et al. Effectiveness of bone suppression imaging in the detection of lung nodules on chest radiographs: relevance to anatomic location and observer's experience. *J Thorac Imaging*. 2017;32:398-405.
- Kodama N, Loc TV, Hai PT, et al. Effectiveness of bone suppression imaging in the diagnosis of tuberculosis from chest radiographs in Vietnam: an observer study. *Clin Imaging*. 2018;51:196-201.
- Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans Pattern Anal Mach Intell*. 1989;674-693.
- Farneback G. Two-frame motion estimation based on polynomial expansion. In: Bigun J, Gustavsson T, eds. *Image Analysis. SCIA 2003. Lecture Notes in Computer Science*. Springer. [https://docs.opencv.org/3.4.3/de/d9e/classcv\\_1\\_1FarnebackOpticalFlow.html](https://docs.opencv.org/3.4.3/de/d9e/classcv_1_1FarnebackOpticalFlow.html)
- Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3:32-35.
- Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. *Biom J*. 2005;47:458-472.
- R Development Core Team (2020). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*; 2020.

**How to cite this article:** Tanaka R, Matsumoto I, Takayama T, Ohkura N, Inoue D. Preoperative assessment of pleural adhesions in patients with lung cancer based on quantitative motion analysis with dynamic chest radiography: A retrospective study. *J Appl Clin Med Phys*. 2023;24:e14036. <https://doi.org/10.1002/acm2.14036>



# Deep learning-based cardiothoracic ratio measurement on chest radiograph: accuracy improvement without self-annotation

Kotaro Yoshida<sup>1^</sup>, Atsushi Takamatsu<sup>1</sup>, Takashi Matsubara<sup>1</sup>, Taichi Kitagawa<sup>1</sup>, Fomihito Toshima<sup>1</sup>, Rie Tanaka<sup>2</sup>, Toshifumi Gabata<sup>1</sup>

<sup>1</sup>Department of Radiology, Kanazawa University Graduate School of Medical Sciences, Kanazawa, Japan; <sup>2</sup>College of Medical, Pharmaceutical & Health Sciences, Kanazawa University, Kanazawa, Japan

**Contributions:** (I) Conception and design: K Yoshida, A Takamatsu; (II) Administrative support: R Tanaka, T Gabata; (III) Provision of study materials or patients: K Yoshida, A Takamatsu; (IV) Collection and assembly of data: K Yoshida, A Takamatsu; (V) Data analysis and interpretation: K Yoshida, A Takamatsu, T Matsubara, T Kitagawa, R Tanaka; (VI) Manuscript writing: All authors; (VII) Final approval of the manuscript: All authors.

**Correspondence to:** Kotaro Yoshida, MD, PhD. Department of Radiology, Kanazawa University Graduate School of Medical Sciences, 13-1 Takaramachi, Kanazawa, Ishikawa 920-8641, Japan. Email: kotaro\_yoshida@staff.kanazawa-u.ac.jp.

**Background:** A reproducible and accurate automated approach to measuring cardiothoracic ratio on chest radiographs is warranted. This study aimed to develop a deep learning-based model for estimating the cardiothoracic ratio on chest radiographs without requiring self-annotation and to compare its results with those of manual measurements.

**Methods:** The U-net architecture was designed to segment the right and left lungs and the cardiac shadow, from chest radiographs. The cardiothoracic ratio was then calculated using these labels by a mathematical algorithm. The initial model of deep learning-based cardiothoracic ratio measurement was developed using open-source 247 chest radiographs that had already been annotated. The advanced model was developed using a training dataset of 729 original chest radiographs, the labels of which were generated by the initial model and then screened. The cardiothoracic ratio of the two models was estimated in an independent test set of 120 original cases, and the results were compared to those obtained through manual measurement by four radiologists and the image-reading reports.

**Results:** The means and standard deviations of the cardiothoracic ratio were 52.4% and 9.8% for the initial model, 51.0% and 9.3% for the advanced model, and 49.8% and 9.4% for the total of four manual measurements, respectively. The intraclass correlation coefficients (ICCs) of the cardiothoracic ratio ranged from 0.91 to 0.93 between the advanced model and the manual measurements, whereas those for the initial model and the manual measurements ranged from 0.77 to 0.82.

**Conclusions:** Deep learning-based cardiothoracic ratio estimation on chest radiographs correlated favorably with the results obtained through manual measurements by radiologists. When the model was trained on additional local images generated by the initial model, the correlation with manual measurement improved even more than the initial model alone.

**Keywords:** Deep learning; cardiothoracic ratio measurement; chest radiograph

Submitted Feb 15, 2023. Accepted for publication Jul 21, 2023. Published online Sep 13, 2023.

doi: 10.21037/qims-23-187

**View this article at:** <https://dx.doi.org/10.21037/qims-23-187>

<sup>^</sup> ORCID: 0000-0002-8398-0969.

## Introduction

The cardiothoracic ratio is a value calculated from a chest radiograph and is used to assess the size of the heart relative to the chest cavity (1). This ratio can vary based on factors such as the individual's body size, imaging conditions, and degree of inspiration during the radiograph. The presence of cardiomegaly, or an enlarged heart, on a chest radiograph can be used to screen for various diseases such as cardiac conditions and mediastinal lesions (2,3). In countries where hemodialysis is readily available, the cardiothoracic ratio is also used as an indicator to maintain the appropriate dry weight in dialysis patients (4).

In current clinical practice, the cardiothoracic ratio is manually measured and documented by the radiologist or attending physician. However, this approach often results in inconsistencies in inter-reader and intra-reader agreement (5). To address this problem, various automatic measurement methods for the cardiothoracic ratio have been developed. However, most of those methods have not been widely adopted, and there are still several issues that need to be addressed to make these methods suitable for use in daily practice, such as accuracy and versatility issues (6). Recently, there has been increased interest in developing an automatic measurement model using deep learning (5). While it is anticipated that deep learning-based models will be highly accurate, such models require a large number of training datasets with annotation or labeling to increase their accuracy (7). While there has been a demand for the development of models that minimize the effort required to create training datasets, there have been few reports on this topic to date (8). In this study, we developed a deep learning-based model using open-source training data, and then expanded the training data with local data to create a more advanced model. The purpose of this study was to develop a deep learning-based model for estimating the cardiothoracic ratio using U-net, without the need for additional manual annotation. We also compared the results of these models with those of manual measurements.

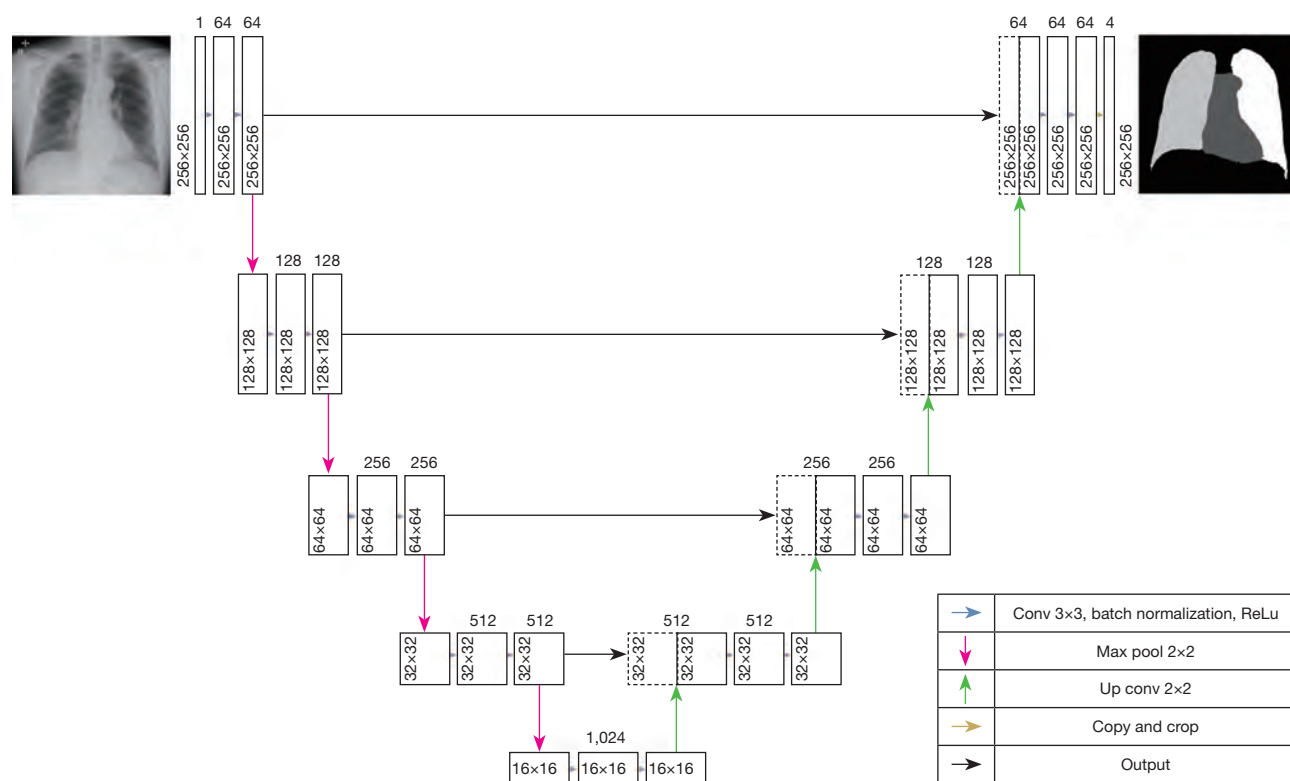
## Methods

This retrospective study was conducted at Kanazawa University Hospital and was approved by the institutional review board of this hospital. The requirement for informed consent was waived due to the retrospective nature of the study. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Chest

radiographs of 247 cases from open-source datasets and 1,495 cases from Kanazawa University Hospital were used to create a training dataset for the development of a deep learning-based model. The 247 open-source cases were obtained from the miniJSRT\_database provided by the Japanese Society of Radiological Technology (9). The images consist of the original image in the BMP file format with a matrix size of 256×256, and a label image in PNG file format with pixel values 255 for lung, 85 for mediastinum, 170 for non-lung area, and 0 for the area outside the body. The images of 1,495 chest radiographs collected from our institutional image database were acquired between January 2020 and December 2020 with the cardiothoracic ratio noted in the imaging report. The 120 images in the test set included 30 randomly selected patients in each of four groups with cardiothoracic ratios of <40%, 40–49%, 50–59%, and ≥60% based on the cardiothoracic ratio in imaging reports from January to June 2021.

Two types of deep learning-based models were developed based on the U-net architecture. U-net is a popular architecture for image segmentation that consists of U-shaped encoder-decoder network (10). Our architecture of U-net is shown in *Figure 1*. The detailed parameters for our U-net model are shown in *Table S1*. The same U-net architecture was used for both the initial and advanced models. All U-net processing was performed on a Windows 10 (×64) workstation with two NVIDIA GeForce 2080Ti GPUs with 24 GB memory and were implemented with Python (version 3.8.8) and PyTorch (version 1.1.0). We trained the models using the Adam optimizer with a minibatch size of 64 for 1,000 epochs and a learning rate of 0.0001. The image data were converted into a BMP file with a matrix size of 256×256 for the input image, and the output image was a PNG file with a matrix size of 256×256 with pixel values 255 for left lung, 170 for right lung, 85 for mediastinum, and 0 for the area outside the body. The initial model was created using the 247 open-source images with segmentation information of the right lung, left lung, and heart. To develop the advanced model, additional training data were prepared by segmentation labels on new images and were screened. In total, 1,495 original images were segmented using the initial model. The images with contour information of segmentation for both lungs and the heart were visually evaluated by two certified radiologists (with 17- and 14-year experience) to select an appropriate training dataset. The images were classified into three groups: excellent, good, and fair. Excellent was defined





**Figure 1** A U-net architecture for segmentation of lung and heart from a chest radiograph.

as good segmentation of the heart and lungs, suitable for measuring the cardiothoracic ratio; good was defined as some defective segmentation of the heart and lungs, which may have a minor effect on measuring the cardiothoracic ratio; and fair was defined as poor segmentation of the heart and lungs, which was not suitable for measuring the cardiothoracic ratio. A flowchart of model development, cardiothoracic ratio measurement output, and an example of segmentation is shown in [Figures S1-S3](#). The cases that were deemed excellent by both two readers were used as a new training dataset for the advanced model. Ultimately, 729 chest radiographs were included in the training dataset for the advanced model. With this training dataset, we developed an advanced deep learning-based model with semantic segmentation using U-net, in the same manner as the initial model.

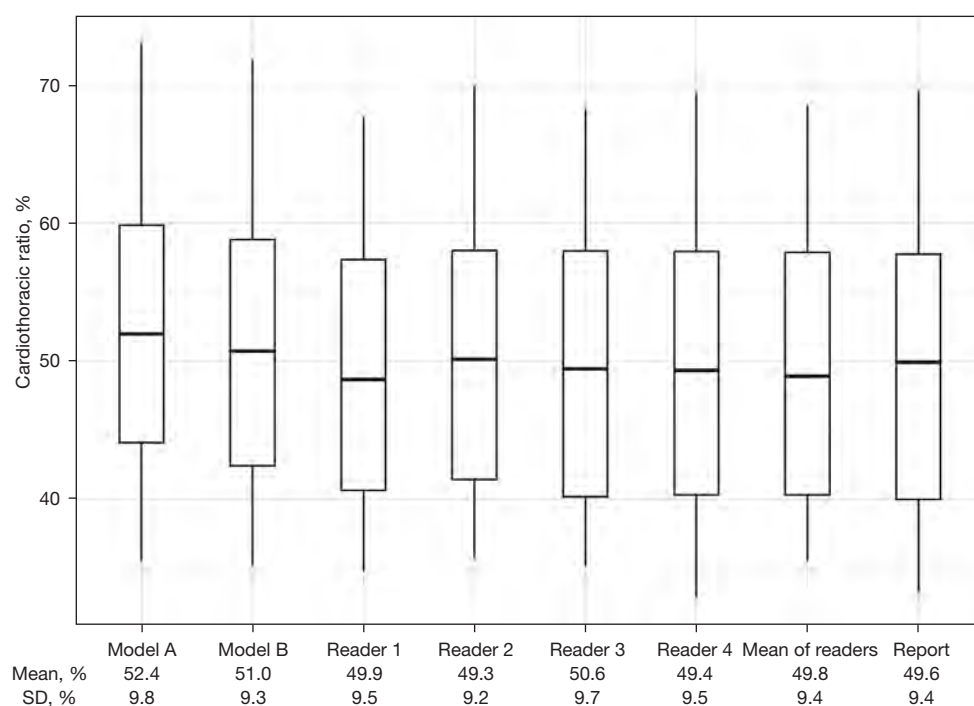
The cardiothoracic ratio was calculated by dividing the transverse axis length of the right margin of the segmented right lung and the left margin of the left lung (representing the thorax) by the transverse axis length of the right and left margins of the heart (representing the heart). Image segmentation and subsequent sequential cardiothoracic

ratio measurements were performed on the independent test dataset using the two models, and cardiothoracic ratios were output for each of the 120 test cases.

The reference for the cardiothoracic ratio was measured by four radiologists (with 18, 15, 5 and 3 years of experience), and the description of cardiothoracic ratio in the radiology report was also used as a reference. The measurement of the cardiothoracic ratio was performed using a medical image viewer (EV Insite, PSP Co., Tokyo, Japan). The results of automatic measurement of the cardiothoracic ratio on the test dataset by the initial and advanced models were compared with the results of manual measurement.

The cardiothoracic ratios estimated by two deep learning-based models were compared by the intraclass correlation coefficient (ICC), and the errors between the two models and the radiologist's manual measurements were examined by a Bland-Altman plot. R version 4.0.3 was used for the graphing and statistical processing. To evaluate the segmentation performance of the two models, we calculated the Dice similarity coefficient (DSC) between each model and the manual segmented data by a radiologist.





**Figure 2** The cardiothoracic ratio in the deep learning-based models and the references in the test set. The box plot shows the medians and first and third quartiles in boxes, as well as the minimum and maximum values of the cardiothoracic ratio for the deep-learning models and references. Model A, cardiothoracic ratio estimated by the initial deep-learning model; Model B, cardiothoracic ratio estimated by the advanced deep-learning model; Reader 1-4, cardiothoracic ratio measured by four radiologists; Report, cardiothoracic ratio measured by radiology reports; SD, standard deviation.

We randomly selected 40 patients from the test set for this evaluation.

## Results

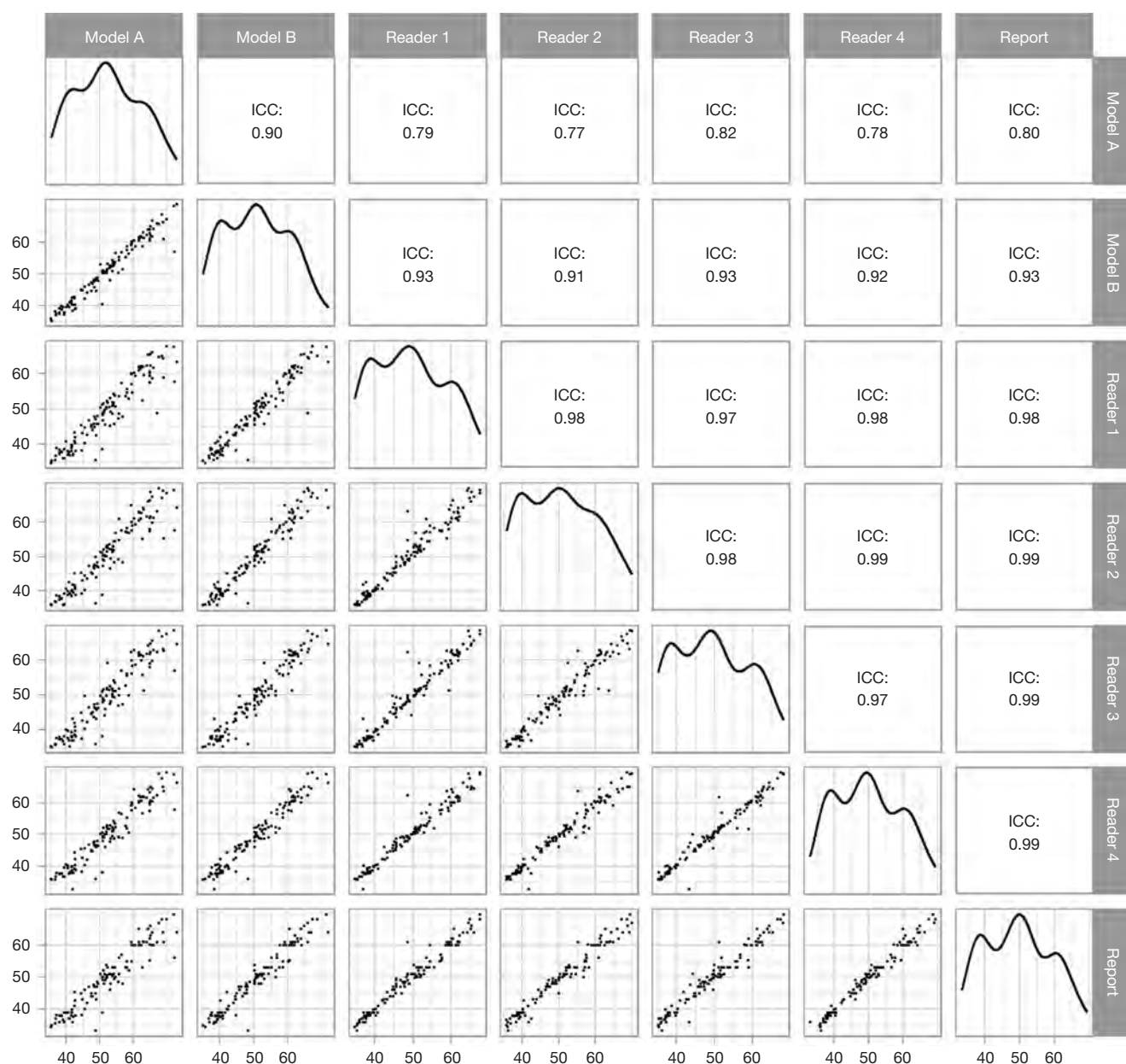
The cardiothoracic ratio in the artificial intelligence (AI) model, manual readings, and reading reports in the test data are shown in *Figures 2,3*. The mean cardiothoracic ratio was slightly larger in the AI model than in the manual readings and reading reports (52.4% and 9.8% for the initial model, 51.0% and 9.3% for the advanced model, and 49.8% and of 9.4% for the total of four manual measurements). Correlation analysis of the cardiothoracic ratio in each case showed a higher correlation with manual measurements in the advanced model than in the initial model (ICC: 0.91 to 0.93 between the advanced model and the manual measurements, ICC: 0.77 to 0.82 between the initial model and the manual measurements).

The relationship between each AI model and manual readings/reading reports was examined using Bland-Altman

plots (*Figure 4*). The advanced model showed a decrease in error compared to the initial model. No systematic errors were found for either model. Examples of segmentation by the initial and advanced models are shown in *Figures 5,6*. The DSCs between each model and the manual segmentation were follows: for the initial model, heart: mean 0.96, standard deviation (SD) 0.02, lung: mean 0.98, SD 0.01; for the advanced model, heart: mean 0.97, SD 0.02, lung: mean 0.99, SD 0.01.

## Discussion

In this study, we attempted to develop deep learning-based models to estimate the cardiothoracic ratio in chest radiographs. The advanced model, which was trained using data acquired without additional segmentation, showed strong correlation with the reference values of cardiothoracic ratio obtained through manual measurement. The advanced model had an ICC of approximately 0.93 with manual measurements, indicating that it could

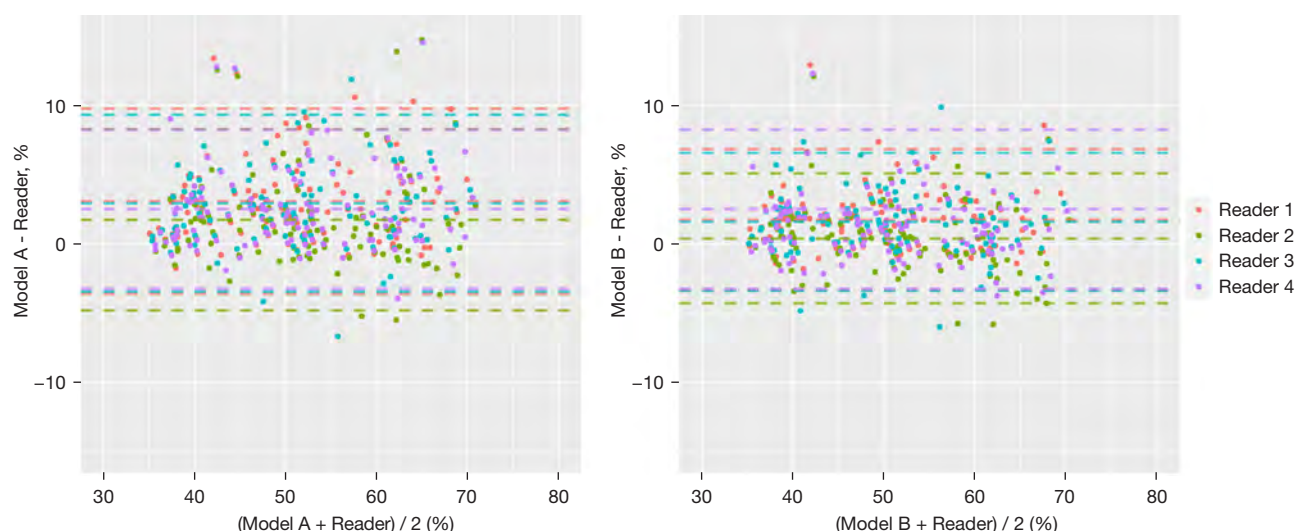


**Figure 3** Correlation of the cardiothoracic ratio between the two deep learning models, manual reading by four radiologists, and radiology report in the test dataset. The correlation matrix shows the relationship between the two deep-learning models, radiology report, and manual measurements by four radiologists. The X-axis represents the cardiothoracic ratio (%) calculated by the upper item, and the Y-axis represents the cardiothoracic ratio (%) calculated by the right-hand item. Model A, cardiothoracic ratio estimated by the initial deep-learning model; Model B, cardiothoracic ratio estimated by the advanced deep-learning model; Reader 1-4, cardiothoracic ratio measured by four radiologists; Report, cardiothoracic ratio measured by radiology reports; ICC, intraclass correlation coefficient.

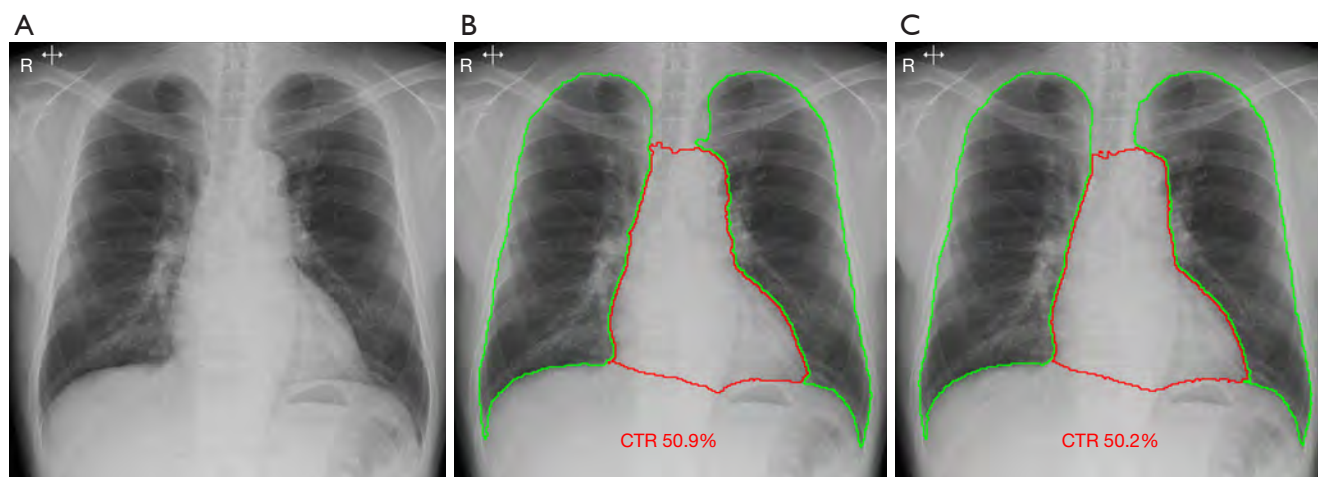
potentially be used in clinical practice.

Measurements of the target lesion in diagnostic imaging including abnormal lesions and normal anatomical structure are used in a variety of situations. The measurement of

abnormal lesions is important in diagnostic imaging because it is related to disease progression and severity. However, until recently, the measurement of normal anatomical structures was not considered significant. Longitudinal



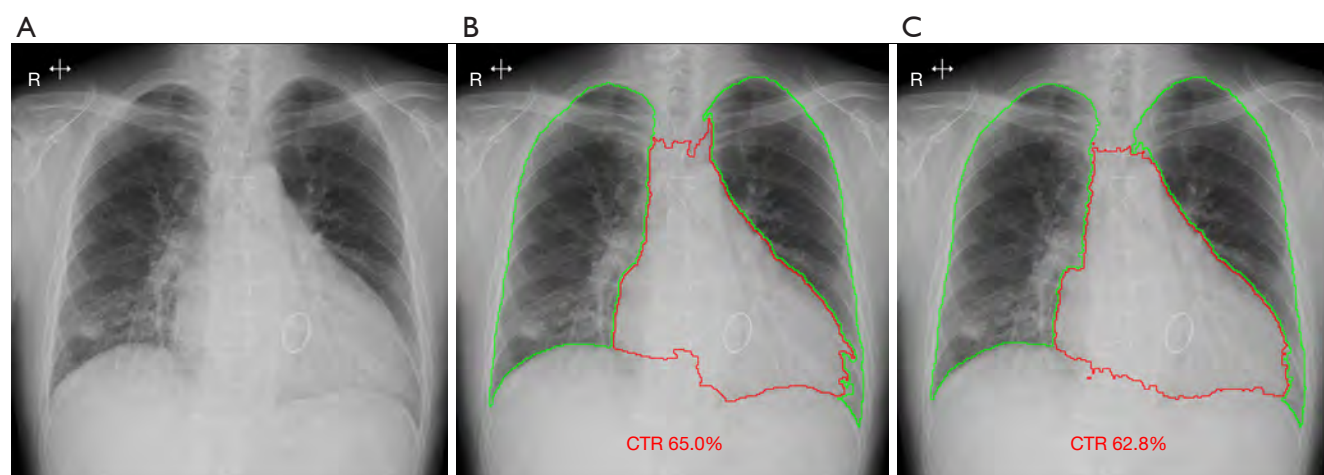
**Figure 4** Bland-Altman plots of the two deep-learning models compared to manual measurements. The left panel shows the Bland-Altman plots of the initial model, and the references, and the right panel shows the plots of the advanced model and the references. Model A, cardiothoracic ratio estimated by the initial deep-learning model; Model B, cardiothoracic ratio estimated by the advanced deep-learning model; Reader 1-4, cardiothoracic ratio measured by four radiologists.



**Figure 5** Example of chest radiograph showing good correlation with the references. (A) Original chest radiograph, (B) CTR estimation by the initial model, and (C) CTR estimation by the advanced model. The green line represents the contour of lung segmentation, and the red line represents the contour of heart segmentation. The mean CTR determined by four radiologists was 48.7% and the radiology report stated a CTR of 49%. CTR, cardiothoracic ratio.

studies have shown that measurement values may change with changes in patient conditions, highlighting the importance of the measurement of normal structures (11). In diagnostic imaging which focuses on detecting abnormalities, it can be a strain on the radiologist to simultaneously measure various normal structures. In

addition, manual measurements performed by a radiologists can result in errors, including inter-reader or intra-reader error (12). Especially in longitudinal observations that investigate temporal changes in the same patient, inter-measurement discrepancies by various radiologists constitute a challenge that needs to be resolved. To reduce



**Figure 6** Example of chest radiograph showing poor correlation with the references. (A) Original chest radiograph, (B) CTR estimation by the initial model and (C) CTR estimation by the advanced model. The green line represents the contour of lung segmentation, and the red line represents the contour of heart segmentation. Compared to the initial model (B), the advanced model is more precise in the segment at the boundary between the left ventricle and the lung. The mean CTR determined by four radiologists was 58.1%, while the radiology report stated a CTR of 60.1%. CTR, cardiothoracic ratio.

workload and errors in measurement, the development of automatic measurement is necessary.

We attempted to construct a model for automatic measurement using U-net, a type of AI that has recently been shown to have high performance (13). To reduce the most effort-intensive segmentation work, we first built an initial model using an open-source training dataset and then constructed an advanced model by adding additional training data produced by the initial model to improve accuracy. While additional training can help the model adapt to the data that it is being trained on, using a limited number of institutions for training may result in overfitting to those specific institutions. However, even if limited to the institutions, this method has the potential to improve accuracy while minimizing the workload required for additional training.

In the two deep learning-based models for estimating the cardiothoracic ratio that were examined in this study, the results of the cardiothoracic ratio in the initial model tended to be slightly larger than that derived from manual measurement. This trend improved in the advanced model, which was trained with additional training data. The large cardiothoracic ratio in the initial model may be due to difficulties in accurately segmenting the boundary between the left ventricle or right atrium and the lung in the region close to the lung. On chest radiographs, faint increased

density due to bilateral pericardial fat can often make it difficult to manually set the boundaries for cardiothoracic ratio measurement. However, the advanced model used in this study was able to make estimations that more closely resemble human manual measurements. The cardiothoracic ratio calculated by the models were slightly higher from that obtained through manual assessment. Consequently, when we used the common cut-off point of a cardiothoracic ratio  $\geq 50\%$  to define cardiomegaly, the results for the presence of cardiac enlargement may differ from those of manual measurements in certain cases.

Several methods for estimating the cardiothoracic ratio have been studied using various classical mathematical algorithm; however, in the current era of AI, automatic measurement by segments using deep learning with higher accuracy is expected to be applied to clinical practice (14-16).

This study has several limitations. First, it does not assess generalizability by studying data from other institutions, because it only attempts to improve accuracy by fitting with additional training data. It is possible that generalizability could be achieved by using this method to learn from data from multiple institutions. Second, it has not been determined whether results comparable to manual measurement can be obtained in a longitudinal study of the same patient. Third, patient clinical information, such as sex, age, and body shape, is different between the chest



radiographs of the open-source dataset and the original, which may have affected the learning process.

## Conclusions

The deep learning-based cardiothoracic ratio estimation model for chest radiographs showed good correlation with the manual measurements made by radiologists. The modified model, which used additional cases from the initial learning model, showed improved correlation with manual measurements compared to the initial model.

## Acknowledgments

*Funding:* None.

## Footnote

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <https://qims.amegroups.com/article/view/10.21037/qims-23-187/coif>). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the institutional board of Kanazawa University Hospital and individual consent for this retrospective analysis was waived.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. Truszkiewicz K, Poręba R, Gać P. Radiological Cardiothoracic Ratio in Evidence-Based Medicine. *J Clin Med* 2021;10:2016.
2. Dimopoulos K, Giannakoulas G, Bendayan I, Liodakis E, Petraco R, Diller GP, Piepoli MF, Swan L, Mullen M, Best N, Poole-Wilson PA, Francis DP, Rubens MB, Gatzoulis MA. Cardiothoracic ratio from postero-anterior chest radiographs: a simple, reproducible and independent marker of disease severity and outcome in adults with congenital heart disease. *Int J Cardiol* 2013;166:453-7.
3. Bradley AJ, Carrington BM, Lawrance JA, Ryder WD, Radford JA. Assessment and significance of mediastinal bulk in Hodgkin's disease: comparison between computed tomography and chest radiography. *J Clin Oncol* 1999;17:2493-8.
4. Yotsueda R, Taniguchi M, Tanaka S, Eriguchi M, Fujisaki K, Torisu K, Masutani K, Hirakata H, Kitazono T, Tsuruya K. Cardiothoracic Ratio and All-Cause Mortality and Cardiovascular Disease Events in Hemodialysis Patients: The Q-Cohort Study. *Am J Kidney Dis* 2017;70:84-92.
5. Saiviroonporn P, Rodbangyang K, Tongdee T, Chaisangmongkon W, Yodprom P, Siriapisith T, Wonglaksanapimon S, Thiravit P. Cardiothoracic ratio measurement using artificial intelligence: observer and method validation studies. *BMC Med Imaging* 2021;21:95.
6. Hasan MA, Lee SL, Kim DH, Lim MK. Automatic evaluation of cardiac hypertrophy using cardiothoracic area ratio in chest radiograph images. *Comput Methods Programs Biomed* 2012;105:95-108.
7. Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, Folio LR, Summers RM, Rubin DL, Lungren MP. Preparing Medical Imaging Data for Machine Learning. *Radiology* 2020;295:4-15.
8. Chen D, Liu S, Kingsbury P, Sohn S, Storlie CB, Habermann EB, Naessens JM, Larson DW, Liu H. Deep learning and alternative learning strategies for retrospective real-world clinical data. *NPJ Digit Med* 2019;2:43.
9. Shiraishi J, Katsuragawa S, Ikezoe J, Matsumoto T, Kobayashi T, Komatsu K, Matsui M, Fujita H, Kodera Y, Doi K. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. *AJR Am J Roentgenol* 2000;174:71-4.
10. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2015*:234-41. Available online: <https://arxiv.org/pdf/1505.04597.pdf>
11. Raghu VK, Weiss J, Hoffmann U, Aerts HJWL, Lu MT. Deep Learning to Estimate Biological Age From Chest



- Radiographs. *JACC Cardiovasc Imaging* 2021;14:2226-36.
12. Zhao B, Tan Y, Bell DJ, Marley SE, Guo P, Mann H, Scott ML, Schwartz LH, Ghiorghiu DC. Exploring intra- and inter-reader variability in uni-dimensional, bi-dimensional, and volumetric measurements of solid tumors on CT scans reconstructed at different slice intervals. *Eur J Radiol* 2013;82:959-68.
  13. Siddique N, Paheding S, Elkin CP, Devabhaktuni V. U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. *IEEE Access* 2021;9:82031-57.
  14. Agrawal T, Choudhary P. Segmentation and classification on chest radiography: a systematic survey. *Vis Comput* 2023;39:875-913.
  15. Lee WL, Chang K, Hsieh KS. Unsupervised segmentation of lung fields in chest radiographs using multiresolution fractal feature vector and deformable models. *Med Biol Eng Comput* 2016;54:1409-22.
  16. Chamveha I, Promwiset T, Tongdee T, Saiviroonporn P, Chaisangmongkon W. Automated cardiothoracic ratio calculation and cardiomegaly detection using deep learning approach. *Arxiv* 2020:1-11. Available online: <https://arxiv.org/pdf/2002.07468.pdf>

**Cite this article as:** Yoshida K, Takamatsu A, Matsubara T, Kitagawa T, Toshima F, Tanaka R, Gabata T. Deep learning-based cardiothoracic ratio measurement on chest radiograph: accuracy improvement without self-annotation. *Quant Imaging Med Surg* 2023;13(10):6546-6554. doi: 10.21037/qims-23-187



## Association between switching prescribed drugs for lower urinary tract symptoms and independence of urination in post-stroke patients: A retrospective cohort study

Koyomi Sumiya<sup>a</sup>, Miho Shogenji<sup>b</sup>, Yasunori Ikenaga<sup>c,d</sup>, Yoru Ogawa<sup>d,e</sup>, Kohei Hirako<sup>f,m</sup>, Arimi Fujita<sup>a,g</sup>, Tsutomu Shimada<sup>a,g,\*</sup>, Masako Hashimoto<sup>h</sup>, Akio Masuda<sup>i</sup>, Takehito Nagamoto<sup>j</sup>, Ikumi Tamai<sup>j,k</sup>, Hisayuki Ogura<sup>k,l</sup>, Tadashi Toyama<sup>k,l</sup>, Takashi Wada<sup>l</sup>, Yoshimichi Sai<sup>a,g,k</sup>

<sup>a</sup> Department of Clinical Pharmacokinetics, Graduate School of Medical Sciences, Kanazawa University, Kanazawa, Ishikawa, Japan

<sup>b</sup> Faculty of Health Sciences, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Ishikawa, Japan

<sup>c</sup> Department of Rehabilitation Medicine, Yawata Medical Center, Ishikawa, Japan

<sup>d</sup> Council of Kaga Local Stroke Network, South Ishikawa, Japan

<sup>e</sup> Department of Pharmacy, Komatsu Municipal Hospital, Ishikawa, Japan

<sup>f</sup> Frontier Science and Social Co-creation Initiative, Kanazawa University, Ishikawa, Japan

<sup>g</sup> Department of Hospital Pharmacy, University Hospital, Kanazawa University, Ishikawa, Japan

<sup>h</sup> Temari Pharmacy, Ishikawa, Japan

<sup>i</sup> Komeya Pharmacy, Ishikawa, Japan

<sup>j</sup> Division of Pharmacy, Graduate School of Pharmaceutical Sciences, Kanazawa University, Ishikawa, Japan

<sup>k</sup> AI Hospital/Macro Signal Dynamics Research and Development Center, Kanazawa University, Ishikawa, Japan

<sup>l</sup> Department of Nephrology and Laboratory Medicine, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Ishikawa, Japan

<sup>m</sup> The Establishment Preparation Office for The Faculty of Interdisciplinary Economics, Kinjo University, Ishikawa, Japan

### ARTICLE INFO

#### Keywords:

Activities of daily living  
Lower urinary tract symptoms  
Stroke rehabilitation  
Urinary medication

### ABSTRACT

**Objectives:** Stroke patients frequently exhibit loss of independence of urination, and their lower urinary tract symptoms change with the phase of stroke. However, it is unclear whether switching prescribed drugs for lower urinary tract symptoms during hospitalization from acute care wards to convalescence rehabilitation wards affects patients' independence of urination at discharge. It is also unclear whether the impact of switching varies by stroke type. This retrospective cohort study aimed to examine these issues.

**Materials and methods:** We analyzed 990 patients registered in the Kaga Regional Cooperation Clinical Pathway for Stroke database during 2015–2019. Prescriptions for lower urinary tract symptoms from pre-onset to convalescence rehabilitation were surveyed. Logistic regression analysis was performed to examine the association between switching drugs and independence of urination based on bladder management and voiding location at discharge. Stroke types were also examined in subgroup analyses.

**Results:** About 21 % of patients had their lower urinary tract symptoms prescriptions switched during hospitalization. Switching was positively associated with independence of bladder management (odds ratio 1.65, 95 % confidence interval 1.07 to 2.49) and voiding location (odds ratio 2.72, 95 % confidence interval 1.72 to 4.37). Similar associations were observed in different stroke types.

**Conclusions:** Approximately 20 % of patients had their lower urinary tract symptoms medications switched upon transfer from acute to convalescence rehabilitation wards. Switching was significantly associated with improved urinary independence at discharge. Consistent results were observed across different stroke types, suggesting that switching medications contributes to urinary independence after stroke, regardless of the etiology or severity of stroke.

\* Corresponding author at: Department of Hospital Pharmacy, University Hospital, Kanazawa University, 13-1 Takara-machi, Kanazawa 920-8641, Japan.  
E-mail address: [t-shimada@staff.kanazawa-u.ac.jp](mailto:t-shimada@staff.kanazawa-u.ac.jp) (T. Shimada).

<https://doi.org/10.1016/j.jstrokecerebrovasdis.2023.107419>

Received 6 April 2023; Received in revised form 24 August 2023; Accepted 6 October 2023

Available online 14 October 2023

1052-3057/© 2023 Elsevier Inc. All rights reserved.

## Introduction

It is estimated that about 220,000 new stroke cases occur annually in Japan, and more than half of the patients require nursing care or die.<sup>1</sup> Lower urinary tract dysfunction (LUTD) frequently occurs in post-stroke survivors, and lower urinary tract symptoms (LUTS) change with the phase of stroke. About 21–47 % of patients experience voiding symptoms in the acute phase of stroke,<sup>2,3</sup> on the other hand 37–79 % experience storage symptoms due to detrusor overactivity after the acute phase.<sup>4–6</sup> In addition, paralysis and higher brain dysfunction due to stroke impair voiding behaviors such as moving to the toilet and lowering underwear,<sup>7</sup> which affect the quality of life of patients and caregivers.<sup>8,9</sup> Poor voiding behavior and LUTS are associated with multiple factors, including comorbidities,<sup>10</sup> impairment in activities of daily living (ADL),<sup>11</sup> and medications.<sup>12</sup> Therefore, it is crucial for multidisciplinary teams treating post-stroke survivors to estimate the patients' level of independence of urination; that is, "the ability to manage urination on one's own, regardless of the voiding method or devices", taking account of improving lower urinary tract function and voiding behavior. Stroke patients in Japan were treated in three phases: acute phase, convalescent phase, and community-based phase. During acute and convalescent phase, most patients are treated in acute care wards (ACWs) and convalescent rehabilitation wards (CRWs), respectively. A previous study has found that a continuous continence self-management program during the stay in ACWs and in CRWs involving professionals, including urologists, was associated with improved independence in voiding behavior.<sup>13</sup>

Pharmacotherapy for LUTS involves the use of medications with opposing mechanisms of action for voiding and storage symptoms. For voiding symptoms, cholinergic drugs or  $\alpha$ 1-blockers are used to promote urination by activating muscarinic receptors in the detrusor muscle and blocking  $\alpha$ -receptors in the urethral sphincter. Conversely, for storage symptoms, muscarinic receptor antagonists and  $\beta$ 3-agonists are used to inhibit overactivity of the detrusor muscle, increase bladder capacity, and promote urine storage.<sup>5,14</sup> Particularly in post-stroke patients, LUTS often fluctuate during the transition from ACWs to CRWs, making drug management a crucial issue.<sup>12,15</sup> This management includes switching LUTS medications in response to changes in symptoms, which can entail initiating treatments or discontinuing LUTS medications. Such LUTS medication switching in the ACWs has been reported to be associated with the improvement of lower urinary tract function during acute hospitalization.<sup>16,17</sup> However, in Japan's current stroke treatment landscape, patients often switch hospitals in the transition from ACWs to CRWs, and not all patients are treated by urologists, which may lead to insufficient review of LUTS medications, with potential adverse effects for patients.<sup>13</sup> However, the actual situation requires further investigation. Furthermore, although the existing stroke treatment guidelines indicate the use of drugs for LUTD in stroke patients according to overactivity/hypofunction of the detrusor and hypertonia/hypotonia of the urethra, there is no explicit mention of the management of switching LUTS medications.<sup>18</sup> Additionally, it is unclear whether different stroke types require different management of LUTS and have different effects on urinary independence.<sup>19</sup>

In this study, by using a stroke database that includes longitudinal information about the phases of stroke, which can be followed even when patients are transferred to different hospitals, we aimed to investigate the patterns of change in prescribed medications for LUTS during hospitalization, including the transition from ACWs to CRWs, and to clarify how these changes relate to urinary independence and functional recovery before patients reach the community-based phase. Furthermore, we investigated how these influences vary among stroke types with different severities and causes. We believe our findings will contribute to the development of appropriate, individually tailored management of LUTS medications after stroke.

## Methods

### Database

This study used data from Kaga Regional Cooperation Clinical Pathway for Stroke (KRCCPS) version 5, which is one of the tools for inter-professional sharing of information about rehabilitation and treatment among ACWs, CRWs, and community-based facilities in the southern part of Ishikawa Prefecture in Japan. The KRCCPS includes check fields or free-comment fields about patients' demographics, stroke-related assessments, and medical information recorded by multidisciplinary medical staff using Microsoft Excel worksheets (Supplementary Table 1). The KRCCPS has been implemented since 2009 in 19 acute care hospitals, 11 hospitals with CRWs, and 466 facilities for community-based rehabilitation. By 2019, KRCCPS contained data for more than 5,000 patients.<sup>13,20</sup>

### Study design and participants

A descriptive study was conducted to clarify the patterns of prescribed drugs for LUTS and the number of patients prescribed drugs for LUTS before stroke onset, during the ACWs, and during their stay in CRWs. A retrospective cohort study was conducted to clarify the relationship between switching prescribed drugs for LUTS and independence of urination at discharge from CRWs.

Patients who were enrolled in the KRCCPS between January 2015 and December 2019 and were admitted to CRWs were included in this study. Patients who had missing data were excluded. Written informed consent for using the KRCCPS database in this study was obtained from the patients or their key persons. All data derived from the KRCCPS database were anonymized, and the patient's personal information was not identifiable. This study was approved by the Kanazawa University Medical Ethics Review Committee (Review No. 3580-2).

### Switching prescribed drugs for LUTS

The drug information in KRCCPS completed by pharmacists consists of contraindications, concomitant or prior-onset medications, discharge prescriptions, and a drug free-comment field, including a description of prescription switching during hospitalization, in the form of unstructured text data. Prescription information on drugs for LUTS was extracted. Transient acute illness medications (medications for urinary tract infections, pneumonia, and other infectious diseases), patch medications, eye drops, nasal sprays, and over-the-counter medications were excluded. The medications for LUTS were categorized into those used to treat voiding symptoms and storage symptoms based on Japanese guidelines for LUTS<sup>21–23</sup> (Supplementary Table 2). Patients were classified into four groups: (1) prescribed drugs for voiding symptoms, (2) prescribed drugs for storage symptoms, (3) combination of (1) and (2), and (4) no drug prescription for LUTS. If there was more than one switch within each phase, the first switch was counted. The prescribing patterns before stroke onset, during the ACWs, and in the CRWs were determined based on pre-onset medication continued in the ACWs, on discharge prescriptions and drug comments in the ACWs, and on discharge prescriptions and drug comments at the CRWs, respectively. In clinical practice, prescribed drugs are switched according to the actual LUTS in patients, but since LUTS symptoms cannot be extracted from the database used in the present study, the above categorization was used instead.

### Independence of urination

In this study, we used two definitions of independence of urination: the functional independence measure (FIM) bladder management score and the voiding location. FIM bladder management was assessed based on the frequency of incontinence, the amount of assistance required, and

the use of medications and equipment. A bladder management score of 6 or higher indicates independence, while a score of 5 or lower indicates dependence on assistance.<sup>24,25</sup> As regards voiding location, we followed a previous study,<sup>13</sup> defining toilet or bedside commode as voiding in the toilet, while diaper or indwelling urinary catheter were defined as voiding on the bed. If more than one voiding location was checked, the location involving the least amount of movement was selected. If the location was unknown, it was treated as a missing value.

#### Other variables

Other variables related to patient characteristics were extracted from the KRCCPS: these were age, sex, complications (hypertension, diabetes, cardiac disease), stroke type (cerebral infarction, cerebral hemorrhage, subarachnoid hemorrhage), National Institutes of Health Stroke Scale (NIHSS) score, FIM score, continence management (continuous continence self-management program or traditional voiding care),<sup>13</sup> and the number of medications taken. The NIHSS is used to assess the severity of acute stroke and consists of 11 items with scores ranging from 0 to 42, with higher scores indicating greater neurological severity.<sup>26</sup> The FIM is used to evaluate ADL and consists of 13 items related to motor function (FIM-M) and five items related to cognitive function (FIM-C), with each item scored on a scale of 1 (total assistance) to 7 (complete independence). A bladder management score is one of the factors in FIM-M. The FIM was assessed by a multidisciplinary team at the time of admission (within three days after admission) and at discharge (within three days before discharge). The number of medications taken was calculated based on the number of oral medications regularly prescribed from the pre-onset medications and discharge prescriptions.

#### Statistical analysis

Continuous variables were summarized using medians and interquartile range (IQR). The chi-square and Mann-Whitney U tests were used to assess the significance of differences between the groups. Multiple logistic regression analysis was used to examine whether the independence of urination (bladder management [ $\geq 6$  points] and voiding location [toilet]) was associated with the switching of prescribed drugs for LUTS. All variables with *p*-values less than 0.15 in the group comparison for the presence or absence of switching LUTS therapeutic drugs were forcibly included in the model as covariates to adjust for confounding factors. While this cutoff value is arbitrary, it is used to minimize the risk of overlooking important covariates.<sup>27</sup> The adjusted odds ratios (AORs) and 95 % confidence intervals (95 % CIs) were calculated for independence of urination. Further, a multiple regression analysis was used to examine whether switching LUTS medications was associated with FIM-M and FIM-C at CRWs discharge. Covariates were selected using the same criteria as above. Multicollinearity was assessed using the variance inflation factor (VIF), with a value of less than 4 indicating no multicollinearity. The data types of the variables included in the model are shown in the footnotes of each table. Statistical analyses were conducted using R version 4.1.2, with the significance level set at 5 %.

#### Sensitivity analysis

To evaluate the robustness of the result, sensitivity analysis was conducted by excluding the most frequently missing item, NIHSS, from the covariates. The missing rate of each item is shown in Supplementary Table 3.

#### Subgroup analysis

We also conducted a subgroup analysis by types of stroke (cerebral infarction, cerebral hemorrhage, subarachnoid hemorrhage), which differ in severity and etiology, to elucidate how the prescription status of LUTS therapeutic drugs and the relationship between LUTS therapeutic

drugs, urination independence, and functional recovery varied among these subgroups. For each type of stroke, we extracted the phase-specific prescription status of LUTS therapeutic drugs and examined the association between the switching of LUTS therapeutic drugs and urination independence and functional recovery in a similar way to that employed in the primary analysis.

## Results

#### Patients' characteristics

The final analysis included 990 patients, after exclusion of 635 patients who had missing survey items (Fig. 1). The median age was 75 years, 55.7 % were male, and 60.8 % had cerebral infarction. Patients for whom the prescribed drugs for LUTS were switched had a higher proportion of aged males, a longer length of hospital stay, lower scores on FIM-M and FIM-C at acute admission, and a higher proportion of patients prescribed drugs for voiding symptoms and storage symptoms before the onset of stroke, compared with the group without drug switching (Table 1).

#### Patterns of change of prescribed drugs for LUTS from the ACWs to CRWs

The percentage of prescriptions for LUTS was 9.6 % before onset of stroke, 14.9 % during the ACWs, and 20.2 % in the CRWs (Table 2). Prescription of drugs for voiding symptoms tended to increase in the ACWs compared with before onset, and then remained unchanged in the CRWs. Prescribed drugs for storage symptoms tended to decrease in the ACWs compared with before onset, and increased in the CRWs compared with the ACWs. Combination prescriptions tended to increase in the CRWs compared with before onset and the ACWs. Among the patients who changed medications, none switched from drugs for storage symptoms to drugs for voiding symptoms, and the initiation of medication for voiding symptoms in the ACWs was the most frequent prescribing pattern (Supplementary Table 4). Examination of the prescribing patterns by phase and mechanism of action of LUTS medications for males and females revealed that the most common prescribing pattern among patients who had switched LUTS medications was the initiation of  $\alpha 1$ -blockers, cholinergic agents, or a combination of both in the ACWs, for both genders. Furthermore, in the CRWs, there was a tendency for more prescriptions to be initiated for muscarinic receptor antagonists or a combination of muscarinic receptor antagonists and  $\alpha 1$ -blockers (Supplementary Tables 5, 6). Among  $\alpha 1$ -blockers, males tended to be prescribed silodosin and tamsulosin and females tended to be prescribed urapidil. Among anticholinergic drugs, imidafenacin and solifenacin tended to be frequently prescribed for both genders (Supplementary Table 7).

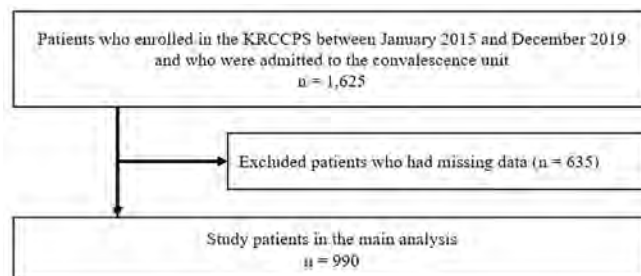


Fig. 1. Flowchart for patient selection. KRCCPS, Kaga Regional Cooperation Clinical Pathway for Stroke.

**Table 1**

Characteristics of patients with and without switched drugs for lower urinary tract symptoms.

	Overall (n = 990)	No switching of drugs for LUTS (n = 783)	Switched drugs for LUTS (n = 207)	p value
Age (year), median [IQR]	75 [66, 82]	74 [66, 82]	77 [70, 84]	<0.001*
Gender n, (%)				0.010*
Female	439 (44.3)	364 (46.5)	75 (36.2)	
Male	551 (55.7)	419 (53.5)	132 (63.8)	
Type of stroke n, (%)				0.120
Subarachnoid hemorrhage	52 (5.3)	47 (6.0)	5 (2.4)	
Cerebral hemorrhage	336 (33.9)	264 (33.7)	72 (34.8)	
Cerebral infarction	602 (60.8)	472 (60.3)	130 (62.8)	
Comorbidities n, (%)				
Hypertension	760 (76.8)	606 (77.4)	154 (74.4)	0.415
Diabetes mellitus	243 (24.5)	183 (23.4)	60 (29.0)	0.114
Cardiac disease	272 (27.5)	207 (26.4)	65 (31.4)	0.182
Length of stay (day), median [IQR]				
Acute	28 [18, 41]	27 [18, 40]	30 [20, 44]	0.014*
Convalescence	88 [52, 134]	86 [48, 134]	94 [63, 139]	0.020*
Condition at acute admission, median [IQR]				
NIHSS score	8 [4, 14]	7 [4, 14]	9 [5, 15]	0.049*
FIM-C (points)	15 [7, 25]	16 [7, 26]	13 [7, 21]	0.006*
FIM-M (points)	15 [13, 34]	15 [13, 36]	13 [13, 26]	0.011*
Independently urinate at acute admission n, (%)				0.012*
Voiding location				
Toilet	287 (29.0)	242 (30.9)	45 (21.7)	
Toilet	224 (22.6)	191 (24.4)	33 (15.9)	
Bedside commode	63 (6.4)	51 (6.5)	12 (5.8)	
On the bed	703 (71.0)	541 (69.1)	162 (78.3)	
Diaper	263 (26.6)	205 (26.2)	58 (28.1)	
Indwelling urinary catheter	440 (44.4)	336 (42.9)	104 (50.2)	
FIM bladder management				0.006*
≥6 points	182 (18.4)	158 (20.2)	24 (11.6)	
≤5 points	808 (81.6)	625 (79.8)	183 (88.4)	
Continence management n, (%)				0.209
Continuous continence self-management program	476 (48.1)	398 (50.8)	116 (56.0)	
Traditional voiding care	514 (51.9)	385 (49.2)	91 (44.0)	
Prescribing status before stroke onset				
Number of drugs, median [IQR]	3 [0, 6]	2 [0, 6]	4 [0, 8]	<0.001*
Prescribing drugs for voiding symptoms n, (%)	64 (6.5)	34 (4.3)	30 (14.5)	<0.001*
Prescribing drugs for storage symptoms n, (%)	38 (3.8)	10 (1.3)	28 (13.5)	<0.001*

**Table 1 (continued)**

	Overall (n = 990)	No switching of drugs for LUTS (n = 783)	Switched drugs for LUTS (n = 207)	p value
Change in the number of drugs from acute to CRWs, median [IQR]	2 [0, 4]	2 [0, 4]	2 [−1, 5]	0.774
Independently urinate at CRWs discharge n, (%)				0.622†
Voiding location				
Toilet	711 (71.8)	559 (71.4)	152 (73.4)	
Toilet	677 (68.4)	543 (69.3)	134 (64.7)	
Bedside commode	34 (3.4)	16 (2.0)	18 (8.7)	
On the bed	279 (28.2)	224 (28.6)	55 (26.6)	
Diaper	254 (25.7)	209 (26.7)	45 (21.7)	
Indwelling urinary catheter	25 (2.5)	15 (1.9)	10 (4.8)	
FIM bladder management				0.076
≥6 points	610 (61.6)	494 (63.1)	116 (56.0)	
≤5 points	380 (38.4)	289 (36.9)	91 (44.0)	

IQR, interquartile rang; NIHSS, National Institutes of Health Stroke scale; FIM, functional independence measure; FIM-M, FIM motor score; FIM-C, FIM cognitive score; LUTS, lower urinary tract symptoms; CRWs, convalescence rehabilitation wards.

† p value comparing the number of patients voiding on the bed to the number voiding on the toilet

\* p value < 0.05.

**Table 2**

Numbers of patients receiving various patterns of prescribed drugs for lower urinary tract symptoms from the acute to convalescence rehabilitation wards.

	Before stroke onset n = 990 (%)	ACWs n = 990 (%)	CRWs n = 990 (%)
Prescription status of drugs for LUTS			
(1) Prescribing drugs for voiding symptoms	57 (5.8)	116 (11.7)	127 (12.8)
(2) Prescribing drugs for storage symptoms	30 (3.0)	20 (2.0)	37 (3.7)
(3) Combination prescription of (1) and (2)	8 (0.8)	12 (1.2)	36 (3.6)
Any drugs for LUTS (total of (1) to (3))	95 (9.6)	148 (14.9)	200 (20.2)
(4) No drug prescription for LUTS	895 (90.4)	842 (85.1)	790 (79.8)

LUTS, lower urinary tract symptoms; ACWs, acute care wards; CRWs, convalescence rehabilitation wards.

The sum of the individual components may not add up to the subtotal due to rounding.

### Switching prescribed drugs for LUTS and independence of urination and functional recovery

In the primary analysis of all eligible patients, 207/990 (20.9 %) patients showed switching of prescribed drugs for LUTS. In the sensitivity analysis, 299/1,460 (20.5 %) patients showed switching of prescribed drugs for LUTS (Supplementary Table 8).

In multivariate analysis, switching was associated with independence of bladder management (≥6 points) (AOR = 1.657, 95 % CI 1.090–2.534) and independence of voiding location (toilet) (AOR = 2.790, 95 % CI 1.763–4.490) (Table 3). In an examination of the association with functional recovery, switching was also independently associated with higher FIM-M (standardized coefficient = 0.064, p =



**Table 3**

Association between switching drugs for lower urinary tract symptoms and urinary independence at convalescence discharge.

	Bladder management Independence vs. non- independence				Voiding location Toilet vs. on the bed			
	AOR	95 %CI Lower Upper	p value		AOR	95 %CI Lower Upper	p value	
Age	0.893	0.873	0.911	<0.001*	0.901	0.880	0.921	<0.001*
Gender	1.005	0.692	1.456	0.981	0.952	0.640	1.413	0.8079
Diabetes mellitus	0.653	0.422	1.007	0.055	0.563	0.354	0.889	0.0140*
Stroke type								
Subarachnoid hemorrhage	Referent				Referent			
Cerebral hemorrhage	0.783	0.350	1.764	0.551	0.457	0.196	1.035	0.0642
Cerebral infarction	0.531	0.235	1.210	0.129	0.317	0.134	0.726	0.0074*
Length of stay	0.989	0.985	0.992	<0.001*	0.989	0.985	0.992	<0.001*
NIHSS at acute admission	0.976	0.949	1.003	0.084	0.960	0.933	0.987	0.0047*
FIM-M at acute admission	1.042	1.011	1.076	0.009*	1.062	1.019	1.111	0.0059*
FIM-C at acute admission	1.097	1.068	1.128	<0.001*	1.105	1.071	1.142	<0.001*
Voiding location at acute admission	1.837	1.059	3.208	0.031*	2.422	1.234	4.931	0.0119*
Bladder management at acute admission	0.950	0.782	1.159	0.606	0.902	0.706	1.167	0.4171
Number of drugs before stroke onset	0.994	0.946	1.044	0.806	1.017	0.964	1.073	0.5453
Prescribing drugs for voiding symptoms before stroke onset	0.840	0.407	1.737	0.637	0.420	0.193	0.914	0.0284*
Prescribing drugs for storage symptoms before stroke onset	0.235	0.091	0.572	0.002*	0.309	0.127	0.745	0.0089*
Switched drugs for LUTS	1.657	1.090	2.534	0.019*	2.790	1.763	4.490	<0.001*

AOR, adjusted odds ratio; CI, confidence interval; NIHSS, National Institutes of Health Stroke scale; FIM, functional independence measure; FIM-M, FIM motor score; FIM-C, FIM cognitive score; LUTS, lower urinary tract symptoms.

Each model was adjusted for variables with a *p*-value less than 0.15 in the univariate analysis for switching LUTS medications as covariates.

Continuous variables: age, length of hospital stay, NIHSS, FIM-C, FIM-M, bladder management at acute admission, and number of prescribed medications.

Categorical variables: bladder management (independence: 1, non-independence:0), voiding location (toilet: 1, on the bed: 0), gender (male: 1, female: 0), diabetes (yes: 1, no: 0), stroke type (subarachnoid hemorrhage, cerebral hemorrhage, cerebral infarction), prescription of medication for LUTS before stroke onset (yes: 1, no: 0), and switching drugs for LUTS (yes: 1, no: 0).

\* *p* value <0.05.

0.009) and FIM-C (standardized coefficient = 0.080, *p* < 0.001) scores at CRWs discharge (Supplementary Table 9). Sensitivity analysis confirmed these findings, with switching also associated with independence of bladder management ( $\geq 6$  points) (AOR = 1.55, 95 % CI 1.10–2.18) and independence of voiding location (toilet) (AOR = 2.19, 95 % CI 1.52–3.18) (Supplementary Table 10), and higher FIM-M (standardized coefficient = 0.052, *p* = 0.011) and FIM-C (standardized coefficient = 0.060, *p* = 0.001) scores at CRWs discharge (Supplementary Table 11).

#### Subgroup analysis by stroke subtype

We performed a subgroup analysis according to types of stroke (cerebral infarction, cerebral hemorrhage, subarachnoid hemorrhage) of the prescribing patterns of LUTS medications by phase and the relationship between switching LUTS medications and urinary independence. The background of patients with cerebral infarction, cerebral hemorrhage, and subarachnoid hemorrhage is shown in Supplementary Table 12. From the survey on the prescribing status by stroke subtype, before the onset of stroke, medications for voiding symptoms were more

frequently prescribed in the cerebral infarction group (7.6 %) compared to the cerebral hemorrhage group (2.7 %) and the subarachnoid hemorrhage group (1.9 %). However, drugs for storage symptoms were more common in the cerebral hemorrhage group (4.2 %). Following the acute phase, hemorrhagic stroke patients displayed a larger increase in the prescription of voiding medications but a smaller increase in storage medications compared to the cerebral infarction group. This pattern held through to the CRWs (Table 4). In patients with cerebral infarction, the switch in LUTS medications was significantly associated with the voiding location (toilet) (AOR 2.827, 95 % CI 1.593–5.144), while in patients with cerebral hemorrhage, it was significantly associated with bladder management independence and voiding location (toilet) (AOR 2.486, 95 % CI 1.228–5.189, *p* = 0.013 and AOR 2.308, 95 % CI 1.149–4.806, *p* = 0.022, respectively). On the other hand, no significant association was observed in patients with subarachnoid hemorrhage (Tables 5 and 6).

#### Discussion

Key findings of this study are: 1) switching of prescribed drugs for

**Table 4**

Numbers of patients per prescribing pattern by phase and by stroke type.

	Cerebral infarction (n == 602)			Cerebral hemorrhage (n == 336)			Subarachnoid hemorrhage (n == 52)		
	Before stroke onset (%)	ACWs (%)	CRWs (%)	Before stroke onset (%)	ACWs (%)	CRWs (%)	Before stroke onset (%)	ACWs (%)	CRWs (%)
Prescription status of drugs for LUTS									
(1) Prescribing drugs for voiding symptoms	46 (7.6)	69 (11.5)	80 (13.3)	9 (2.7)	40 (11.9)	42 (12.5)	1 (1.9)	5 (9.6)	4 (7.7)
(2) Prescribing drugs for storage symptoms	15 (2.5)	13 (2.2)	23 (3.8)	14 (4.2)	6 (1.8)	12 (3.6)	1 (1.9)	0 (0.0)	0 (0.0)
(3) Combination prescription of (1) and (2)	5 (0.8)	7 (1.2)	27 (4.5)	3 (0.9)	5 (1.5)	6 (1.8)	0 (0.0)	0 (0.0)	0 (0.0)
Any drugs for LUTS (total of (1) to (3))	66 (11.0)	89 (14.8)	130 (21.6)	26 (7.7)	51 (15.2)	60 (17.9)	2 (3.8)	5 (9.6)	4 (7.7)
(4) No drug prescription for LUTS	536 (89.0)	513 (85.2)	472 (78.4)	310 (92.3)	285 (84.8)	276 (82.1)	50 (96.2)	47 (90.4)	48 (92.3)

LUTS, lower urinary tract symptoms; ACWs, acute care wards; CRWs, convalescence rehabilitation wards.

The sum of the individual components may not add up to the subtotal due to rounding.

**Table 5**

Association between switching drugs for lower urinary tract symptoms and bladder management by stroke type.

	Bladder management Independence vs. non-independence				Cerebral hemorrhage				Subarachnoid hemorrhage			
	Cerebral infarction		<i>p</i> value		Cerebral hemorrhage		<i>p</i> value		Subarachnoid hemorrhage		<i>p</i> value	
	AOR	95 %CI			AOR	95 %CI			AOR	95 %CI		
	Lower	Upper			Lower	Upper			Lower	Upper		
Age	0.896	0.869	0.922	<0.001*	0.905	0.877	0.930	<0.001*	0.906	0.836	0.966	0.007*
Gender					0.947	0.531	1.679	0.853				
Diabetes mellitus					0.681	0.329	1.403	0.297				
Length of stay	0.991	0.986	0.996	<0.001*								
NIHSS at acute admission	1.030	0.987	1.075	0.173								
FIM-C at acute admission	1.082	1.047	1.118	<0.001**								
FIM-M at acute admission	1.046	1.013	1.083	0.007*								
Bladder management at acute admission	0.963	0.782	1.191	0.727	3.047	1.837	6.913	0.001*				
Voiding location at acute admission	2.645	1.444	4.887	0.002*								
Number of drugs before stroke onset	0.886	0.809	0.968	0.008*	0.920	0.820	1.033	0.159	1.146	0.876	1.559	0.345
Change in the number of drugs from acute to CRWs	0.870	0.788	0.958	0.005*	0.869	0.779	0.966	0.010*				
Continuous continence self-management program					0.996	0.568	1.749	0.987				
Prescribing drugs for voiding symptoms before stroke onset	0.892	0.392	2.031	0.785	1.286	0.302	5.827	0.734				
Prescribing drugs for storage symptoms before stroke onset	0.221	0.059	0.735	0.017*	0.374	0.098	1.265	0.127				
Switched drugs for LUTS	1.587	0.937	2.708	0.088	2.486	1.228	5.189	0.013*	0.312	0.012	3.678	0.397

AOR, adjusted odds ratio; CI, confidence interval; NIHSS, National Institutes of Health Stroke scale; FIM, functional independence measure; FIM-M, FIM motor score; FIM-C, FIM cognitive score; LUTS, lower urinary tract symptoms.

Each model was adjusted for variables with a *p*-value less than 0.15 in the univariate analysis for switching LUTS medications as covariates.

Continuous variables: age, length of hospital stay, NIHSS, FIM-C, FIM-M, bladder management at acute admission, and number of prescribed medications.

Categorical variables: bladder management (independence: 1, non-independence:0), voiding location (toilet: 1, on the bed: 0), gender (male: 1, female: 0), diabetes (yes: 1, no: 0), stroke type (subarachnoid hemorrhage, cerebral hemorrhage, cerebral infarction), prescription of medication for LUTS before stroke onset (yes: 1, no: 0), and switching drugs for LUTS (yes: 1, no: 0).

\* *p* value <0.05.

**Table 6**

Association between switching drugs for lower urinary tract symptoms and voiding location by stroke type.

	Voiding location Toilet vs. on the bed				Cerebral hemorrhage				Subarachnoid hemorrhage			
	Cerebral infarction		<i>p</i> value		Cerebral hemorrhage		<i>p</i> value		Subarachnoid hemorrhage		<i>p</i> value	
	AOR	95 %CI			AOR	95 %CI			AOR	95 %CI		
	Lower	Upper			Lower	Upper			Lower	Upper		
Age	0.893	0.863	0.922	<0.001*	0.923	0.896	0.949	<0.001*	0.940	0.875	0.998	0.062
Gender					0.712	0.396	1.264	0.249				
Diabetes mellitus					0.589	0.288	1.208	0.147				
Length of stay	0.990	0.984	0.994	<0.001*								
NIHSS at acute admission	1.024	0.979	1.071	0.300								
FIM-C at acute admission	1.098	1.058	1.141	<0.001*								
FIM-M at acute admission	1.081	1.033	1.136	0.001*								
Bladder management at acute admission	0.875	0.672	1.146	0.321	4.412	1.989	18.278	0.008*				
Voiding location at acute admission	3.046	1.466	6.591	0.004*								
Number of drugs before stroke onset	1.033	0.967	1.104	0.341	0.922	0.823	1.033	0.159	1.291	0.953	1.941	0.144
Change in the number of drugs from acute to CRWs					0.930	0.836	1.033	0.176				
Continuous continence self-management program					0.885	0.504	1.554	0.668				
Prescribing drugs for voiding symptoms before stroke onset	0.354	0.144	0.855	0.022*	1.368	0.334	6.442	0.672				
Prescribing drugs for storage symptoms before stroke onset	0.252	0.073	0.858	0.027*	0.883	0.276	2.843	0.833				
Switched drugs for LUTS	2.827	1.593	5.144	0.001*	2.308	1.149	4.806	0.022*	1.479	0.103	38.646	0.778

AOR, adjusted odds ratio; CI, confidence interval; NIHSS, National Institutes of Health Stroke scale; FIM, functional independence measure; FIM-M, FIM motor score; FIM-C, FIM cognitive score; LUTS, lower urinary tract symptoms.

Each model was adjusted for variables with a *p*-value less than 0.15 in the univariate analysis for switching LUTS medications as covariates.

Continuous variables: age, length of hospital stay, NIHSS, FIM-C, FIM-M, bladder management at acute admission, and number of prescribed medications.

Categorical variables: bladder management (independence: 1, non-independence:0), voiding location (toilet: 1, on the bed: 0), gender (male: 1, female: 0), diabetes (yes: 1, no: 0), stroke type (subarachnoid hemorrhage, cerebral hemorrhage, cerebral infarction), prescription of medication for LUTS before stroke onset (yes: 1, no: 0), and switching drugs for LUTS (yes: 1, no: 0).

\* *p* value <0.05.

LUTS during the transition from the ACWs to the CRWs occurred in about 20 % of patients, and 2) switching was associated with independence of urination at discharge from CRWs. The findings were stable in the sensitivity and subgroup analysis.

The age, gender, and comorbidity rates of the patients in this study population were consistent with the results of a national survey of CRWs<sup>28</sup> and a previous report in Japan.<sup>29</sup> On the other hand, the proportion of patients with cerebral hemorrhage was about 5 % higher than that in the Japanese national survey of CRWs.<sup>28</sup>

Prescribed drugs for LUTS were switched in about 20 % of the patients during the hospitalization from the ACWs to the CRWs, which is smaller than the reported incidence of LUTS after stroke, which ranges from 21 % to 79 %.<sup>2,4-6</sup> This result suggested that urologists and other experts in continence care are not sufficiently involved in the continence care of post-stroke patients. However, it is also possible that non-pharmacological interventions, such as indwelling urinary catheters<sup>30</sup> and behavioral treatment<sup>5</sup> have been used for post-stroke LUTD, while the use of anticholinergics and cholinergics, which are used to treat LUTS, may have been approached cautiously in the elderly population due to concerns about possible adverse effects.<sup>31</sup> These non-pharmacological interventions and cautious approach to medication may have contributed to the small percentage of patients who switched their prescriptions.

This study showed that prescription drugs for voiding symptoms tended to increase during the ACWs, while prescription drugs for storage symptoms decreased during the ACWs and increased in the CRWs. Combination prescriptions tended to increase in the CRWs. Post-stroke, LUTS tends to feature voiding symptoms in the ACWs<sup>2,3</sup> and storage symptoms in the CRWs.<sup>4-6</sup> The patterns of prescribed drugs based on their mechanisms of action revealed a trend of initiating  $\alpha$ 1-blockers, cholinergic drugs, and their combination in the ACWs, and muscarinic receptor antagonists, and a combination of muscarinic receptor antagonists and  $\alpha$ 1-blockers in the CRWs. The combination of muscarinic receptor antagonists and  $\alpha$ 1-blockers is effective in treating overactive bladder associated with benign prostatic hyperplasia and neurogenic bladder.<sup>32,33</sup> When looking at the selection of individual drugs, among  $\alpha$ 1-blockers, particularly in males, the prescription of subtype-selective silodosin and tamsulosin, which are associated with lower risk of orthostatic hypotension and lower risk of falling in the elderly,<sup>34</sup> was common. In females, the prescription of urapidil, the only urethra resistance-reducing drug available for females,<sup>14</sup> was commonly observed. Among anticholinergic drugs, imidafenacin and solifenacin, which are selective for the M3 receptors involved in detrusor muscle contraction and are associated with fewer side effects such as dry mouth and cognitive decline compared to other anticholinergic drugs,<sup>35</sup> were commonly prescribed for both males and females. These observations suggest that LUTS therapeutic drugs were selected with appropriate consideration of gender and risk factors such as falling risk and cognitive impairment in the elderly. In the KRCCPS, the reasons for switching prescriptions were not recorded for all patients. However, in the records of some patients, the reasons for switching LUTS therapeutic drugs were given as changes in the patient's LUTS or the emergence of side effects. Furthermore, the prescription patterns of patients who switched LUTS therapeutic drugs matched the trend of changes in LUTS after stroke, suggesting that pharmacotherapy was given in line with the guidelines.<sup>21-23</sup> This implies that the switches in this study would have been made in response to individual LUTS and systemic conditions after stroke.

Our findings represent the first evidence of an association between switching prescribed drugs for LUTS during the hospitalization from the ACWs to the CRWs and independence of urination at discharge from CRWs in stroke patients. In a previous study, a continuous continence self-management program organized by a team led by a nurse was reported to improve toileting at discharge in stroke patients.<sup>13</sup> Pharmacist intervention, including recommendations for appropriate switching of drugs for LUTS, was associated with improved lower urinary tract

function in the ACWs.<sup>16</sup> In addition, urinary incontinence in CRWs was one of the factors that inhibited successful rehabilitation of stroke patients.<sup>36</sup> These reports are consistent with our present findings. We used two indices to assess independence of urination: FIM bladder management score and voiding location. The bladder management score is evaluated by considering the frequency of incontinence, the amount of assistance, and the use of medications and equipment. However, it is based on the amount of assistance, and not the actual voiding location. We considered that adding the voiding location allowed us to assess the independence of urination from the viewpoints of both caregivers and patients themselves. Switching medications for LUTS in post-stroke patients was also associated with FIM-M and FIM-C at convalescence discharge. It has been reported that LUTS in stroke patients is associated with lower cognitive and functional status, lower quality of life, and lower return-to-home rates.<sup>37,38</sup> This suggests that switching LUTS medications may have contributed to functional recovery by improving LUTS and actively promoting toileting-related behaviors.

Subgroup analysis was done to examine the association between the switching of LUTS therapeutic drugs and urinary independence, as well as the prescription status of LUTS therapeutic drugs by phase for patients with the three types of stroke, who differed in terms of age, severity, sex ratio, and history of LUTD. The prescription status of LUTS therapeutic drugs showed a different tendency depending on the type of stroke. The pre-onset prescription status of LUTS therapeutic drugs was characterized by more prescriptions for voiding symptoms in cerebral infarction and more prescriptions for storage symptoms in cerebral hemorrhage. This may be due to the differences in male-to-female ratios between patients with cerebral infarction and cerebral hemorrhage. After the acute phase, hemorrhagic stroke showed a greater increase in the proportion of therapeutic drug for voiding symptoms from before the onset compared to ischemic stroke, and a smaller increase in therapeutic drug for storage symptoms. This result is consistent with previous reports indicating a higher proportion of detrusor underactivity and a lower proportion of detrusor overactivity in patients with hemorrhagic stroke compared to those with ischemic stroke.<sup>19</sup> These results suggest that there may be different trends in the changes in LUTS by phase of stroke depending on gender and type of stroke. When the association between the switching of LUTS therapeutic drugs and urinary independence was examined, it was clear that the switching of LUTS therapeutic drugs was associated with excretion in the toilet for patients with cerebral infarction and cerebral hemorrhage. This suggests that while there may be different trends in LUTS depending on the type of stroke, the switching of LUTS therapeutic drugs may contribute to urinary independence after stroke, regardless of the cause or severity of the stroke. However, the KRCCPS data used in this study did not provide detailed information on the lesion site or size of the stroke, and subarachnoid hemorrhage had a small sample size and could not be adequately evaluated. Therefore, further research is needed to understand the trends in changes in LUTS and the effects of the switching of LUTS therapeutic drugs depending on the type of stroke.

Additional factors, such as older age, diabetes, extended length of stay, stroke severity, low ADL at admission, voiding location on the bed at admission, and a history of LUTD, were associated with a lack of improvement in independent urination at CRWs discharge. This aligns with previous reports on risk factors for post-stroke urinary incontinence<sup>39-41</sup> and the increased need for assistance with voiding behaviors.<sup>42</sup> In addition, our study found that, compared to hemorrhagic stroke, cerebral infarction was associated with voiding on the bed and a decline in FIM-C. This result is also consistent with prior studies indicating that patients with hemorrhagic stroke have a higher potential for functional recovery than those with ischemic stroke.<sup>43,44</sup>

Our study has some limitations. First, as the data were collected retrospectively, there may be confounding factors that were not analyzed, such as indices of LUTS function (presence or absence of urinary incontinence, frequency, and volume of urination) and the amount of rehabilitation time. However, the tendency of LUTS

transition after stroke and the change of prescription pattern of drugs for LUTS in this study were consistent. In addition, the rehabilitation unit per day for convalescent rehabilitation in Japan is defined by the medical reimbursement guidelines, so there should be no significant differences among individuals in Japan, which has a universal health insurance system.<sup>45</sup> Another possible limitation is that the unstructured data from KRCCPS may have omissions and thus may underestimate the number of patients whose medication was switched. In addition, the effects of dosage, duration of prescriptions, and different patterns of switching on urinary independence were not considered. Further studies using larger sample sizes or different databases would be needed to overcome this limitation. Moreover, the database used in this study covered only a part of Japan, though background factors of the study population were consistent with the results of a national survey of CRWs in Japan.<sup>28</sup> Finally, many patients were excluded because of missing data. However, the sensitivity analysis for 1460 patients showed similar results, supporting the robustness of our findings.

In conclusion, prescribed drugs for LUTS were switched in about 20 % of post-stroke patients during the hospitalization from the ACWs to the CRWs, and switching was associated with improved independence of urination at discharge from the CRWs. Consistent results were observed across different stroke types, suggesting that switching LUTS medications contributes to urinary independence after stroke, regardless of the etiology or severity of the stroke. Our results should be helpful in promoting seamless coordination of urinary care in stroke patients.

### Approval of the research protocol by an institutional reviewer board

The protocol for this research project has been approved by a suitably constituted Ethics Committee of the institution, and it conforms to the provisions of the Declaration of Helsinki. Committee of Kanazawa University, Approval No. 3580-2.

### Funding

This study was supported by research grant 20-03-184 from the Univers Foundation in 2020–2021 and JST SPRING, Grant number JPMJSP2135.

### Declaration of Competing Interest

The authors declare that they have no conflict of interest.

### Acknowledgments

We thank all the study participants for their cooperation and the Council of Kaga Local Stroke Network, South Ishikawa, Japan, for their support. This study was supported by research grant 20-03-184 from the Univers Foundation in 2020–2021 and JST SPRING, Grant number JPMJSP2135.

### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.jstrokecerebrovasdis.2023.107419](https://doi.org/10.1016/j.jstrokecerebrovasdis.2023.107419).

### References

- Takashima N, Arima H, Kita Y, et al. Incidence, management and short-term outcome of stroke in a general population of 1.4 million Japanese - Shiga stroke registry. *Circ J*. 2017;81(11):1636–1646.
- Doshi VS, Say JH, Young SH-Y, et al. Complications in stroke patients: a study carried out at the rehabilitation medicine service, Changi general hospital. *Singapore Med J*. 2003;44(12):643–652.
- Burney TL, Senapati M, Desai S, et al. Acute cerebrovascular accident and lower urinary tract dysfunction: a prospective correlation of the site of brain injury with urodynamic findings. *J Urol*. 1996;156(5):1748–1750.
- Cai W, Wang J, Wang L, et al. Prevalence and risk factors of urinary incontinence for post-stroke inpatients in Southern China. *Neurol Urodyn*. 2015;34(3):231–235.
- Thomas LH, Coupe J, Cross LD, et al. Interventions for treating urinary incontinence after stroke in adults. *Cochrane Database Syst Rev*. 2019;2, CD004462.
- Pizzi A, Falsini C, Martini M, et al. Urinary incontinence after ischemic stroke: clinical and urodynamic studies. *Neurol Urodyn*. 2014;33(4):420–425.
- Kawanabe E, Suzuki M, Tanaka S. Impairment in toileting behavior after a stroke. *Geriatr Gerontol Int*. 2018;18(8):1166–1172.
- Tseng C-N, Huang G-S, Yu P-J, et al. A qualitative study of family caregiver experiences of managing incontinence in stroke survivors. *PLoS One*. 2015;10(6), e0129540.
- Di Rosa M, Lamura G. The impact of incontinence management on informal caregivers' quality of life. *Aging Clin Exp Res*. 2016;28(1):89–97.
- Coyne KS, Kaplan SA, Chapple CR, et al. Risk factors and comorbid conditions associated with lower urinary tract symptoms: EpiLUTS. *BJU Int*. 2009;103(Suppl 3):24–32.
- Magari T, Sekiguchi Y, Furuya R, et al. What factors are associated with changes in lower urinary tract dysfunction in cerebral infarction patients? *JCS*. 2019;30(2):487–493.
- Hashimoto M, Hashimoto K, Ando F, et al. Prescription rate of medications potentially contributing to lower urinary tract symptoms and detection of adverse reactions by prescription sequence symmetry analysis. *J Pharm Health Care Sci*. 2015;1:7.
- Shogenji M, Yoshida M, Sumiya K, et al. Association of a continuous continence self-management program with independence in voiding behavior among stroke patients: a retrospective cohort study. *Neurol Urodyn*. 2022;41:1109–1120.
- Sekido N, Igawa Y, Kakizaki H, et al. Clinical guidelines for the diagnosis and treatment of lower urinary tract dysfunction in patients with spinal cord injury. *Int J Urol*. 2020;27(4):276–288.
- Yamada S, Ito Y, Nishijima S, et al. Basic and clinical aspects of antimuscarinic agents used to treat overactive bladder. *Pharmacol Ther*. 2018;189:130–148.
- Umemura T, Wakita E, Asano M, et al. Effectiveness of pharmaceutical support by pharmacists in urinary care teams. *J Pharm Health Care Sci*. 2019;5:11.
- Ogawa Y. The clinical role of pharmacists in supporting urinary continence care. *JCS*. 2019;30:384–387.
- Miyamoto S, Ogasawara K, Kuroda S, et al. Japan stroke society guideline 2021 for the treatment of stroke. *Int J Stroke*. 2022;17(9):1039–1049.
- Han K-S, Heo SH, Lee S-J, et al. Comparison of urodynamics between ischemic and hemorrhagic stroke patients; can we suggest the category of urinary dysfunction in patients with cerebrovascular accident according to type of stroke? *Neurol Urodyn*. 2010;29(3):387–390.
- Ikenaga Y, Kusunoki T, Yamaguchi H. Percutaneous endoscopic gastrostomy reduces aspiration pneumonia rate in stroke patients with enteral feeding in convalescent rehabilitation wards. *Prog Rehabil Med*. 2021;6, 20210031.
- The Japanese Continence Society. *Clinical Guidelines for Overactive Bladder*. 3rd ed. RichHill Medical; 2022.
- The Japanese Continence Society. *Japanese Clinical Guideline for Female Lower Urinary Tract Symptoms 2nd edition*. RichHill Medical; 2019.
- The Japanese Urological Association. *Clinical Guidelines for Male Lower Urinary Tract Symptoms and Benign Prostatic Hyperplasia*. 2017.
- Linacre JM, Heinemann AW, Wright BD, et al. The structure and stability of the Functional Independence Measure. *Arch Phys Med Rehabil*. 1994;75(2):127–132.
- Mizrachi EH, Waitzman A, Arad M, et al. Bladder management and the functional outcome of elderly ischemic stroke patients. *Arch Gerontol Geriatr*. 2011;53(2):e125–e128.
- Brott T, Adams Jr HP, Olinger CP, et al. Measurements of acute cerebral infarction: a clinical examination scale. *Stroke*. 1989;20(7):864–870.
- Bursac Z, Gauss CH, Williams DK, et al. Purposeful selection of variables in logistic regression. *Source Code Biol Med*. 2008;3:17.
- Association of Convalescence Phase Rehabilitation Units General Incorporated Association. Survey report on the current status and issues of convalescence phase rehabilitation wards, 2019, [http://plus1co.net/d\\_data/2019\\_zitai\\_book.pdf](http://plus1co.net/d_data/2019_zitai_book.pdf).
- Liu M, Tsuji T, Tsujiuchi K, et al. Comorbidities in stroke patients as assessed with a newly developed comorbidity scale. *Am J Phys Med Rehabil*. 1999;78(5):416–424.
- Shogenji M, Yuno C, Nakada H, et al. Continence care can prevent urinary tract infections among elderly patients by reducing the duration of indwelling urinary catheter use. *J Jpn WOCM*. 2015;19(3):336–345.
- Kouladjian O'Donnell L, Gnjdric D, Nahas R, et al. Anticholinergic burden: considerations for older adults. *Am J Pharmacogenomics*. 2017;47(1):67–77.
- Filson CP, Hollingsworth JM, Clemens JQ, et al. The efficacy and safety of combined therapy with  $\alpha$ -blockers and anticholinergics for men with benign prostatic hyperplasia: a meta-analysis. *J Urol*. 2013;190(6):2153–2160.
- Cameron AP, Clemens JQ, Latini JM, et al. Combination drug therapy improves compliance of the neurogenic bladder. *J Urol*. 2009;182(3):1062–1067.
- Kojima T, Mizukami K, Tomita N, et al. Screening tool for older persons' appropriate prescriptions for Japanese: report of the Japan geriatrics society working group on "Guidelines for medical treatment and its safety in the elderly. *Geriatr Gerontol Int*. 2016;16(9):983–1001.
- Sakakibara R, Hamano H, Yagi H. Cognitive safety and overall tolerability of imidafenacin in clinical use: a long-term, open-label, post-marketing surveillance study. *Low Urin Tract Symptoms*. 2014;6(3):138–144.
- Turhan N, Atalay A, Atabek HK. Impact of stroke etiology, lesion location and aging on post-stroke urinary incontinence as a predictor of functional recovery. *Int J Rehabil Res*. 2006;29(4):335–338.




- 37 Akkoç Y, Bardak AN, Ersöz M, et al. Post-stroke lower urinary system dysfunction and its relation with functional and mental status: a multicenter cross-sectional study. *Top Stroke Rehabil.* 2019;26(2):136–141.
- 38 Shogenji M, Yoshida M, Sumiya K, et al. Relationship between bowel/bladder function and discharge in older stroke patients in convalescent rehabilitation wards: a retrospective cohort study. *Prog Rehabil Med.* 2022;7, 20220028.
- 39 Patel M, Coshall C, Rudd AG, et al. Natural history and effects on 2-year outcomes of urinary incontinence after stroke. *Stroke.* 2001;32(1):122–127.
- 40 Nakayama H, Jørgensen HS, Pedersen PM, et al. Prevalence and risk factors of incontinence after stroke. The Copenhagen stroke study. *Stroke.* 1997;28(1):58–62.
- 41 Williams MP, Srikanth V, Bird M, et al. Urinary symptoms and natural history of urinary continence after first-ever stroke—A longitudinal population-based study. *Age Ageing.* 2012;41(3):371–376.
- 42 Sato A, Okuda Y, Fujita T, et al. Cognitive and physical functions related to the level of supervision and dependence in the toileting of stroke patients. *Phys Ther Res.* 2016; 19(1):32–38.
- 43 Chu C-L, Chen Y-P, Chen CCP, et al. Functional recovery patterns of hemorrhagic and ischemic stroke patients under post-acute care rehabilitation program. *Neuropsychiatr Dis Treat.* 2020;16:1975–1985.
- 44 Salvadori E, Papi G, Insalata G, et al. Comparison between ischemic and hemorrhagic strokes in functional outcome at discharge from an intensive rehabilitation hospital. *Diagnostics.* 2020;11(1):38.
- 45 Miyai I, Sonoda S, Nagai S, et al. Results of new policies for inpatient rehabilitation coverage in Japan. *Neurorehabil Neural Repair.* 2011;25:540–547.

RESEARCH ARTICLE

Open Access



# Validation of an automated sample preparation module directly connected to LC-MS/MS (CLAM-LC-MS/MS system) and comparison with conventional immunoassays for quantitation of tacrolimus and cyclosporin A in a clinical setting

Tsutomu Shimada<sup>1,2\*</sup> , Daisuke Kawakami<sup>3,4</sup>, Arimi Fujita<sup>1,2</sup>, Rintaro Yamamoto<sup>3</sup>, Satoshi Hara<sup>5</sup>, Kiyoaki Ito<sup>5</sup>, Ichiro Mizushima<sup>5</sup>, Shinji Kitajima<sup>6</sup>, Yasunori Iwata<sup>6</sup>, Norihiko Sakai<sup>6</sup>, Mitsuhiro Kawano<sup>5</sup>, Takashi Wada<sup>6</sup> and Yoshimichi Sai<sup>1,2,7</sup>

## Abstract

**Background** Therapeutic drug monitoring (TDM) systems generally use either liquid chromatography/tandem mass spectrometry (LC-MS/MS) or immunoassay, though both methodologies have disadvantages. In this study, we aimed to evaluate whether a CLAM-LC-MS/MS system, which consists of a sample preparation module directly connected to LC-MS/MS, could be used for clinical TDM work for immunosuppressive drugs in whole blood, which requires a hemolytic process. For this purpose, we prospectively validated this system for clinical measurement of tacrolimus and cyclosporin A in patients' whole blood. The results were also compared with those of commercial immunoassays.

**Methods** Whole blood from patients treated with tacrolimus or cyclosporin A at the Department of Nephrology and Departments of Rheumatology, Kanazawa University Hospital, from May 2018 to July 2019 was collected with informed consent, and drug concentrations were measured by CLAM-LC-MS/MS and by chemiluminescence immunoassay (CLIA) for tacrolimus and affinity column-mediated immunoassay (ACMIA) for cyclosporin A. Correlations between the CLAM-LC-MS/MS and immunoassay results were analyzed.

**Results** Two hundred and twenty-four blood samples from 80 patients were used for tacrolimus measurement, and 76 samples from 21 patients were used for cyclosporin A. Intra- and inter-assay precision values of quality controls were less than 7%. There were significant correlations between CLAM-LC-MS/MS and the immunoassays for tacrolimus and cyclosporin A (Spearman rank correlation coefficients: 0.861, 0.941,  $P < 0.00001$  in each case). The drug concentrations measured by CLAM-LC-MS/MS were about 20% lower than those obtained using the immunoassays.

\*Correspondence:

Tsutomu Shimada  
t-shimada@staff.kanazawa-u.ac.jp

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

CLAM-LC-MS/MS maintenance requirements did not interfere with clinical operations. Compared to manual pretreatment, automated pretreatment by CLAM showed lower inter-assay precision values and greatly reduced the pretreatment time.

**Conclusions** The results obtained by CLAM-LC-MS/MS were highly correlated with those of commercial immunoassay methods. CLAM-LC-MS/MS offers advantages in clinical TDM practice, including simple, automatic pretreatment, low maintenance requirement, and avoidance of interference.

**Keywords** CLAM-LC-MS/MS, Automatic pretreatment, Immunoassay, Tacrolimus, Cyclosporin A

## Background

Therapeutic drug monitoring (TDM) in a clinical setting enables patient-specific dosing, and is critical for drugs such as immunosuppressants, antiepileptic drugs, and anticancer drugs, for which the therapeutic range is narrow, and which have serious, concentration-dependent side effects and large inter-individual variability. There are two main types of analytical methods for general TDM: immunoassay methods that utilize specific antibodies and separation-analysis methods such as liquid chromatography/tandem mass spectrometry (LC-MS/MS). The advantages of commercial immunoassay include rapidity and automation, as well as a high level of technical support from manufacturers [1]. Disadvantages include potential for cross-reactions, high cost per sample, inability to simultaneously measure multiple compounds, and inter-laboratory variability [2, 3]. However, in recent years, new immunoassay methods have ameliorated some of these problems [4, 5]. In contrast, LC-MS/MS is considered a worldwide gold standard method for TDM because of its high sensitivity, high specificity, and suitability for multiplexing including metabolites [6–8]. Nevertheless, it has disadvantages such as unsuitability for direct injection, limited automation, time-consuming procedure, the need for manual deproteinization of samples, and the requirement of special training for users. To overcome some of these disadvantages, one-step protein precipitation methods such as Rapid-fire, liquid-handling platforms, on-line column extraction, and automated systems have been developed [9–13]. Nevertheless, to extend the availability of LC-MS/MS for clinical TDM, it is necessary to implement pretreatment automation, provide a user-friendly system, and ensure the availability of rapid technical support.

The CLAM-LC-MS/MS system, which is composed of an automated pretreatment device for blood (CLAM) directly connected to an LC-MS/MS instrument, has already been employed to measure drugs of abuse [14, 15], metabolite biomarkers [16], anti-HIV drugs [17], uracil and dihydrouracil [18], organic acids [19], and immunosuppressive drugs [20].

Immunosuppressive drugs bind extensively to red blood cells [21, 22], so a hemolysis process is required for clinical TDM work with whole blood samples. Therefore,

in this work we aimed to establish whether CLAM-LC-MS/MS could be employed for this purpose by using it for the quantitation of tacrolimus and cyclosporin A in whole blood. We also compared the results with those of conventional immunoassays.

## Methods

### Clinical samples

Inpatients and outpatients who were treated with tacrolimus (Prograf Capsules or Graceptor Capsules; Astellas Pharma US) or cyclosporin A (Neoral; Novartis Pharma) at the Department of Nephrology and Department of Rheumatology, Kanazawa University Hospital, from May 2018 to July 2019 were recruited with informed consent. The present study was performed after receiving approval from the Medical Ethics Committee of Kanazawa University (protocol no. 2017–195).

### Equipment

The CLAM-LC-MS/MS system consists of a CLAM-2000 CL automated sample pretreatment device, CBM-20A CL system controller, CTO-20AC CL column oven, SIL-30AC CL autosampler, two LC-30AD CL pumps, FCV-20AH2 CL valve, DGU-20A5R CL degasser and LCMS-8050 CL (Shimadzu Corporation, Kyoto, Japan), all of which are approved for use as medical equipment. A DOSIMMUNE trap column (trap column) and a DOSIMMUNE analytical column (Alsachim, France) were used in the LC system. Commercial DOSIMMUNE kits (Alsachim, France) including calibrators (6 concentrations), 4 QC samples, and stable isotope-labeled internal standards (ISs; [ $^{13}\text{C}_2$ ,  $^2\text{H}_4$ ]-everolimus, [ $^{13}\text{C}$ ,  $^2\text{H}_3$ ]-sirolimus, [ $^{13}\text{C}$ ,  $^2\text{H}_4$ ]-tacrolimus, and [ $^2\text{H}_{12}$ ]-cyclosporin A) for everolimus, sirolimus, tacrolimus and cyclosporin A were used respectively. The validation and acceptable ranges for QC samples are shown in supplemental Table 1.

A chemiluminescence immunoassay (CLIA) method was applied to measure tacrolimus using the Architect system (ARCHITECT i1000SR, Abbott, Illinois, U.S.A) with an ARCHITECT Tacrolimus Reagent Kit, ARCHITECT Tacrolimus Calibrators, ARCHITECT Tacrolimus Whole Blood Precipitation Reagent, and Multichem WBT (Technopath Clinical Diagnostics,

Ireland). An affinity column-mediated immune assay (ACMIA) method was applied to measure cyclosporin A, using Dimension® Xpand Plus (Siemens, Germany) with Dimension Systems CSA Assay, Dimension CSA Calibrator, and MORE RAP/Tac/CsA Control (More Diagnostics, CA, USA).

### Sample preparation

Whole blood collected from patients was divided into two tubes containing EDTA-Na. One was stored at -30°C for CLAM-LC-MS/MS measurement and the other was stored at 4°C for immunoassay. For application to the CLAM-LC-MS/MS system, the frozen whole blood samples (95 µL) were thawed at room temperature for hemolysis and placed in the CLAM unit. The CLAM unit was programmed to perform pretreatment, including sample extraction and protein precipitation followed by filtration and sample collection. First, 20 µL of 75% 2-propanol was dispensed onto the filter to activate the dedicated hydrophobic filter. Next, 20 µL of whole blood sample, 150 µL of extraction buffer, and 12.5 µL of ISs were dispensed into a dedicated vial and stirred for 60 s. The samples were vacuum-filtered (approximately 50 to 60 kPa) for 60 s in a dedicated filter consisting of a polytetrafluoroethylene membrane with a pore size of 0.45 µm and collected in a collection vial. Finally, the filtrate was automatically transported to the HPLC unit for LC-MS/MS analysis. The MS/MS conditions are shown in Table 1.

For measuring tacrolimus with CLIA, sample preparation was conducted according to the manufacturer's protocol. Briefly, Whole Blood Precipitation Reagent was added into the same volume (200 µL) of whole blood samples, QCs, or calibrators. After vortexing and centrifugation, the supernatant was immediately transferred to the Architect system. To measure cyclosporine A with ACMIA, whole blood samples (250 µL) were directly placed in the carousel of the Dimension system.

### Maintenance

Regular maintenance of the CLAM-LC-MS/MS system was performed every 6 months as follows. The flow path of LC, lens system of the mass spectrometer, and sample probe of CLAM were cleaned, and the capillary, desolvation line, and PEEK tube were replaced. Technical support was promptly available if required at other times.

### Statistical analysis

Statistical analyses were performed with MedCalc statistical software for Windows (Ostend, Belgium). Data collected by different assay methods were compared by means of Passing and Bablok regression analysis [23] and calculation of Spearman correlation coefficients. The Bland-Altman approach was used to compare immunological assay and CLAM-LC-MS/MS assay results by plotting the relative differences [24]. A p-value of 0.05 was considered statistically significant.

### Results

#### Blood samples

Between May 2018 and July 2019, 101 Japanese patients (18 in the Department of Nephrology and 83 in the Department of Rheumatology at Kanazawa University Hospital) gave informed consent to participate in this study. Finally, 224 whole blood samples were obtained from 80 patients treated with tacrolimus; 73 with Prograf (Astellas Pharma, Japan), and 7 with Graceptor (Astellas Pharma, Japan). Among these patients, 18 were inpatients and 62 were outpatients. In addition, 76 whole blood samples were obtained from 21 patients treated with cyclosporin A (8 were inpatients, 13 were outpatients), all of whom received Neoral (Novartis Japan). All whole blood samples were collected at times corresponding to a trough concentration.

#### CLAM-LC-MS/MS assay

The analysis time from placing the sample in the CLAM unit to obtaining the result with CLAM-LC-MS/MS was about 10 min; 6 min for pretreatment by CLAM, 2.6 min

**Table 1** Optimised mass spectrometry parameters of tacrolimus, cyclosporin A, and ISs in CLAM-LC-MS/MS

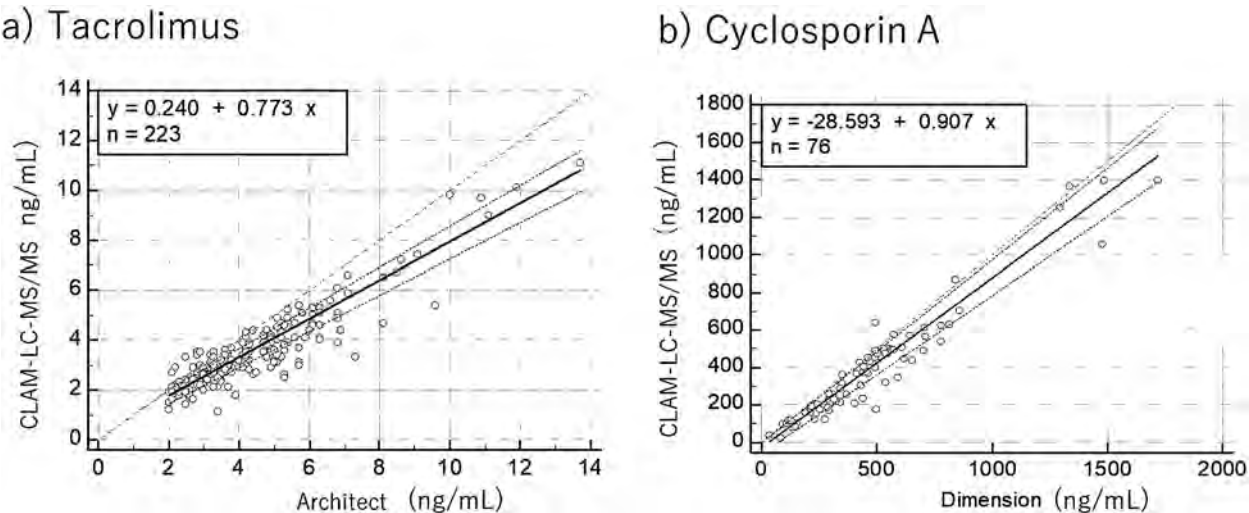
Compounds	Precursor ion		Quantitation			Product ion		Reference		Retention time (min)
	m/z	m/z	Q1 pre-bias (V)	Collision energy (V)	Q3 pre-bias (V)	m/z	Q1 pre-bias (V)	Collision energy (V)	Q3 pre-bias (V)	
Tacrolimus	821.5	768.3	-30	-22	-22	576.2	-30	-22	-30	0.735
[ <sup>13</sup> C, <sup>2</sup> H <sub>4</sub> ]-Tacrolimus	826.5	773.4	-30	-22	-22	581.4	-30	-22	-30	0.735
Cyclosporin A	1219.8	1202.9	-36	-19	-32	1184.6	-38	-35	-36	0.783
[ <sup>2</sup> H <sub>12</sub> ]-Cyclosporin A	1231.7	1214.9	-36	-19	-32	1196.8	-38	-35	-36	0.780



**Table 2** Results of the validation study of QCs (tacrolimus and cyclosporin A) in CLAM-LC-MS/MS

Compounds	Calibration curve Dynamic range (ng/mL)	r <sup>2</sup>	LLOQ (ng/mL)	Intra-assay precision (%) (%RSD) (n = 10)				Inter-assay precision (%) (%RSD) (10 days)			
				C1	C2	C3	C4	C1	C2	C3	C4
Tacrolimus	2.0-35.6	0.999	2.0	97.3 (6.5)	100.4 (3.5)	100.4 (2.6)	105.5 (3.6)	94.8 (3.1)	99.8 (4.7)	101.7 (4.4)	105.9 (5.1)
Cyclosporin A	24.3-1838.6	0.999	24.3	95.3 (2.5)	97.7 (1.6)	97.7 (1.9)	104.3 (1.6)	98.8 (4.0)	99.0 (3.7)	96.6 (3.6)	98.7 (4.5)

LLOQ: low limit of quantification, RSD: relative standard deviation



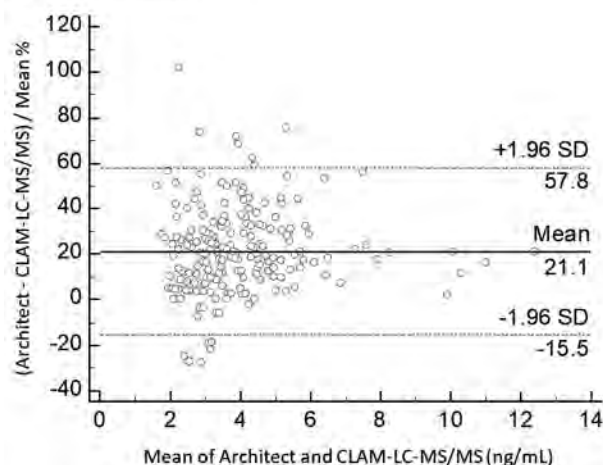
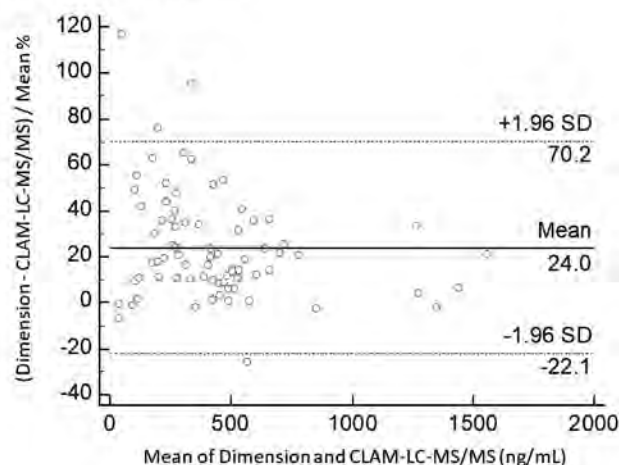
**Fig. 1** Passing-Bablok correlations between the CLAM-LC-MS/MS and immunoassays. **a)** Tacrolimus (n = 223). **b)** Cyclosporin A (n = 76). The regression lines are shown as solid lines, and the 95% confidence limits are shown as dashed lines. The dotted lines are the lines of identity

for LC-MS/MS analysis, and 30 s for result output. In the case of multiple continuous measurements, since pretreatment of each subsequent sample with CLAM was performed during the LC-MS/MS analysis of the previous sample, a measurement could be performed about every 4 min from the second sample on. For example, the measurement of a total of 40 samples including 20 actual samples, 6 calibration standards and 4 QCs (measured twice, before and after sample measurement) took approximately 180 min, while CLIA took about 60 min and ACMIA took about 40 min to measure the same sample number. Pretreatment of whole blood manually instead with CLAM in the LC-MS/MS method took 90 min longer (20 samples). The intra- and inter-assay precision values of four QCs of tacrolimus and cyclosporin A were less than 7% (Table 2). The values of carry-over of LLOQ (lower limit of quantitation) were 4.4% for tacrolimus and 9.6% for cyclosporin A. The measured concentrations of QCs in DOSIMMUNE kits of tacrolimus measured by the CLIA method were within the acceptable range ( $\pm 20\%$ ). The inter-assay precision values of QCs of tacrolimus in the case of manual pretreatment methods were higher (QC1 15.3%, QC2 13.4, QC3 15.4%, QC4 17.2%) than those obtained in the case of automated pretreatment by CLAM (Table 2).

**Correlation between CLAM-LC-MS/MS and immunoassay results**

The correlations between CLAM-LC-MS/MS and immunoassay results for tacrolimus and cyclosporin A were analyzed in terms of the Passing-Bablok correlation and Bland-Altman plots (Figs. 1 and 2). A significant correlation was obtained between CLAM-LC-MS/MS and CLIA for tacrolimus (Spearman rank correlation coefficient: 0.861,  $P < 0.00001$ ). The Passing-Bablok intercept of the linear regression was 0.2398 (95% CI: 0.01818 to 0.3828), and the slope was 0.773 (95% CI: 0.7241 to 0.8182). Constant error and proportional error were observed, and the concentration data obtained by CLAM-LC-MS/MS were 21.1% (from  $-15.5$  to  $57.8\%$ ) lower than those measured by CLIA. Large variations were observed, especially in the low concentration range.

A significant correlation was also obtained between CLAM-LC-MS/MS and ACMIA for cyclosporin A (Spearman rank correlation coefficient: 0.941,  $P < 0.00001$ ). The Passing-Bablok intercept of the linear regression was  $-28.5933$  (95% CI:  $-58.5836$  to  $-6.6225$ ), and the slope was 0.9066 (95% CI: 0.8425 to 0.9786). Constant error and proportional error were observed, and the concentration data obtained by CLAM-LC-MS/MS were 24% (from  $-22.1$  to  $70.2\%$ ) lower than those

**a) Tacrolimus****b) Cyclosporin A**

**Fig. 2** Bland-Altman plots to identify relative differences between CLAM-LC-MS/MS and immunoassays. **a)** Tacrolimus ( $n=223$ ). **b)** Cyclosporin A ( $n=76$ ). Mean differences are shown as solid lines and 95% confidence limits are shown as dashed lines

obtained by ACMIA. Large variations were observed, especially in the low concentration range.

**Maintenance**

Regular maintenance was performed every 6 months and completed within 2 days. Incidental problems, such as clogging of the sampling probe or position adjustment error of the sampling probe, that occurred outside scheduled maintenance were dealt with by technical support within one day.

**Discussion**

Measurement of the whole blood concentration of immunosuppressive drugs in clinical samples requires hemolysis and pretreatment processes. Our results in this study show that CLAM-LC-MS/MS offers advantages in clinical TDM practice, including simple, automatic pretreatment, low maintenance requirements, and avoidance of interference. We also confirmed that the values measured by CLAM-LC-MS/MS were highly correlated with those obtained by immunological assays.

Compared to manual pretreatment of whole blood samples, automated pretreatment by CLAM showed lower inter-assay precision values of the QCs and greatly reduced the pretreatment time. Thus, automated pretreatment by CLAM should be advantageous in the application of LC-MS/MS for clinical TDM. The CLAM-LC-MS/MS system requires hemolysis of blood samples by pre-freezing and thawing, but once the sample is placed in the CLAM unit, all subsequent processing is automatically performed, eliminating variability due to differences in the operator's experience. The CLIA method for measurement of tacrolimus requires a pretreatment process that includes the addition of deproteinizing agent, mixing, and centrifugation. On the

other hand, the ACMIA method enables measurement of cyclosporin A concentration in whole blood, so that samples can be directly placed in the carousel. However, ACMIA has disadvantages, such as high measurement error and low minimum detection sensitivity (about 5 times higher than CLIA), suggesting that the elimination of hemolysis and deproteinization processes may not be appropriate for TDM [2, 25]. In this study, we conducted hemolysis by freezing the whole blood samples at  $-30^{\circ}\text{C}$  overnight and thawing them at room temperature, but this might not be suitable for clinical samples, where rapid feedback to clinical staff is important for prompt dosing design. Recently, a hemolysis method for whole blood that involves freezing at  $-80^{\circ}\text{C}$  for just 10 min and thawing in running water for 5 min was reported and was employed to measure tacrolimus by CLAM-LC-MS/MS [20]. Furthermore, since there are several hemolysis methods [26], it will be necessary to optimize and validate the hemolysis method for the measurement of tacrolimus and cyclosporin A by CLAM-LC-MS/MS in clinical practice.

We found that the blood levels of tacrolimus and cyclosporin A in clinical samples measured by CLAM-LC-MS/MS and immunoassay were significantly correlated. However, the Bland-Altman plot for both tacrolimus and cyclosporin A indicated that the concentrations measured by CLAM-LC-MS/MS were 20% lower than those by the immunoassays. It is well known that immunoassays often show cross-reactivity with metabolites of parent compounds and the extent of the cross-reactivity is variable, depending on the antibodies used [27]. In previous studies, LC-MS/MS also showed a 10–20% lower value than immunoassay [28, 29], in agreement with the present finding. This difference should be noted when updating from immunoassay to LC-MS/MS in a clinical

setting, or when a patient moves to another facility where a different method is in use.

Appropriate maintenance is also a critical point for TDM, since continuous availability of analysis is clinically important. In this study using whole blood samples, increased column pressure and clogging in the sampling probe each occurred once, but did not occur after modifying the washing process to include alkali detergent, indicating that this modification successfully enabled the use of CLAM-LC-MS/MS for whole blood samples. The combination of regular maintenance and prompt technical support to deal with incidental problems was effective to maintain availability of the instrument.

Furthermore, the commercial CLAM-LC-MS/MS system is provided with standards, QCs, and stable isotope ISs, but it remains necessary to conduct multi-facility studies to confirm consistency across multiple facilities in the future.

## Conclusions

CLAM-LC-MS/MS showed high robustness for the measurement of tacrolimus and cyclosporin A in whole blood samples, which require hemolysis and pretreatment, and the results were highly correlated with those of conventional immunoassay systems. Importantly, the commercial CLAM-LC-MS/MS system, which includes standards, QCs, ISs, and technical support, is user-friendly, and pretreatment is automatic, so that the required technical expertise is minimal. Our results suggest that CLAM-LC-MS/MS would be suitable for routine TDM operation, and would reduce the burden on clinical pharmacists, though further multi-facility validation remains necessary.

## List of abbreviations

LC-MS/MS	Liquid chromatography/tandem mass spectrometry
CLAM	Automated pretreatment device for blood
TDM	Therapeutic drug monitoring
QC	Quality control
IS	Internal standard
CLIA	Chemiluminescence immunoassay
ACMIA	Affinity column-mediated immune assay
LLOQ	Low limit of quantitation
LOD	Limit of detection
RSD	Relative standard deviation

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40780-023-00318-6>.

Supplementary Material 1

## Acknowledgements

We thank Ms. Masami Matsuo and Ms. Mariko Yasuda for collecting patients' samples and for assistance with measurements.

## Authors' contributions

Tsutomu Shimada: Conceptualization, Methodology, Writing—Original Draft, Daisuke Kawakami: Conceptualization, Methodology, Arimi Fujita:

Investigation, Rintaro Yamamoto: Conceptualization, Satoshi Hara: Resources, Kiyooki Ito: Resources, Ichiro Mizushima: Resources, Shinji Kitajima: Resources, Yasunori Iwata: Resources, Norihiko Sakai: Resources, Mitsuhiro Kawano: Resources, Supervision, Takashi Wada: Resources, Supervision, Yoshimichi Sai: Conceptualization, Supervision.

## Funding

This study was funded by Shimadzu Corporation.

## Data Availability

Not applicable.

## Declarations

### Ethics approval and consent to participate

All patients were recruited with informed consent. The present study was performed after receiving approval from the Medical Ethics Committee of Kanazawa University (protocol no. 2017–195).

### Consent for publication

The consent for publication has been obtained from all patients.

### Competing interests

This study was funded by Shimadzu Corporation.

### Author details

<sup>1</sup>Department of Clinical Pharmacokinetics, Graduate School of Medical Sciences, Kanazawa University, Kanazawa, Ishikawa, Japan

<sup>2</sup>Department of Hospital Pharmacy, University Hospital, Kanazawa University, 13-1 Takara-machi, Kanazawa, Ishikawa 920-8641, Japan

<sup>3</sup>Shimadzu Corporation, Kyoto, Japan

<sup>4</sup>Shimadzu Europa GmbH, Duisburg, Germany

<sup>5</sup>Department of Rheumatology, Graduate School of Medical Sciences, Kanazawa University, Kanazawa, Ishikawa, Japan

<sup>6</sup>Department of Nephrology and Laboratory Medicine, Faculty of Medicine, Institute of Medical, Pharmaceutical and Health Sciences, Graduate School of Medical Sciences, Kanazawa University, Kanazawa, Ishikawa, Japan

<sup>7</sup>AI Hospital/Macro Signal Dynamics Research and Development Center, Kanazawa University, Kanazawa, Ishikawa, Japan

Received: 24 July 2023 / Accepted: 13 November 2023

Published online: 08 January 2024

## References

1. Freudenberger K, Hilbig U, Gauglitz G. Recent advances in therapeutic drug monitoring of immunosuppressive Drugs. *Trac-Trend Anal Chem.* 2016;79:257–68.
2. Wallemacq P, Armstrong VW, Brunet M, et al. Opportunities to optimize tacrolimus therapy in solid organ transplantation: report of the European consensus conference. *Ther Drug Monit.* 2009;31:139–52.
3. Knorr JP, Grewal KS, Balasubramanian M, et al. Falsely elevated tacrolimus levels caused by immunoassay interference secondary to beta-galactosidase antibodies in an infected liver transplant recipient. *Pharmacotherapy.* 2010;30:954.
4. Zhang L, Long H, Li X, et al. Production of a sensitive antibody against sirolimus for chemiluminescence immunoassay potential in its therapeutic drug monitoring. *Anal Methods.* 2016;8:6298–304.
5. Menotta M, Biagiotti S, Streppa L, et al. Label-free quantification of Tacrolimus in biological samples by atomic force microscopy. *Anal Chim Acta.* 2015;884:90–6.
6. Christians U, Vinks AA, Langman LJ, et al. Impact of Laboratory practices on Interlaboratory variability in therapeutic drug monitoring of immunosuppressive Drugs. *Ther Drug Monit.* 2015;37:718–24.
7. McShane AJ, Bunch DR, Wang S. Therapeutic drug monitoring of immunosuppressants by liquid chromatography-mass spectrometry. *Clin Chim Acta.* 2016;454:1–5.
8. Polledri E, Mercadante R, Fusarini CF, et al. Immunosuppressive Drugs in whole blood: validation of a commercially available liquid chromatography/

- tandem mass spectrometry kit and comparison with immunochemical assays. *Rapid Commun Mass Spectrom*. 2017;31:1111–20.
9. Grote-Koska D, Czajkowski S, Brand K. Performance of the New RapidFire System for Therapeutic Monitoring of immunosuppressants. *Ther Drug Monit*. 2015;37:400–4.
  10. Marinova M, Artusi C, Brugnolo L. Immunosuppressant therapeutic drug monitoring by LC-MS/MS: workflow optimization through automated processing of whole blood samples. *Clin Biochem*. 2013;46:1723–7.
  11. Bittersohl H, Schniedewind B, Christians U, et al. A simple and highly sensitive on-line column extraction liquid chromatography-tandem mass spectrometry method for the determination of protein-unbound tacrolimus in human plasma samples. *J Chrom A*. 2018;1547:45–52.
  12. Gong ZS, Wu ZH, Xu SX, et al. A high-throughput LC-MS/MS method for the quantification of four immunosuppressants Drugs in whole blood. *Clin Chim Acta*. 2019;498:21–6.
  13. Hörber S, Peter A, Lehmann R, et al. Evaluation of the first immunosuppressive drug assay available on a fully automated LC-MS/MS-based clinical analyzer suggests a new era in laboratory medicine. *Clin Chem Lab Med*. 2020;59:913–20.
  14. Robin T, Barnes A, Dulaurent S, et al. Fully automated sample preparation procedure to measure Drugs of abuse in plasma by liquid chromatography tandem mass spectrometry. *Anal Bioanal Chem*. 2018;410:5071–83.
  15. Robin T, El Balkhi S, Dulaurent S, et al. First-line toxicological screening with fully automated extraction. *J Anal Toxicol*. 2021;45:252–68.
  16. Nishiumi S, Shima K, Azuma T, et al. Evaluation of a novel system for analyzing hydrophilic blood metabolites. *J Biosci Bioeng*. 2017;123:754–9.
  17. Masuda J, Kato Y. Simultaneous determination of Raltegravir, Dolutegravir, Elvitegravir, and Bictegravir in Human plasma using high-performance liquid chromatography-tandem Mass Spectrometry. *Showa Univ J Med Sci*. 2020;32:91–102.
  18. Robin T, Saint-Marcoux F, Toinon D, et al. Automatic quantification of uracil and dihydrouracil in plasma. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2020;1142:122038.
  19. Ueyanagi Y, Setoyama D, Kawakami D, et al. Fully automated quantitative measurement of serum Organic acids via LC-MS/MS for the diagnosis of Organic Acidemias: establishment of an automation system and a proof-of-Concept Validation. *Diagnostics*. 2021;11:2195.
  20. Takane M, Tanabiki A, Kaneko Y, et al. Performance evaluation of the LC-MS/MS method for measuring immunosuppressive Drugs using the CLAM TM-2030. *Jap J Transplant*. 2021;56:15–23.
  21. Atkinson K, Britton K, Biggs J. Distribution and concentration of cyclosporin in human blood. *J Clin Pathol*. 1984;37:1167–71.
  22. Zahir H, Nand RA, Brown KF, et al. Validation of methods to study the distribution and protein binding of tacrolimus in human blood. *J Pharmacol Toxicol Methods*. 2001;46:27–35.
  23. Passing H, Bablok A. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, part I. *J Clin Chem Clin Biochem*. 1983;21:709–20.
  24. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–10.
  25. Amann S, Parker TS, Levine DM. Evaluation of 2 immunoassays for monitoring low blood levels of tacrolimus. *Ther Drug Monit*. 2009;31:273–6.
  26. Kalaria T, Gill H, Sharrod-Cole H, Ford C, Gama R. Conflicting effects of haemolysis on plasma sodium and chloride are due to different haemolysis study protocols: a case for standardisation. *Ann Clin Biochem*. 2022;59:101–9. <https://doi.org/10.1177/00045632211040691>.
  27. Hashi S, Masuda S, Kikuchi M, et al. Assessment of four methodologies (microparticle enzyme immunoassay, chemiluminescent enzyme immunoassay, affinity column-mediated immunoassay, and flow injection assay-tandem mass spectrometry) for measuring tacrolimus blood concentration in Japanese liver transplant recipients. *Transpl Proc*. 2014;46:758–60.
  28. Kaneko T, Fujioka T, Suzuki Y, et al. Comparison of whole-blood tacrolimus concentrations measured by different immunoassay systems. *J Clin Lab Anal*. 2018;32:e22587.
  29. Shigematsu T, Suetsugu K, Yamamoto N, et al. Comparison of 4 commercial immunoassays used in measuring the concentration of Tacrolimus in blood and their cross-reactivity to its metabolites. *Ther Drug Monit*. 2020;42:400–6.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



## Review

# Research on the Development of Methods for Detection of Substandard and Falsified Medicines by Clarifying Their Pharmaceutical Characteristics Using Modern Technology

Naoko Yoshida

*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.*

Received October 31, 2023

The existence of substandard and falsified medicines threatens people's health and causes economic losses as well as a loss of trust in medicines. As the distribution of pharmaceuticals becomes more globalized and the spread of substandard and falsified medicines continues worldwide, pharmaceutical security measures must be strengthened. To eradicate substandard and falsified medicines, our group is conducting fact-finding investigations of medicines distributed in lower middle-income countries (LMICs) and on the Internet. From the perspective of pharmaceuticals, such as physical assessment of medicines, we are working to clarify the actual situation and develop methods to detect substandard and falsified medicines. We have collected substandard and falsified medicines distributed in LMICs and on the Internet and performed pharmacopoeial tests, mainly using HPLC, which is a basic analytic method. In addition to quality evaluation, we have evaluated the applicability of various analytic methods, including observation of pharmaceuticals using an electron microscope, Raman scattering analysis, near-IR spectroscopic analysis, chemical imaging, and X-ray computed tomography (CT) to detect substandard and falsified medicines, and we have clarified their limitations. We also developed a small-scale quality screening method using statistical techniques. We are engaged in the development of methods to monitor the distribution of illegal medicines and evolve research in forensic and policy science. These efforts will contribute to the eradication of substandard and falsified medicines. Herein, I describe our experience in the development of detection methods and elucidation of the pharmaceutical status of substandard and falsified medicines using novel technologies.

**Key words** substandard medicine, falsified medicine, pharmaceutical security, X-ray computed tomography, non-destructive spectroscopy, chemical imaging

## 1. INTRODUCTION

The problem of substandard and falsified medicines is an issue that must be solved on a global scale. WHO defines substandard medicines, also called “out of specification,” as authorized medical products that fail to meet their quality standards or specifications, or both. Falsified medical products are those with deliberately misrepresented or fraudulent identity, composition, or source.<sup>1)</sup> Substandard and falsified medicines are present in every region of the world. Their existence causes not only health damage but also economic damage, and the situation is becoming more serious with the globalization of supply chains.

The distribution of substandard and falsified medicines is increasing globally, especially in lower middle-income countries (LMICs), where the proportion of substandard and falsified medicines is as high as 10.5% of the pharmaceutical market.<sup>2)</sup> Pharmaceutical regulatory authorities, pharmaceutical companies, and researchers in LMICs are working to deter the distribution of substandard and falsified medicines,<sup>3)</sup> but the current situation regarding the distribution of pharmaceutical products requires further improvement. Japan also faces

the threat of falsified medicines. Worryingly, substandard and falsified medicines that have already been identified only represent the tip of the iceberg.<sup>3)</sup>

To exploit the wishes of consumers and patients, the targets of falsification change daily, and falsified medicines have become more sophisticated. Additionally, the profits from falsified medicines are considered to be a source of funds for criminals, and there is concern that this will lead to further expansion of criminal activity in this area. As the distribution of pharmaceuticals becomes more globalized and the spread of substandard and falsified medicines continues worldwide, the development of methods to detect substandard and falsified medicines and deter their inappropriate distribution is needed to strengthen pharmaceutical security measures. Under these circumstances, I am working to strengthen pharmaceutical security measures with the aim of protecting patients from these risks, focusing on the development of detection technology (Fig. 1). To develop technology for the detection of substandard and falsified medicines, we analyzed pharmaceutical products that were actually distributed using a variety of methods, including observation that does not require equipment, non-destructive analysis that is possible even after analysis, and destructive analysis in which we mainly use HPLC. From the perspectives of evaluating the physical properties of pharmaceuticals, forensic science, and regulatory science, I aim to eliminate substandard and falsified medicines, the in-

This review of the author's work was written by the author upon receiving the 2023 Pharmaceutical Society of Japan Incentive Award for Women Scientists.

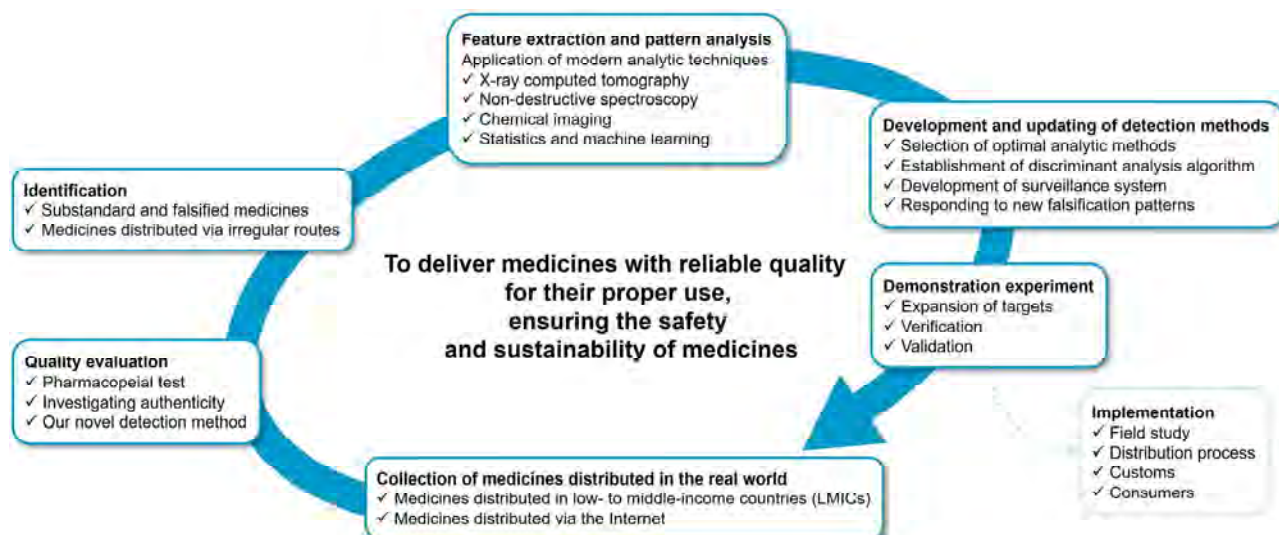


Fig. 1. Schematic Diagram of the Strengthening of Pharmaceutical Security Measures on Which I Am Working

I am working to strengthen pharmaceutical security measures, focusing on the development of detection technology with the aim to deliver medicines with reliable quality, ensuring their proper use and realizing the sustainability of medicines.

appropriate distribution of medicines, and to meet the needs of this field. My goal is to realize the sustainability of medicines, deliver medicines with reliable quality, and ensure their proper use. This will ultimately lead to the proper use of medicines and greater health benefits from more effective and safer pharmaceutical treatment.

My research on substandard and falsified medicines began with an investigation of the status of medicines available on the market in LMICs as well as proposals for improvement measures based on that investigation.<sup>4-7)</sup> To solve the problem of falsified medicines in Japan, our group also investigated the status of medicines distributed on the Internet.<sup>8,9)</sup> Our work clarified the fact that “health products” obtained *via* the Internet advertising a weight-loss effect, which claimed to only contain ingredients derived from natural products, actually contained the unapproved pharmaceutical ingredient sibutramine whose approval was withdrawn owing to serious side effects. Our findings called attention to the easy personal import of such health products.

Through the studies described above, it was possible to grasp the epidemiological situation regarding the problem of substandard and falsified medicines. However, the nature of

low-quality and falsified medicines remained unclear. Moreover, techniques to detect substandard and falsified medicines, which is essential for deterring their distribution and preventing harm to health, had not yet been developed. This was the impetus for research on the development of methods to detect substandard and falsified medicines based on clarifying their pharmaceutical characteristics using modern technology. To accomplish this, it is essential to grasp the actual situation regarding medicines that are available on the market, which requires great effort. We continued to investigate the quality of medicines distributed in LMICs and personally imported medicines<sup>10-19)</sup> and to develop detection methods.<sup>20-30)</sup> In addition to demonstrating the applicability of spectroscopic analysis, some results obtained in these studies could be extremely valuable from the perspective of contributing to planning countermeasures that match the actual situation by analyzing currently available medicines that are substandard and falsified. The main research outcomes are described below.

### Biography

Dr. Naoko Yoshida graduated from the Toyama Medical and Pharmaceutical University Faculty of Pharmacy and Graduate School of Pharmaceutical Sciences, where she received her master's degree in 2005 and Ph.D. in Clinical Pharmacy in 2008. She worked as a pharmacist at the Department of Hospital Pharmacy, Hamamatsu University Hospital and Drug Management and Policy, Faculty of Pharmacy, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, where she undertook research for countermeasures against substandard and falsified medicines. Dr. Yoshida transferred to Clinical Pharmacy and Healthcare Sciences in 2019 and subsequently to AI Hospital/Macro Signal Dynamics Research and Development Center, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, where she is currently an Assistant Professor. She is board-certified as a pharmacist and supervisor pharmacist by the Japan Society of Clinical Pharmacology and Therapeutics. She received the Nakamura Prize for Women Researchers in Kanazawa University in 2015 and Pharmaceutical Society of Japan Incentive Award for Women Scientists in 2023. Her research focus is eradicating substandard and falsified medicines and deterring irregular distribution of pharmaceuticals to ensure pharmaceutical safety and security.



Naoko Yoshida

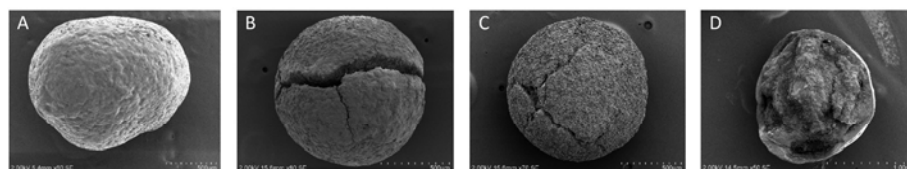


Fig. 2. Electron Microscopy Image of Intracapsular Granules with Poor Dissolution in Omeprazole Enteric-Coated Capsules

In the genuine innovator product (A), a smooth film was observed on the surface. In contrast, the three samples with poor dissolution that were distributed in a lower middle-income country (B–D) showed no gloss on the surface and it had cracks (B, C); although there appeared to be a film coating, the granules were broken (D).

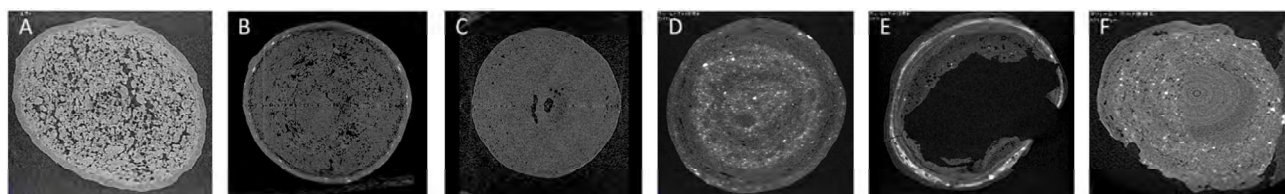


Fig. 3. Microfocus X-Ray Computed Tomography (CT) Cross-Sectional Image of Intracapsular Granules with Poor Dissolution in Omeprazole Enteric-Coated Capsules

A) Surface of granules in the genuine innovator product; B–F) surface of granules in the two generic samples with poor dissolution distributed in lower middle-income countries. Granules B and C were in a capsule from one sample; D, E, and F were in a capsule from another sample.

## 2. CLARIFICATION OF THE PHARMACEUTICAL CHARACTERISTICS OF SUBSTANDARD MEDICINES

**2.1. Observation of Enteric-Coated Granules Using Electron Microscopy and X-Ray Computed Tomography (CT)** To elucidate the pharmaceutical status of substandard medicines, we applied electron microscopy and X-ray CT in pharmaceutical analysis. We sought to visualize the formulation structure of pharmaceutical products with poor dissolution that are distributed in LMICs.<sup>7,13)</sup> We clarified that omeprazole capsules with severe dissolution failure lacked the enteric coating that should have been present in the granules inside the capsule.<sup>20)</sup> As a result of observing the intracapsular granules of enteric-coated omeprazole formulations with poor dissolution using an electron microscope, a smooth film was observed on the surface in the genuine innovator product (Fig. 2A). In contrast, there were no glossy surfaces and there were cracks found, among the intracapsular granules with poor dissolution that were distributed in an LMIC (Figs. 2B, C). Additionally, although there appeared to be a film coating, the granules were broken (Fig. 2D). In observing the granules inside the capsule of enteric-coated omeprazole using X-ray CT, we also found that the enteric coating, which can be seen in the genuine innovator product, was missing in some samples with poor dissolution distributed in LMICs. In the innovator product, the color of granules was uniform and a film coating was observed on the surface of the granule (Fig. 3A). In samples with poor dissolution distributed in LMICs, the color of granules within the capsules was not uniform. One capsule contained pale yellow granules with a film coating (Fig. 3B) and white granules without a film coating (Fig. 3C). In another capsule, we found white granule that appeared to have a film coating (Fig. 3D), white granules with no film coating and that were broken (Fig. 3E), and pale yellow granules without a film coating (Fig. 3F). Although differences in the components present on the surface of the granules inside the capsule of omeprazole tablets could be detected using Raman scattering analysis, X-ray CT clarified the cause of the difference in Raman spectra. Additionally, voids and electron-dense

aggregates such as metal elements that were not observed in authentic candesartan tablets, could be seen in falsified tablets using X-ray CT.<sup>21)</sup> As a result, the existence of pharmaceutical defects that had been overlooked was revealed, making it possible to propose quality improvement measures based on scientific evidence. By accumulating this information and clarifying the relationship with falsified medicines, even when the authenticity of a product is unknown, this approach can be used to detect falsified medicines.

**2.2. Visualization of the Ingredient Distribution in Tablets Using Chemical Imaging** The actual physical properties of substandard and falsified medicines cannot be determined by only identifying differences from genuine products. We thought that if we could determine differences from the genuine product owing to the manufacturing process and further identify the raw materials used, we would be able to obtain clues for detecting substandard and falsified products. Therefore, we applied chemical imaging to visualize the component distribution and the internal structure of tablets. Thus, near-IR (NIR) imaging was used to visualize the ingredient distribution in tablets. In roxithromycin tablets with poor dissolution and poor disintegration,<sup>11)</sup> we found that although the content of the active ingredient was appropriate, some ingredients were aggregated and dispersed unevenly.<sup>26)</sup> We scanned the NIR spectrum of a flat cross section obtained by shaving the surface of roxithromycin tablets, and images were generated based on the first principal component score obtained using principal component analysis (PCA) of the standardized NIR spectrum data (Fig. 4). PCA was used because the excipients contained in tablets distributed in LMICs are not informed. In the innovator product, which had no problems in quality tests, the difference was small when comparing the NIR spectra of the red part with a high PCA score and the blue part with a low PCA score; this suggested no agglomeration of ingredients and uniform dispersion (Fig. 4A). However, in some generic medicines with poor dissolution that were distributed in one LMIC, large differences were observed in the NIR spectra of each part, indicating that the dispersion of ingredients was non-uniform and that there were parts where the components



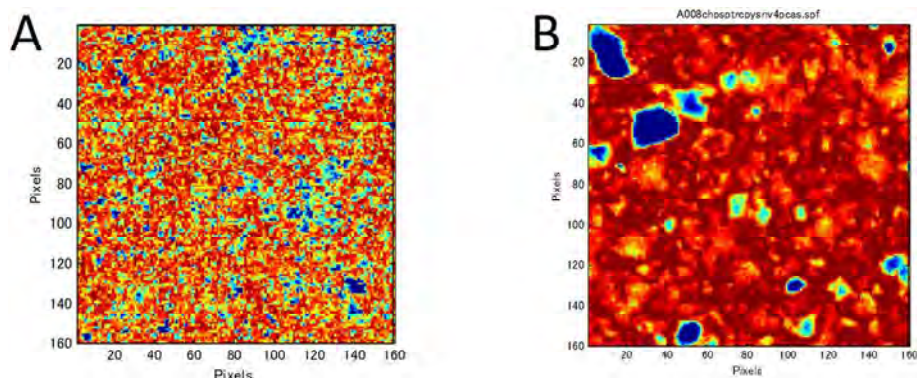


Fig. 4. The Inside of a Roxithromycin Tablet Visualized Using Near-IR (NIR) Imaging

An innovator product in which the components are finely dispersed (A), and a generic product in which the components are aggregated (B). The principal component analysis (PCA) score of the NIR spectrum obtained from each pixel is shown as the hue saturation value between red and blue. Higher PCA scores are redder in color, and lower PCA scores are bluer in color. The greater the color contrast, the clearer the bias in component dispersion.

aggregated. The aggregates were confirmed by comparing the NIR spectrum of the part where aggregation was observed with the NIR spectra of the active pharmaceutical ingredient and the major excipients, revealing the existence of large aggregates of magnesium stearate and cornstarch (Fig. 4B). We also used Raman imaging to identify the aggregated components that were difficult to identify with NIR spectra.

The above findings highlighted the need to review the manufacturing process. These results also indicated the possibility of evaluating substandard and falsified medicines based on their formulation characteristics. Our research results yielded valuable information on the pharmaceutical situation in LMICs that can greatly contribute to the maintenance and improvement of pharmaceutical quality in these regions.

### 3. IDENTIFICATION OF FALSIFIED MEDICINES USING RAMAN SCATTERING ANALYSIS

In research on the development of detection methods for substandard and falsified medicines, we focused on non-destructive analysis from the viewpoint of speed and simplicity, and we evaluated the applicability of NIR spectroscopy and Raman scattering analysis in the identification of falsified medicines. We examined tablets distributed in LMICs and on the Internet using several portable devices. Multivariate analysis of the obtained spectra clarified the possibility of this approach for identifying falsified medicines.<sup>21–24,27)</sup> The reason for using a portable device was to facilitate on-site inspection. Whereas evaluation with a high degree of accuracy using a large device is the mainstream approach, our research yielded a revolutionary approach of detection using a portable device. Later, with the development of an inexpensive, domestically produced, ultra-compact Raman spectroscopic module, we established a falsified medicine identification system using Raman scattering analysis and this device. The stability of data acquisition was improved by making a prototype focus guide and improving the device side.<sup>25)</sup> As samples, I used four kinds of erectile dysfunction therapeutic agent, Cialis, Viagra, Levitra and Diflucan. These were obtained by personal import *via* the Internet and included falsified medicines that were actually distributed. The obtained Raman spectra are shown in Fig. 5. In these spectra, some differences and obvious saturation could be observed, but it was difficult to visually evaluate their similarity and discriminate them all, even

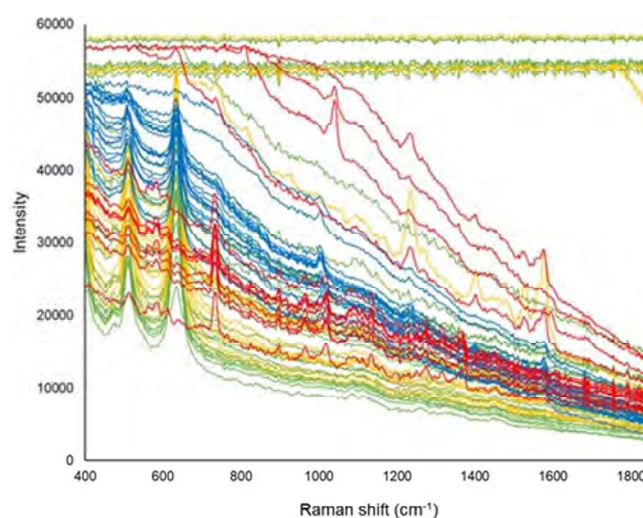


Fig. 5. Cialis Samples ( $n=32$ ) Including 23 Falsified Samples; Yellow Lines: Levitra Samples ( $n=22$ ) Including 13 Falsified Samples; Blue Lines: Viagra Samples ( $n=22$ ) Including 18 Falsified Samples; Red Lines: Diflucan Samples ( $n=13$ ) Including Two Falsified Samples

Two Diflucan falsified samples were measured using both the front and back side of the tablet, and four spectral data were obtained.

in spectra where sharp peaks appeared. I therefore tried using chemometric analysis. We examined discriminant analysis methods, considering the characteristics of existing falsified medicines. We investigated the applicability of two representative discriminant analysis methods, partial least squares discriminant analysis (PLS-DA) and soft independent modeling of class analogies (SIMCA). Using PLS-DA, if the prediction is successful, only a genuine sample will have a predicted value of 0.5 or more compared with the genuine model. In this analysis, falsified medicines and different products were falsely distinguished as genuine models (Fig. 6A). The reasons for this are thought to be because it is inappropriate to define falsified medicines from which various spectra are obtained in one category as well as owing to the small number of samples used in PLS-DA, which is supervised learning. However, in SIMCA using a PCA model (PCA-SIMCA), samples that fall within the limit of both the sample distance to the model ( $S_i$ ) and leverage ( $H_i$ ), which shows the 95% confidence interval for the standard PCA model, are considered to belong to the model. In this analysis, all samples were correctly dis-



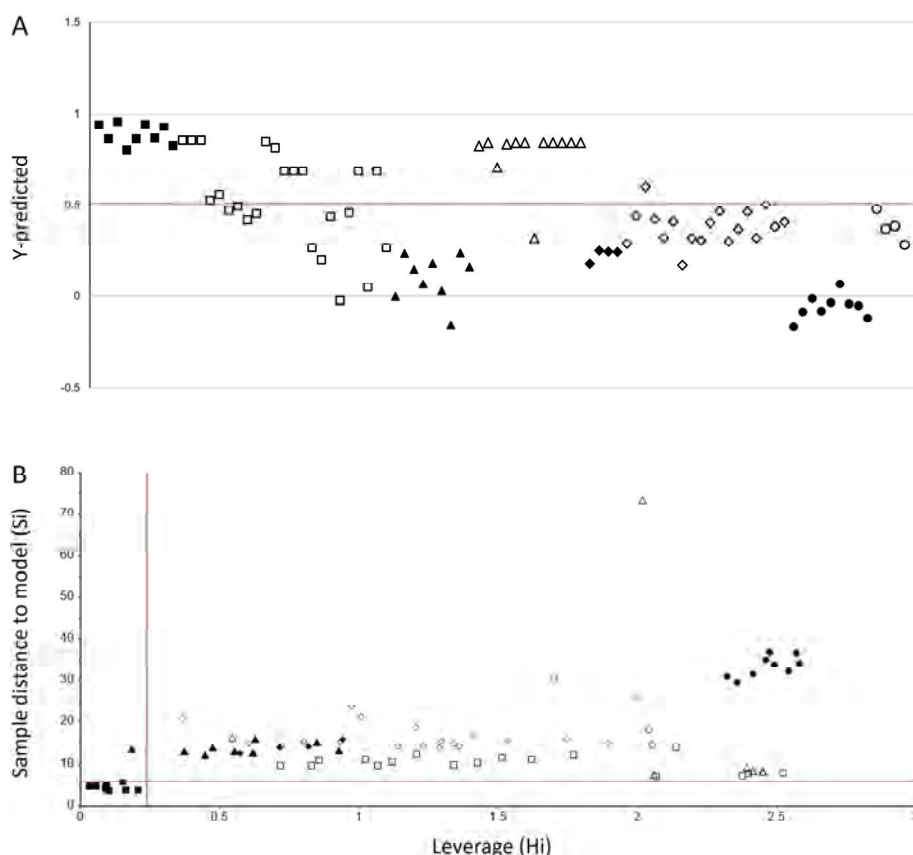


Fig. 6. Partial Least Squares Discriminant Analysis (PLS-DA) Prediction Results and the Sample Distance to the Model (Si) versus Leverage (Hi) Plot in the Standard Cialis Model

A) Results of PLS-DA; a predicted value of 0.5 is shown as a red line, compared with the genuine Cialis model. B) Results of principal component analysis-soft independent modeling of class analogies (PCA-SIMCA); the 95% confidence interval for the standard PCA model is shown as a red line. The results of analysis using the genuine Cialis model are shown. Closed and open markers represent genuine and falsified samples, respectively. Squares, diamonds, triangles, and circles represent Cialis, Viagra, Levitra, and Diflucan, respectively.

tinguished, revealing that PCA-SIMCA is more suitable for identifying falsified medicines<sup>28)</sup> (Fig. 6B). As a result, we established a discrimination algorithm for detecting falsified medicines, which is essential for social implementation, and developed a system of identifying falsified medicines using an ultra-compact Raman spectroscopic module. In the future, social implementation of this system will become even more realistic through building a spectrum library of genuine products and developing an application that executes SIMCA in real time.

#### 4. DEVELOPING A SMALL-SCALE QUALITY TEST METHOD

Another approach is the development of small-scale quality test methods.<sup>29)</sup> Without the eradication of substandard and falsified medicines, it is impossible to solve the problem of access to medicines and achieve universal health coverage as stated in Sustainable Development Goal 3.8. However, LMICs with limited resources are unable to conduct pharmacopoeial tests. I attempted to develop a small-scale quality test method using a small number of samples and statistical techniques, as a quality testing method for pharmaceuticals that can be implemented in LMICs. Many substandard medicines that have been previously identified have poor dissolution, which is often found in falsified medicines. We therefore considered downsizing the dissolution test.

The first stage of the United States Pharmacopeia dissolution test requires a minimum of six tablets. We aimed to develop a screening method using three tablets. We first derived the criteria using metronidazole and validated these criteria using cimetidine. The medicines tested were collected in Cambodia. We sought to determine a reference value that could be used to absolutely judge acceptability or unacceptability of samples in the first stage.

As a first method, the lower limit of the 95% confidence interval for the average dissolution rate of three metronidazole tablets, which had a Q value of 85%, was used to examine the acceptance criteria. We derived an acceptance criterion based on a lower limit of the 95% confidence interval of 85% or more, that is, the Q value or more.

As a second method, we examined the criteria using the average dissolution rate of three tablets and the minimum dissolution rate. The criteria derived involved an average dissolution rate of 91% or more and a minimum dissolution rate of 87% or more; that is, the average dissolution rate is the Q value +6% or more and the minimum dissolution rate is the Q value +2% or more.

As a result of verification with cimetidine tablets, we concluded that the criteria for judging 100% of acceptable samples as being acceptable in the first stage and 100% of unacceptable samples as being unacceptable are that the average dissolution rate of three tablets is the Q value +6% or more and the minimum dissolution rate is Q value +2% or more.

In this small-scale quality test method, dissolution tests are performed using three tablets, against full-scale quality tests in accordance with the pharmacopoeia. This approach is positioned as a screening method for assessing drug quality. It can be expected to function as a test method for monitoring and guiding the quality control of pharmaceuticals in LMICs.

## 5. FUTURE ISSUES IN ERADICATING SUBSTANDARD AND FALSIFIED MEDICINES

For social implementation of the falsified medicine identification system and small-scale quality test methods developed in these studies, it is necessary to expand the sample and evaluate the external validity. Currently, we are proceeding with the verification of medicines distributed in LMICs and medicines obtained *via* personal import from the Internet. Obtaining genuine products marketed in foreign countries is another challenge, and we are requesting the cooperation of manufacturers and distributors in our research.

We have also started to evaluate the applicability of novel methods for detecting substandard and falsified medicines using various analytic approaches that are not bound by conventional methods. Efforts to apply modern technology to the physical property evaluation of pharmaceuticals can contribute not only to finding new patterns of substandard and falsified medicines but also to expanding applications of this technology. At present, the basis for discriminating a falsified medicine is to distinguish it from the genuine medicine. If it is difficult to obtain genuine products, such as foreign products, the authenticity of substandard medicines cannot be confirmed even if they can be detected by referring to the pharmacopoeia. However, if there is no genuine product available, it is useful to request the manufacturer to authenticate their product. We have previously requested that manufacturers authenticate their products. Although we have been able to confirm the authenticity of some samples, the authenticity of other samples remains unknown.

In our investigation to assess the quality of medicines, we have found that the process of mixing excipients or the quality and selection of excipients themselves may be inappropriate for pharmaceutical products.<sup>26)</sup> To prioritize profit, falsified medicines might be composed of low-quality raw materials that are readily available rather than high-quality, expensive raw materials that are only suitable for manufacturing medicines. Additionally, falsified medicines are different from authentic ones in that management of the manufacturing process is insufficient, including in terms of hygiene, and there is a possibility of contamination with foreign substances in a poor manufacturing environment. It has become easier to obtain materials and equipment to manufacture falsified medicines, which are continually becoming more sophisticated. However, some characteristic or property always exists that cannot be imitated, and this can be used to distinguish medicines that are falsified from those that are genuine. Aiming to improve the differential accuracy, we are evaluating the applicability of various analytic methods and expanding the range of medicines to be verified. In this approach, detecting substandard medicines according to pharmaceutical characteristics that reveal differences in manufacturing processes could overcome the limits of detection methods that rely on comparisons with authentic products. In the future, I would like to clarify the

limits of Raman scattering analysis, which is expected to be the most implementable approach, and develop a discrimination method that can compensate for those limits, thereby improving the accuracy of this method for discriminating substandard and falsified medicines.

It is the responsibility of pharmaceutical researchers to ensure the safety and security of pharmaceuticals. Preventing damage to health owing to substandard and falsified medicines and optimizing the distribution of pharmaceuticals are other issues that should be addressed by pharmaceutical researchers. Seven years have passed since Society 5.0 was proposed as a future society that Japan should aspire to in January 2016.<sup>31)</sup> The distribution of pharmaceuticals may change in the future with smart medical care. However, with the persistence and availability of substandard and falsified medicines and medicines that are improperly distributed without quality assurance, smart medical care will not be possible and the ideal future envisioned for society will not be realized. I believe achievement of the future medical care that society seeks is only possible by providing patients with high-quality pharmaceuticals to treat disease.

## 6. CONCLUSION

Through the research activities described above, we have demonstrated the possibility of detecting substandard and falsified medicines using various analytic methods that are not bound by conventional methods. Our efforts to apply modern technology to evaluation of the physical properties of pharmaceuticals can contribute not only to identifying newly available substandard and falsified medicines but also to expanding the application of these technologies. I will continue to contribute to the enhancement of security measures for medicines based on scientific evidence to help protect individuals from the risks of substandard and falsified medicines so that patients can receive the benefits of more effective and safe pharmaceutical treatment.

**Acknowledgments** I would like to express my deep gratitude to Prof. Yoshimichi Sai (AI Hospital/Macro Signal Dynamics Research and Development Center, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University; Department of Hospital Pharmacy, Kanazawa University Hospital) for support and encouragement; Dr. Kazuko Kimura (previously professor of Drug Management and Policy, Faculty of Pharmacy, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University; Medicine Security Workshop) for lending her experience and expertise; Dr. Tsuyoshi Tanimoto (previously professor of Department of Analytical Chemistry, Faculty of Pharmaceutical Sciences, Department of Clinical Pharmacy, Doshisha Women's College of Liberal Arts), Dr. Tatsuo Koide (Division of Drugs, National Institute of Health Sciences), Prof. Mikio Koyano (Japan Advanced Institute of Science and Technology), and Dr. Makoto Watanabe (Research Center for Structural Materials, Bonding and Manufacturing Field, National Institute for Materials Science) for technical assistance with the experiments and discussion; and Dr. Mohammad Sofiqur Rahman (Department of Clinical Pharmacy, School of Pharmacy, University of California, San Francisco), Dr. Shu Zhu (AI Hospital/Macro Signal Dynamics Research and Development Center, Institute

of Medical, Pharmaceutical and Health Sciences, Kanazawa University), Dr. Mirai Sakuda (Plus Pharmacy), and Dr. Tomoko Sanada (graduate of Drug Management and Policy, Kanazawa University) for collaboration from the early stages of this work. I am deeply grateful to my colleagues who co-operated in the field studies. This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Nos. JP26870220 and JP19KK0236, Shibuya Science Culture and Sports Foundation in FY 2019, and Heiwa Nakajima Foundation in FY 2019. Sample collection of Viagra, Cialis, Levitra, and Diflucan tablets was supported by Health and Labor Sciences Research Grants from the Ministry of Health, Labour and Welfare, Japan [Grant Nos. H23-chikyukibo-shitei-006 and H26-chikyukiboA-shitei-003]. The RXM tablets used in this study were collected with financial support from WHO Reference 2015/491209.0. The author thanks Analisa Avila for editing a draft of this manuscript.

**Conflict of Interest** The author declares no conflict of interest.

## REFERENCES

- World Health Organization. Member State mechanism on substandard/spurious/falsely-labelled/falsified/counterfeit medical products. Report by the Director-General. The Seventieth World Health Assembly: provisional agenda item 13.6, WHA70/23, 2017.
- World Health Organization. "Substandard and falsified medical products. Fact sheet, 31 January, 2018." <https://www.who.int/news-room/fact-sheets/detail/substandard-and-falsified-medical-products>, accessed 13 September, 2023.
- 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](http://who.int/entity/medicines/regulation/ssffc/publications/GSMS_Report.pdf), accessed 13 September, 2023.
- Khan MH, Akazawa M, Dararath E, Kiet HB, Sovannarith T, Nivanna N, Yoshida N, Kimura K. Perceptions and practices of pharmaceutical wholesalers surrounding counterfeit medicines in a developing country: a baseline survey. *BMC Health Serv. Res.*, **11**, 306 (2011).
- Khan MH, Okumura J, Sovannarith T, Nivanna N, Nagai H, Taga M, Yoshida N, Akazawa M, Tanimoto T, Kimura K. Counterfeit medicines in Cambodia—possible causes. *Pharm. Res.*, **28**, 484–489 (2011).
- Khan MH, Hatanaka K, Sovannarith T, Nivanna N, Casas LC, Yoshida N, Tsuboi H, Tanimoto T, Kimura K. Effects of packaging and storage conditions on the quality of amoxicillin-clavulanic acid—an analysis of Cambodian samples. *BMC Pharmacol. Toxicol.*, **14**, 33 (2013).
- Yoshida N, Khan MH, Tabata H, Dararath E, Sovannarith T, Kiet HB, Nivanna N, Akazawa M, Tsuboi H, Tanimoto T, Kimura K. A cross-sectional investigation of the quality of selected medicines in Cambodia in 2010. *BMC Pharmacol. Toxicol.*, **15**, 13 (2014).
- 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).
- Takahashi N, Tsuboi H, Yoshida N, Tanimoto T, Khan MH, Kimura K. Investigation into the antinfluenza agent oseltamivir distributed via the internet in Japan. *Ther. Innov. Regul. Sci.*, **47**, 699–705 (2013).
- 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).
- Islam MR, Yoshida N, Tsuboi H, Sovannarith T, Dararath E, Kiet HB, Sokchamroeun U, Keila T, Tanimoto T, Kimura K. Four-year survey of the quality of antimicrobial medicines in Cambodia. *Kokusai Hoken Iryo (Journal of International Health)*, **32**, 233–242 (2017).
- Rahman MS, Yoshida N, Sugiura S, Tsuboi H, Keila T, Kiet HB, Zin T, Tanimoto T, Kimura K. Quality of omeprazole purchased via the internet and personally imported into Japan: comparison with products sampled in other Asian countries. *Trop. Med. Int. Health*, **23**, 263–269 (2018).
- Islam MR, Yoshida N, Kimura K, Uwatoko C, Rahman MS, Kumada S, Endo J, Ito K, Tanimoto T, Zin T, Tsuboi H. An investigation into the quality of medicines in Yangon, Myanmar. *Pharmacy (Basel)*, **6**, 96 (2018).
- Yoshida N, Yuasa M, Sovannarith T, Dararath E, Keila T, Kiet HB, Tsuboi H, Tanimoto T, Kimura K. A cross-sectional investigation for verification of globalization of falsified medicines in Cambodia, indicated by tablets of sildenafil citrate. *Pharmacy (Basel)*, **7**, 111 (2019).
- Rahman MS, Yoshida N, Tsuboi H, Sokchamroeun U, Keila T, Sovannarith T, Kiet HB, Dararath E, Akimoto Y, Tanimoto T, Kimura K. A cross-sectional investigation of the quality of selected medicines for non-communicable diseases in private community drug outlets in Cambodia during 2011–2013. *Am. J. Trop. Med. Hyg.*, **101**, 1018–1026 (2019).
- Sakuda M, Yoshida N, Takaoka T, Sanada T, Rahman MS, Tanimoto T, Zin T, Kimura K, Tsuboi H. Substandard and Falsified Medicines in Myanmar. *Pharmacy (Basel)*, **8**, 45 (2020).
- Rahman MS, Yoshida N, Tsuboi H, Maeda E, Ibarra AVV, Zin T, Akimoto Y, Tanimoto T, Kimura K. Patient safety and public health concerns: poor dissolution rate of pioglitazone tablets obtained from China, Myanmar and internet sites. *BMC Pharmacol. Toxicol.*, **22**, 12 (2021).
- 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).
- Rahman MS, Yoshida N, Tsuboi H, Karmoker JR, Kabir N, Schaeffermann S, Akimoto Y, Bhuiyan MA, Reza MS, Kimura K. A comprehensive analysis of selected medicines collected from private drug outlets of Dhaka city, Bangladesh in a simple random survey. *Sci. Rep.*, **12**, 234 (2022).
- Rahman MS, Yoshida N, Tsuboi H, Keila T, Sovannarith T, Kiet HB, Dararath E, Zin T, Tanimoto T, Kimura K. Erroneous formulation of delayed-release omeprazole capsules: alert for importing countries. *BMC Pharmacol. Toxicol.*, **18**, 31 (2017).
- Kakio T, Yoshida N, Macha S, Moriguchi K, Hiroshima T, Ikeda Y, Tsuboi H, Kimura K. Classification and visualization of physical and chemical properties of falsified medicines with handheld Raman spectroscopy and X-ray computed tomography. *Am. J. Trop. Med. Hyg.*, **97**, 684–689 (2017).
- Kakio T, Nagase H, Takaoka T, Yoshida N, Hirakawa J, Macha S, Hiroshima T, Ikeda Y, Tsuboi H, Kimura K. Survey to identify substandard and falsified tablets in several Asian countries with pharmacopeial quality control tests and principal component analysis of handheld Raman spectroscopy. *Am. J. Trop. Med. Hyg.*, **98**, 1643–1652 (2018).
- Zhu S, Yoshida N, Kimura K, Matsushita R, Tsuboi H. Falsified vardenafil tablets available online. *J. Pharm. Biomed. Anal.*, **177**, 112872 (2020).
- 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).
- 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).
- 26) Sakuda M, Yoshida N, Koide T, Keila T, Kimura K, Tsuboi H. Clarification of the internal structure and factors of poor dissolution of substandard roxithromycin tablets by near-infrared chemical imaging. *Int. J. Pharm.*, **596**, 120232 (2021).
- 27) Sanada T, Ohnishi M, Yoshida N, Kimura K, Tsuboi H. Quality assessment of Diflucan® tablets distributed online: Diflucan® distributed online. *J. Med. Access*, **5**, 1–8 (2021).
- 28) 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).
- 29) Rahman MS, Yoshida N, Tsuboi H, Ishii Y, Akimoto Y, Kimura K. Small scale dissolution test screening tool to select potentially substandard and falsified (SF) medicines requiring full pharmacopoeial analysis. *Sci. Rep.*, **11**, 12145 (2021).
- 30) Rahman MS, Yoshida N, Hanafusa M, Matsuo A, Zhu S, Stub Y, Takahashi C, Tsuboi H, Matsushita R, Mackawa K, Kimura K. Screening and quantification of undeclared PGF2 $\alpha$  analogs in eyelash-enhancing cosmetic serums using LC-MS/MS. *J. Pharm. Biomed. Anal.*, **219**, 114940 (2022).
- 31) Cabinet Office, Government of Japan. Society 5.0.: [https://www8.cao.go.jp/cstp/english/society5\\_0/index.html](https://www8.cao.go.jp/cstp/english/society5_0/index.html), accessed 18 January, 2024.




RESEARCH

Open Access



# Falsified and problematic methandienone products available online: active pharmaceutical ingredient identification by portable Raman spectrometers and quantification by ultra-high-performance liquid chromatography–Fourier transform mass spectrometry

Robin Schreiber<sup>1</sup>, Manami Hori<sup>2</sup>, Chisato Takahashi<sup>2</sup>, Mohammad Sofiqur Rahman<sup>3</sup>, Ayane Nakao<sup>2</sup>, Shu Zhu<sup>3</sup>, Feiyu Zhu<sup>1</sup>, Naoko Yoshida<sup>4,5\*</sup> , Keiko Maekawa<sup>2</sup> and Kazuko Kimura<sup>3,5</sup>

## Abstract

This study aimed on the one hand to clarify the quality, authenticity, safety, and other issues related to products of the anabolic-androgenic steroid methandienone advertised on the Internet and personally imported to Japan and on the other hand to evaluate the use of two portable Raman spectrometers in identifying the active pharmaceutical ingredient (API). The study found that all  $n = 15$  samples purchased from 14 websites were problematic regarding their package, labeling, and/or content. Specifically, one sample (6.7%) was confirmed falsified, twelve samples (80%) were found either to be falsified or unlicensed as pharmaceutical product, and two samples (13.3%) were received without information on the manufacturers' physical address or country of origin, with one sample (6.7%) having no labeling or other accompanying information at all. Both Raman spectrometers were able to identify the API in all samples as confirmed and quantified by ultra-high-performance liquid chromatography–Fourier transform mass spectrometry. Twelve samples contained on average less than 90% of the declared API content. By contacting national regulatory authorities in 44 countries, methandienone products were found to be approved in 1 country and not approved in 21 countries. To prevent health hazards and abuse, measures against the acquisition of anabolic-androgenic steroids from unknown sources are required. Portable Raman spectrometers may be suitable for the non-destructive and quick identification of methandienone in tablets.

**Keywords** Falsified, Anabolic-androgenic steroids (AASs), Methandienone, Personal import, Raman spectroscopy, Portable, Handheld, Raman spectrometer, Ultra-high-performance liquid chromatography–Fourier transform mass spectrometry

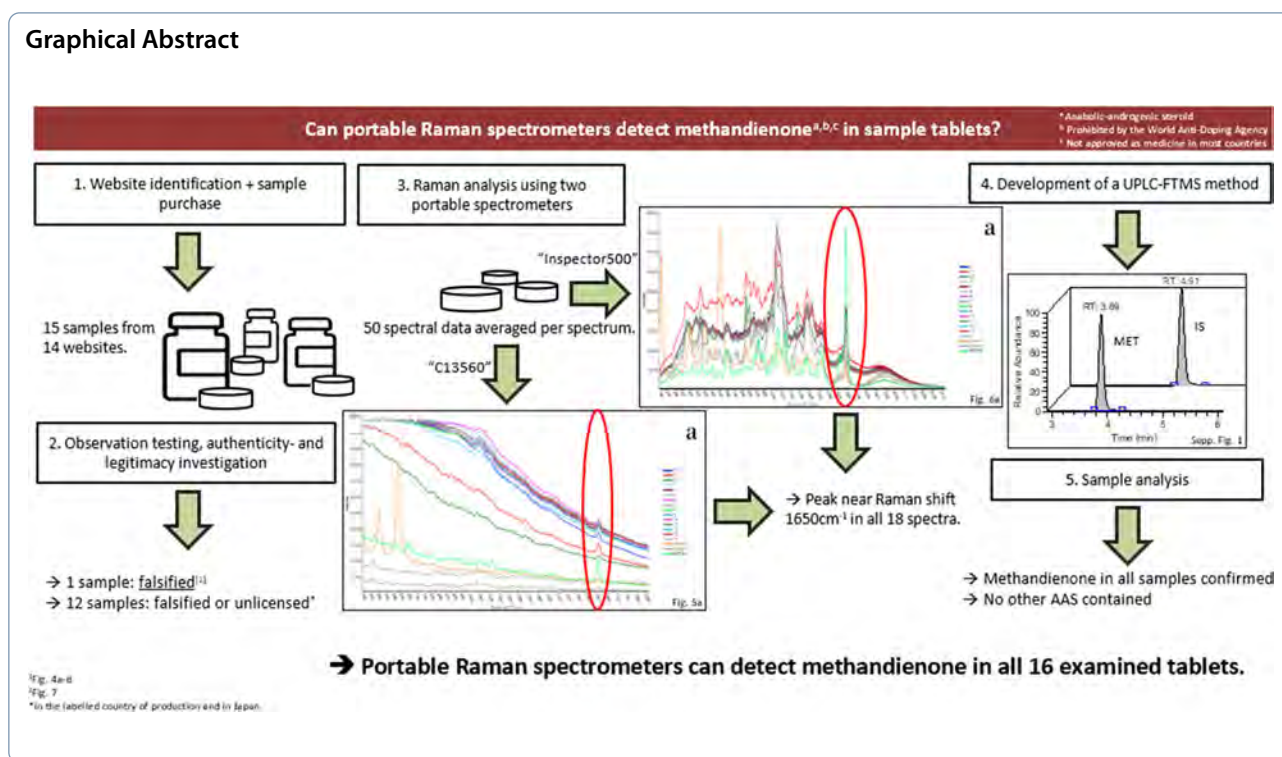
\*Correspondence:

Naoko Yoshida

naoko@p.kanazawa-u.ac.jp

Full list of author information is available at the end of the article

## Graphical Abstract



## Introduction

Anabolic-androgenic steroids (AASs) are synthetic steroids structurally related to testosterone, the male sex hormone, and well known for their androgenic effects (Ganesan et al. 2023; Bhasin et al. 1996; National Institute on Drug Abuse 2018; Barceloux and Palmer 2013). As prescription medicines, they are used in treating hypogonadism and bone marrow stimulation in leukemia, kidney failure, and aplastic anemia (Ganesan et al. 2023). However, the non-medical use of AASs is a serious global public health concern (Sagoe et al. 2014; Frude et al. 2020) in that AASs are widely abused to increase muscle mass, not only among bodybuilders and athletes for doping purposes but also increasingly among individuals to achieve masculine cosmetic benefits (Magnolini et al. 2022; Rahnema et al. 2014; McBride et al. 2018; Storer et al. 2003).

Non-medical dosages are up to 100 times the therapeutic dose (National Institute on Drug Abuse 2018; Barceloux and Palmer 2013; Solimini et al. 2017) and known to cause psychological changes, addiction, and severe physiological changes (Bhasin et al. 1996; Magnolini et al. 2022; Rahnema et al. 2014; Solimini et al. 2017; van Amsterdam et al. 2010; Liu and Wu 2019; Takayanagi et al. 2008; Nieschlag and Vorona 2015) including hepatotoxicity (Solimini et al. 2017) and AAS-induced hypogonadotropic hypogonadism, which does not improve after the discontinuation of AASs (Rahnema et al. 2014).

Substandard and falsified (SF) medicines have been defined by the World Health Organization since 2017 (World Health Organization 2018). Substandard medicines fail to meet either their quality standards, specifications, or both. Falsified medicines deliberately/fraudulently misrepresent their identity, composition, or source. Unregistered/unlicensed medicines have not been approved for sale under relevant regulations and legislation (World Health Organization 2018). Falsified medicines might contain an altered amount of the active pharmaceutical ingredient (API), no API, or the wrong API and can therefore have serious health effects up to death of the consumer (Rahman et al. 2018).

The online availability of SF medicines is a global challenge and there is a high demand for suitable analytical detection methods and tools for identifying SF medicines (Ahmed et al. 2022), with Japan being no exception to this development (Zhu et al. 2020; Sanada et al. 2020; Rahman et al. 2022; Sanada et al. 2021; Yoshida et al. 2015; Khan et al. 2012). The widespread international online distribution of SF AASs is well documented (Magnolini et al. 2022; McBride et al. 2018; Fabresse et al. 2021; Coopman and Cordonnier 2012), and a recent systematic review estimated the overall mean prevalence of falsified AASs on the black market to be more than one in three (Magnolini et al. 2022). SF medicines are entering Japan through personal importation, often without the

requirement for a prescription (Zhu et al. 2020; Sanada et al. 2020; Rahman et al. 2018).

Although reports of health hazards of AAS products posted by consumers on online forums reduce the consumers' risk (Frude et al. 2020), analytical methods for identification, quantification, and determining quality-related parameters are urgently required to protect consumer health (Piatkowski et al. 2023).

Raman scattering analysis identifies the constituents of a mixture (Rebierre et al. 2016) and is thus a technology for the quick and non-destructive qualitative analysis of APIs (Rebierre et al. 2016; Opuni et al. 2019; Sanada et al. 2020). Portable and handheld Raman spectrometers are comparably inexpensive devices that can be used on site to detect SF medicines, reducing the number of samples requiring expensive laboratory analysis (Sanada et al. 2020; Hajjou et al. 2013; Dégardin et al. 2017).

The present study aimed on the one hand to investigate the quality, harmfulness, and other problems relating to methandienone (also known as methandrostenedione; MET) products advertised and sold over the Internet and personally imported to Japan and to assess whether samples are falsified. MET is an orally available AAS that is explicitly prohibited for use in sports by the World Anti-Doping Agency (WADA) (World anti-doping agency 2023). The availability of officially licensed MET products was investigated in a worldwide search covering 44 countries on five continents. This study aimed on the other hand to explore the use of two portable Raman spectrometers for non-destructive MET analysis that could be useful for the quick on-site detection of MET containing products, e.g. at airports or at customs. To verify the Raman scattering analysis results, a liquid chromatography–mass spectrometry (LC-MS) method was developed for MET identification, quantification, and impurity testing.

## Materials

### Chemicals

MET (#H1193, Lot H4FHF, purity 97.5%) and methyltestosterone (#M0435, Lot YNV3C) were purchased from Tokyo Chemical Industry (Tokyo, Japan) and used as reference standards. DANABOL 10 mg, an MET product approved in the Republic of Moldova, was imported under an agreement between SC Balkan Pharmaceuticals SRL and Kanazawa University and used as the standard formulation and control.

LC-MS grade ultrapure water was purchased from Kanto Chemical Co., Inc. (Tokyo, Japan). Methanol (MeOH), acetonitrile (MeCN), isopropanol (IPA), and 0.1% acetic acid were used in LC-MS grade only and

purchased from Wako Pure Chemical Industries, Ltd. (Kyoto, Japan).

### Devices

Images of the samples were taken with an IXY 650 compact digital camera manufactured by Canon (Tokyo, Japan).

Raman scattering analysis was conducted using a C13560 ultra-compact portable Raman spectrometer (96 mm × 14.5 mm × 60 mm; 90 g) manufactured by Hamamatsu Photonics K.K. (Shizuoka, Japan) with a silicon substrate provided by the manufacturer (Sanada et al. 2020) and an Inspector500 portable Raman spectrometer manufactured by SciAps Inc. (Laramie, WY, USA) with a polystyrene standard provided by the manufacturer. The supplied software and drivers were installed on a personal computer prior to the analysis.

LC-MS experiments were performed adopting ultra-high-performance liquid chromatography–Fourier transform mass spectrometry (UPLC-FTMS; Q Exactive, Thermo Fisher Scientific, Waltham, MA) interfaced with an UltiMate 3000 RS UPLC system.

## Methods

### Source selection and sample collection

Websites of personal import agents offering MET products to consumers in Japan were searched via the Google Japan search engine (Mountain View, CA, USA) using the term “methandienone AND personal import”. The minimum unit number per sample was 60 tablets. All identified samples were purchased and personally imported using a private address as the recipient address. After receipt, samples were stored under suitable conditions and sample identification numbers (IDs) were assigned according to the website of purchase.

### Visual observation, authenticity investigation, and legitimacy status request

#### Visual observation test

For each sample, the packaging, labeling, and dosage units were visually examined and evaluated using the International Pharmaceutical Federation's Tool for Visual Inspection of Medicines (International Pharmaceutical Federation n.d.), and detailed information was collected and documented. This information is presented anonymized in Supplemental Table 1. Images of problematic samples were taken.

### Authenticity investigation

Each labeled manufacturer was requested to authenticate their samples. A questionnaire with photographs of the

samples and relevant information, including the labeled trade name, batch number, manufacturing and expiry dates, manufacturer name and address, API name, and dosage strength, was sent to the labeled email addresses. If no email address was available, the request was sent via the inquiry form of the labeled website. The authenticity investigation was not possible for unlabeled samples.

#### **Legitimacy status request**

The national regulatory authorities (NRAs) of the countries indicated on the sample and of the countries of origin of the shipments were queried about the legitimacy of the samples and whether the labeled manufacturer holds a manufacturing license for the sample product or was generally allowed to handle MET-containing products.

#### **Raman scattering analysis**

##### **Measurement conditions**

Both the C13560 and Inspector500 Raman spectrometers were set up according to the manufacturers' instructions.

In analysis using the C13560 Raman spectrometer, the output was set at 15 mW ("High"), the excitation laser wavelength was fixed at 785 nm, the scanning time was 1,000 ms/scan, and the spectral X-axis wavenumber interval was defined as 403–1852  $\text{cm}^{-1}$ . Before taking measurements, the dark signal was measured by inserting a silicon substrate provided by Hamamatsu Photonics K.K. Subsequent calibration was performed using the Raman shift peak of the substrate near 521  $\text{cm}^{-1}$ . Following instructions, the emitted laser wavelength was corrected manually to 785 nm.

In analysis using the Inspector500 Raman spectrometer, the output was set at 300 mW ("High"), the excitation laser wavelength was fixed at 1,030 nm, the scanning time was set to the default "automatic" setup (maximum of 8 s), and the spectral X-axis wavenumber interval was defined as 150–2,450  $\text{cm}^{-1}$ . Before taking measurements, a self-test was run to check the calibration status. The device was continually calibrated until it passed this test.

##### **Measurement procedure**

The measurement procedure was the same for the two Raman spectrometers. Each spectrometer's attachment for analysis was placed in direct contact with the AAS tablet such that the tablet surface was fully covered. One measurement was the average of five spectral data in the Raman scattering analysis. A total of 10 measurements per tablet were performed on randomly chosen spots of the sample surface. Five measurements were taken from the upside and five measurements from the downside of the tablet. Therefore, a total of 50 spectral data were averaged to give one spectrum for each sample. During

analysis, black opaque fabric covered the setup to prevent light contamination.

A methandienone reference standard (MET-RS), only available as powder, was placed in a small plastic film bag and analyzed in the same manner as the sample after smoothing the powder to create a smooth, consistent powder surface. DANABOL 10 mg, an MET product approved in the Republic of Moldova and identified in this study, was twice analyzed by scanning the surface of the coating (Control 1) and scanning the product surface after sufficiently removing the coating (Control 2) to eliminate the effect of the coating.

#### **Determination of AASs by LC-MS**

##### **Preparation of standard solutions**

MET and methyltestosterone standards were separately dissolved in 100% methanol at 10 mM, and the standard solution of MET was stepwise diluted to 10, 5, 2, 1, 0.5, and 0.2 mM. To quantify the MET contents of the sample tablets, calibration standards were prepared in the same manner as the samples as outlined below. A 2-mL quantity of each diluted MET standard solution was added to a mixture of 1 mL distilled water with 3 mL MeOH, the mixture was shaken for 10 min and centrifuged at 3,000 rpm, and 5 mL of the supernatant liquid was separated. This procedure was conducted thrice and the 3×5 mL was combined for each supernatant solution. A 1-mL quantity of 4 mM methyltestosterone was added to this solution as an internal standard, and the total volume was made up to 20 mL using MeOH. Exactly 1 mL of this solution was taken and diluted with MeOH to 20 mL. Again, exactly 1 mL was taken and MeOH and distilled water at a ratio of 65:35 were added to obtain a total volume of 10 mL. Finally, this solution was filtered through a 0.22- $\mu\text{m}$  filter and used as the standard solution. The final standard solutions for the MET contained 5.0, 2.5, 1.0, 0.5, 0.25, and 0.1  $\mu\text{M}$  MET-RS and 1.0  $\mu\text{M}$  methyltestosterone as an internal standard. Quality control solutions (4, 2, and 0.8  $\mu\text{M}$  MET-RS) were prepared in the same manner as the calibration standards.

##### **Preparation of sample solutions**

The MET extraction for each tablet including DANABOL, the standard product obtained from the Republic of Moldova, was performed through the assay pretreatment of estriol tablets and prednisolone tablets listed in the Japanese Pharmacopoeia (The Japanese Pharmacopoeia 18th edition - Official Monographs (A to L) 2021; The Japanese Pharmacopoeia 18th edition - Official Monographs (M to Z) 2021).

The mass of one tablet of each product was weighed before the tablet was homogenized using a mortar. A quantity corresponding to approximately 2 mg of the



labeled MET content was precisely weighed, exactly 1 mL of distilled water was added, and the solution was sonicated. A 5-mL quantity of 100% MeOH was added, and the solution was then shaken for 10 min and subsequently centrifuged. After centrifugation, 5 mL of the supernatant liquid was separated and another 5 mL of MeOH was added to the remaining solution. This extraction procedure was conducted thrice. The obtained 3×5 mL of each supernatant was combined, and 1 mL of a MeOH solution of 4 mM methyltestosterone was added as an internal standard. MeOH was added to obtain a total volume of 20 mL. Exactly 1 mL of this solution was taken and diluted with MeOH to 20 mL. Again, exactly 1 mL was taken, and MeOH and distilled water at a ratio of 65:35 were added to obtain a total volume of 10 mL. Finally, this solution was filtered through a 0.22-μm filter (Millipore, Bedford, MA, USA) and used as the sample solution.

To evaluate the recovery of the extraction, a positive control sample was prepared as follows. A 2-mg quantity of the MET-RS was weighed precisely and extracted three times in the same manner as the sample, and 3×5 mL of the resulting supernatant solution was combined. The internal standard solution was added and MeOH was used to make up to a total volume of 20 mL. This solution was diluted in the same manner as the sample solution to obtain the positive control solution.

**Screening of testosterone analogous substances**

First, all samples were screened for 25 AASs including MET (Table 1) through LC-FTMS following Tircova et al. (2019).

The injection volume of sample solution was fixed at 2 μL, and a Shim-pack FC-ODS column (2.0 mm i.d. × 75 mm, 3-μm particles) was used for separation. The column temperature was maintained at 40 °C. The flow rate of the mobile phase was 300 μL/min. Mobile phase A comprised MeCN/water at a ratio of 3:7 (including 0.1% formic acid) whereas mobile phase B comprised MeCN/isopropanol at a ratio of 9:1. During separation in liquid chromatography (LC), the mobile phases A and B were used in gradient elution. Mobile phase A was initially set at 100% for 0.5 min. The proportion of mobile phase B was then gradually increased from 0 to 95% over 6.5 min. For the next 3 min, the 95% proportion of mobile phase B was maintained. Blank runs were carried out randomly between samples to check that there was no significant chromatographic carryover. Mass spectrometry was performed in synchronous full-scan positive mode with a resolution of 70,000 and data dependent MS/MS with a resolution of 17,500 (full-scan MS<sup>1</sup>/dd-MS<sup>2</sup>). The scan range of the instrument was set at a mass-to-charge

**Table 1** 25 AASs used in a study conducted in the Czech Republic and Slovakia (Tircova et al. 2019)

AAS	Molecular formula	m/z (positive ion mode)
Boldenone cypionate	C <sub>27</sub> H <sub>38</sub> O <sub>3</sub>	411.2894
Boldenone undecylenate	C <sub>30</sub> H <sub>44</sub> O <sub>3</sub>	453.3363
Clenbuterol	C <sub>12</sub> H <sub>18</sub> Cl <sub>2</sub> N <sub>2</sub> O	277.0869
Drastanolone propionate	C <sub>23</sub> H <sub>36</sub> O <sub>3</sub>	361.2737
Fluoxymesterone	C <sub>20</sub> H <sub>29</sub> FO <sub>3</sub>	337.2173
Chlorodehydromethyltestosterone	C <sub>20</sub> H <sub>27</sub> ClO <sub>2</sub>	335.1772
Mesterolone	C <sub>20</sub> H <sub>32</sub> O <sub>2</sub>	305.2475
Methandienone	C <sub>20</sub> H <sub>28</sub> O <sub>2</sub>	301.2162
Methenolone enanthate	C <sub>27</sub> H <sub>42</sub> O <sub>3</sub>	415.3207
Nandrolone decanoate	C <sub>28</sub> H <sub>44</sub> O <sub>3</sub>	429.3363
Nandrolone phenylpropionate	C <sub>27</sub> H <sub>34</sub> O <sub>3</sub>	407.2581
Nandrolone undebanoate	C <sub>29</sub> H <sub>46</sub> O <sub>3</sub>	443.352
Oxandrolone	C <sub>19</sub> H <sub>30</sub> O <sub>3</sub>	307.2268
Oxymetholone	C <sub>21</sub> H <sub>32</sub> O <sub>3</sub>	333.2424
Stanozolol	C <sub>21</sub> H <sub>32</sub> N <sub>2</sub> O	329.2587
Testosterone cypionate	C <sub>27</sub> H <sub>40</sub> O <sub>3</sub>	413.305
Testosterone decanoate	C <sub>29</sub> H <sub>46</sub> O <sub>3</sub>	443.352
Testosterone enanthate	C <sub>26</sub> H <sub>40</sub> O <sub>3</sub>	401.305
Testosterone isocaproate	C <sub>25</sub> H <sub>38</sub> O <sub>3</sub>	387.2894
Testosterone phenylpropionate	C <sub>28</sub> H <sub>36</sub> O <sub>3</sub>	421.2737
Testosterone propionate	C <sub>22</sub> H <sub>32</sub> O <sub>3</sub>	345.2424
Testosterone undecanoate	C <sub>30</sub> H <sub>48</sub> O <sub>3</sub>	457.3676
Trenbolone acetate	C <sub>20</sub> H <sub>24</sub> O <sub>3</sub>	313.1798
Trenbolone enanthate	C <sub>25</sub> H <sub>34</sub> O <sub>3</sub>	383.2581
Trenbolone hexahydrobenzylcar-bonate	C <sub>26</sub> H <sub>34</sub> O <sub>4</sub>	411.253

m/zmass to charge ratio

ratio *m/z* of 100–1,000, and the top three ions were fragmented at a collision energy of 30 eV with a dynamic exclusion time of 10.0 s. Heated electrospray ionization was applied at a spray voltage in positive ion mode of 3.5 kV. The sheath gas (nitrogen) was set at 50 arbitrary units; the auxiliary gas (nitrogen) was set at 10 arbitrary units; and the vaporizer and capillary temperatures were set at 300 and 250 °C, respectively. Data were collected in profile mode using Xcalibur software (Thermo Fisher Scientific).

The UPLC-FTMS data were processed using Compound Discoverer 3.1 software (Thermo Fisher Scientific), and *m/z* values of the extracted ion peaks were compared with the known *m/z* values of AASs (Table 1) to determine whether the products contained these AASs.

**Quantification of MET by LC-MS**

MET contents in sample solutions were quantified also using calibration standards by LC-FTMS but applying a

gradient program and data acquisition mode as follows. During separation in LC, the mobile phases A and B were used in gradient elution. The proportion of mobile phase B was gradually increased from 5 to 50% for the first 7 min and then rapidly raised to 95% and maintained for the next 3 min. Mass spectrometry was performed in target-single ion monitoring mode at a resolution of 70,000. MET ( $C_{20}H_{28}O_2$ ) and methyltestosterone ( $C_{20}H_{30}O_2$ ) were monitored at an  $[M+H]$  ion of  $m/z$  303.2319 and  $m/z$  301.2162, respectively. The UPLC-FTMS data were processed using the Quan Browser (Thermo Fisher Scientific). Areas of the extracted ion peak chromatograms were obtained with a mass tolerance of 5 ppm.

#### Availability of approved MET products

A total of  $n=44$  NRAs on five continents (Africa ( $n=1$ ), Asia ( $n=9$ ), Australia/Oceania ( $n=1$ ), Europe ( $n=31$ ), and North America ( $n=2$ )) and  $n=33$  institutions (including international chemical and pharmaceutical industry associations, the world and Japanese anti-doping agencies WADA and JADA, and the Japanese Ministry of Economy, Trade, and Industry) were emailed about the availability of officially approved, registered, or authorized MET products. In addition, MET products and manufacturers producing MET products were searched for using the Google search engine (Mountain View, CA, USA) for text and images by combining search terms including “Methandienone”, “Methandienon”, “Methandrostenolone”, “medicine”, “product”, “leaflet”, “buying”, “manufacturing”, “approval/approved”, “registration/registered”, “supplement”, and related terms. The search results were documented, and the companies were queried by email for specific information on the product identified. Concurrently, the NRAs of the manufacturers’ countries of location were asked for information on product approval and the results were recorded. All responses are reported in Supplemental Table 2.

#### Data analysis and visualization

Data were analyzed using Microsoft Excel for Microsoft 365 MSO (Microsoft Corporation; WA, USA). Raman spectra were line plotted and colorized using The Unscrambler X 10.5 (CAMO Software AS; Oslo, Norway). Images were formatted by applying Microsoft Word “Format” equally across each entire image.

## Results

#### Sample collection

$n=15$  samples of four MET products were purchased between December 25, 2019 and January 6, 2020 from 14 identified websites of personal import agents and received by January 31, 2020. Sample 2.2 was provided by a personal import agent as a replacement

product for sample 2.1 with a statement that the product of sample 2.1 had sold out: however, the two samples arrived together, with payment claimed for sample 2.1 but not for sample 2.2.  $n=12$  (80%) samples were labeled with trade name A, one (6.7%) with trade name B, and one (6.7%) with trade name C, and one sample (6.7%) was received unlabeled. Upon arrival, no sample had exceeded the expiry date indicated on the label; however, no expiry date was given for the unlabeled sample (Supplemental Table 1).

The postal labels on the packages indicated that 12 of the 15 samples (80%) were shipped from Taiwan and the other three (20%) from Thailand (Supplemental Table 1).

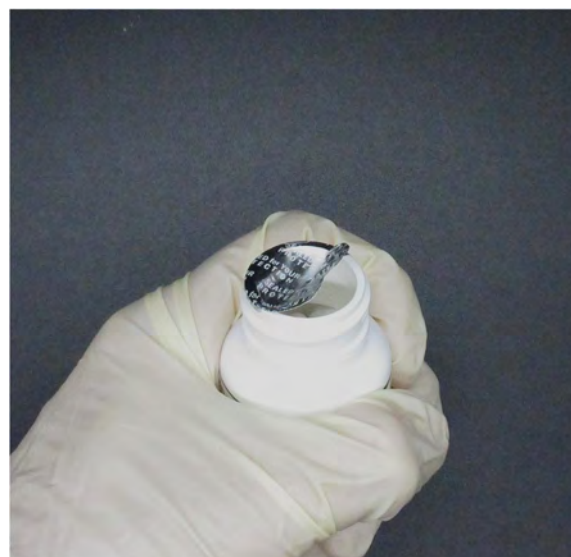
Twelve of the 14 websites (85.7%) described and advertised their product as a pharmaceutical product, with eight indicating their product was as an “anabolic steroid” and four that their product was a “pharmaceutical product”. Two websites did not describe the purpose, use, or intention of their product (Supplemental Table 1).

Notably, none of the 14 orders required a prescription or similar documentation.

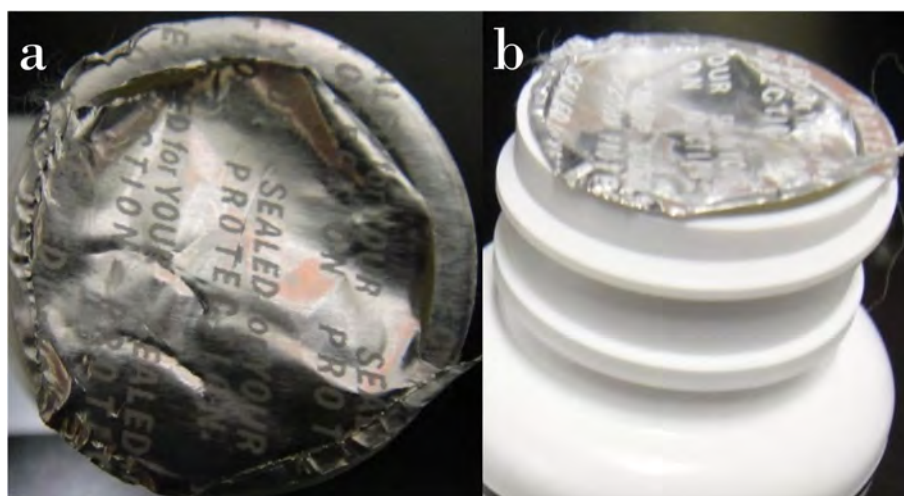
#### Sample evaluation

##### Sample observation

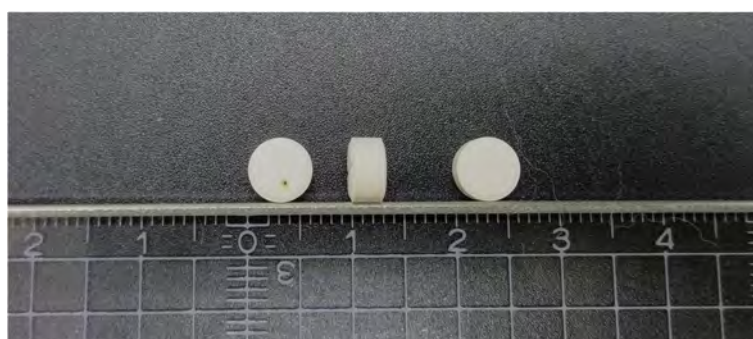
All 15 samples (100%) failed the visual observation test. The packaging of two samples (samples 1 and 3) was opened or damaged (Figs. 1 and 2a, b). One sample



**Fig. 1** Opened seal of sample 1. Fig. 1 shows an image of the primary packaging and seal of sample 1, which was found opened when the sample was collected. In Fig. 1, the brightness of the original was reduced by ~30%



**Fig. 2** Damaged seal of sample 3 viewed from two angles. Fig. 2 shows an image of the primary packaging and seal of sample 3, which was found damaged when the sample was collected. **a** top view. **b** side view



**Fig. 3** Tablets of sample 9. Fig. 3 shows an image of tablets of sample 9. Notably, one tablet has a dark spot on the surface. In Fig. 3, the brightness of the original was reduced by ~30%

gave a fictitious manufacturer's address, had grammatical and spelling errors, and did not provide a batch number or dosage strength. Twelve samples only had the text "Manufactured for: [Manufacturer A]", leaving the actual manufacturer unclear. Furthermore, one sample was received unlabeled and another did not give the manufacturer's address or country of origin.

In addition, two samples contained more tablets than indicated on the label, and one had a dark spot on the surface of one tablet (Fig. 3).

Two samples did not contain any leaflet. Twelve of the 15 samples had physical leaflets attached, and one sample presented a quick response (QR) code for downloading the leaflet. Notably, in August 2023, it was discovered that this QR code no longer linked to an accessible website, and therefore, the leaflet could not be accessed or downloaded.

All 13 collected leaflets were in English, and none had a Japanese translation.

#### **Authenticity investigation**

Repeatedly contacting the manufacturers to inquire about the authenticity did not lead to any response from the manufacturers. Two samples were labeled with a QR code for online authenticity.

These QR codes linked to two websites stating the authenticity of the product. Notably, one of these two samples was confirmed falsified (see [Identification of falsified MET](#) section). Remarkably, the website claiming authenticity for the falsified product included the spelling mistake "authenticity" (*sic!*).

Owing to the absence of a label or accompanying information, there was no contact information for one sample.

#### **Legitimacy and registration status**

There were no MET products approved in Japan and therefore all 15 samples were prohibited for distribution. The 12 samples labeled "Manufactured for: [Manufacturer A]" were confirmed to be unlicensed in the country



labeled and the country of distribution. Notably, one sample was confirmed falsified (see [Identification of falsified MET](#) section). For two samples, it was impossible to inquire about legitimacy owing to the absence of information on the manufacturer’s address or country.

**Identification of falsified MET**

Sample 2.2 did not have a batch number, dosage strength, or registration number. In addition, it gave a fictitious German address, which was confirmed not to exist by the German municipality (83674 Gaißach) responsible for the labeled address (Fig. 4b).

Notably, the label indicated the name of the German municipality using what is believed to be the Latin letter uppercase “B” or the Greek letter uppercase beta “B”, whereas the municipality would be correctly written using the German letter “Eszett; sharp S” expressed as “ß”, which is different from the Greek letter lowercase beta “β” and therefore a spelling

mistake. Furthermore, the storage instruction on the label stated “Donot [...]” (*sic!*), which is grammatically incorrect (Fig. 4b).

The sample was confirmed to be a falsified product by the responsible German regulation authority; i.e., no manufacturing license, marketing authorization, or export license was issued in the European Union, including Germany, for any MET-containing product. In addition, the labeled manufacturer was found to be fictitious. Images of falsified sample 2.2 are presented in Fig. 4a, b, c and d. A summary of the information labeled on falsified methandienone product EP.Dbol-10 is presented in Table 2.

**Raman scattering analysis**

Spectra of all 15 samples, the standard formulation (Control 1 and Control 2), and the MET-RS obtained by Raman scattering analysis using the two spectrometers



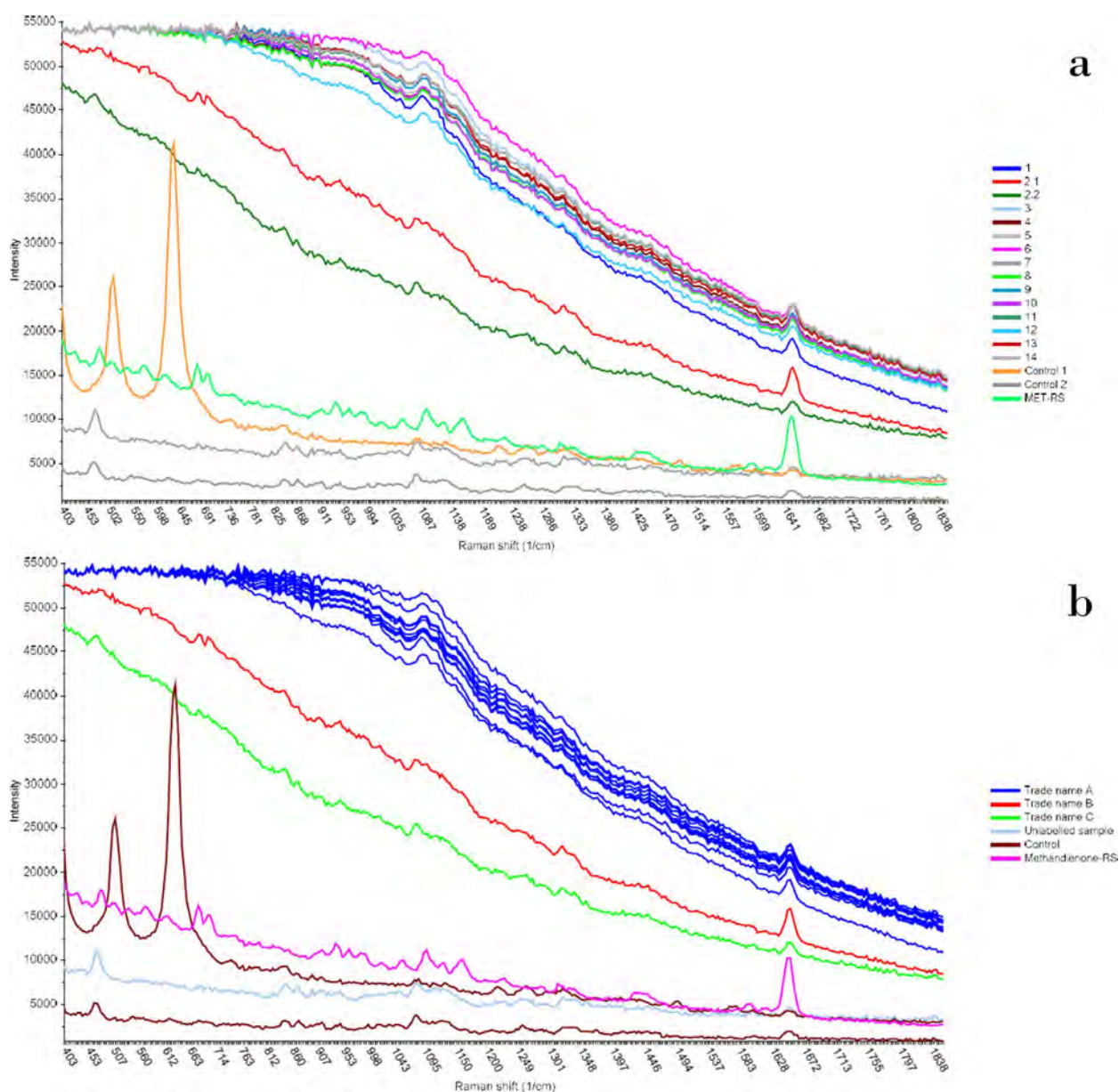
**Fig. 4** Primary packaging of falsified sample 2.2 viewed from four angles. Fig. 4 shows images taken of the primary packaging and labeling of the falsified sample 2.2 viewed from four angles **a** Front, **b** Product information, **c** Authentication Codes, **d** Hologram and website information. In Fig. 4d, the brightness of the original was increased by + 20%

**Table 2** Information on falsified methandienone product EP.Dbol-10 (sample 2.2)

Sample code	Labeled trade name	Labeled strength	Labeled batch number	Labeled Mfg.date,	Labeled Exp.date	Labeled manufacturer	Labeled manufacturer address
Sample 2.2	EP.Dbol-10	N/A	N/A	01. June 2018	01. June 2022	Confirmed fictitious	Confirmed fictitious

N/A Not available / Not labeled





**Fig. 5** Raman spectra obtained using a C13560 ultra-compact Raman spectrometer. Fig. 5 shows Raman spectra of the 15 sample products, the standard formulation without and with the sufficient removal of the coating (Control 1, Control 2), and the methandienone-reference standard (MET-RS). For each spectrum, 50 spectral data taken from five randomly selected positions each on the tablet upside and tablet downside were averaged. The MET-RS, available as powder, was placed in a plastic bag and piled up and 50 spectral data were taken with five randomly selected positions from each of the two sides. A peak near a Raman shift of 1650  $\text{cm}^{-1}$  is visible in all spectra. The spectra are presented with (a) coloring according to the sample, control, and MET-RS and (b) coloring according to the labeled trade name

show a clear peak near the Raman shift wavenumber of 1,650  $\text{cm}^{-1}$  as presented in Figs. 5a, b and 6a, b.

#### LC-MS analysis

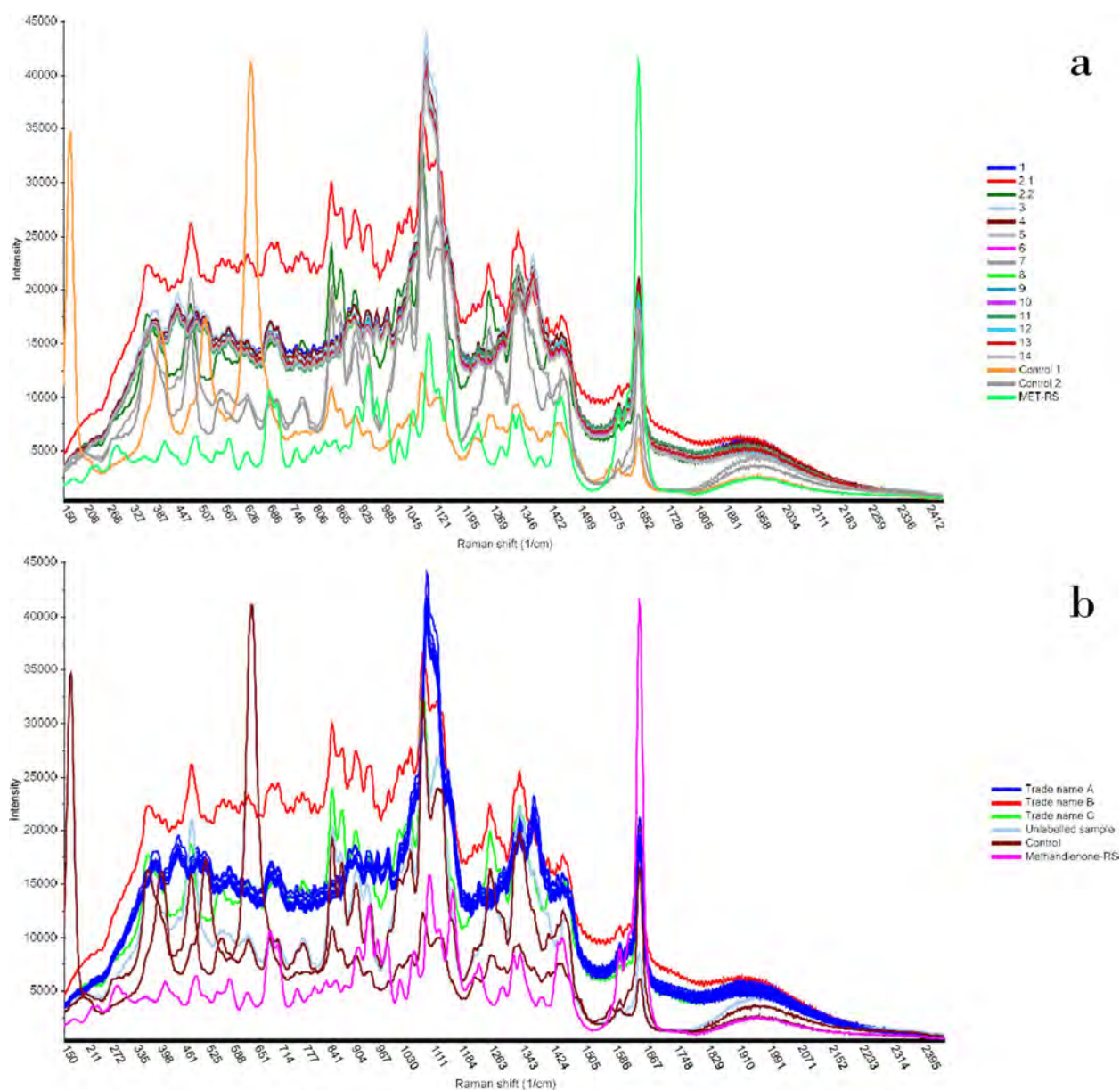
##### Detection of analogues

Twenty-five AASs (Table 1) considered in the investigation of AAS counterfeits in the Czech Republic and

Slovakia (Tircova et al. 2019) were screened for through LC-MS with full-scan  $\text{MS}^1/\text{dd-MS}^2$  mode. Among these 25 AASs, none other than MET was detected.

##### Quantitative analysis of MET

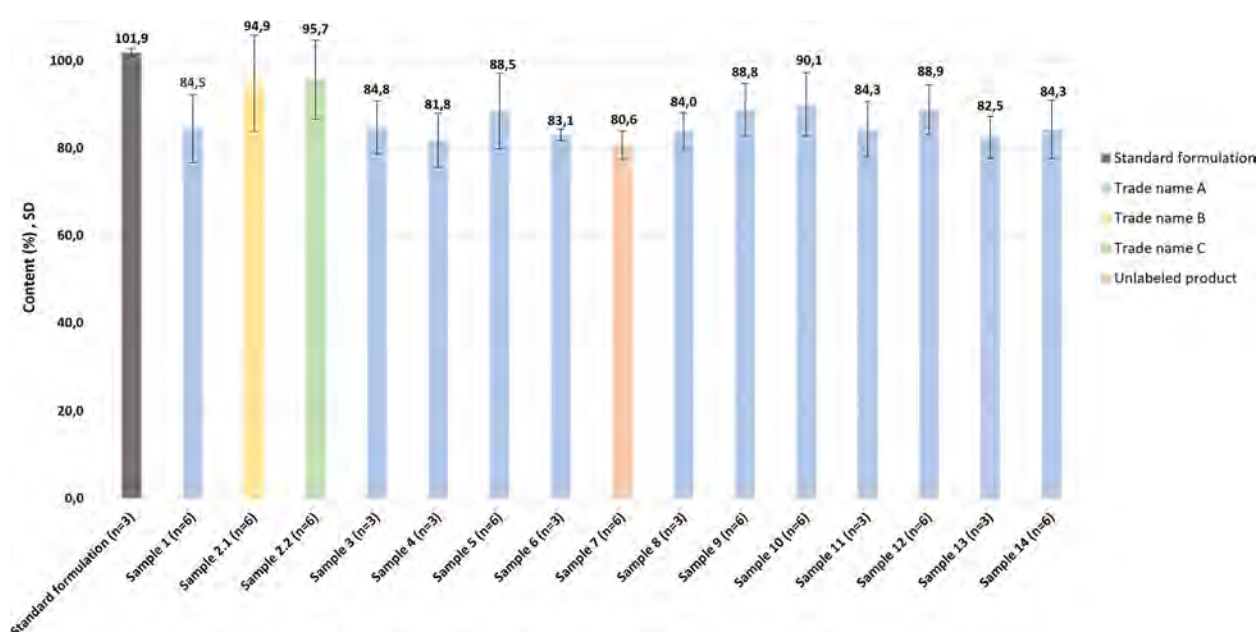
Under the high-performance liquid chromatography condition, MET eluted at 3.9 min and methyltestosterone



**Fig. 6** Raman spectra obtained using an Inspector500 ultra-compact Raman spectrometer. Fig. 6 shows Raman spectra of the 15 sample products, the standard formulation without and with the sufficient removal of the coating (Control 1, Control 2), and the methandienone-reference standard (MET-RS). For each spectrum, 50 spectral data taken from five randomly selected positions each on the tablet upside and tablet downside were averaged. The MET-RS, available as powder, was placed in a plastic bag and piled up, and 50 spectral data were taken at five randomly selected positions from each of the two sides. A peak near a Raman shift of  $1650\text{ cm}^{-1}$  is visible in all spectra. The spectra are presented with (a) coloring according to the sample, control, and MET-RS and (b) coloring according to the labeled trade name

at 4.9 min (Supplemental Fig. 1). Validation data are summarized in Supplemental Table 3. In view of the lack of labeling of sample 7, the MET content (5 mg) indicated on website 7 was chosen as the reference value. The recovery for the MET-RS of the positive control solution was calculated to be  $100.7\% \pm 5.4\%$  ( $n=9$ ). The MET content in the standard prescribed product obtained from

the Republic of Moldova was  $101.6\% \pm 2.5\%$  ( $n=3$ ) of the labeled amount. The average API content of 12 samples was less than 90%, and all 15 samples contained at least one tablet with a content of less than 90%, with 11 samples having at least one tablet with a content below 80%. Among the 15 samples, the product with the lowest average content per tablet was sample 7 with  $80.6\% \pm 3.3\%$



**Fig. 7** Content (%) and standard deviation (SD) of methandienone samples and the standard formulation determined by UPLC-FTMS. Fig. 7 shows the average content and the standard deviation (both in percent) of  $n$  tablets of each sample and of the standard formulation. The results were determined using the UPLC-FTMS method developed in this study. Bars are colored according to the product trade names

( $n=6$ ). The product with the highest average content was confirmed falsified sample 2.2 with  $95.7\% \pm 9.1\%$  ( $n=6$ ). The results of the quantification analysis are presented in Fig. 7.

#### Availability of approved MET products

Twenty-two of 44 (50%) NRAs responded to the information request, among which one, the Medicines and Medical Devices Agency of the Republic of Moldova, confirmed that it had approved an MET-containing product. The other 21 NRAs stated that no registered MET-containing pharmaceutical product was authorized or produced in their country. Furthermore, eight of the 33 institutions other than the NRAs replied to the information inquiry, with five institutions responding that their accessible databases did not show any manufacturing or exporting of MET products and the other three institutions stating non-responsibility in this matter (Supplemental Table 2).

#### Discussion

This study revealed that unapproved MET products could enter Japan through personal importation via the Internet. All  $n=15$  collected samples that were personally imported to Japan in this study were found to have defects, and one sample (6.7%) was confirmed as falsified (Fig. 4a–d). All fifteen samples (100%) failed the visual observation test, and two samples (13.3%) did not give information on the origin of manufacture. Two different portable Raman

spectrometers were found to identify MET in all  $n=16$  tested products, as verified using a developed UPLC-FTMS (Lim et al. 2016) method, which has sufficient mass accuracy for the identification and quantification of MET. This study thus adds an LC-FTMS method to a low number of available quantitative LC-MS AAS analysis methods; i.e., LC-MS/MS with triple-quadrupole mass spectrometers and time-of-flight mass spectrometry (Tircova et al. 2019; Cho et al. 2015; Van Poucke et al. 2007; Prokudina et al. 2015; Mesmer and Satzger 1997).

Raman scattering analysis was performed prior to UPLC-FTMS analysis to explore its possibilities as a rapid screening technology. UPLC-FTMS analysis was required to identify AASs, impurities other than MET, or related compounds to demonstrate that the Raman scattering methods correctly detected MET and have not falsely judged the samples. Furthermore, it is well known that SF medicines may contain no API, the wrong API, or an altered API content among individual units or impurities owing to incorrect manufacturing or storage (Rahman et al. 2018), and the API content or ingredients may differ even between units from the same blister pack or units packaged together. Therefore, the API contents of individual units were accurately determined using LC-MS rather than by performing an assay to report average contents, but the sample material was insufficient to perform further analyses.

Despite not being approved as pharmaceutical products or being licensed for distribution in Japan, all 15



collected samples were ordered without a prescription, as previously observed for other medicines (Zhu et al. 2020; Sanada et al. 2020; Rahman et al. 2018), indicating that personal importation readily enables the bypassing of medical prescriptions and the acquisition of prohibited and unauthorized products. All fifteen samples had problematic packaging, labeling, and/or content, putting consumers at high risk of accidental or unknown misuse and health hazards. Furthermore, opened, damaged or inappropriate primary packaging may lead to insufficient protection of the product from environmental effects and therefore reduced quality, efficacy, and safety, but no impurities that would provide evidence of such degradation processes were detected in this study. Inadequate or missing labeling and leaflets lead to a lack of product-related and health-specific information and can result in medication errors including incorrect application, incorrect dosing, and consumption despite contraindications. Furthermore, it was found that samples having a leaflet that is only downloadable via a labeled QR code are prone to the risk that information may be removed from websites or become inaccessible, as happened for one sample.

Although the presence of API was confirmed, the low and varying API content on average and across individual tablets of the samples (Fig. 7) suggests violations of good manufacturing practice; however, the effect of inadequate primary packaging and the storage conditions in the international distribution chain was not assessable, and therefore, it cannot be completely ruled out that product degradation reduced the content.

Notably, the online authentication message received after scanning the QR code on the label of the falsified product (Fig. 4c) demonstrates the efforts criminals are prepared to take to deceive customers and highlights that QR code authentication is not a trustable method per se. Verifying the authenticity of a sample thus requires direct information from the manufacturer or the responsible pharmaceutical entrepreneur. Remarkably, the website of the falsified product changed its layout between July and October 2023, confirming continuous activities in online product presentation and advertisement and therefore suggesting the continuous trade of falsified products.

As MET is not listed in the main pharmacopoeias (i.e., USP, BP, Int. Ph., EP, and JP), authentic products are crucial for comparison in quality analysis and differentiation from falsified products in Raman scattering analysis and especially in chemical analysis. However, owing to the low response rate of the labeled manufacturers, it was not possible to clarify the authenticity of all samples or acquire authentic products matching the samples. Regardless, problematic samples including a falsified product (Fig. 4a–d) were identified, reaffirming the importance of visual observation testing and legitimacy investigation.

In Raman scattering analysis, there was a peak near a Raman shift wavenumber of  $1,650\text{ cm}^{-1}$  for the samples, reference standard, and control (Figs. 5a, b and 6a, b), similar to previous results (Rebiere et al. 2016). Non-destructive Raman scattering analysis can thus detect the presence of MET as confirmed by LC-FTMS (Fig. 5). Notably, the spectrum of the coated standard formulation (Control 1) had a peak near  $1,650\text{ cm}^{-1}$  for both spectrometers (Figs. 5a, b and 6a, b), suggesting that MET identification is possible despite the presence of a coating, which is known to interfere often with Raman scattering analysis. Raman scattering analysis using C13560 and Inspector500 spectrometers may thus be suitable for the quick and non-destructive API detection of MET and used to screen for SF medicines having a wrong or no API or considerably low API contents, but the detection limit of the API content has not been assessed in this study. As AASs by definition have similar structures (Ganesan et al. 2023; Bhasin et al. 1996; National Institute on Drug Abuse 2018; Barceloux and Palmer 2013), identification of other AASs through portable Raman scattering analysis seems plausible. The differential identification of AASs needs to be evaluated in future work.

In comparison with LC-MS analysis, the Raman scattering methods used in this study are non-destructive, do not require organic solvents during the analytical process, can be used on-site owing to low weight and small size of the apparatuses, and can be learned time effectively by analysts in training as the analytical process is simple and fast to carry out. The methods may thus cost-effectively contribute to supply chain integrity and protection by supporting costly and time-consuming LC-MS analysis through pre-screening products of interest, or even replacing LC-MS analysis in certain scenarios such as the qualitative analysis of known MET products.

Considering that only half of the NRAs responded regarding the availability of approved products (Supplemental Table 2), better cooperation between NRAs, governmental institutions and researchers regarding requests for information on the availability and legitimacy of products should be encouraged. MET was not approved by NRAs in most countries investigated.

Limitations of the present study include that only one AAS was analyzed and evaluated with a limited number of samples collected from websites identified only per generic name and only within a short period of time. The low response rate of manufacturers prevented clarification of the authenticity of all samples, and authentic products could not be obtained for comparison in visual observation testing and analysis.

This study highlights that potent products not registered or approved as medicines in most countries are being distributed globally, in that MET products are

readily available online without a prescription for personal import to Japan, as has been reported for other medicines (McBride et al. 2018; Zhu et al. 2020; Coopman and Cordonnier 2012; Rahman et al. 2018). Substances with strong effects on human physiology, such as AASs (Bhasin et al. 1996; Rahnema et al. 2014; Solimini et al. 2017; van Amsterdam et al. 2010; Liu and Wu 2019; Takayanagi et al. 2008), must be controlled to reduce inadvertent misuse and abuse.

## Conclusion

Fifteen samples (100%) of four MET (ASS) products were obtained via personal importation over the Internet without prescription and found to have problematic packaging, labeling, and/or contents, with one sample (6.7%) confirmed as being falsified. Two portable Raman spectrometers identified the presence of MET in all samples, as subsequently confirmed adopting an LC-MS method, suggesting potential on-site use. The API was detected with a content below 90% on average for each of 12 samples. Further research is needed to differentiate AASs using portable Raman spectrometers. The results underline the importance of visual observation testing and legitimacy investigation in addition to chemical analysis. To avoid health hazards, measures of preventing the personal import and abuse of MET and other AASs are suggested for implementation in Japan and internationally.

## Abbreviations

AAS	Anabolic-androgenic steroid
API	Active pharmaceutical ingredient
FTMS	Fourier transform mass spectrometry
IPA	Isopropyl alcohol
LC	Liquid chromatography
MeCN	Acetonitrile
MeOH	Methanol
MET	Methandienone
MET-RS	Methandienone reference standard
MS	Mass spectrometry
MS/MS	Liquid chromatography–tandem mass spectrometry
<i>m/z</i>	Mass-to-charge ratio
NRA	National regulatory authority
N/A	Not available
QR code	Quick response code
SD	Standard deviation
SF medicines	Substandard and falsified medicines
UPLC	Ultra-high-performance liquid chromatography

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s41120-024-00093-0>.

**Supplementary Material 1.**

**Supplementary Material 2.**

**Supplementary Material 3.**

**Supplementary Material 4.**

## Authors' contributions

RS carried out the sample investigation, observation testing, authenticity and legitimacy investigations, Raman scattering analysis, data analysis and data interpretation, and visualization of results, wrote the original draft and reviewed and edited the manuscript. MH was majorly involved in the LC-MS method development, validation, sample analysis and visualization of results. CT, MSR, and AN were majorly involved in the LC-MS method development and sample analysis. SZ and FZ collected the samples and carried out observation testing and authenticity investigations. NY supervised and conceptualized the study and its methodology, was involved in the investigation, data curation, and resource provision, and reviewed and edited the manuscript. KM supervised the LC-MS investigation, provided resources, and reviewed and edited the manuscript. KK supervised and conceptualized the study and its methodology, was involved in the investigation and data curation, reviewed and edited the manuscript, and was responsible for funding acquisition and project administration. All authors read and approved the final manuscript.

## Declarations

### Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

### Competing interests

The authors declare that they have no competing interests.

### Funding

This work was funded by the Ministry of Health, Labour and Welfare of the Japan Pharmaceuticals and the Medical Devices Regulatory Science Policy Research Program under grant number H30-iyaku-ippan-001.

### Acknowledgements

This study was made possible by the generous support of the Japanese Ministry of Health, Labour and Welfare, for which the authors express their gratitude. The authors thank SC Balkan Pharmaceuticals SRL (Singer, Republic of Moldova) for providing DANABOL 10 mg for purchase; DANABOL 10 mg was used as the standard formulation and control in this study. The authors thank the national regulatory authorities, in particular FDA Philippines, FDA Taiwan, and RP Tübingen, Germany as well as the municipality of Gaißach, Germany for their cooperation. The authors thank Edanz (<https://jp.edanz.com/ac>) for editing a draft of this manuscript. The contribution of RS was supported by the PhD "Monbukagakusho" scholarship from the Japanese Ministry of Education, Culture, Sports, Science and Technology in Japan.

### Author details

<sup>1</sup>Clinical Pharmacy and Healthcare Sciences, Division of Pharmaceutical Sciences, Graduate School of Medical Sciences, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan. <sup>2</sup>Department of Analytical Chemistry, Faculty of Pharmaceutical Sciences, Doshisha Women's College of Liberal Arts, Kyotanabe, Kyoto 610-0395, Japan. <sup>3</sup>Medi-Quality Security Institute, Graduate School of Medical Sciences, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan. <sup>4</sup>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. <sup>5</sup>Medicine Security Workshop, 4F Venture Business Laboratory, Kanazawa University Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan.

**Received: 20 December 2023 Accepted: 24 April 2024**  
**Published online: 03 June 2024**

## References

- Ahmed J, Modica de Mohac L, Mackey TK, Raimi-Abraham BT (2022) A critical review on the availability of substandard and falsified medicines online: incidence, challenges and perspectives. *J Med Access* 6:23992026221074548. <https://doi.org/10.1177/23992026221074548>



- Barceloux DG, Palmer RB (2013) Anabolic-androgenic steroids. *Dis Mon* 59(6):226–248. <https://doi.org/10.1016/j.disamonth.2013.03.010>
- Bhasin S, Storer TW, Berman N, Callegari C, Clevenger B, Phillips J et al (1996) The effects of supraphysiologic doses of testosterone on muscle size and strength in normal men. *N Engl J Med* 335(1):1–7. <https://doi.org/10.1056/NEJM199607043350101>
- Cho SH, Park HJ, Lee JH, Do JA, Heo S, Jo JH et al (2015) Determination of anabolic-androgenic steroid adulterants in counterfeit drugs by UHPLC-MS/MS. *J Pharm Biomed Anal* 111:138–146. <https://doi.org/10.1016/j.jpba.2015.03.018>
- Coopman V, Cordonnier J (2012) Counterfeit drugs and pharmaceutical preparations seized from the black market among bodybuilders. *Ann Toxicol Anal* 24(2):73–80. <https://doi.org/10.1051/ata/2012012>
- Dégardin K, Guillemin A, Roggo Y (2017) Comprehensive Study of a Hand-held Raman Spectrometer for the analysis of counterfeits of solid-dosage form Medicines. *J Spectrosc (Hindawi)* 1:1–13 Article ID 3154035. <https://doi.org/10.1155/2017/3154035>
- Fabresse N, Gheddar L, Kintz P, Knapp A, Larabi IA, Alvarez JC (2021) Analysis of pharmaceutical products and dietary supplements seized from the black market among bodybuilders. *Forensic Sci Int* 322:110771. <https://doi.org/10.1016/j.forsciint.2021.110771>
- Frude E, McKay FH, Dunn M (2020) A focused netnographic study exploring experiences associated with counterfeit and contaminated anabolic-androgenic steroids. *Harm Reduct J* 17(1):42. <https://doi.org/10.1186/s12954-020-00387-y>
- Ganesan K, Rahman S, Zito PM (2023) Anabolic steroids. StatPearls. StatPearls, Treasure Island (FL). Available from: <https://www.ncbi.nlm.nih.gov/books/NBK482418>. Accessed 16 Aug 2023
- Hajjou M, Qin Y, Bradby S, Bempong D, Lukulay P (2013) Assessment of the performance of a handheld Raman device for potential use as a screening tool in evaluating medicines quality. *J Pharm Biomed Anal* 74:47–55. <https://doi.org/10.1016/j.jpba.2012.09.016>
- Khan MH, Tanimoto T, Nakanishi Y, Yoshida N, Tsuboi H, Kimura K (2012) Public health concerns for anti-obesity medicines imported for personal use through the internet: a cross-sectional study. *BMJ Open* 2:e000854. <https://doi.org/10.1136/bmjopen-2012-000854>
- Lim L, Yan F, Bach S, Pihakari K, Klein D (2016) Fourier Transform Mass Spectrometry: the Transformation of Modern Environmental analyses. *Int J Mol Sci* 17(1):104. <https://doi.org/10.3390/ijms17010104>
- Liu JD, Wu YQ (2019) Anabolic-androgenic steroids and cardiovascular risk. *Chin Med J (Engl)* 132(18):2229–2236. <https://doi.org/10.1097/CM9.0000000000000407>
- Magnolini R, Falcato L, Cremonesi A, Schori D, Bruggmann P (2022) Fake anabolic androgenic steroids on the black market - a systematic review and meta-analysis on qualitative and quantitative analytical results found within the literature. *BMC Public Health* 22(1):1371. <https://doi.org/10.1186/s12889-022-13734-4>
- McBride JA, Carson CC 3rd, Coward RM (2018) The availability and Acquisition of Illicit anabolic androgenic steroids and Testosterone preparations on the internet. *Am J Mens Health* 12(5):1352–1357. <https://doi.org/10.1177/1557988316648704>
- Mesmer MZ, Satzger RD (1997) Determination of anabolic steroids by HPLC with UV-vis-particle beam mass spectrometry. *J Chromatogr Sci* 35(1):38–42. <https://doi.org/10.1093/chromsci/35.1.38>
- National Institute on Drug Abuse, National Institutes of Health (2018) Anabolic Steroids. In: U.S. Department of Health and Human Services. <https://nida.nih.gov/sites/default/files/drugfacts-steroids.pdf>. Accessed 25 Aug 2023
- Nieschlag E, Vorona E (2015) Doping with anabolic androgenic steroids (AAS): adverse effects on non-reproductive organs and functions. *Rev Endocr Metab Disord* 16(3):199–211. <https://doi.org/10.1007/s11154-015-9320-5>
- Opuni KF, Nettey H, Larbi MA, Amartey SNA, Nti G, Dzidonu A et al (2019) Usefulness of combined screening methods for rapid detection of falsified and/or substandard medicines in the absence of a confirmatory method. *Malar J* 18(1):403. <https://doi.org/10.1186/s12936-019-3045-y>
- Piatkowski T, Puljevic C, Francis C, Ferris J, Dunn M (2023) They sent it away for testing and it was all bunk: exploring perspectives on drug checking among steroid consumers in Queensland, Australia. *Int J Drug Policy* 119:104139. <https://doi.org/10.1016/j.drugpo.2023.104139>
- Prokudina EA, Prchalová J, Vyšatová E, Kuchař M, Rajchl A, Lapčík O (2015) Analysis of anabolic androgenic steroids by direct analysis in real time ionization with time-of-flight mass spectrometry. *Int J Mass Spectrom* 392:28–33. <https://doi.org/10.1016/j.ijms.2015.08.022>
- Rahnema CD, Lipshultz LI, Crosnoe LE, Kovac JR, Kim ED (2014) Anabolic steroid-induced hypogonadism: diagnosis and treatment. *Fertil Steril* 101(5):1271–1279. <https://doi.org/10.1016/j.fertnstert.2014.02.002>
- Rahman MS, Yoshida N, Tsuboi H, Tomizu N, Endo J, Miyu O et al (2018) The health consequences of falsified medicines- A study of the published literature. *Trop Med Int Health* 23(12):1294–1303. <https://doi.org/10.1111/tmi.13161>
- Rahman MS, Yoshida N, Hanafusa M, Matsuo A, Zhu S, Stub Y et al (2022) Screening and quantification of undeclared PGF2α analogs in eyelash-enhancing cosmetic serums using LC-MS/MS. *J Pharm Biomed Anal* 219:114940. <https://doi.org/10.1016/j.jpba.2022.114940>
- Rahman MS, Yoshida N, Sugiura S, Tsuboi H, Keila T, Kiet HB et al (2018) Quality of omeprazole purchased via the internet and personally imported into Japan: comparison with products sampled in other Asian countries. *Trop Med Int Health* 23(3):263–269. <https://doi.org/10.1111/tmi.13028>
- Rebieri H, Ghyselinck C, Lempereur L, Brenier C (2016) Investigation of the composition of anabolic tablets using near infrared spectroscopy and Raman chemical imaging. *Drug Test Anal* 8(3–4):370–377. <https://doi.org/10.1002/dta.1843>
- Sagoe D, Molde H, Andreassen CS, Torsheim T, Pallesen S (2014) The global epidemiology of anabolic-androgenic steroid use: a meta-analysis and meta-regression analysis. *Ann Epidemiol* 24(5):383–398. <https://doi.org/10.1016/j.annepidem.2014.01.009>
- Sanada T, Ohnishi M, Yoshida N, Kimura K, Tsuboi H (2021) Quality assessment of Diflucan® tablets distributed online: Diflucan® distributed online. *Med Access Point Care* 5:23992026211002089. <https://doi.org/10.1177/23992026211002089>
- Sanada T, Yoshida N, Kimura K, Tsuboi H (2020) Discrimination of falsified erectile dysfunction medicines by use of an ultra-compact Raman scattering spectrometer. *Pharm (Basel)* 9(1):3. <https://doi.org/10.3390/pharm9010003>
- Sanada T, Yoshida N, Matsushita R, Kimura K, Tsuboi H (2020) Falsified tadalafil tablets distributed in Japan via the internet. *Forensic Sci Int* 307:110143. <https://doi.org/10.1016/j.forsciint.2020.10143>
- Storer TW, Magliano L, Woodhouse L, Lee ML, Dzekov C, Dzekov J et al (2003) Testosterone dose-dependently increases maximal voluntary strength and leg power, but does not affect fatigability or specific tension. *J Clin Endocrinol Metab* 88(4):1478–1485. <https://doi.org/10.1210/jc.2002-021231>
- Solimini R, Rotolo MC, Mastrobattista L, Mortali C, Minutillo A, Pichini S et al (2017) Hepatotoxicity associated with illicit use of anabolic androgenic steroids in doping. *Eur Rev Med Pharmacol Sci* 21(1 Suppl):7–16
- Takayanagi A, Kobayashi K, Hashimoto K, Kato R, Masumori N, Itoh N et al (2008) Case of androgenic anabolic steroid abuse caused hypogonadotropic hypogonadism. *Nihon Hinyokika Gakkai Zasshi* 99(7):729–32. <https://doi.org/10.5980/jpnjuro1989.99.729>. Japanese
- The Ministry of Health, Labour and Welfare (2021) Official Monographs (A to L), The Japanese Pharmacopeia 18th edition (JP XVIII). 7:960–961. <https://www.mhlw.go.jp/content/11120000/000912386.pdf>. Accessed 05 Sept 2023
- The Ministry of Health, Labour and Welfare (2021) Official Monographs (M to Z), The Japanese Pharmacopeia 18th edition (JP XVIII). 7:1572–1573. <https://www.mhlw.go.jp/content/11120000/000912386.pdf>. Accessed 05 Sept 2023
- Tool for Visual Inspection of Medicines (n.d.) International Pharmaceutical Federation <https://www.fip.org/files/fip/counterfeit/VisualInspection/A%20tool%20for%20visual%20inspection%20of%20medicines%20EN.pdf>. Accessed 25 Aug 2023
- Tircova B, Bosakova Z, Kozlik P (2019) Development of an ultra-high performance liquid chromatography-tandem mass spectrometry method for the determination of anabolic steroids currently available on the black market in the Czech Republic and Slovakia. *Drug Test Anal* 11(2):355–360. <https://doi.org/10.1002/dta.2541>
- van Amsterdam J, Opperhuizen A, Hartgens F (2010) Adverse health effects of anabolic-androgenic steroids. *Regul Toxicol Pharmacol* 57(1):117–123. <https://doi.org/10.1016/j.yrtph.2010.02.001>
- Van Poucke C, Detavernier C, Van Cauwenbergh R, Van Peteghem C (2007) Determination of anabolic steroids in dietary supplements by liquid chromatography-tandem mass spectrometry. *Anal Chim Acta* 586(1–2):35–42. <https://doi.org/10.1016/j.aca.2006.09.050>

- World Health Organization (2018) Substandard and falsified medical products. 2018 Jan 31. <https://www.who.int/news-room/fact-sheets/detail/substandard-and-falsified-medical-products>. Accessed 25 Aug 2023
- World anti-doping agency (2023) Prohibited List, international standard. In: World anti-doping code. [https://www.wada-ama.org/sites/default/files/2022-09/2023list\\_en\\_final\\_9\\_september\\_2022.pdf](https://www.wada-ama.org/sites/default/files/2022-09/2023list_en_final_9_september_2022.pdf). Accessed 25 Aug 2023
- Yoshida N, Numano M, Nagasaka Y, Ueda K, Tsuboi H, Tanimoto T et al (2015) 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(1):430. <https://doi.org/10.1186/s12906-015-0955-2>
- Zhu S, Yoshida N, Kimura K, Matsushita R, Tsuboi H (2020) Falsified vardenafil tablets available online. *J Pharm Biomed Anal* 177:112872. <https://doi.org/10.1016/j.jpba.2019.112872>

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## RESEARCH ARTICLE

## MEDICAL PHYSICS

# Estimation of threshold thickness of residual normal tissue in lung dysfunction detectable by dynamic chest radiography: A virtual imaging trial

Shunya Yamaguchi<sup>1</sup> | Rie Tanaka<sup>2</sup> | Isao Matsumoto<sup>3</sup> | Noriyuki Ohkura<sup>4</sup> | William Paul Segars<sup>5</sup> | Ehsan Abadi<sup>5</sup> | Ehsan Samei<sup>5</sup>

<sup>1</sup>Graduate School of Health Sciences, Kanazawa University, Kanazawa, Ishikawa, Japan

<sup>2</sup>College of Medical, Pharmaceutical & Health Sciences, Kanazawa University, Kanazawa, Ishikawa, Japan

<sup>3</sup>Department of Thoracic Surgery, Kanazawa University, Kanazawa, Ishikawa, Japan

<sup>4</sup>Department of Respiratory Medicine, Kanazawa University Hospital, Kanazawa, Ishikawa, Japan

<sup>5</sup>Carl E Ravin Advanced Imaging Labs, Department of Radiology, Duke University, Durham, North Carolina, USA

## Correspondence

Rie Tanaka, College of Medical, Pharmaceutical & Health Sciences, Kanazawa University; 5-11-80 Kodatsuno, Kanazawa 920-0942, Japan.  
Email: rie44@mhs.mp.kanazawa-u.ac.jp

## Funding information

JSPS KAKENHI Grant-in-aid for Scientific Research, Grant/Award Number: 21H02866; JSPS KAKENHI Fostering Joint International Research, Grant/Award Number: 21KK0286; the National Institute of Health, Grant/Award Number: 5P41EB028744; Tateishi Science and Technology Foundation

## Abstract

**Background:** Dynamic chest radiography (DCR) is a recently developed functional x-ray imaging technique that detects pulmonary ventilation impairment as a decrease in changes in lung density during respiration. However, the diagnostic performance of DCR is uncertain owing to an insufficient number of clinical cases. One solution is virtual imaging trials (VITs), which is an emerging alternative method for efficiently evaluating medical imaging technology via computer simulation techniques.

**Purpose:** This study aimed to estimate the typical threshold thickness of residual normal tissue below which the presence of emphysema may be detected by DCR via VITs using virtual patients with different physiques and a user-defined ground truth.

**Methods:** Twenty extended cardiac-torso (XCAT) phantoms that exhibited changes in lung density during respiration were generated to simulate virtual patients. To simulate a locally collapsed lung, an air sphere was inserted into each lung regions in the phantom. The XCAT phantom was virtually projected using an x-ray simulator. The respiratory changes in pixel value ( $\Delta PV$ ) were measured on the projected air spheres (simulated lesions) to calculate the percentage of decrease ( $\Delta PV\%$ ) relative to  $\Delta PV_{\text{exp-ins}}$  in the absence of an air sphere. The relationship between the amount of residual normal tissue and  $\Delta PV\%$  was fitted to a cubic approximation curve (hereafter, performance curve), and the threshold at which the  $\Delta PV\%$  began to decrease ( $\text{normal-tissue}_{\text{thre}}$ ) was determined. The goodness of fit for each performance curve was evaluated according to the coefficient of determination ( $R^2$ ) and the 95% confidence interval derived from the standard errors between the measured and theoretical values corresponding to each performance curve. The  $\Delta PV\%$  was also visualized as a color scaling to validate the results of the VITs in both virtual and clinical patients.

**Results:** For each lung region in all body sizes, the  $\Delta PV\%$  decreased as the amount of residual normal tissue decreased and could be defined as a function of the amount of residual normal tissue in front of and behind the simulated lesions with high  $R^2$  values. Meanwhile, the difference between the measured and theoretical values corresponding to each performance curve was only partially included in the 95% confidence interval. The  $\text{normal-tissue}_{\text{thre}}$  values

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Medical Physics* published by Wiley Periodicals LLC on behalf of American Association of Physicists in Medicine.

were 146.0, 179.5, and 170.9 mm for the upper, middle, and lower lungs, respectively, which were demonstrated in virtual patients and one real patient, where the value of the residual normal tissue was less than that of normal-tissue<sub>thre</sub>; any reduction in the residual normal tissue was reflected as a reduced  $\Delta PV$  and depicted as a reduced color intensity.

**Conclusions:** The performance of DCR-based pulmonary impairment assessment depends on the amount of residual normal tissue in front of and behind the lesion rather than on the lesion size. The performance curve can be defined as a function of the amount of residual normal tissue in each lung region with a specific threshold of normal tissue remaining where lesions become detectable, shown as a decrease in  $\Delta PV$ . The results of VITs are expected to accelerate future clinical trials for DCR-based pulmonary function assessment.

#### KEYWORDS

dynamic chest radiography, pulmonary ventilation, virtual imaging trial

## 1 | INTRODUCTION

Respiratory disease is one of the leading causes of death worldwide. Specifically, chronic obstructive pulmonary disease is the third leading cause,<sup>1</sup> and pneumonia and other lower respiratory tract infections are the fourth leading cause, with 5.42 million excess deaths due to COVID-19 (Jan 2020–Dec 2021).<sup>2,3</sup> To determine the severity of respiratory diseases and the treatment options, pulmonary function is generally evaluated using pulmonary function tests (PFTs), lung scintigraphy, four-dimensional computed tomography (4DCT), and cine magnetic resonance imaging (MRI).<sup>4–9</sup> Although these methods are effective for diagnosing pulmonary function, PFTs cannot be performed in the presence of COVID-19 because of infection control measures, and access to diagnostic equipment such as 4DCT and cine MRI is limited in routine clinical practice. A simpler and more accessible examination method would allow more medical facilities to diagnose pulmonary function and follow up on the results. Dynamic chest radiography (DCR) is such a method, as it can be performed in general radiography rooms.<sup>10</sup>

DCR is a flat panel detector (FPD)-based functional x-ray imaging technique proposed in the early 2000s and commercialized in 2019.<sup>10,11</sup> The total entrance surface dose delivered by DCR is lower than that recommended by the International Atomic Energy Agency (1.9 mGy) for the frontal and lateral views produced by conventional chest radiography.<sup>12,13</sup> Sequential images taken during respiration reveal the motion of the ribs, diaphragm, and heart wall, as well as the changes in lung density resulting from respiration and cardiac pumping. Furthermore, previous animal and clinical studies have indicated that DCR can be used to quantify and visualize pulmonary ventilation and circulation through temporal changes in pixel values ( $\Delta PV$ ), and that it can also be used to detect pulmonary impairments as decreased  $\Delta PV$  even without contrast media.<sup>14–21</sup>

Although several clinical studies have performed pulmonary function evaluations using DCR, its diagnostic performance is uncertain because of an insufficient number of clinical cases with a sufficiently wide variety of pulmonary ventilation impairments. In animal studies, airway obstruction was conducted at the lobe bronchi level; therefore, the detection performance of DCR for pulmonary dysfunction was demonstrated only in the lobe unit.<sup>17,18</sup>

One way to efficiently acquire more data on the performance of DCR is VITs. In VITs, human subjects are replaced by virtual patients, for which medical images are simulated. VITs are an efficient and non-invasive alternative to clinical trials and are useful for optimizing imaging conditions and multimodality performance evaluations.<sup>22</sup> Consequently, the results of VITs are expected to accelerate future clinical trials that use DCR-based pulmonary function assessments. One means of generating virtual patients is 4D XCAT phantoms, which are computational phantoms that have been developed based on 4DCT data of real human bodies. An XCAT phantom simulates the changes in thousands of organ structures and their associated characteristics during cardiac and respiratory motions, including the changes in lung density that occur during respiration.<sup>23–26</sup> In addition, lesions of any size, location, and composition can be simulated anywhere in an XCAT phantom, which is not possible in animal experiments. A previous preliminary study based on one reference-size XCAT phantom reported that the degree to which the  $\Delta PV$  decreased depended on the amount of normal lung tissue in front of and behind the area with pulmonary impairments.<sup>27</sup> However, the influence of body size, gender, and the changes in the body thickness and thorax motion that occur during respiration remains to be determined. In addition, the incremental diameter of the air sphere was 20 mm, which limited the number of air sphere measurements that could be performed. Because of these limitations, the maximum residual



**TABLE 1** Characteristics of the 4D XCAT phantoms.

Sex	Body size	Number	Age (years)	Height (cm)	Weight (kg)	BMI <sup>a</sup> (kg/m <sup>2</sup> )
Male	Obese	5	50.6 ± 15.8	176.7 ± 3.4	102.9 ± 3.5	32.9 ± 0.9
	Overweight	5	56.8 ± 16.2	179.2 ± 4.9	89.3 ± 4.8	27.8 ± 0.4
Female	Normal	5	56.6 ± 14.3	178.4 ± 6.0	76.0 ± 6.7	23.8 ± 0.6
	Normal	5	55.4 ± 5.2	167.0 ± 8.4	64.8 ± 4.7	23.2 ± 0.7

Note: The data are expressed as mean ± standard deviation.

Abbreviation: XCAT, extended cardiac-torso.

<sup>a</sup>BMI: body mass index.

normal tissue at which emphysema can be detected as decreased  $\Delta PV$  by DCR remains unresolved. If a VIT was conducted on an increasing number of virtual patients with various physiques under simplified conditions that do not involve movement of the rib cage and diaphragm, these issues seen in the previous study could be overcome. Elucidation of these factors would contribute to more efficient clinical trials and thereby help in the early establishment and dissemination of DCR-based pulmonary function assessment.

This study aimed to estimate the threshold thickness of residual normal tissue below which the presence of emphysema can be detected by DCR through VITs of virtual patients with different physiques and a user-defined ground truth.

## 2 | MATERIALS AND METHODS

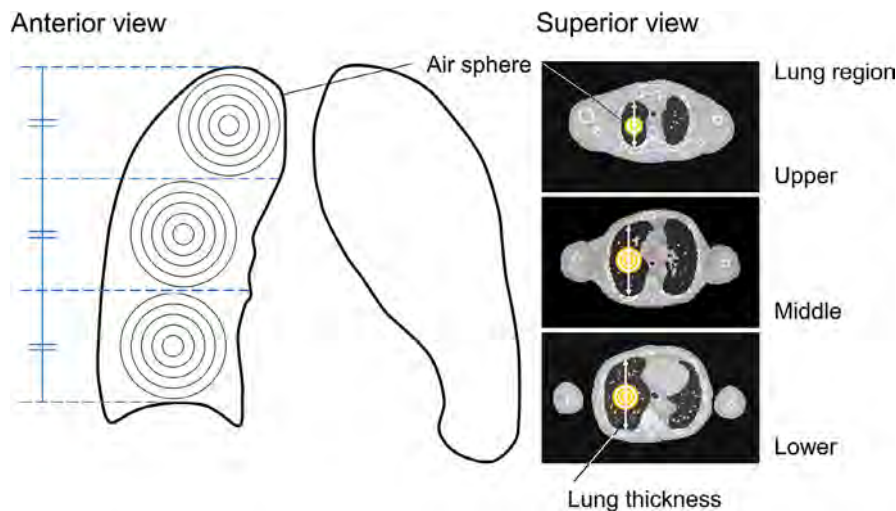
### 2.1 | Phantom preparation

A total of 20 XCAT phantoms, of which 15 were males (five obese, five overweight, and five with normal body sizes) and five were females (all with normal body sizes), were generated by the XCAT program.<sup>23–25</sup> The characteristics of the XCAT phantoms are shown in Table 1. Using the World Health Organization's body mass index (BMI [kg/m<sup>2</sup>]) classification criteria,<sup>28</sup> the body sizes were classified into three groups: obese (30.0–34.9 kg/m<sup>2</sup>), overweight (25.0–29.9 kg/m<sup>2</sup>), and normal (18.5–24.9 kg/m<sup>2</sup>). To evaluate how small pulmonary ventilation impairments can be detected in the form of decreased  $\Delta PV$  in ideal conditions without the displacement of lung structures, the XCAT phantoms were designed such that the breathing caused changes in lung density equivalent to a 4 cm movement of the diaphragm but did not move the rib cage and diaphragm. The pulmonary vessels and bronchi of the XCAT phantom were reproduced up to the fifth generation.<sup>25</sup> The density of the inflated lung was set to 0.26 g/cm<sup>3</sup>, the average lung density of a healthy adult,<sup>29</sup> and increased to 0.40 g/cm<sup>3</sup> at the deflated lung. To simulate a locally collapsed lung, that is, that with emphysema, an air sphere was inserted into the center of each of three lung regions (upper, middle, and lower), using computed

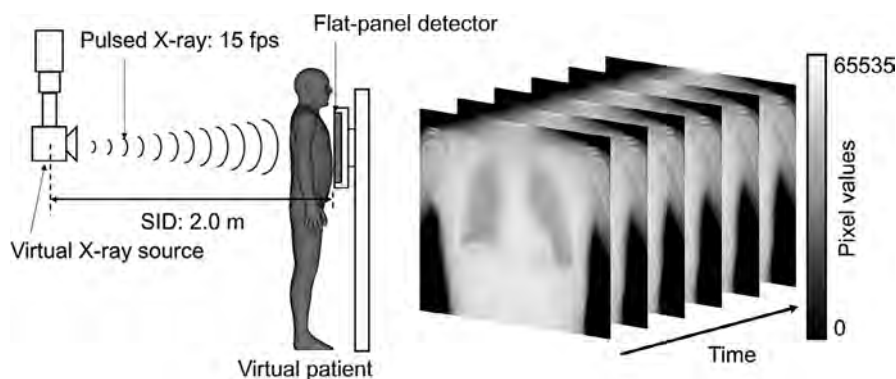
tomography (CT) images as references (Figure 1). For the upper lung region, the diameter of the air sphere ranged from 10 to 60 mm in 10 mm increments, and for the middle and lower lung regions, it ranged from 20 to 120 mm in 20 mm increments. The thickness of each lung region was measured at the center of each air sphere (see Figure 1 and Table 1). The residual normal tissue in each lung region was then derived by subtracting the diameter of each air sphere from the measured lung thickness. The bone suppression (BS) technique is essential for analyzing very small  $\Delta PV$  of a lung field; however, in this study, the bone shadows can be a source of error in the  $\Delta PV$  measurements due to inter-frame differences in BS processing effect. Therefore, assuming that the BS technique was perfectly applied to the resulting projection images, bone structures (e.g., the ribs, clavicles, and scapula) were removed by deleting the descriptions identified by unique organ IDs from the XCAT phantoms before the projection images were generated.

### 2.2 | Creation of projection images

The XCAT phantoms were virtually projected using an x-ray simulator, which was set to model a dynamic FPD system (PaxScan, 4343CB, Varex Imaging, Salt Lake City, Utah, USA). The projection images were created under the same imaging conditions as the actual DCR imaging conditions (100 kV, 0.20 mAs/pulse, 15 frames per second, Fluoro mode: 3 × 3 binning, source-to-detector distance (SID): 2.0 m, 0.2 mm Cu additional filter), without electric noise, using an energy spectrum file produced by a free spectrum processing software (Spectrum processor SRS-78),<sup>30</sup> and a 10 s respiratory cycle composed of 150 phases was obtained (see Figure 2). The image size was 427 × 427 mm, the matrix size was 512 × 512 pixels (half of the normal matrix size of 1024 × 1024 pixels to save computation time), the pixel size was 834 × 834  $\mu\text{m}$ ,<sup>2</sup> and the grayscale was 32-bit real. The projection images were rescaled from 32-bit real grayscale images to 16-bit unsigned grayscale images, and the histograms were matched to that of actual DCR images. DCR creates images using a linear representation of  $\Delta PV$  and evaluates the



**FIGURE 1** Locations of air spheres inserted into the right lung of each 4D XCAT phantom. The blue dashed lines indicate the three equal regions into which the lung fields were divided: upper, middle, and lower. The black (on the left) and orange (on the right) circles indicate the locations of the inserted air spheres. The white arrows on the right indicate the thickness of the lung for each region. XCAT, extended cardiac-torso.



**FIGURE 2** Virtual setup (left) for generating projection images (right) of the 4D XCAT phantom SID. SID, source-to-image distance; XCAT, extended cardiac-torso.

relative pulmonary function based on relative changes in  $\Delta PV$ s due to respiration. Therefore, the simple rescaling process of dividing the original value by  $2^{16}$  followed by adjusting the  $\Delta PV$ s was sufficient for this study, in which the measurement target was changes in  $\Delta PV$ . High  $\Delta PV$ s corresponded to white areas in the images, and these in turn corresponded to areas of high x-ray absorption or low x-ray translucency through the body. Although the shape of the inserted air sphere was simple and did not simulate the actual shape of the lesion on the projected image, the lung area where the air sphere was placed was simulated to exhibit increased x-ray translucency and fewer changes during breathing.

## 2.3 | Measurement of $\Delta PV$ and analysis

To measure the average  $\Delta PV$  in both the presence and absence of the projected air spheres at the same coordi-

nate position, regions of interest (ROIs) 10 mm in diameter were manually located in each lung region using the projected images. Then, for each ROI, the difference between the average  $\Delta PV$ s of the maximum expiration and maximum inspiration frames (hereafter,  $\Delta PV_{\text{exp-ins}}$ ) was calculated. Next, the relative rate at which  $\Delta PV_{\text{exp-ins}}$  decreased (hereafter,  $\Delta PV\%$ ) was calculated by dividing  $\Delta PV_{\text{exp-ins}}$  corresponding to the presence of an air sphere (simulated lesion) by  $\Delta PV_{\text{exp-ins}}$  corresponding to the absence of an air sphere (normal region). In equation form, this is expressed as<sup>31</sup>

$$\Delta PV\% = \frac{\Delta PV_{\text{exp-ins}} \text{ on simulated lesion}}{\Delta PV_{\text{exp-ins}} \text{ on normal region}} \times 100. \quad (1)$$

For each lung region in each body size, the relationship between the diameter of the air sphere and  $\Delta PV\%$  was determined. In addition, the relationship between the amount of residual normal tissue and  $\Delta PV\%$  was

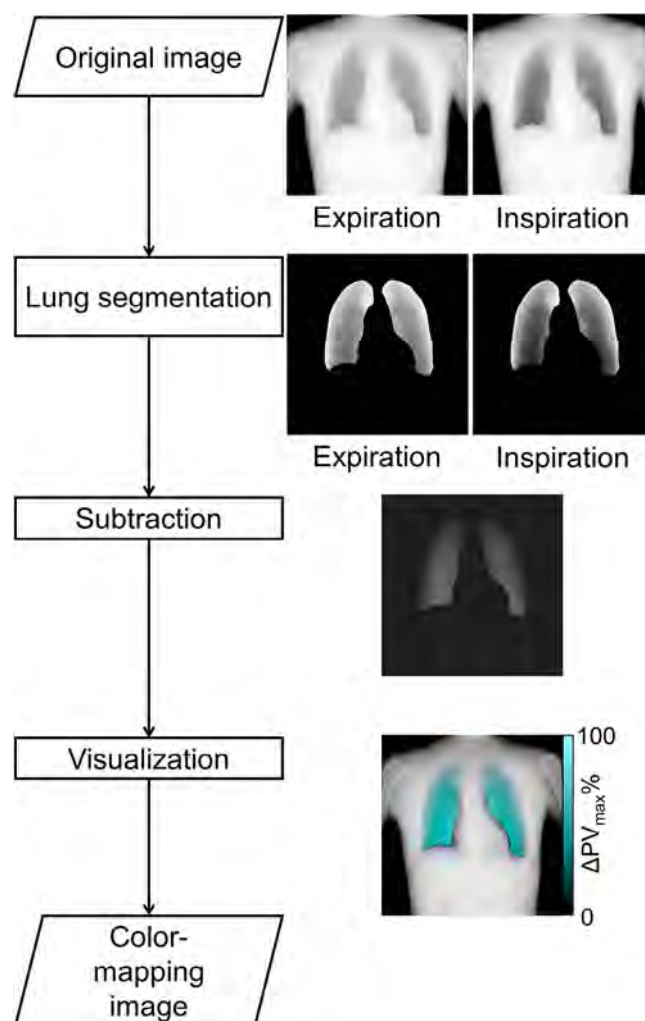
established and fitted to a cubic approximation curve (hereafter, performance curve). The coefficient of determination ( $R^2$ ) was used to evaluate the goodness of fit of each performance curve. In addition, the threshold value at which  $\Delta PV\%$  began to decrease (hereafter,  $\text{normal-tissue}_{\text{thre}}$ ) as the remaining normal tissue decreases was determined as a real solution to the first derivative of the performance curve. When a real solution could not be obtained for the first derivative, the thickness of the remaining normal tissue with the smallest slope of the performance curve was used as the threshold value which can be obtained as a real solution of the second derivative of the performance curve. As previously reported, changes in lung density during respiration depend on the relative increase or decrease in the amount of air and the pulmonary vascular volume per unit lung volume.<sup>32</sup> Thus, the lung regions with impaired pulmonary ventilation less than  $\text{normal-tissue}_{\text{thre}}$  would be represented as a decrease in  $\Delta PV\%$  regardless of their size, but those more than  $\text{normal-tissue}_{\text{thre}}$  would not be represented as a decrease in  $\Delta PV\%$ .

## 2.4 | Statistical analysis

Statistical analysis was performed to determine the goodness of fit of the performance curves for each lung region. The theoretical values were calculated by substituting the residual normal tissue (mm) into the performance curve, whereas the measured values were the average  $\Delta PV$ s measured in ROIs located in projection images. Subsequently, the standard errors between the measured values and the theoretical values of the performance curves were calculated. The mean of each standard error was then calculated, and twice its value was deemed the 95% confidence interval. The performance curves obtained from the 20 XCAT phantoms for each lung region were examined to determine whether the difference between the theoretical values for the two performance curves, that is, the difference in  $\Delta PV\%$  when the same amount of residual normal tissue is present, was included in the 95% confidence interval.

## 2.5 | Visualization of the changes in the $\Delta PV$ s of the virtual patients

To facilitate visual evaluation, the  $\Delta PV$  were visualized as color scales, in which larger  $\Delta PV$  were indicated by higher color intensities (hereafter, color-mapping images) (Figure 3).<sup>19,27</sup> Color-mapping images were created using a home-built program (C++BuilderXE8; Embarcadero Technologies, Austin, TX, USA), as in previous studies.<sup>19,27</sup> The degree of color transparency in the color-mapping images was determined based on

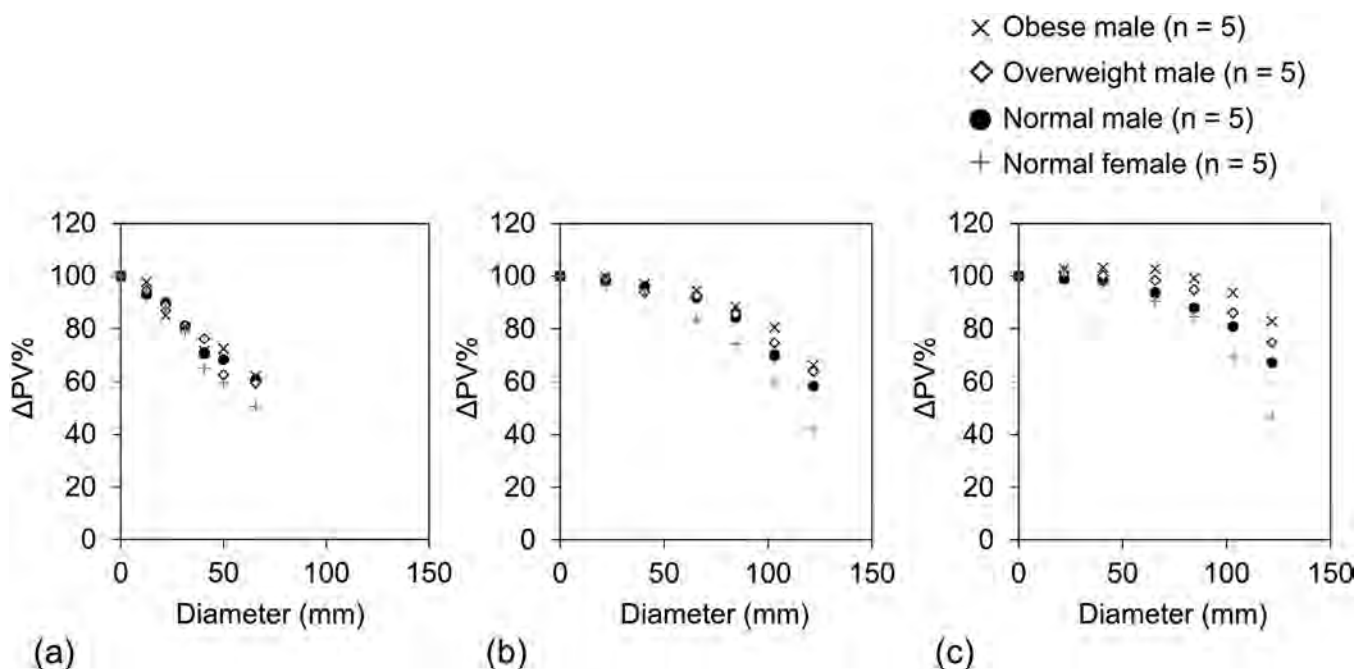


**FIGURE 3** Overall scheme of the algorithm used for visualizing changes in the lung density in the 4D XCAT phantoms. XCAT, extended cardiac-torso.

normalized  $\Delta PV$ , which were derived by dividing the  $\Delta PV$  for each pixel by the maximum  $\Delta PV$  ( $\Delta PV_{\text{max}}$ ) in the whole lung. The normalized  $\Delta PV$  ranged from 0.0 to 1.0, which corresponded to transparent and a solid color, respectively, in the color scale. Visual evaluation was performed by one diagnostic medical physicist (R.T.) with 20 years of experience in chest imaging research (R.T.). Images were displayed on a 2 M general-purpose LCD monitor under normal room lighting, similar to actual clinical settings.

## 2.6 | Clinical validation

The VIT results were validated by comparing them to those from a real clinical case of a patient with giant lung cysts (age: 40 years; sex: male; BMI: 15.7 kg/m<sup>2</sup>). DCR images were captured using an x-ray imaging system (prototype model; Konica Minolta, Tokyo, Japan) consisting of an x-ray generator/tube (DHF-155H II/



**FIGURE 4** Relationship between the  $\Delta PV\%$  for the 4D XCAT phantoms (five obese males, five overweight males, five normal males, and five normal females) and air sphere diameter for the (a) upper, (b) middle, and (c) lower lung regions. Each plot indicates, for a given air sphere diameter, the average  $\Delta PV\%$  of the five XCAT phantoms for each body size. XCAT, extended cardiac-torso.

UH-6QC-07E; Hitachi Healthcare, Ltd., Chiyoda, Japan) and a dynamic FPD (PaxScan, 4343CB; Varex Imaging Corporation, Salt Lake City, Utah, USA). The imaging was performed in the standing position and the posteroanterior direction, and approximately 210 frames of DCR images (100 kV, 0.2 mA/pulse, SID: 2.0 m, field-of-view: 43 cm  $\times$  43 cm) were obtained while the patient engaged in forced breathing for approximately 14 s. The matrix size of the DCR images was 1024  $\times$  1024 pixels, the pixel size was 417  $\times$  417  $\mu\text{m}$ ,<sup>2</sup> and the image range was 16-bit grayscale. In previous studies, color-mapping images of clinical cases were created using the same visualization algorithm.<sup>19,27</sup> Utilizing CT images (Revolution CT, GE Healthcare, Milwaukee, WI, USA), the thickness of the residual normal tissue in the anteroposterior direction was measured in each slice based on CT value thresholds (−950 HU and −700 HU).<sup>33–36</sup> The matrix size of the chest CT images was 512  $\times$  512 pixels, the slice thickness was 2.5 mm, and the image range was 16-bit grayscale. The area covered by the residual normal tissue was measured using the open-source 3D Slicer software (version 5.1.0, <https://www.slicer.org>). This study was approved by our hospital's institutional review board (registration number 114188).

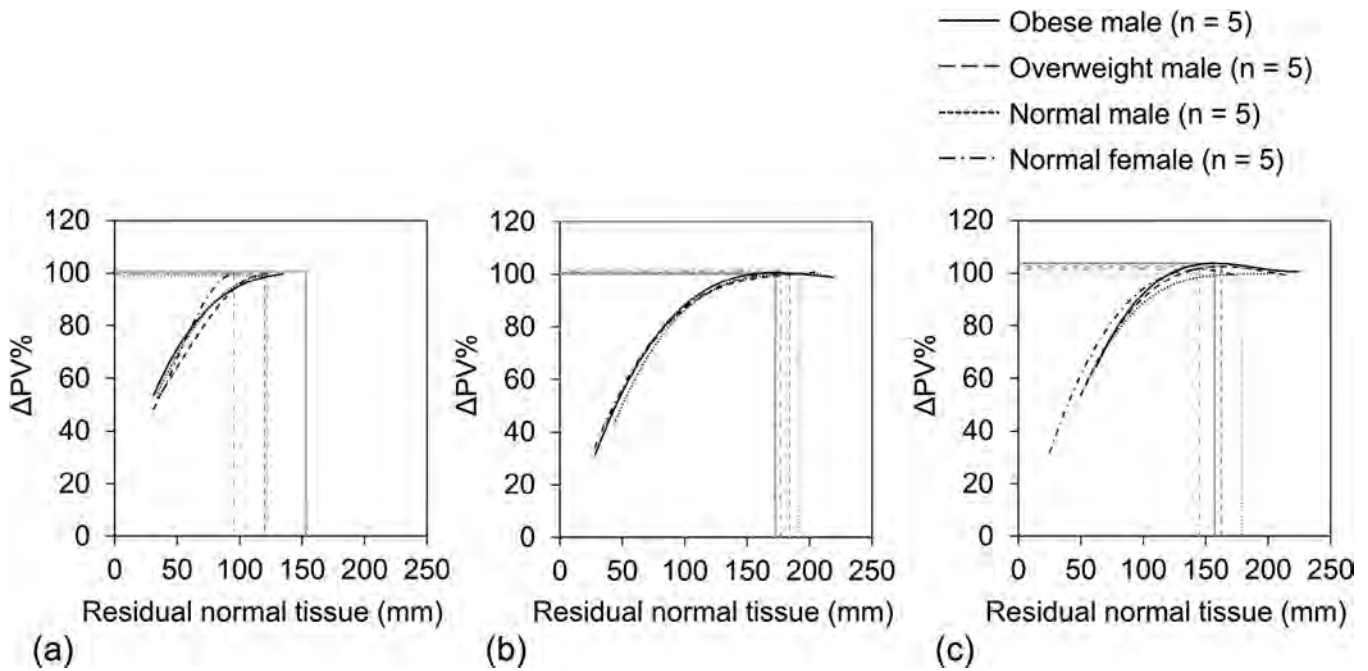
### 3 | RESULTS

Figure 4 shows the relationship between the diameter of the air sphere and  $\Delta PV\%$  in each lung region for the five

XCAT phantoms of each body size. For all body sizes and both sexes, the  $\Delta PV\%$  decreased as the size of the air sphere increased. There was no body size trend for the upper lung. However, in the middle and lower lungs (except for the 40 mm air sphere in the middle lung and the 20 mm air sphere in the lower lung), for the same air sphere size,  $\Delta PV\%$  decreased in the following order: obese male, overweight male, normal male, and normal female.

Figure 5 shows the performance curves derived from the relationship between the residual normal tissue and  $\Delta PV\%$  in each lung region for the five XCAT phantoms of each body size. The  $\Delta PV\%$  for all body sizes and both sexes decreased as the amount of the residual normal tissue decreased and could be defined as a function of the residual normal tissue in front of and behind the simulated lesion. The  $R^2$  values for the upper/middle/lower lung regions for the obese males, overweight males, normal males, and normal females were 0.84/0.99/0.88, 0.88/0.98/0.97, 0.84/0.98/0.97, and 0.93/0.99/0.82, respectively. As shown in Table 2, the normal-tissue<sub>thre</sub> values in the upper/middle/lower lung regions for the obese males, overweight males, normal males, and normal females were 153.2 mm/191.3 mm/157.0 mm, 120.1 mm/184.1 mm/162.3 mm, 121.9 mm/172.4 mm/178.5 mm, and 95.4 mm/176.5 mm/144.7 mm, respectively (see Table 2). The order of patient physique in the normal-tissue<sub>thre</sub> values was different in each lung field, with no consistent trend. Figure 6 shows the performance





**FIGURE 5** Performance curves derived from the relationship between the  $\Delta PV\%$  of the 4D XCAT phantoms and the amount of residual normal tissue in each of the five body sizes. The  $R^2$  values for the obese males/overweight males/normal males/normal females were (a) 0.84/0.88/0.84/0.93 in the upper lung region, (b) 0.99/0.98/0.98/0.99 in the middle lung region, and (c) 0.88/0.97/0.97/0.82 in the lower lung region. The gray vertical and horizontal lines indicate the amount of residual normal tissue at which  $\Delta PV\%$  began to decrease (i.e.,  $\text{normal-tissue}_{\text{thre}}$ ) for each performance curve. XCAT, extended cardiac-torso.

**TABLE 2** Lung thickness and  $\text{normal-tissue}_{\text{thre}}$  for each body size in each lung region.

Sex	Body size	Number	Lung thickness (mm)			Normal-tissue <sub>thre</sub> <sup>a</sup> (mm)		
			Upper	Middle	Lower	Upper	Middle	Lower
Male	Obese	5	97.5 ± 23.8	191.9 ± 16.5	206.3 ± 15.7	153.2	191.3	157.0
	Overweight	5	105.6 ± 16.7	182.5 ± 15.9	195.0 ± 15.3	120.1	184.1	162.3
Female	Normal	5	99.4 ± 18.9	178.8 ± 24.6	187.5 ± 12.0	121.9	172.4	178.5
	Normal	5	91.3 ± 8.7	157.5 ± 6.4	162.5 ± 9.3	95.4	176.5	144.7
Total		20	98.4 ± 18.6	177.7 ± 21.2	187.8 ± 20.9	146.0	179.5	170.9

Note: Data are expressed as mean ± standard deviation.

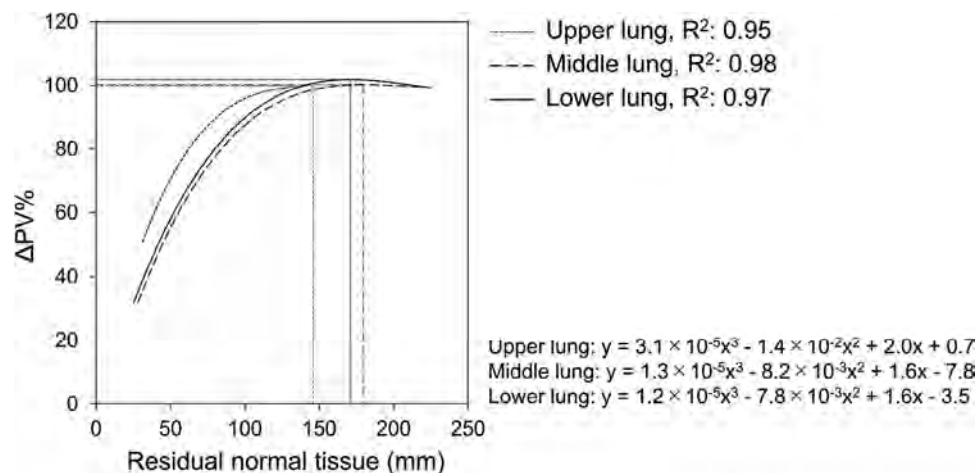
<sup>a</sup>Normal-tissue<sub>thre</sub>: the amount of residual normal tissue at which pulmonary ventilation impairment was detected as a decrease in  $\Delta PV\%$ .

curves derived from the relationship between the amount of residual normal tissue and  $\Delta PV\%$  obtained from all 20 XCAT phantoms for each lung region. In all lung regions, the  $\Delta PV\%$  decreased as the amount of residual normal tissue decreased and could also be defined as a function of the residual normal tissue in front of and behind the simulated lesion. The  $R^2$  values of the performance curves for the upper, middle, and lower lung regions were 0.97, 0.98, and 0.95, respectively, and the  $\text{normal-tissue}_{\text{thre}}$  values for the upper, middle, and lower lung regions were 146.0, 179.5, and 170.9 mm, respectively (see Table 2). The amount of residual normal tissue was partially included in the 95% interval; 132–187 mm was between the upper and middle lung regions, 163–219 mm was between the

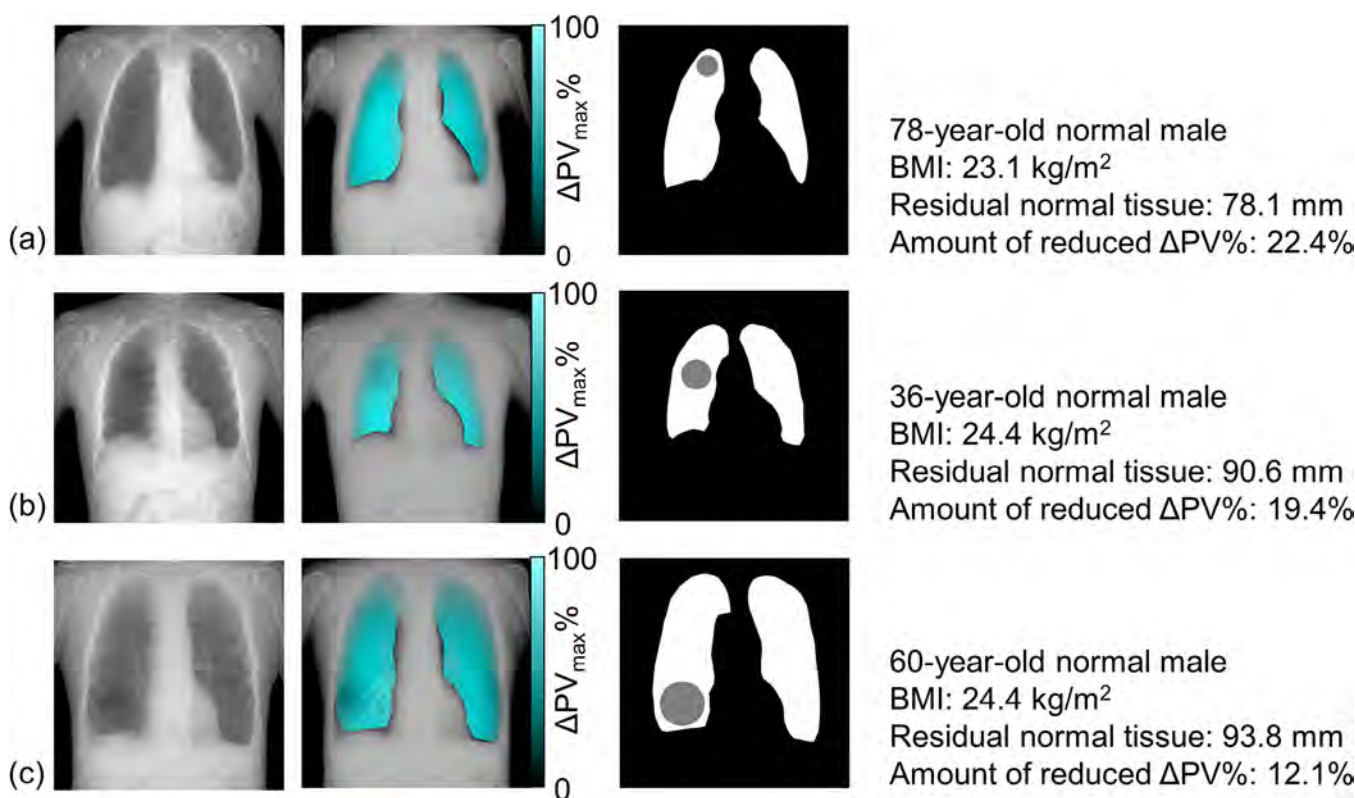
middle and lower lung regions, and 120–191 mm was between the upper and lower lung regions.

Figure 7 shows examples of virtual patients with an amount of residual normal tissue less than  $\text{normal-tissue}_{\text{thre}}$  on the simulated lesion. In all virtual patients, the simulated lesion indicated a local reduction in the  $\Delta PV\%$  (22.4%, 19.4%, and 12.1%), resulting in decreased color intensity in the color-mapping images.

Figures 8 shows the verification results for the clinical case. Sagittal CT images indicated the presence of giant lung cysts throughout the lungs and a reduction in normal tissue along the lung height except for the left lower lung (Figure 8a,e). The maximum amount of residual normal tissue in this patient



**FIGURE 6** Performance curves derived from the relationship between the  $\Delta PV\%$  of the 20 4D XCAT phantoms and the amount of residual normal tissue for each lung region. The gray vertical and horizontal lines indicate the amount of residual normal tissue at which  $\Delta PV\%$  began to decrease (i.e.,  $\text{normal-tissue}_{\text{thre}}$ ) for each performance curve. XCAT, extended cardiac-torso.

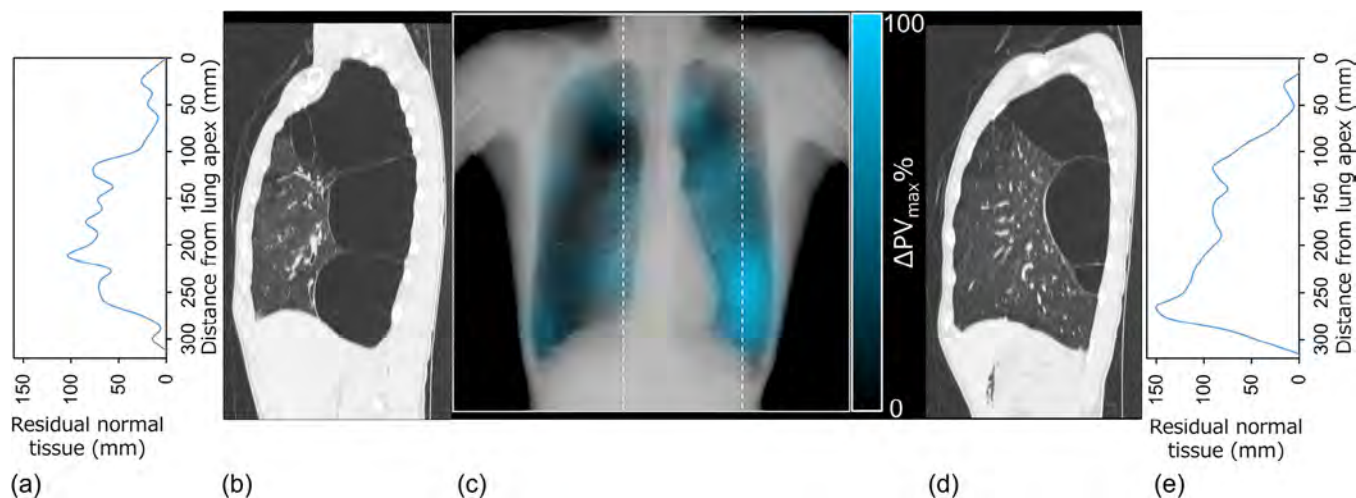


**FIGURE 7** Examples of virtual patients with amounts of residual normal tissue less than the threshold value of residual normal tissue ( $\text{normal-tissue}_{\text{thre}}$ ) on a simulated lesion in the (a) upper, (b) middle, and (c) lower lung regions. The left column shows the original images, the middle column shows the color-mapping images, and the right column shows the locations of the inserted air spheres (simulated lesions), as indicated by the gray circles BMI, body mass index.

was approximately 150 mm, which was less than  $\text{normal-tissue}_{\text{thre}}$  for every lung region. As shown in Figure 8c, any amount of normal lung tissue below 150 mm is displayed as a region of decreased color intensity.

## 4 | DISCUSSION

In this study, the typical threshold thickness of residual normal tissue at which the presence of emphysema can be detected by DCR was investigated using virtual



**FIGURE 8** Verification results for the clinical case (40-year-old male patient with giant lung cysts, BMI: 15.7 kg/m<sup>2</sup>). The plots on the left and right show the amount of residual normal tissue along the height of the (a) right and (b) left lungs, respectively. Sagittal CT images for the (b) right and (d) left lungs, respectively, and the corresponding (c) color-mapping image are shown. BMI, body mass index; CT, computed tomography.

patients with different physiques and sexes and a user-defined ground truth. For each lung region of each body size and sex, the  $\Delta PV\%$  decreased as the amount of residual normal tissue decreased and could be defined as a function of the amount of residual normal tissue with high  $R^2$  values. Meanwhile, the difference between the theoretical values of each performance curve and the corresponding measured values was only partially included in the 95% confidence interval. In addition, the normal-tissue<sub>thre</sub> values for upper, middle, and lower lung regions were 146.0, 179.5, and 170.9 mm, respectively; these values were validated in both virtual and real patients where the amounts of residual normal tissue was lower than the normal-tissue<sub>thre</sub>; any reduction in the residual normal tissue was revealed as a reduction in the  $\Delta PV$  and was depicted as a reduction in the color intensity in the color-mapping image. These results indicate that the performance of DCR-based pulmonary functional evaluation depends on the amount of residual normal lung tissues rather than on lesion size. In contrast, the order of patient physique in the normal-tissue<sub>thre</sub> values was different in each lung field, with no consistent trend. Therefore, this study did not provide insight into how the patient physique affects the  $\Delta PV\%$ .

The VIT conducted in this study provided additional evidence and perspectives on DCR-based pulmonary functional evaluation. In the middle and lower lungs, for the same lesion size,  $\Delta PV\%$  decreased in the following order: obese male, overweight male, normal male, and normal female. This order was consistent with the order of lung thickness, which indicated that detecting pulmonary ventilation impairment via DCR was more difficult for patients with thicker lungs than those with thinner lungs for the same pulmonary ventilation impairment size. This finding was refined by

further investigation based on the amount of residual normal tissue. As shown in Figures 5 and 6, the  $\Delta PV\%$  decreased as the amount of residual normal tissue decreased in each lung region for each body size and sex, suggesting that the detection performance of DCR depended on the amount of residual normal tissue in front of and behind the pulmonary ventilation impairment. The findings were also consistent with the fact that changes in lung density during respiration depend on the relative increase or decrease in the amount of air and the pulmonary vascular volume per unit lung volume.<sup>32</sup> Although each performance curve was only partially included in the 95% confidence interval, the  $R^2$  values of the performance curves for each body size ( $n = 5$ ) were relatively high, ranging from 0.84 to 0.99, and the  $R^2$  values for each lung region ( $n = 20$ ) were even higher. These results suggest that the detection performance of DCR also depends on patient anatomy, for example, differences in the volume of pulmonary vessels and bronchi included in the lung region. Furthermore, this study used normal-tissue<sub>thre</sub>, the threshold value at which  $\Delta PV\%$  began to decrease, as an indicator for pulmonary ventilation impairment that can be detected by DCR. In the validation experiments on virtual patients and one real patient, lung regions with impaired pulmonary ventilation less than normal-tissue<sub>thre</sub> were detected in the form of a decrease in  $\Delta PV_{\max}\%$  regardless of lesion size, suggesting the clinical utility of normal-tissue<sub>thre</sub>. This VIT was performed under special conditions without rib or diaphragm motion. However, deep learning-based BS and deformable image registration techniques have been reported by several researchers. If these techniques are applied to clinical images, the effects of rib and diaphragm motion can be reduced. Therefore, the



results of this study can be utilized in actual clinical practice, even if they were obtained without rib or diaphragm movement during breathing.

Although this study provided new insights into the diagnostic performance of DCR, it had several limitations. First, it was conducted using a restricted range of lesion morphology and pathology; the only pulmonary impairment tested was emphysema, which was simulated by inserting a simple air sphere into the lung with the pulmonary vessels and bronchi reproduced to the fifth generation. However, a variety of respiratory diseases require testing, and more realistic shapes and appearances need to be used to accurately simulate pulmonary diseases, along with more realistic simulation images. Second, this study revealed the typical threshold thickness of residual normal tissue below which the presence of emphysema can be detected by DCR when the lung structures were not displaced as a result of respiration. Color-mapping images of real patients visualize not only changes in lung density during respiration but also the displacement of the lung structures, which may degrade the performance of DCR in detecting pulmonary impairment. Additional VITs are required to investigate the effect of the respiratory motion of the thorax and diaphragm on the performance of DCR-based pulmonary function evaluation. Third,  $\Delta PV\%$  cannot be directly applied to an actual patient since the baseline of  $\Delta PV\%$  for each lung region is unknown. However, in this VIT, typical normal-tissue<sub>thre</sub> values for various thickness of normal lung tissue were determined using various baselines of  $\Delta PV\%$ . Therefore, they are expected to serve as a guide to the residual normal tissue where the presence of emphysema can be detected by DCR in actual patients. In contrast, if a patient has not had a prior CT scan, the normal tissue thickness (mm) for that patient is unknown, and therefore, the normal-tissue<sub>thre</sub> would not be used to determine which patients are eligible for DCR. In addition, DCR evaluates the relative pulmonary function based on changes in lung density. In fact, in both virtual and real patients, the lung regions with the amounts of residual normal tissue less than the normal-tissue<sub>thre</sub> which surrounded by those more than the normal-tissue<sub>thre</sub> were validated to represent a reduction in the delta  $\Delta PV$ s and was depicted as a reduction in color intensity in the color-mapping image. Given these characteristics, the role of DCR in patient care would be to assess disease progression as a monitoring tool in follow-up. Fourth, because of the limited number of virtual patients and measured lung regions, this study could not elucidate how the patient physique and lung anatomy affect the normal-tissue<sub>thre</sub> value. Furthermore, the effect of breast shadows was not revealed in the setting without rib cage and diaphragm motion. Further studies with more virtual patients are needed, along with investigations of various lung regions with the same normal tissue but different locations. On the

other hand, this VIT was conducted under fixed tube voltage. The effect of beam spectrum on  $\Delta PV$  should also be investigated, since different beam spectra may affect  $\Delta PV$ , the functional image contrast due to respiration. Finally, because VITs are an emerging methodology, their validation methods remain to be established. This VIT focused on the amount of residual normal tissue in front of and behind the lesion and the resulting normal-tissue<sub>thre</sub> values were validated in one real clinical case. Although the findings obtained by the VITs conducted in this study were confirmed by the clinical case, further clinical studies should be conducted to fully utilize the results of VITs.

Despite having certain limitations in assessing the diagnostic performance of DCR, the results of VITs can be used to improve the efficiency of clinical trials. Because DCR is a recently developed functional x-ray imaging technique and was launched as a commercial system in 2018 and approved by the US Food and Drug Administration in 2019, the number of clinical cases with specific ground truth for pulmonary function is limited and its diagnostic performance has not yet been fully determined. In this study, VITs were performed on virtual patients using a user-defined ground truth, and they demonstrated without irradiating patients that the amount of residual normal lung tissue determines the performance of DCR-based pulmonary functional evaluation. Thus, the results of VITs may contribute to more efficient clinical trials.

## 5 | CONCLUSIONS

This study discovered that the performance of DCR-based pulmonary impairment assessment depended on the amount of residual normal tissue in front of and behind the lesion rather than lesion size. The performance curve could be defined as a function of the amount of residual normal tissue in each lung region, and the curve exhibited a threshold of residual normal tissue at which lesions could be detected as a decrease in  $\Delta PV$ . The findings of VITs are expected to accelerate future clinical trials for DCR-based pulmonary function assessment. Further clinical studies should be conducted to fully utilize the results of VITs, along with simulating more realistic virtual patients and lesions.

## ACKNOWLEDGMENTS

This work was partly supported by JSPS KAKENHI Grant-in-aid for Scientific Research (B) [21H02866], Fostering Joint International Research (A) [21KK0286], the National Institute of Health [5P41EB028744], and Tateishi Science and Technology Foundation. The authors sincerely thank the staff from the departments of respiratory medicine, thoracic surgery, and radiology at Kanazawa University Hospital for providing the clinical data described above, Hiroyuki Okamoto from college



of medical, pharmaceutical health sciences, Kanazawa University for advice on statistical analysis.

## CONFLICT OF INTEREST STATEMENT

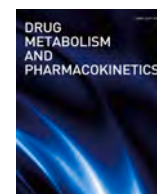
This research was supported in part by Konica Minolta, Inc., Tokyo, Japan. The device used in our clinical study was also provided by Konica Minolta, Inc., Tokyo, Japan.

## REFERENCES

- World Health Organization (WHO). The top 10 causes of death. Accessed February 12, 2024. <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>
- WHO. WHO reveals leading causes of death and disability worldwide: 2000–2019. Accessed February 12, 2024. <https://www.who.int/news/item/09-12-2020-who-reveals-leading-causes-of-death-and-disability-worldwide-2000-2019>
- WHO. Global excess deaths associated with COVID-19, January 2020–December 2021. Accessed February 12, 2024. <https://www.who.int/data/stories/global-excess-deaths-associated-with-covid-19-january-2020-december-2021>
- Sheikh K, Coxson HO, Parraga G. This is what COPD looks like. *Respirology*. 2016;21(2):224–236. doi:10.1111/resp.12611
- Shimada A, Kawata N, Sato H, et al. Dynamic quantitative magnetic resonance imaging assessment of areas of the lung during free-breathing of patients with chronic obstructive pulmonary disease. *Acad Radiol*. 2022;29(2):S215–S225. doi:10.1016/j.acra.2021.03.034
- Ding K, Bayouth JE, Buatti JM, Christensen GE, Reinhardt JM. 4DCT-based measurement of changes in pulmonary function following a course of radiation therapy. *Med Phys*. 2010;37(3):1261–1272. doi:10.1118/1.3312210
- Yamashiro T, Moriya H, Matsuoka S, et al. Asynchrony in respiratory movements between the pulmonary lobes in patients with COPD: continuous measurement of lung density by 4-dimensional dynamic-ventilation CT. *Int J Chron Obstruct Pulmon Dis*. 2017;12:2101–2109. doi:10.2147/COPD.S140247
- Sato H, Kawata N, Shimada A, et al. Semiautomatic assessment of respiratory dynamics using cine MRI in chronic obstructive pulmonary disease. *Eur J Radiol Open*. 2022;9:100442. doi:10.1016/j.ejro.2022.100442
- Mochizuki E, Kawai Y, Morikawa K, et al. Difference in local lung movement during tidal breathing between COPD patients and asthma patients assessed by four-dimensional dynamic-ventilation CT scan. *Int J Chron Obstruct Pulmon Dis*. 2020;15:3013–3023. doi:10.2147/COPD.S273425
- Tanaka R. Dynamic chest radiography: flat-panel detector (FPD) based functional x-ray imaging. *Radiol Phys Technol*. 2016;9(2):139–153. doi:10.1007/s12194-016-0361-6
- Tanaka R, Sanada S, Suzuki M, et al. Breathing chest radiography using a dynamic flat-panel detector combined with computer analysis. *Med Phys*. 2004;31(8):2254–2262. doi:10.1118/1.1769351
- International Atomic Energy Agency (IAEA). *International Basic Safety Standards for Protection Against Ionizing Radiation and for the Safety of Radiation Sources*. IAEA Safety Series No. 115. IAEA; 1996: 279.
- Annex IV, Schedule III. Guidance levels of dose, dose rate and activity for medical exposure. *Radiological Protection for Medical Exposure to Ionizing Radiation, Safety Guide*. IAEA Safety Standards Series No. SSG-46. International Atomic Energy Agency (IAEA); 2018.
- Tanaka R, Sanada S, Okazaki N, et al. Evaluation of pulmonary function using breathing chest radiography with a dynamic flat panel detector: primary results in pulmonary diseases. *Invest Radiol*. 2006;41(10):735–745. doi:10.1097/01.rli.0000236904.79265.68
- Tanaka R, Sanada S, Fujimura M, et al. Development of functional chest imaging with a dynamic flat-panel detector (FPD). *Radiol Phys Technol*. 2008;1(2):137–143. doi:10.1007/s12194-008-0020-7
- Tanaka R, Sanada S, Fujimura M, et al. Ventilatory impairment detection based on distribution of respiratory-induced changes in pixel values in dynamic chest radiography: a feasibility study. *Int J Comput Assist Radiol Surg*. 2011;6(1):103–110. doi:10.1007/s11548-010-0491-y
- Tanaka R, Tani T, Nitta N, et al. Pulmonary function diagnosis based on respiratory changes in lung density with dynamic flat-panel detector imaging: an animal-based study. *Invest Radiol*. 2018;53:417–423. doi:10.21037/qims-20-1217
- Tanaka R, Tani T, Nitta N, et al. Detection of pulmonary embolism based on reduced changes in radiographic lung density during cardiac beating using dynamic flat-panel detector: an animal-based study. *Acad Radiol*. 2019;26(10):1301–1308. doi:10.1016/j.acra.2018.12.012
- Tanaka R, Matsumoto I, Tamura M, et al. Comparison of dynamic flat-panel detector-based chest radiography with nuclear medicine ventilation-perfusion imaging for the evaluation of pulmonary function: a clinical validation study. *Med Phys*. 2020;47(10):4800–4809. doi:10.1002/mp.14407
- Tanaka R, Tani T, Yamada A, et al. Correlations between cardiovascular parameters and image parameters on dynamic chest radiographs in a porcine model under fluid loading. *Radiol Phys Technol*. 2021;14(3):288–296. doi:10.1007/s12194-021-00626-2
- Tanaka R, Matsumoto I, Tamura M, et al. Dynamic chest radiography: clinical validation of ventilation and perfusion metrics derived from changes in radiographic lung density compared to nuclear medicine imaging. *Quant Imaging Med Surg*. 2021;11(9):4016–4027. doi:10.21037/qims-20-1217
- Abadi E, Segars WP, Tsui BMW, et al. Virtual clinical trials in medical imaging: a review. *J Med Imaging*. 2020;7(4):042805. doi:10.1117/1.JMI.7.4.042805
- Segars WP, Mahesh M, Beck TJ, Frey EC, Tsui BM. Realistic CT simulation using the 4D XCAT phantom. *Med Phys*. 2008;35(8):3800–3808. doi:10.1118/1.2955743
- Segars WP, Bond J, Frush J, et al. Population of anatomically variable 4D XCAT adult phantoms for imaging research and optimization. *Med Phys*. 2013;40(4):043701. doi:10.1118/1.4794178
- Segars WP, Sturgeon G, Mendonca S, Grimes J, Tsui BM. 4D XCAT phantom for multimodality imaging research. *Med Phys*. 2010;37(9):4902–4915. doi:10.1118/1.3480985
- Visible Human Male and Female datasets, National Library of Medicine. Accessed February 12, 2024. [http://www.nlm.nih.gov/research/visible/visible\\_human.html](http://www.nlm.nih.gov/research/visible/visible_human.html)
- Tanaka R, Samei E, Segars P, et al. Dynamic chest radiography for pulmonary function diagnosis: a validation study using 4D extended cardiac-torso (XCAT) phantom. *Med Imaging*. 2019;1094831:859–865. doi:10.1117/12.2512332
- Obesity: preventing and managing the global epidemic. Report of a WHO consultation. *World Health Organ Tech Rep Ser*. 2000;894:i-253.
- ICRU. *International Commission on Radiation Units and Measurement, Photon, Electron, Proton and Neutron Interaction Data for Body Tissues*, ICRU Report No 46, ICRU; 1992.
- Spectrum processor SRS-78. Accessed Oct 16, 2020. Accessed February 12, 2024. <http://linux.fjfi.cvut.cz/%7Emadlenka/medphys.htm>
- Tanaka R, Segars WP, Abadi E, Minami S, Samei E, Optimization of imaging conditions in pediatric dynamic chest radiography: a virtual imaging trial. In: *Proc. SPIE 12031, Medical Imaging 2022: Physics of Medical Imaging*. SPIE; 2022:1203141. doi:10.1117/12.2612720

32. Squire LF, Novelline RA. Overexpansion and collapse of the lung. *Fundamentals of Radiology*. 4th ed. Harvard University Press; 1988:88-89.
33. Lynch DA, Al-Qaisi MA. Quantitative computed tomography in chronic obstructive pulmonary disease. *J Thorac Imaging*. 2013;28(5):284-290. doi:10.1097/RTI.0b013e318298733c
34. Coxson HO, Dirksen A, Edwards LD, et al. The presence and progression of emphysema in COPD as determined by CT scanning and biomarker expression: a prospective analysis from the ECLIPSE study. *Lancet Respir Med*. 2013;1(2):129-136. doi:10.1016/S2213-2600(13)70006-7
35. Shin KE, Chung MJ, Jung MP, Choe BK, Lee KS. Quantitative computed tomographic indexes in diffuse interstitial lung disease: correlation with physiologic tests and computed tomography visual scores. *J Comput Assist Tomogr*. 2011;35(2):266-271. doi:10.1097/RCT.0b013e31820ccf18
36. Ohkubo H, Kanemitsu Y, Uemura T, et al. Normal lung quantification in usual interstitial pneumonia pattern: the impact of threshold-based volumetric ct analysis for the staging of idiopathic pulmonary fibrosis. *PLoS One*. 2016;11(3):e0152505. doi:10.1371/journal.pone.0152505

**How to cite this article:** Yamaguchi S, Tanaka R, Matsumoto I, et al. Estimation of threshold thickness of residual normal tissue in lung dysfunction detectable by dynamic chest radiography: A virtual imaging trial. *Med Phys*. 2024;51:5978–5989.  
<https://doi.org/10.1002/mp.17271>



## Regular Article

# Effect of obesity on pharmacokinetics of transdermal fentanyl: Single-center retrospective study and animal study

Satoshi Mizuno<sup>a,1</sup>, Shintaro Gake<sup>a,1</sup>, Makiko Takabayashi<sup>b</sup>, Yuriko Ito<sup>b</sup>, Hiroko Sanada<sup>a</sup>, Natsumi Sugimoto<sup>a</sup>, Akari Maeda<sup>a</sup>, Takuto Tamamura<sup>a</sup>, Kazuki Sawamoto<sup>b</sup>, Yusuke Hara<sup>b</sup>, Yoshiko Ohi<sup>b</sup>, Chiaki Tsuji<sup>b</sup>, Yukiko Shiomoto<sup>b</sup>, Yukio Kato<sup>c</sup>, Arimi Fujita<sup>a,b</sup>, Tsutomu Shimada<sup>a,b,d</sup>, Ken-ichi Miyamoto<sup>b</sup>, Yoshimichi Sai<sup>a,b,e,\*</sup>

<sup>a</sup> Department of Clinical Pharmacokinetics, Graduate School of Medical Sciences, Kanazawa University, Kanazawa, Japan

<sup>b</sup> Department of Hospital Pharmacy, University Hospital, Kanazawa University, Kanazawa, Japan

<sup>c</sup> Department of Molecular Pharmacotherapeutics, Faculty of Pharmacy, Kanazawa University, Kanazawa, Japan

<sup>d</sup> Department of Clinical Pharmacy and Healthcare Science, Faculty of Pharmacy, Institute of Medical, Pharmaceutical and Health Science, Kanazawa University, Kanazawa, Japan

<sup>e</sup> AI Hospital/Macro Signal Dynamics Research and Development Center, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Kanazawa, Japan



## ARTICLE INFO

## Keywords:

Transdermal fentanyl  
Opioid switching  
Pain relief  
Obesity  
Zucker rats

## ABSTRACT

A retrospective study and an animal study were conducted to investigate factors affecting the transdermal fentanyl dose to achieve adequate pain relief in patients switched from other opioids. In the retrospective study, patient factors were included as gender, age, body mass index (BMI), and serum albumin concentration. In obese (BMI  $\geq 25$ ) patients, the post-titration dose of transdermal fentanyl was significantly lower than in normal (BMI 18.5–25) patients despite the initial dose was the same. To support this unexpected finding, fentanyl was administered intravenously and transdermally to Zucker (*fa/fa*) rats as an obese model and Zucker (+/+) rats as a control. No difference in systemic clearance ( $CL_{tot}$ ) after intravenous administration was observed between the two groups. However, transdermal bioavailability ( $F$ ) and fentanyl release ratio from the formulation ( $F_a$ ) were significantly increased in Zucker (*fa/fa*) rats compared to Zucker (+/+) rats. Skin availability ( $F_{skin}$ ), calculated as  $F$  divided by  $F_a$ , was also significantly increased. These results indicated that obesity should be considered as a factor in the titration of transdermal fentanyl dose.

## 1. Introduction

Transdermal fentanyl, an opioid, is used to relieve moderate to severe cancer pain and chronic pain [1,2], but the therapeutic window of blood concentration is narrow [3–5], and there are large inter-individual variabilities in fentanyl release from transdermal formulations [6,7]. Therefore, to prevent severe adverse events such as somnolence and respiratory depression, transdermal fentanyl is titrated from a low dose until adequate analgesia is achieved [8–11]. Clarifying the factors that influence the required dose will contribute to relieving pain more quickly and safely, helping to optimize pain treatment.

In general, drug doses are determined with reference to body size. Higher total body weight increases the intravenous fentanyl dose for

postoperative analgesia [12]. Some factors that influence transdermal fentanyl dose have been investigated. As an extrinsic factor related to the variability of transdermal fentanyl, heating of the formulation increases blood fentanyl concentration [13,14]. Some intrinsic factors have been investigated such as gender, elderly, low body mass index (BMI) and hypoalbuminemia [15–18].

This study aimed to clarify factors affecting the transdermal fentanyl dose to achieve adequate pain relief in patients switched from other opioids. A retrospective study was conducted to investigate the influence of factors such as gender, age, BMI, and serum albumin concentration. In order to test a hypothesis that the post-titration dose was changed in obese patients, fentanyl was administered intravenously and transdermally to an obese animal model, Zucker (*fa/fa*) rats [19,20],

\* Corresponding author. Department of Hospital Pharmacy, University Hospital, Kanazawa University, 13-1 Takara-machi, Kanazawa, 920-8641, Japan.

E-mail address: [sai-ys@staff.kanazawa-u.ac.jp](mailto:sai-ys@staff.kanazawa-u.ac.jp) (Y. Sai).

<sup>1</sup> These two authors contributed equally to this work.

and the skin availability ( $F_{\text{skin}}$ ), fraction through the skin, was compared to that of Zucker (+/+) rats.

## 2. Materials and methods

### 2.1. Retrospective study

Patients who were initiated on treatment with Fentos® Tape (24 h fentanyl citrate matrix formulation) between Apr 1, 2015, and Mar 31, 2019, at Kanazawa University Hospital in Japan were enrolled. Exclusion criteria are shown in Fig. 1. Gender, age, height, weight, BMI, and serum albumin concentration were surveyed as patient background factors. Data were collected at the time before treatment. However, for BMI, data during treatment period were also included due to the infrequency of height measurements. The initial dose was defined as the dose of transdermal fentanyl on day 1, which was calculated from the dose of the previous opioid using approximate potency relative to morphine [21]. The post-titration dose was defined as the dose when both of the following criteria were met.

1. No dose change for 5 days; Because the half-life of fentanyl after removal of the transdermal formulation is  $21.9 \pm 8.9$  h (mean  $\pm$  S. D.) [22], the blood concentration was estimated to reach a steady state at about 5 days after dose changes.
2. Rescue opioids within 3 times per day; Rescue opioids with short-acting duration are used for transient exacerbations of pain and breakthrough pain. The European Society for Medical Oncology (ESMO) guidelines recommend baseline opioid adaptations if rescue opioids are needed 4 times or more per day [2].

This retrospective study was approved by the Committee on Medical Ethics of Kanazawa University (2019-007 (3045)).

### 2.2. Chemicals and animals

Fentanyl citrate injection and fentanyl citrate formulation (Fentos® Tape) were purchased from Janssen Pharmaceutical K.K. (Tokyo, Japan) and Kyowa Hakko Kirin Co. Ltd. (Tokyo, Japan), respectively. Diethyl ether was purchased from Nacalai Tesque, Inc. (Kyoto, Japan). Epilat hair-removing cream was purchased from Kracie, Ltd (Tokyo, Japan). All other chemicals were of analytical or high-performance liquid chromatography (HPLC) grade. Zucker (*fa/fa*) obese rats and Zucker (+/+) control rats were purchased from Japan SLC (Hamamatsu, Japan). Animals were treated in accordance with the Fundamental Guidelines for Proper Conduct of Animal Experiment and Related Activities in Academic Research Institutions, under the jurisdiction of the Ministry of Education, Culture, Sports, Science and Technology of

Japan. All animal experiments were approved by the Committee on Animal Experimentation of Kanazawa University.

### 2.3. Intravenous and transdermal administration of fentanyl

Zucker rats, which lack leptin receptor, were used as an obese animal model. Male Zucker (*fa/fa*) and Zucker (+/+) rats were purchased at age of 8–9 weeks and allowed free access to standard diet and water, then used for experiments at the age of 12 weeks. The liver, and epididymal fat were collected and weighed to confirm the obese phenotype. The rats were anesthetized by diethyl ether. For intravenous administration, fentanyl citrate injection diluted in saline ( $10 \mu\text{g}$  fentanyl/2 mL/kg body weight) was injected via the jugular vein in 10 s. Blood samples were collected from the orbital vein via a heparin-coated hematocrit tube at 5, 10, 20, 30, 60, 120, 180 and 300 min. For transdermal administration, the dorsolateral hair of rats was trimmed using an electric shaver and then removed with a hair-removing cream. The rats were allowed to recover for 3 days, then one-fourth of a 1 mg fentanyl citrate formulation ( $160 \mu\text{g}$  as fentanyl) was applied to the back skin of each rat. The formulation was secured by applying surgical tape over the formulation and using an abdominal bandage loop to ensure a constant skin interface and to prevent the rats from disturbing the formulation. Blood samples were collected via the orbital vein at 0.5, 1, 2, 4, 8, 12, 24, 31, 48 and 72 h. The blood samples were centrifuged, and plasma samples were stored at  $-80^\circ\text{C}$  until analysis. The fentanyl formulation was removed at 24 h and stored at  $4^\circ\text{C}$  until analysis.

### 2.4. Analysis of fentanyl in blood samples using LC-MS

Plasma samples ( $150 \mu\text{L}$ ) were mixed with  $15 \mu\text{L}$  of 1 M potassium phosphate (pH 8.0) and  $10 \mu\text{L}$  of 5 ng/mL papaverine (internal standard), and extracted with 4 mL of diethyl ether for 10 min. After centrifugation at  $1670 \times g$  for 10 min, the ether layer ( $3.5 \text{ mL}$ ) was collected and evaporated to dryness under a nitrogen stream. The residue was reconstituted with  $100 \mu\text{L}$  of methanol and transferred to auto-sampler vials.

A Shimadzu HPLC system coupled to an LCMS-2020 mass spectrometer (Shimadzu, Kyoto, Japan) was used with a  $\mu\text{-BONDASPHERE}$ , PH 5  $\mu\text{m}$ ,  $300 \text{ A}$ ,  $150 \times 3.9 \text{ mm}$  analytical column (Waters, Milford, USA). The column temperature was maintained at  $40^\circ\text{C}$  and eluted in isocratic mode with 0.1 % formic acid water/acetonitrile (70:30). The selected-ion-monitoring (SIM) mode was used for quantitation by simultaneously monitoring  $\text{H}^+$ -adduct ions; mass-to-charge ratio ( $m/z$ ) 337.2 for  $[\text{fentanyl} + \text{H}]^+$  and  $m/z$  340.1 for  $[\text{papaverine} + \text{H}]^+$ . The limit of the quantification was 0.1 ng/mL. The inter-day coefficient of variation (CV) was 14.5 %, 12.4 % and 0.5 % at 0.1 ng/mL, 3.0 ng/mL and 30 ng/mL, respectively. LCMSsolution software (ver. 5) was used for acquisition of the LC-MS data, construction of the internal calibration curves and derivation of the regression equations.

### 2.5. HPLC analysis of fentanyl in the formulation

The fentanyl formulations removed from rats or the unused formulation as control were cut into small pieces and the remaining fentanyl was extracted with 40 mL of methanol for 5 h at  $70^\circ\text{C}$ . The total volume was adjusted to 100 mL with methanol. A 10 mL portion was mixed with  $10 \mu\text{L}$  of 0.5 mg/mL papaverine and 15 mL of 10 mM sulfuric acid. Fentanyl was quantitated with a Shimadzu HPLC system coupled to SPD-20A UV/VIS spectrophotometer, as described above except for the mass spectrometer. UV absorption was monitored at 195 nm. The retention times were 8.3 min and 6.5 min for fentanyl and papaverine, respectively.

### 2.6. Pharmacokinetic analysis

Pharmacokinetic parameters of fentanyl after intravenous and

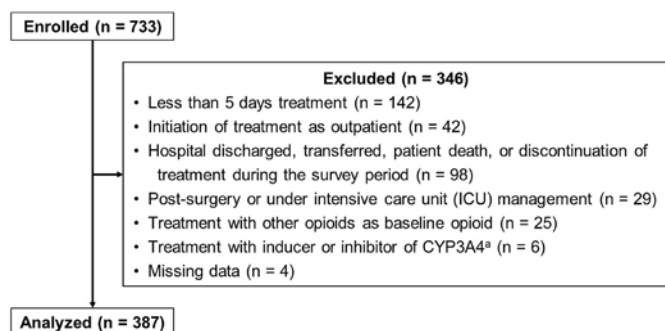


Fig. 1. Analysis flowchart in the retrospective study

<sup>a</sup>Inducer or inhibitor of CYP3A4 as follows: carbamazepine, enzalutamide, mitotane, phenobarbital, phenytoin, rifampicin, amiodarone, cobicistat, clarithromycin, diltiazem, fluvoxamine, indinavir, itraconazole, nelfinavir, saquinavir, ritonavir, and voriconazole.



transdermal administration were estimated by model-independent moment analysis, non-compartmental analysis, using Napp version 2.01 software [23]. Transdermal bioavailability ( $F$ ) was calculated by dividing the  $AUC$  after transdermal administration ( $AUC_{td,0-\infty}$ ) obtained from each animal by the mean  $AUC$  after intravenous administration ( $AUC_{iv,0-\infty}$ ), corrected for the dose of fentanyl. Fentanyl release ratio from the formulation (the fraction absorbed,  $F_a$ ) and  $F_{skin}$  were calculated as follows:

$$F_a = \frac{X_{unused} - X_{used}}{X_{unused}}$$

$$F_{skin} = \frac{F}{F_a}$$

where  $X_{used}$  and  $X_{unused}$  represent the residual amount of fentanyl in the formulation removed from the rats and the amount of fentanyl in the unused formulation, respectively.

## 2.7. Statistical analysis

Statistical analysis was performed with EZR software [24]. In the retrospective study, age was divided into two groups (<65 and ≥65 years), and BMI was divided into three groups (<18.5, 18.5–25, ≥25) based on Japanese population characteristics [25]. Serum albumin concentration was divided into two groups (<3.5 and ≥3.5 g/dL) based on criteria for hypoalbuminemia [18]. The Mann-Whitney  $U$  test was performed for comparison of gender groups, age groups or serum albumin concentration groups. The Kruskal-Wallis test and Steel-Dwass test were performed for BMI groups. In the animal model study, the unpaired Student's  $t$ -test was employed for comparison between the two groups. The criterion of significance was taken to be  $p < 0.05$ .

## 3. Results

### 3.1. Patient factors related to post-titration dose of transdermal fentanyl in the retrospective study

In total, 733 patients were enrolled and 387 were finally analyzed in this retrospective study (Fig. 1). Table 1 summarizes patient background factors. Among the study patients, 239 (61.8 %) were male. The median age, BMI, and serum albumin concentration were 66 years, 20.3, and 2.9 g/dL, respectively. Albumin data were missing in 34 patients. There were only two chronic pain patients, and the others had cancer pain.

The initial dose and the post-titration dose of transdermal fentanyl are shown in Fig. 2. Patients were given an initial dose of 3 mg transdermal fentanyl or less, except for 13 patients who received an initial dose of 4 mg or more, which was calculated based on the dose of the previous opioids (oxycodone in 12 patients and tramadol in 1 patient). Cancer sites were the pancreas in 6 patients, lung in 2 patients, and other (pharynx, stomach, prostate, ovary, and primary site unknown) in 1 patient each. Four patients were undergoing radiotherapy or chemotherapy, and 4 patients were receiving best supportive care (BSC) at the terminal stage. Gender differences were not observed for both the initial dose and the post-titration dose (Fig. 2A). In the ≥65 years group, the

**Table 1**  
Background factors of patients in the retrospective study.

Total	387
Male/Female	239/148
Age (years)	66 [58–73]
Height (cm)	162.3 [155.4–168.2]
Weight (kg)	52.9 [45.7–60.7]
BMI (kg/m <sup>2</sup> )	20.3 [18.3–22.5]
Serum albumin concentration (g/dL)	2.9 [2.4–3.4] <sup>a</sup>

Data are expressed as number (n) or median [25–75 percentile]. <sup>a</sup>Data for 353 patients.

post-titration dose was significantly lower than in the <65 years group, but the initial dose was also significantly lower (Fig. 2B). Regarding BMI, the post-titration dose in the ≥25 group was significantly lower compared to the 18.5–25 group, despite no difference in the initial dose (Fig. 2C). Under 3.5 g/dL serum albumin concentration was not related to both the initial dose and the post-titration dose (Fig. 2D).

### 3.2. Obese phenotype in Zucker ( $fa/fa$ ) rats

Table 2 summarizes the body, liver, and epididymis fat weights in 12-week-old Zucker (+/+) and Zucker ( $fa/fa$ ) rats. The body, liver, and epididymis fat weight in Zucker ( $fa/fa$ ) rats were 1.7-, 2.0-, and 5.6-fold higher than those in Zucker (+/+) rats, indicating that the Zucker ( $fa/fa$ ) rats were indeed obese.

Pharmacokinetic parameters of fentanyl after intravenous and transdermal administration in Zucker (+/+) and Zucker ( $fa/fa$ ) rats.

The plasma concentration-time profile and pharmacokinetic parameters after intravenous administration of fentanyl (10 µg/kg) in Zucker (+/+) and Zucker ( $fa/fa$ ) rats are shown in Fig. 3 and Table 3, respectively. After intravenous administration, no significant difference was observed in plasma concentration-time profile,  $AUC_{0-\infty}$ , half-life ( $t_{1/2}$ ), systemic clearance ( $CL_{tot}$ ), or volume of distribution at steady state ( $Vd_{ss}$ ) between Zucker (+/+) and Zucker ( $fa/fa$ ) rats.

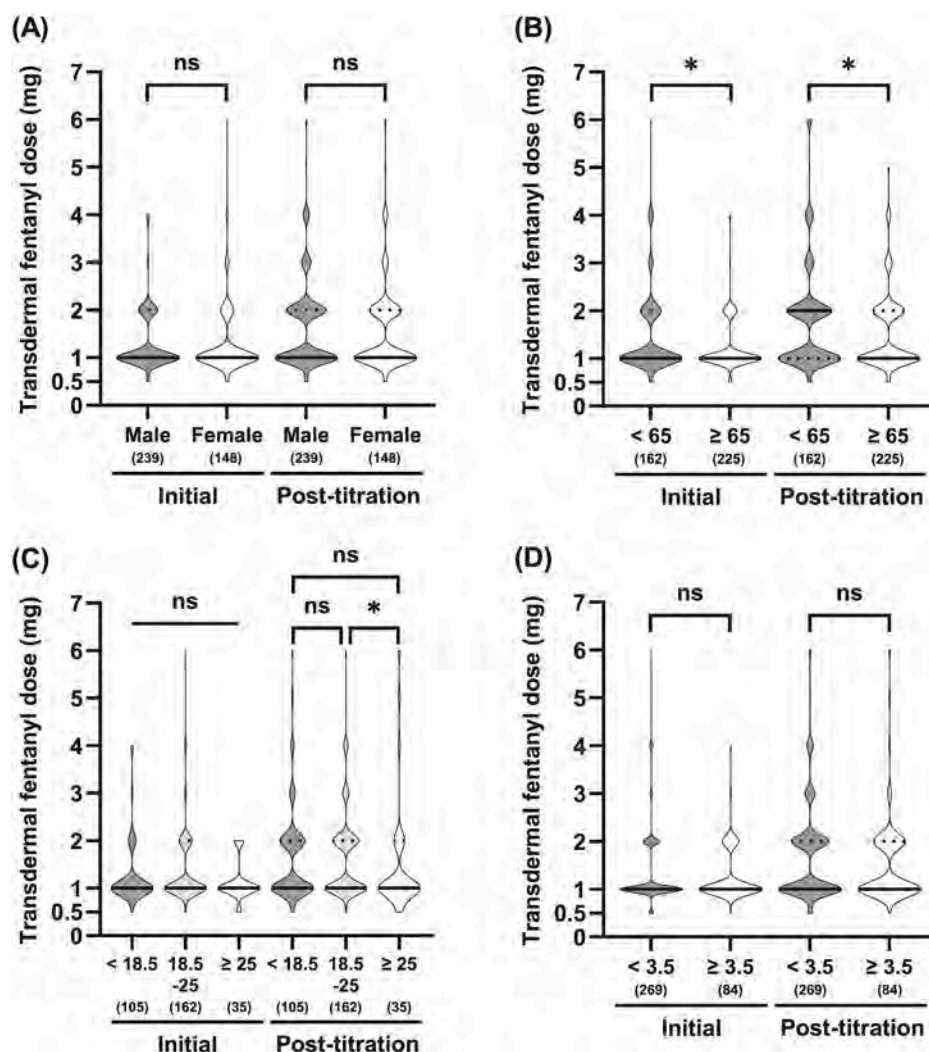
The plasma concentration-time profile and pharmacokinetic parameters were also obtained after transdermal administration of fentanyl (160 µg/body) in Zucker (+/+) and Zucker ( $fa/fa$ ) rats for 24 h, and are shown in Fig. 4 and Table 4, respectively. The peak concentration ( $C_{max}$ ) and  $AUC_{0-\infty}$  were significantly increased, and peak time ( $T_{max}$ ) was significantly shorter in Zucker ( $fa/fa$ ) rats compared to those in Zucker (+/+) rats. Transdermal clearance ( $CL_{tot}/F$ ) was significantly decreased in Zucker ( $fa/fa$ ) rats compared to that in Zucker (+/+) rats.  $F$ ,  $F_a$ , and  $F_{skin}$  were significantly increased by 2.2-, 1.2-, 1.9-fold, respectively, in Zucker ( $fa/fa$ ) rats compared to those in Zucker (+/+) rats.

## 4. Discussion

Relieving pain quickly and safely contributes to improving patients' quality of life. In this study, we investigated the factors affecting the transdermal fentanyl dose to achieve adequate pain relief in patients switched from other opioids, by means of a retrospective study and an animal study. Patient-rated pain scales such as visual analogue scales (VAS), numerical rating scales (NRS), and verbal rating scales (VRS) are usually used to assess pain [2,26]. However, because the target pain intensity varied for individual patients, it was difficult to assess whether pain relief was achieved. Therefore, the retrospective study evaluated the dose stability.

Gender differences in transdermal fentanyl dose were not observed (Fig. 2A), and the ≥65 years patients showed decreases not only in the post-titration dose but also in the initial dose compared to the <65 years patients (Fig. 2B). The influence of gender on serum fentanyl concentration is small, explaining less than 1 % of variability in its concentration in patients receiving transdermal fentanyl [15,27]. Regarding age group, the different initial dose suggested that patient backgrounds including pain intensity were not same. However, drug doses are generally adjusted lower in elderly patients, which could be associated with both lower the initial dose and the post-titration doses of transdermal fentanyl.

Unexpectedly, obese (BMI ≥25) patients showed a decrease in the post-titration dose of transdermal fentanyl compared to normal (BMI 18.5–25) patients (Fig. 2C). No difference in the initial dose among the BMI groups suggested that pain intensity in pre-titration was equivalent, and that obesity itself is a factor that influences titration of transdermal fentanyl dose. Previous reports, in the steady-state phase at arbitrary time, have indicated that there is no association between BMI and fentanyl release ratio from the formulation or between BMI and the dose [7,



**Fig. 2.** Initial dose and post-titration dose of transdermal fentanyl in gender, age, BMI, and serum albumin concentration groups in the retrospective study. Truncated violin plots show the initial dose and the post-titration dose of transdermal fentanyl in gender (A), age (B), BMI (C), and serum albumin concentration (D) groups. Each plots present the median (solid line), the interquartile range (dashed line), the maximum/minimum values (truncation), and number of patients (numbers in parentheses). The 1 mg formulation contains 0.64 mg as fentanyl, and the estimated average absorption is 0.3 mg/day as fentanyl. \*Significant difference at  $p < 0.05$ . ns: not significant.

**Table 2**

Characteristics of Zucker (+/+) and Zucker (*fa/fa*) rats.

	Zucker (+/+)	Zucker ( <i>fa/fa</i> )
Body weight (g)	275 ± 18	460 ± 44*
Liver weight (g)	8.3 ± 0.6	17.0 ± 1.2*
Epididymal fat weight (g)	1.9 ± 0.4	10.6 ± 1.5*

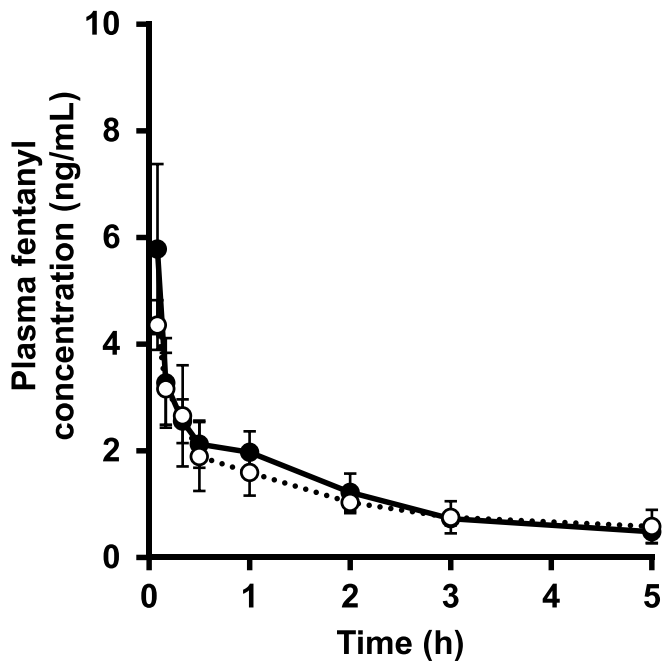
The body and tissue weights are expressed as mean ± S.D. (n = 7). \*Significant difference compared to Zucker (+/+) rats at  $p < 0.05$ .

28,29]. We consider that this apparent discrepancy is likely due to differences in the timing of assessment and the classification of BMI. On the other hand, low BMI (mean BMI: 16) cachectic patients show lower dose-adjusted plasma fentanyl concentration after transdermal administration than normal weight (mean BMI: 23) patients [17]. However, this phenomenon has been suggested to be due to hypoalbuminemia [18]. In this study, the effect of hypoalbuminemia was not observed (Fig. 2D), which could have contributed to the lack of differences in the low BMI group.

The retrospective study in patients indicated that obesity is a factor influencing the post-titration transdermal fentanyl dose. However, there

are limitations to this retrospective study, including lack of validation of the post-titration dose evaluation criteria, the absence of plasma fentanyl concentration data, the entry of patients from a single center, and the lack of adjustment for patient background factors. In order to test a hypothesis that plasma fentanyl exposure is greater in obese patients compared to normal patients, an animal study was conducted using Zucker (*fa/fa*) obese rats and Zucker (+/+) control rats.

After intravenous administration of fentanyl in rats, the effects of obesity on the plasma concentration-time profile,  $t_{1/2}$ , or  $CL_{tot}$  were not observed (Fig. 3 and Table 3), suggesting that obesity does not affect the elimination of fentanyl. The major metabolic enzyme of fentanyl is hepatic CYP3A2 in rats and CYP3A4 in humans [30,31]. Although CYP3A2 mRNA and protein level in the liver of Zucker (*fa/fa*) rats were decreased compared to those of Zucker (+/+) rats, overall CYP3A2 activity was not significantly different due to the liver enlargement in the obese rats [32]. In addition, since hepatic clearance of fentanyl is limited by liver blood flow [33], the effect of the difference in CYP3A2 activity on the systemic clearance of fentanyl may have been relatively small. On the other hand, there was a tendency for  $Vd_{ss}$  to be increased in obese rats, though this was not statistically significant (Table 3). This may be consistent with the finding that sufentanil was similarly distributed in excess body mass and lean tissues [34].



**Fig. 3.** Plasma concentration of fentanyl after intravenous administration of fentanyl citrate in Zucker (+/+) and Zucker (*fa/fa*) rats. Fentanyl citrate injection was administered intravenously to Zucker (+/+) (●) and Zucker (*fa/fa*) (○) rats at 10  $\mu$ g/kg body weight as fentanyl. Data are expressed as mean  $\pm$  S.D. ( $n = 3$ ).

**Table 3**  
Pharmacokinetic parameters of fentanyl after intravenous administration of fentanyl citrate to Zucker (+/+) and Zucker (*fa/fa*) rats.

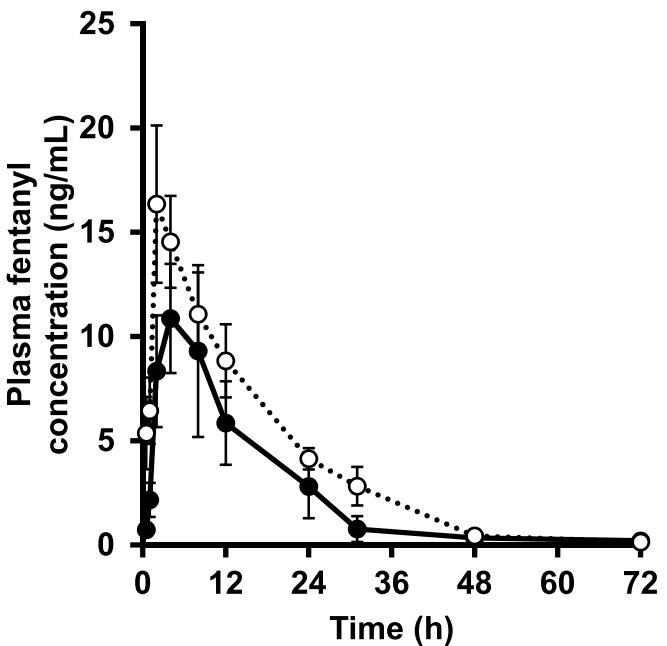
	Zucker (+/+)	Zucker ( <i>fa/fa</i> )
$AUC_{0-\infty}$ (ng $\cdot$ h/mL)	8.25 $\pm$ 1.12	9.97 $\pm$ 4.83
$t_{1/2}$ (hr)	2.70 $\pm$ 1.73	4.32 $\pm$ 2.45
$CL_{tot}$ (L/h/kg)	1.23 $\pm$ 0.18	1.28 $\pm$ 0.86
$Vd_{ss}$ (L/kg)	4.14 $\pm$ 2.14	5.47 $\pm$ 0.89

Fentanyl citrate injection was administered intravenously to Zucker (+/+) and Zucker (*fa/fa*) rats at 10  $\mu$ g/kg body weight as fentanyl. Data are expressed as mean  $\pm$  S.D. ( $n = 3$ ). No significant difference was observed between Zucker (+/+) and Zucker (*fa/fa*) rats at  $p < 0.05$ .

Transdermal administration of fentanyl in obese rats showed significantly increased  $F$ ,  $F_a$ , and  $F_{skin}$ , shorter  $T_{max}$ , and decreased  $CL_{tot}/F$ , compared to control rats (Fig. 4 and Table 4). The disposition of fentanyl after transdermal administration is described by a 1-compartment distribution model with first-order absorption and elimination [35], so the steady-state blood concentration depends on  $F$ . In this study, the contribution to  $F$  was greater for  $F_{skin}$  (1.9-fold increase) than for  $F_a$  (1.2-fold increase) in obese rats (Table 4).  $F_{skin}$  reflects the drug loss during passage through the skin, which could have been influenced by reservoir effects, binding, and metabolism of drugs in the skin [36–39]. Changes in skin physiology and impaired skin barrier function due to obesity [40–42] could have affected the disposition of fentanyl through the skin. These results support the findings in obese patients in the retrospective study, and suggest that a decrease in the post-titration transdermal fentanyl dose was caused by an increase in the transdermal bioavailability of fentanyl.

### 5. Conclusions

Our findings indicated that higher  $F$ , particularly  $F_{skin}$ , due to obesity could contribute to the lower post-titration dose of transdermal fentanyl. Although further investigation is needed, obesity should be considered



**Fig. 4.** Plasma concentration of fentanyl after transdermal administration of fentanyl citrate in Zucker (+/+) and Zucker (*fa/fa*) rats. Fentanyl citrate formulation (160  $\mu$ g as fentanyl) was applied to the back skin of Zucker (+/+) (●) and Zucker (*fa/fa*) (○) rats for 24 h. Data are expressed as mean  $\pm$  S.D. ( $n = 4$ ).

**Table 4**  
Pharmacokinetic parameters of fentanyl after transdermal administration of fentanyl citrate to Zucker (+/+) and Zucker (*fa/fa*) rats.

	Zucker (+/+)	Zucker ( <i>fa/fa</i> )
$C_{max}$ (ng/mL)	11.6 $\pm$ 2.3	17.3 $\pm$ 3.1*
$AUC_{0-\infty}$ (ng $\cdot$ h/mL)	183 $\pm$ 62	277 $\pm$ 35*
$T_{max}$ (h)	7.00 $\pm$ 2.00	3.00 $\pm$ 1.15*
$CL_{tot}/F$ (L/h/kg)	3.60 $\pm$ 1.24	1.29 $\pm$ 0.34*
$F$ (%)	36.2 $\pm$ 10.2	81.2 $\pm$ 18.9*
$F_a$ (%)	73.9 $\pm$ 9.3	87.7 $\pm$ 3.3*
$F_{skin}$ (%)	48.8 $\pm$ 11.0	92.1 $\pm$ 18.9*

Fentanyl citrate formulation (160  $\mu$ g as fentanyl) was applied to the back skin of Zucker (+/+) and Zucker (*fa/fa*) rats for 24 h. Data are expressed as mean  $\pm$  S.D. ( $n = 4$ ). \*Significant difference compared to Zucker (+/+) rats at  $p < 0.05$ .

as one of factors in the titration of transdermal fentanyl dose. In obese patients, management of transdermal fentanyl dose should be conservative when calculating the initial dose and also when increasing the dose.

### CRediT authorship contribution statement

**Satoshi Mizuno:** Writing – original draft, Investigation, Funding acquisition, Conceptualization. **Shintaro Gake:** Methodology, Investigation. **Makiko Takabayashi:** Methodology, Conceptualization. **Yuriko Ito:** Investigation. **Hiroko Sanada:** Investigation. **Natsumi Sugimoto:** Investigation. **Akari Maeda:** Investigation. **Takuto Tamamura:** Investigation. **Kazuki Sawamoto:** Methodology, Investigation, Conceptualization. **Yusuke Hara:** Methodology, Conceptualization. **Yoshiko Ohi:** Methodology. **Chiaki Tsuji:** Methodology. **Yukiko Shiomoto:** Methodology. **Yukio Kato:** Resources. **Arimi Fujita:** Writing – review & editing, Conceptualization. **Tsutomu Shimada:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Ken-ichi Miyamoto:** Conceptualization. **Yoshimichi Sai:** Writing – review & editing, Project administration, Funding acquisition, Conceptualization.

## Declaration of interest

The authors declare no conflict of interest.

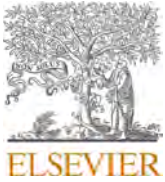
## Acknowledgments

The present study was supported in part by JSPS KAKENHI [Grant Numbers JP23590173, JP15K08092, JP19K07217, JP20K07195, JP23K06209, JP23K06253] and JST SPRING [Grant Number JPMJSP2135].

## References

- Mercadante S, Bruera E. Opioid switching: a systematic and critical review. *Cancer Treat Rev* 2006;32:304–15. <https://doi.org/10.1016/j.ctrv.2006.03.001>.
- Fallon M, Giusti R, Aielli F, Hoskin P, Rolke R, Sharma M, et al. Management of cancer pain in adult patients: ESMO Clinical Practice Guidelines. *Ann Oncol* 2018;29:iv166–1191. <https://doi.org/10.1093/annonc/mdy152>.
- Nakamura A, Hasegawa M, Ito H, Minami K, Koike K, Habu-Tomita N, et al. Distinct relations among plasma concentrations required for different pharmacological effects in oxycodone, morphine, and fentanyl. *J Pain Palliat Care Pharmacother* 2011;25:318–34. <https://doi.org/10.3109/15360288.2011.620689>.
- Mystakidou K, Katsouda E, Tsilika E, Parpa E, Vlahos L. Transdermal therapeutic fentanyl-system (TTS-F). *In Vivo* 2004;18:633–42.
- Stanley TH. The fentanyl story. *J Pain* 2014;15:1215–26. <https://doi.org/10.1016/j.jpain.2014.08.010>.
- Fiset P, Cohane C, Browne S, Brand SC, Shafer SL. Biopharmaceutics of a new transdermal fentanyl device. *Anesthesiology* 1995;83:459–69. <https://doi.org/10.1097/00000542-199509000-00004>.
- Solassol I, Caumette L, Bressolle F, Garcia F, Thézenas S, Astre C, et al. Inter- and intra-individual variability in transdermal fentanyl absorption in cancer pain patients. *Oncol Rep* 2005;14:1029–36. <https://doi.org/10.3892/or.14.4.1029>.
- Donner B, Zenz M, Tryba M, Strumpf M. Direct conversion from oral morphine to transdermal fentanyl: a multicenter study in patients with cancer pain. *Pain* 1996;64:527–34. [https://doi.org/10.1016/0304-3959\(95\)00180-8](https://doi.org/10.1016/0304-3959(95)00180-8).
- Morita T, Takigawa C, Onishi H, Tajima T, Tani K, Matsubara T, et al. Opioid rotation from morphine to fentanyl in delirious cancer patients: an open-label trial. *J Pain Symptom Manag* 2005;30:96–103. <https://doi.org/10.1016/j.jpainsymman.2004.12.010>.
- Mystakidou K, Tsilika E, Parpa E, Kouloulis V, Kouvaris I, Georgaki S, et al. Long-term cancer pain management in morphine pre-treated and opioid naive patients with transdermal fentanyl. *Int J Cancer* 2003;107:486–92. <https://doi.org/10.1002/ijc.11416>.
- Korte W, de Stoutz N, Morant R. Day-to-day titration to initiate transdermal fentanyl in patients with cancer pain: short- and long-term experiences in a prospective study of 39 patients. *J Pain Symptom Manag* 1996;11:139–46. [https://doi.org/10.1016/0885-3924\(95\)00162-x](https://doi.org/10.1016/0885-3924(95)00162-x).
- Shibutani K, Inchiosa MA, Sawada K, Bairamian M. Pharmacokinetic mass of fentanyl for postoperative analgesia in lean and obese patients † Presented in abstract form at the Annual Meeting of the American Society of Anesthesiologists, Las Vegas, October 26, 2004. *Br J Anaesth* 2005;95:377–83. <https://doi.org/10.1093/bja/aei195>.
- Shomaker TS, Zhang J, Ashburn MA. Assessing the impact of heat on the systemic delivery of fentanyl through the transdermal fentanyl delivery system. *Pain Med* 2000;1:225–30. <https://doi.org/10.1046/j.1526-4637.2000.00030.x>.
- Ashburn MA, Ogden LL, Zhang J, Love G, Basta SV. The pharmacokinetics of transdermal fentanyl delivered with and without controlled heat. *J Pain* 2003;4:291–7. [https://doi.org/10.1016/s1526-5900\(03\)00618-7](https://doi.org/10.1016/s1526-5900(03)00618-7).
- Barratt DT, Bandak B, Klepstad P, Dale O, Kaasa S, Christrup LL, et al. Genetic, pathological and physiological determinants of transdermal fentanyl pharmacokinetics in 620 cancer patients of the EPOS study. *Pharmacogenetics Genom* 2014;24:185–94. <https://doi.org/10.1097/FPC.0000000000000032>.
- Holdsworth MT, Forman WB, Killilea TA, Nystrom KM, Paul R, Brand SC, et al. Transdermal fentanyl disposition in elderly subjects. *Gerontology* 1994;40:32–7. <https://doi.org/10.1159/000213572>.
- Heiskanen T, Mätzke S, Haakana S, Gergov M, Vuori E, Kalso E. Transdermal fentanyl in cachectic cancer patients. *Pain* 2009;144:218–22. <https://doi.org/10.1016/j.pain.2009.04.012>.
- Nomura M, Inoue K, Matsushita S, Takahari D, Kondoh C, Shitara K, et al. Serum concentration of fentanyl during conversion from intravenous to transdermal administration to patients with chronic cancer pain. *Clin J Pain* 2013;29:487–91. <https://doi.org/10.1097/AJP.0b013e318266f6a5>.
- Bray GA. The Zucker-fatty rat: a review. *Fed Proc* 1977;36:148–53.
- Kava R, Greenwood MRC, Johnson PR. Zucker (fa/fa) Rat. *ILAR J* 1990;32:4–8. <https://doi.org/10.1093/ilar.32.3.4>.
- World Health Organization. WHO Guidelines for the pharmacological and radiotherapeutic management of cancer pain in adults and adolescents. 2018.
- Portenoy RK, Southam MA, Gupta SK, Lapin J, Layman M, Inturrisi CE, et al. Transdermal fentanyl for cancer pain. Repeated dose pharmacokinetics. *Anesthesiology* 1993;78:36–43. <https://doi.org/10.1097/00000542-199301000-00007>.
- Hisaka A, Sugiyama Y. Analysis of nonlinear and nonsteady state hepatic extraction with the dispersion model using the finite difference method. *J Pharmacokinet Biopharm* 1998;26:495–519. <https://doi.org/10.1023/a:1023294632129>.
- Kanda Y. Investigation of the freely available easy-to-use software “EZ” for medical statistics. *Bone Marrow Transplant* 2013;48:452–8. <https://doi.org/10.1038/bmt.2012.244>.
- Examination committee of criteria for “obesity disease” in Japan; Japan society for the study of obesity. New criteria for “obesity disease” in Japan. *Circ J* 2002;66:987–92. <https://doi.org/10.1253/circj.66.987>.
- Breivik H, Borchgrevink PC, Allen SM, Rosseland LA, Romundstad L, Hals EKB, et al. Assessment of pain. *Br J Anaesth* 2008;101:17–24. <https://doi.org/10.1093/bja/aen103>.
- Kuip EJM, Zandvliet ML, Koolen SLW, Mathijssen RHJ, van der Rijt CCD. A review of factors explaining variability in fentanyl pharmacokinetics; focus on implications for cancer patients. *Br J Clin Pharmacol* 2017;83:294–313. <https://doi.org/10.1111/bcp.13129>.
- Kuip EJM, Oldenmenger WH, Thijs-Visser MF, de Bruijn P, Oosten AW, Oomen-de Hoop E, et al. Effects of smoking and body mass index on the exposure of fentanyl in patients with cancer. *PLoS One* 2018;13:e0198289. <https://doi.org/10.1371/journal.pone.0198289>.
- Moryl N, Bokhari A, Griffo Y, Hadler R, Koranteng L, Filkins A, et al. Does transdermal fentanyl work in patients with low BMI? Patient-reported outcomes of pain and percent pain relief in cancer patients on transdermal fentanyl. *Cancer Med* 2019;8:7516–22. <https://doi.org/10.1002/cam4.2479>.
- Feierman DE. Identification of cytochrome P450 3A1/2 as the major P450 isoform responsible for the metabolism of fentanyl by rat liver microsomes. *Anesth Analg* 1996;82:936–41. <https://doi.org/10.1097/0000539-199605000-00008>.
- Labroo RB, Paine MF, Thummel KE, Kharasch ED. Fentanyl metabolism by human hepatic and intestinal cytochrome P450 3A4: implications for interindividual variability in disposition, efficacy, and drug interactions. *Drug Metab Dispos* 1997;25:1072–80.
- Sawamoto K, Huang TT, Sugimoto N, Mizutani Y, Sai Y, Miyamoto K. Mechanisms of lower maintenance dose of tacrolimus in obese patients. *Drug Metabol Pharmacokinet* 2014;29:341–7. <https://doi.org/10.2133/dmpk.dmpk-13-rg-110>.
- Ibrahim AE, Feldman J, Karim A, Kharasch ED. Simultaneous assessment of drug interactions with low- and high-extraction opioids: application to parecoxib effects on the pharmacokinetics and pharmacodynamics of fentanyl and alfentanil. *Anesthesiology* 2003;98:853–61. <https://doi.org/10.1097/00000542-200304000-00011>.
- Casati A, Putzu M. Anesthesia in the obese patient: pharmacokinetic considerations. *J Clin Anesth* 2005;17:134–45. <https://doi.org/10.1016/j.jclinane.2004.01.009>.
- Bista SR, Haywood A, Hardy J, Norris R, Hennig S. Exposure to fentanyl after transdermal patch administration for cancer pain management. *J Clin Pharmacol* 2016;56:705–13. <https://doi.org/10.1002/jcph.641>.
- Dumont C, Prieto P, Asturiol D, Worth A. Review of the availability of in vitro and in silico methods for assessing dermal bioavailability. *Appl In Vitro Toxicol* 2015;1:147–64. <https://doi.org/10.1089/aivt.2015.0003>.
- Walter K, Kurz H. Binding of drugs to human skin: influencing factors and the role of tissue lipids. *J Pharm Pharmacol* 1988;40:689–93. <https://doi.org/10.1111/j.2042-7158.1988.tb06996.x>.
- Yourick JJ, Koenig ML, Yourick DL, Bronaugh RL. Fate of chemicals in skin after dermal application: does the in vitro skin reservoir affect the estimate of systemic absorption? *Toxicol Appl Pharmacol* 2004;195:309–20. <https://doi.org/10.1016/j.taap.2003.07.015>.
- van Eijl S, Zhu Z, Cupitt J, Gierula M, Götz C, Fritsche E, et al. Elucidation of xenobiotic metabolism pathways in human skin and human skin models by proteomic profiling. *PLoS One* 2012;7:e41721. <https://doi.org/10.1371/journal.pone.0041721>.
- Löffler H, Aramaki JUN, Effendy I. The influence of body mass index on skin susceptibility to sodium lauryl sulphate. *Skin Res Technol* 2002;8:19–22. <https://doi.org/10.1046/j.0909-752x>.
- Yosipovitch G, DeVore A, Dawn A. Obesity and the skin: skin physiology and skin manifestations of obesity. *J Am Acad Dermatol* 2007;56:901–16. <https://doi.org/10.1016/j.jaad.2006.12.004>. quiz 917–916.
- Mori S, Shiraishi A, Epplen K, Butcher D, Murase D, Yasuda Y, et al. Characterization of skin function associated with obesity and specific correlation to local/systemic parameters in American women. *Lipids Health Dis* 2017;16:214. <https://doi.org/10.1186/s12944-017-0608-1>.





# Evaluation of pediatric diaphragmatic eventration using dynamic chest radiography: A case report

Takatoshi Abe<sup>a,\*</sup>, Kozo Nomura<sup>a</sup>, Seisho Sakai<sup>a</sup>, Rie Tanaka<sup>b</sup>, Shuhei Minami<sup>c</sup>,  
Hidenori Iwasaki<sup>d</sup>, Yasuhiro Ikawa<sup>e</sup>

<sup>a</sup> Department of Pediatric Surgery, Kanazawa University Hospital, Ishikawa, Japan

<sup>b</sup> College of Medical, Pharmaceutical & Health Sciences, Kanazawa University, Ishikawa, Japan

<sup>c</sup> Division of Radiological Technology, Kanazawa University Hospital, Ishikawa, Japan

<sup>d</sup> Department of Pediatrics, Graduate School of Medical Science, Kanazawa University, Ishikawa, Japan

<sup>e</sup> Department of Pediatrics, Kanazawa Medical University, Ishikawa, Japan

## ARTICLE INFO

### Keywords:

Dynamic chest radiography  
Radiation exposure  
Diaphragm eventration  
Case report

## ABSTRACT

**Introduction:** Diagnosis of diaphragmatic disease using dynamic chest radiography (DCR), a novel imaging testing modality has been previously reported in adults; however, there have been no prior reports of its use in pediatric patients. Herein, we report the novel case of a 14-year-old girl diagnosed with diaphragmatic eventration (DE) using DCR with extremely low-dose pulsed X-rays.

**Case series:** A 14-year-old girl with pectus excavatum was referred to our hospital for surgery. Pre-operative chest radiography revealed that the right diaphragm was higher than the left however, this difference was judged to be within the normal range. Computed tomography (CT) revealed a Haller index of 3.00. Although the patient had no symptoms, she complained of an unsatisfactory appearance, which was considered an indication for surgery. The Nuss procedure was performed under artificial pneumothorax, and DE was not suspected. Six months after surgery, we performed DCR instead of chest radiographs to evaluate the position of the bar and the patient's respiratory status. Utilizing the captured images, we further measured and analyzed the diaphragm movement. DCR revealed a clear impairment of diaphragmic movement on the affected side, and the patient was subsequently diagnosed with DE. However, as she was asymptomatic, we decided to follow up conservatively.

**Conclusion:** This is the first report of DCR in a pediatric patient. DCR appears to be a useful tool for the diagnosis of diaphragmatic eventration in children.

## 1. Introduction

Diaphragmatic eventration (DE) is defined as an abnormal elevation of the diaphragm caused by muscular hypoplasia following injury to the phrenic nerve [1]. It is diagnosed based on the elevation of the hemidiaphragm on chest radiography; however, it is often difficult to differentiate DE from a late-presenting diaphragmatic hernia.

**Abbreviations:** DE, diaphragmatic eventration; DCR, dynamic chest radiography; FPD, flat-panel detector.

\* Corresponding author. Department of Pediatric Surgery, Kanazawa University Hospital, Takaramachi 13-1, Kanazawa, Ishikawa, Japan.

E-mail address: [tabe3327@gmail.com](mailto:tabe3327@gmail.com) (T. Abe).

<https://doi.org/10.1016/j.epsc.2024.102932>

Received 7 October 2024; Received in revised form 26 November 2024; Accepted 29 November 2024

Available online 30 November 2024

2213-5766/© 2024 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Herein, we present the first case of DE diagnosed using dynamic chest radiography (DCR) and imaging protocols in a pediatric patient. In this report, we focused on the utility of DCR in the diagnosis of diaphragmatic disease. While the diagnosis of diaphragmatic disease using DCR has been previously reported in adult cases [2], there have been no prior reports of its use in pediatric cases.

This manuscript was prepared in accordance with the CARE guidelines (<https://www.care-statement.org>).

## 2. Case report

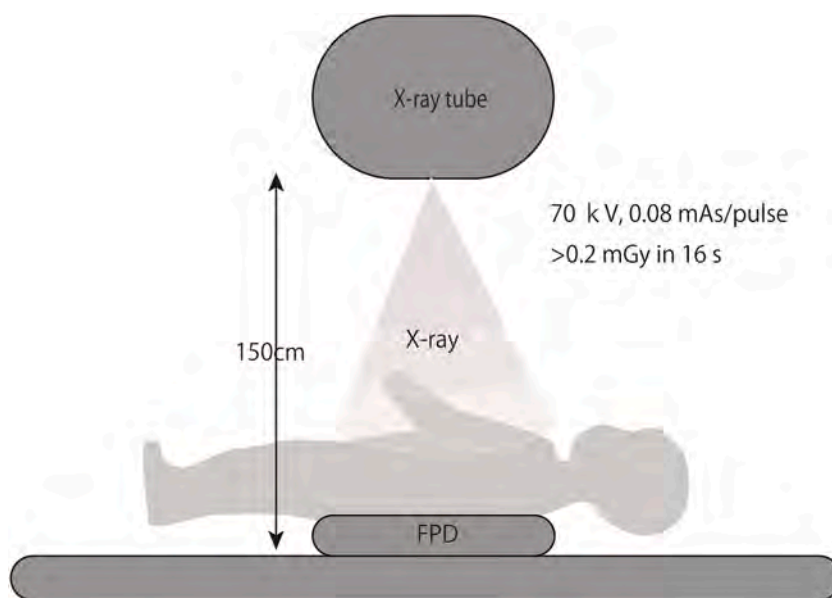
A 14-year-old girl with pectus excavatum was referred to our hospital for surgery. Depression of the chest wall was first noted when she was 12 years, and had since progressed. She expressed dissatisfaction with her appearance, but no chest pain or respiratory symptoms. Pre-operative chest radiography revealed that the right diaphragm was higher than the left, but this was judged to be within the normal range. Computed tomography (CT) revealed a Haller index of 3.00. Pulmonary function tests were within normal limits. We diagnosed pectus excavatum and decided to perform the Nuss procedure. The surgery was performed under artificial pneumothorax, at which point DE was not suspected at that time. The surgery was completed successfully. Although there were no major postoperative problems, the patient expressed such significant anxiety about pain that she was discharged on post-operative day 11.

Six months later, we further performed DCR instead of chest radiographs to evaluate the position of the bar and the patient's respiratory status after the Nuss procedure. The patients underwent DCR using a flat-panel detector (FPD) (Aero DR fine; Konica Minolta) and a pulsed X-ray generator (RADspeed Pro; Shimadzu), employing the same basic method as used in adults [3]. Given that the recommended entrance surface dose for chest radiographs in children is 0.2 mGy according to Japan's National diagnostic reference levels (2020) [4], imaging conditions were adjusted to ensure exposure doses of  $\leq 0.2$  mGy (tube voltage 70kV; tube current-exposure time product 0.08 mAs/pulse), additional copper filter (Cu-filter:0.2mm). Under these conditions, the patient underwent DCR during breathing for 16 seconds (Fig. 1).

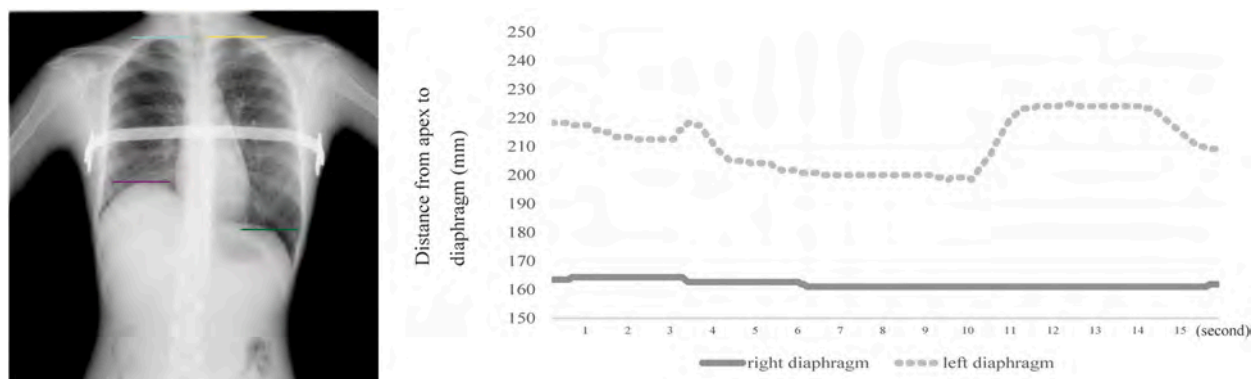
DCR revealed clear impairments of right diaphragmatic movement (Fig. 2). The left lung field had a maximum of 133 cm<sup>2</sup> and a minimum of 112 cm<sup>2</sup>, whereas the right lung field had a maximum of 119 cm<sup>2</sup> and a minimum of 113 cm<sup>2</sup>. We calculated the proportion of time the left and right diaphragms moved in the same direction during breathing, yielding a result of 1.33 %. No paradoxical movement of the hemidiaphragm was observed. We subsequently diagnosed the patient with a right DE associated with pectus excavatum. As the height of the diaphragm was unchanged from that before surgery, we assumed that DE had existed before surgery and was thus not a postoperative complication. As the patient was asymptomatic, the decision was made to perform observation on the patient without surgery.

## 3. Discussion

DE, defined as an abnormal elevation of the diaphragm, is a rare disease with an incidence of 1 in 10,000 live births [1]. DE is often difficult to diagnose as diaphragmatic hernia must be ruled out with as little invasion as possible. Radiographic and fluoroscopic examinations are commonly performed to diagnose abnormal movement of the paralyzed hemidiaphragm; however, the diagnostic criteria are not very clear, and sometimes experts must make judgments based on their experience.



**Fig. 1.** Imaging protocol for DCR for children. DCR is performed using a generator that can continuously emits pulsed X-rays and an FPD that receives the pulsed X-rays to form images.



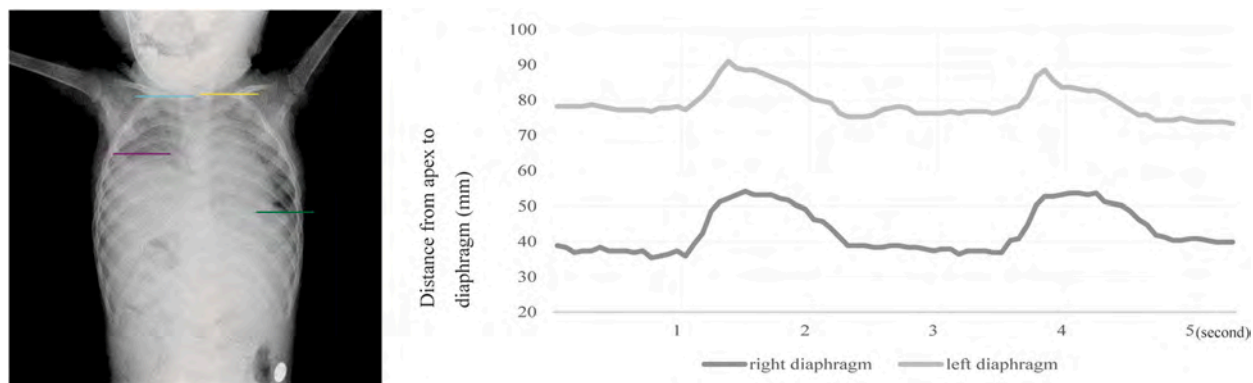
**Fig. 2.** Diaphragm movement of the diaphragmatic eventration (right side): A 14-year-old patient underwent DCR to evaluate her respiratory status following the Nuss procedure, which showed impairment of the movement of the right diaphragm.

In this study, we focused on DCR to facilitate diagnosis. DCR, a novel imaging technique that combines a pulsed X-ray generator and FPD, was introduced into clinical practice in Japan in 2018. The DCR video captures the patient's breathing state and illustrates the chest movements associated with breathing and pulmonary circulation, thereby providing valuable diagnostic information [5]. DCR is used to diagnose various diseases in adults, and it has also been reported to be useful in the evaluation of respiratory function and diagnosis of chest diseases [2]. However, there have been no reports of its use in pediatric cases. It is possible that DCR, owing to its simplicity and low radiation dose, may be beneficial in the pediatric field. Indeed, this study demonstrates the feasibility of using DCR in pediatric patients by adjusting the exposure dose.

DCR and fluoroscopic examination are similar in that they both use continuous X-rays to observe the movement of an object, however, DCR has two advantages, particularly in the diagnosis of DE. Firstly, in DCR, digitized information regarding the movement of the left and right diaphragms can be graphed instantly, allowing rapid objective evaluation. Thus, DCR can be used to analyze and graph even slight changes in dynamic image findings, such as diaphragm movement, that are beyond the human ability to discern and process. Second, the amount of radiation received by the patient during a 5-s fluoroscopy examination is the same as the amount of radiation received to perform a DCR lasting for 16 seconds [4]. Thus, DCR can obtain more diagnostic materials with lower radiation exposure.

In the present case, DCR examination revealed severe motion impairment of the right diaphragm, and the percentage of synchronous movements of the left and right diaphragms during the test was 1.33 %. Conversely, in the case of a late-presenting diaphragmatic hernia, left and right diaphragmatic movements were synchronous, and the proportion of time the left and right diaphragms moved in the same direction during breathing was 49.3 % (Fig. 3). By performing DCR, both radiograph and fluoroscopy are obtained, which are necessary tests for DE, allowing evaluation of the diaphragm movement with lower radiation exposure.

It has been reported that ultrasound examination is useful for differentiating between DE and diaphragmatic hernia [6]. Of course, diagnosis should be made without unnecessary exposure of children to radiation, but there are very few pediatric surgeons or pediatric radiologists who are skilled enough to differentiate between these using ultrasound examinations. In fact, due to a lack of experience, we do not diagnose DE using ultrasound at our institution. DCR exposes patients to radiation equivalent to that of a chest radiograph, but it is possible to make a diagnosis without a direct examination by a specialist. Depending on the patient's region and situation, DCR may be given priority over ultrasound.



**Fig. 3.** Diaphragm movement of diaphragmatic hernia (right side): A 1-year-old patient underwent DCR for pre-operative examination. DCR showed diaphragm movements were synchronous between the left and right sides.

#### 4. Conclusion

In conclusion, DCR is useful for the diagnosis of DE through dose adjustments, even in pediatric patients. By accumulating cases and establishing analysis protocols for children, DCR appears to be a useful tool for the diagnosis of diaphragmatic eventration in children.

#### CRedit authorship contribution statement

**Takatoshi Abe:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Kozo Nomura:** Data curation. **Seisho Sakai:** Data curation. **Rie Tanaka:** Writing – review & editing, Resources, Methodology, Formal analysis, Conceptualization. **Shuhei Minami:** Methodology. **Hidenori Iwasaki:** Data curation. **Yasuhiro Ikawa:** Writing – review & editing, Data curation.

#### Informed consent

Informed consent was obtained from the patient or guardian.

#### Authorship

All authors attest that they meet the current ICMJE criteria for Authorship.

#### Financial interests

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### References

- [1] Groth SS, Andrade RS. Diaphragmatic eventration. *Thorac Surg Clin* 2009;19:511–9. <https://doi.org/10.1016/j.thorsurg.2009.08.003>.
- [2] FitzMaurice TS, McCann C, Nazareth DS, Walshaw MJ. Characterisation of hemidiaphragm dysfunction using dynamic chest radiography: a pilot study. *ERJ open research* 2022;8:343. <https://doi.org/10.1183/23120541.00343-2021>.
- [3] Tanaka R, Matsumoto I, Tamura M, et al. Dynamic chest radiography: clinical validation of ventilation and perfusion metrics derived from changes in radiographic lung density compared to nuclear medicine imaging. *Quant Imag Med Surg* 2021;11:4016–27. <https://doi.org/10.21037/qims-20-1217>.
- [4] Japan Network for Research and Information on Medical Exposure (J-RIME). National diagnostic reference levels in Japan (2020) Japan DRLs 2020. et al, [https://www.radher.jp/J-RIME/report/DRL2020\\_Engver.pdf](https://www.radher.jp/J-RIME/report/DRL2020_Engver.pdf). [Accessed 12 October 2020].
- [5] Tanaka R. Dynamic chest radiography: flat-panel detector (FPD) based functional X-ray imaging. *Radiol Phys Technol* 2016;9:139–53. <https://doi.org/10.1007/s12194-016-0361-6>.
- [6] Karmazyn B, Shold AJ, Delaney LR, et al. Ultrasound evaluation of right diaphragmatic eventration and hernia. *Pediatr Radiol* 2019;49:1010–7. <https://doi.org/10.1007/s00247-019-04417-1>.




RESEARCH ARTICLE

Open Access



# Effect of changes in skin properties due to diabetes mellitus on the titration period of transdermal fentanyl: single-center retrospective study and diabetic animal model study

Satoshi Mizuno<sup>1</sup>, Makiko Takabayashi<sup>2</sup>, Hiroko Makihara<sup>3,4</sup>, Kazuhiro Ogai<sup>5</sup>, Kei Tsukui<sup>6</sup>, Yuriko Ito<sup>2</sup>, Takahiro Kawakami<sup>2</sup>, Yusuke Hara<sup>2</sup>, Arimi Fujita<sup>1,2</sup>, Yoshihiro Tokudome<sup>7</sup>, Tomoko Akase<sup>4</sup>, Yukio Kato<sup>8</sup>, Tsutomu Shimada<sup>9</sup> and Yoshimichi Sai<sup>1,2,10\*</sup> 

## Abstract

**Background** In the dose titration of transdermal fentanyl to prevent unrelieved pain, it is important to consider not only dose adjustment, but also the titration period, which is influenced by the time required to reach the steady state. Many patients with cancer pain experience comorbidities that might affect the skin properties and influence transdermal absorption. We hypothesized that skin changes due to diabetes mellitus (DM) would affect the titration period of transdermal fentanyl. We conducted a retrospective study and diabetic animal model study to test this hypothesis.

**Methods** In the retrospective study, the titration period was defined in terms of “dose change” and “number of rescue opioids” in patients initiated on transdermal fentanyl. Multiple logistic regression analysis was performed to analyze the relation between the titration period and comorbidities, including DM. In the diabetic animal model study, intercellular lipids of stratum corneum (SC) were analyzed in Goto-Kakizaki (GK) rats, a model of DM, and the pharmacokinetics of intravenously or transdermally administered fentanyl was examined.

**Results** In the retrospective study, the titration period ranged from 5 to 39 days ( $n = 387$ ), and the patients taking a longer period (6 days or more) was significantly related to in patients with unspecified DM: AOR (95% confidence interval), 0.438 (0.217–0.884). In the diabetic animal model study, the ceramides (CERs) content in the SC was decreased by approximately 30% in GK rats compared to Wistar rats. The absorption rate constant ( $k_a$ ) of fentanyl administered transdermally was increased approximately 1.4-fold in GK rats, though there was no difference in transdermal bioavailability ( $F$ ) or systemic clearance ( $CL_{tot}$ ).

**Conclusion** Our results suggest that the steady state of transdermally administered fentanyl is reached sooner in cancer patients with DM as a comorbidity. Earlier pain assessment and dose adjustment may be possible in these patients.

**Keywords** Transdermal fentanyl, Opioid titration, GK rats, Diabetes mellitus, Stratum corneum, Intercellular lipids, Ceramides

\*Correspondence:

Yoshimichi Sai

sai-ys@staff.kanazawa-u.ac.jp

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Transdermal fentanyl relieves moderate to severe cancer pain and chronic pain [1, 2]. Unrelieved pain impacts all dimensions of quality of life [3], so needs to be controlled rapidly after switching from other opioids to transdermal fentanyl. However, due to the opioid-specific adverse effects [1, 2] and a large inter-individual variability in drug release from the transdermal formulation [4, 5], the dose must be carefully titrated. Therefore, factors related to transdermal fentanyl dose at the steady state, including gender, age, body mass index (BMI), and serum albumin, have been investigated [6–10], but so far there is little information on factors affecting the titration period.

There is a relationship between blood fentanyl concentration and the extent of pain relief [11, 12], and the dose titration period of transdermal fentanyl is influenced by the time required for the blood concentration to reach a steady state. The disposition of transdermal fentanyl is described by a one-compartment distribution model with first-order absorption and elimination, and follows flip-flop kinetics [13–16], where the half-life depends on the absorption rate constant ( $k_a$ ). The rate-limiting step in transdermal absorption is diffusion through the stratum corneum (SC), the outermost layer of the skin [17, 18]. The SC consists of corneocytes embedded in a lipid matrix, which acts as the main barrier to diffusion of substances through the skin [19].

More than half of cancer patients have at least one comorbidity [20, 21]. Cancer and diabetes mellitus (DM) are prevalent globally and are frequently diagnosed in the same individual [22, 23]. In addition, many diabetic patients experience cutaneous complications [24], such as diabetic ulcers due to impaired healing [25]. Furthermore, diabetic animal models, such as Otsuka Long-Evans Tokushima Fatty (OLETF) rats and db/db mice, show decreased levels of intercellular lipids in the SC even in non-injured skin [26, 27]. Such lipid changes may affect the permeation of drugs.

We hypothesized that skin changes due to comorbidities, including DM, would affect the dose titration period of transdermal fentanyl. To test this hypothesis, we carried out a retrospective study in cancer patients and a diabetic animal model study to investigate the titration period and the pharmacokinetics of transdermal fentanyl.

## Methods

### Retrospective study

Patients who initiated treatment with 24 h-fentanyl citrate formulation (Fentos tape) between Apr 1, 2015, and Mar 31, 2019, at Kanazawa University Hospital were enrolled. Exclusion criteria are shown in Fig. 1. The survey items were gender, age, height, weight, BMI, serum albumin (before treatment), comorbidities listed on the

physician's problem list (including history), initial dose of transdermal fentanyl, and the titration period. Comorbidities were coded using the International Classification of Diseases 10th Revision (ICD-10) ver. 2014 three-character categories (A00-Q99). For cancer metastases, four-character subcategories were used (C78.0-C79.9). The titration period was defined as the time until two criteria, "no dose change for 5 days" and "rescue opioids within 3 times per day" were both met. The half-life of fentanyl after removal of the transdermal formulation is  $21.9 \pm 8.9$  h (mean  $\pm$  S.D.) [28], so that the steady state should be reached in about 5 days. Furthermore, the European Society for Medical Oncology (ESMO) guidelines recommends baseline opioid adaptations if rescue opioids are needed more than 4 times per day [2].

### Chemicals

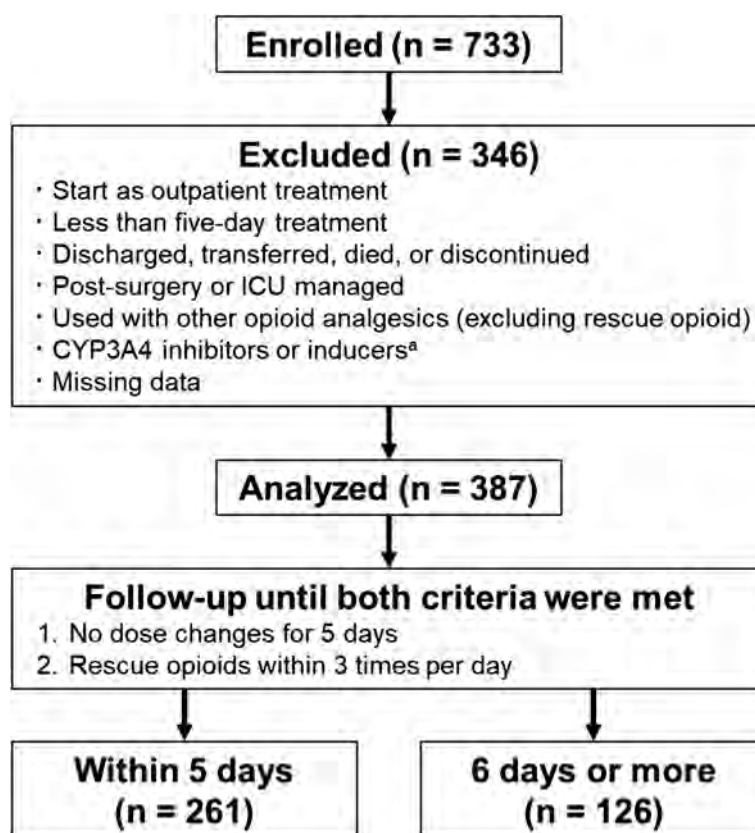
Fentanyl citrate injection and fentanyl citrate formulation (Fentos tape) were purchased from Terumo Corporation (Tokyo, Japan), and Kyowa Kirin Co., Ltd. (Tokyo, Japan), respectively. Isoflurane was purchased from Viatris Inc. (Canonsburg, PA, USA). Epilat hair removing cream was purchased from Kracie Ltd. (Tokyo, Japan). Trypsin from bovine pancreas was purchased from Sigma-Aldrich Co. LLC (St. Louis, MO, USA). Norfentanyl-D5 oxalate and papaverine hydrochloride were purchased from Sigma-Aldrich Co. LLC (St. Louis, MO, USA) and Nacalai Tesque (Kyoto, Japan), respectively. All other chemicals were of analytical or high-performance liquid chromatography (HPLC) grade.

### Animals

Goto-Kakizaki (GK) rats as a diabetic animal model [29] and Wistar rats as controls were purchased at 10 weeks old from Japan SLC (Hamamatsu, Japan). GK rats were fed a low protein/high carbohydrate diet (LABO MR-DBT, Nossan Corporation, Yokohama, Japan) and Wistar rats were fed a standard diet until experiments at 12–14 weeks old. Diets were provided ad libitum. Non-fasting blood glucose levels were measured using Glutest Neo Alpha (Sanwa Kagaku Kenkyusho, Nagoya, Japan). The hairs on the lateral abdominal skin were removed by treatment with hair removal cream at least 1 h before the experiments.

### Analysis of skin parameters and intercellular lipids of SC in rats

Transepidermal water loss (TEWL), SC hydration, and surface pH in the lateral abdominal skin were measured under isoflurane anesthesia using a Tewameter TM300MP, Corneometer CM825, and Skin-pH-Meter



**Fig. 1** Flowchart of analysis in the retrospective study. In total, 733 patients were enrolled, and 387 patients were analyzed. Among them 261 patients showed a titration period of within 5 days, and 261 showed a titration period of 6 days or more. <sup>a</sup> Inducers or inhibitors of CYP3A4 used in excluded patients were as follows: carbamazepine, enzalutamide, mitotane, phenobarbital, phenytoin, rifampicin, amiodarone, cobicistat, clarithromycin, diltiazem, fluvoxamine, indinavir, itraconazole, nelfinavir, saquinavir, ritonavir, and voriconazole

PH905, respectively (Courage + Khazaka electronic GmbH, Köln, Germany). Each value was calculated as the mean of three individual measurements. The SC sheet was separated from the lateral abdominal skin by treatment with 0.1% (w/v) trypsin in water for 21 h at 4°C and 5 h at 37°C. Free fatty acids (FFAs), ceramides (CERs), and cholesterol (CHOL) in the SC were measured as described previously [30]. CERs was calculated for a total of 12 CER subclasses [19].

#### Intravenous and transdermal administration of fentanyl in rats

For intravenous administration, fentanyl citrate injection diluted in saline (10 µg fentanyl/2 mL/kg body weight) was injected via the jugular vein over a period of 10 s under isoflurane anesthesia. Blood samples (100 µL) were collected from the jugular vein using a heparin-coated syringe at 1, 5, 10, 20, 30, 60, 120, 180, and 300 min. For transdermal administration, one-fourth of 1 mg fentanyl citrate formulation (160 µg fentanyl) was applied to the lateral abdominal skin and blood samples were collected

at 1, 2, 4, 8, 12, 24, 30, 48 h. The formulation was removed at 24 h and stored at 4°C until analysis. The blood samples were centrifuged, and plasma samples were stored at -80°C until analysis.

#### Analysis of fentanyl in the blood samples and the formulation

Plasma fentanyl concentration was measured as described, with some modifications [31]. Briefly, a mixture of plasma sample, internal standard (IS, norfentanyl-D5), and acetonitrile was centrifuged, then the supernatant was evaporated to dryness under a nitrogen stream and reconstituted with the mobile phase, 0.1% (v/v) formic acid in water (solvent A)/0.1% (v/v) formic acid in acetonitrile (solvent B) (90:10). A Shimadzu HPLC system and LCMS-8045 (Shimadzu, Tokyo, Japan) were used and the analytical column (TSKgel ODS-100 V, 3 µm, 75×2.0 mm I.D., Tosoh, Tokyo, Japan) was maintained at 40°C. The gradient was 0 min (10% solvent B)-2 (10%)-8 (70%)-8.01 (100%)-11 (100%)-11.01 (10%)-14 (10%), for 14 min per sample.

Fentanyl and IS were monitored at 337.49/188.20 and 238.20/84.15, respectively.

Fentanyl in the formulation was measured by HPLC–UV. The formulation was soaked in methanol/acetonitrile (1:3) containing IS (papaverine) and sonicated at 36°C, 38 kHz for 70 min. A Shimadzu HPLC–UV system was used with the analytical column ( $\mu$ Bondasphere, PH 5  $\mu$ m, 300 Å, 150  $\times$  3.9 mm, Waters Corporation, Milford, MA, USA) maintained at 40°C, and isocratic mode was used in 0.1% (v/v) formic acid water/acetonitrile (70:30). UV absorption was monitored at 195 nm.

### Estimation of pharmacokinetic parameters

Pharmacokinetic parameters of fentanyl were estimated by model-independent moment analysis using Napp version 2.01 software [32]. Transdermal bioavailability ( $F$ ) was calculated by dividing the area under the curve after transdermal administration ( $AUC_{td,0-\infty}$ ) obtained from each animal by the mean  $AUC$  after intravenous administration ( $AUC_{iv,0-\infty}$ ) corrected for the dose of fentanyl. The  $k_a$  of fentanyl administered transdermally was optimized by fitting data obtained after intravenous administration (elimination rate constant,  $k_e$ ; and volume of distribution at steady state,  $Vd_{ss}$ ) and transdermal administration (dose; concentration–time profile; and  $F$ ) to a one-compartment model provided in the software. Fentanyl release ratio from the formulation ( $F_a$ ) and skin availability ( $F_{skin}$ ) were calculated as follows:

$$F_a = \frac{X_{unused} - X_{used}}{X_{unused}}$$

$$F_{skin} = \frac{F}{F_a}$$

where  $X_{used}$  and  $X_{unused}$  represent the residual amount of fentanyl in the formulation removed from the rats and the amount of fentanyl in the unused formulation, respectively.

### Statistical analysis

Patients were classified into two groups: those with a titration period within 5 days and those with a titration period of 6 days or more. Univariate logistic regression analysis was performed for gender, age, BMI, albumin concentration, initial dose, and comorbidities. The adjusted odds ratio (AOR) was calculated by multivariate logistic regression analysis for factors showing  $p < 0.1$  in the univariate analysis. In the diabetic animal model study, the two groups were compared with an unpaired t-test. The analysis software was IBM SPSS

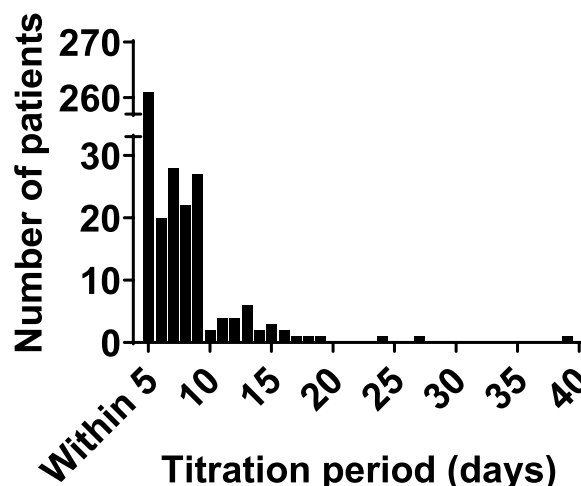
ver. 27 (IBM, Armonk, NY, USA) and the significance level was set at  $p < 0.05$ .

## Results

### Relation between the titration period of transdermal fentanyl and DM as a comorbidity in the retrospective study

In the 387 patients analyzed (Fig. 1), the distribution of the titration period is shown in Fig. 2. Patients who showed the titration period of more than 5, 10, and 15 days amounted to 32.6%, 7.0%, and 2.1%, respectively, and the longest titration period was 39 days. Table 1 shows the patient backgrounds. The 150 comorbidities extracted are shown in Table S1. The most frequent cancer types involved the digestive organs (48.6%), lip, oral cavity and pharynx (18.1%), and respiratory and intrathoracic organs (12.7%), and there were only two non-cancer patients. The most frequent comorbidity groups were circulatory system diseases (41.1%), digestive system diseases (38.0%), and endocrine, nutritional and metabolic diseases (36.4%). DM was classified as type 2 (2.8%) or unspecified (17.1%).

In univariate analysis, no relation was observed between the titration period of 6 days or more and gender, age, BMI, serum albumin, or initial dose (Table 2). Twelve comorbidities with  $p < 0.1$  in univariate analysis were included in the multivariate analysis, which confirmed that AOR for a titration period of 6 days or more was decreased in patients with unspecified DM (E14), AOR (95% confidence interval), 0.438 (0.217–0.884), while it was increased in patients with aortic aneurysm and dissection (I71), 4.590 (1.008–20.905).



**Fig. 2** Titration period distribution of transdermal fentanyl in the retrospective study. The X-axis shows the titration period of transdermal fentanyl, and the Y-axis shows the number of patients by period ( $n = 387$ )



**Table 1** Patient backgrounds in the retrospective study

	Total	Titration period	
		Within 5 days	6 days or more
n	387	261	126
Male	239 (61.8)	158 (60.5)	81 (64.3)
Female	148 (38.2)	103 (39.5)	45 (35.7)
Age, years	66 [58–73]	66 [58–73]	66 [59–73]
Height, cm	162.3 [155.4–168.2]	162.0 [155.1–168.0]	162.4 [156.4–169.0]
Weight, kg	52.9 [45.7–60.7]	53.1 [45.7–61.6]	52.0 [45.7–59.0]
BMI, kg/m <sup>2</sup>	20.3 [18.3–22.5]	20.4 [18.4–22.9]	19.9 [18.2–21.8]
Serum albumin <sup>a</sup> , g/dL	2.9 [2.4–3.4]	2.9 [2.3–3.5]	2.9 [2.5–3.4]
Initial dose <sup>b</sup> , mg	1 [1–1]	1 [1–1]	1 [1–2]

Data are expressed as n (%) or median [25–75 percentile]

<sup>a</sup> Data for 353 (within 5 days: 237 and 6 days or more: 116) patients

<sup>b</sup> The 1 mg formulation contains 0.64 mg as fentanyl, and the estimated average absorption is 0.3 mg/day as fentanyl

### Biochemical parameters, skin parameters, and intercellular lipids of SC in GK and Wistar rats

Non-fasting blood glucose level and body weight in GK rats (mean  $\pm$  S.D.,  $437 \pm 89$  mg/dL and  $339 \pm 17$  g) were significantly higher than those in Wistar rats ( $165 \pm 29$  mg/dL and  $305 \pm 10$  g), respectively ( $n = 10$ ,  $p < 0.05$ ).

No significant difference was observed in non-invasive skin parameters such as TEWL, SC hydration and surface pH between GK and Wistar rats (Figs. 3A–C). However, the composition of intercellular lipids of SC was significantly different, with the CERs content being approximately 30% lower in GK rats than in Wistar rats (Figs. 3D–F).

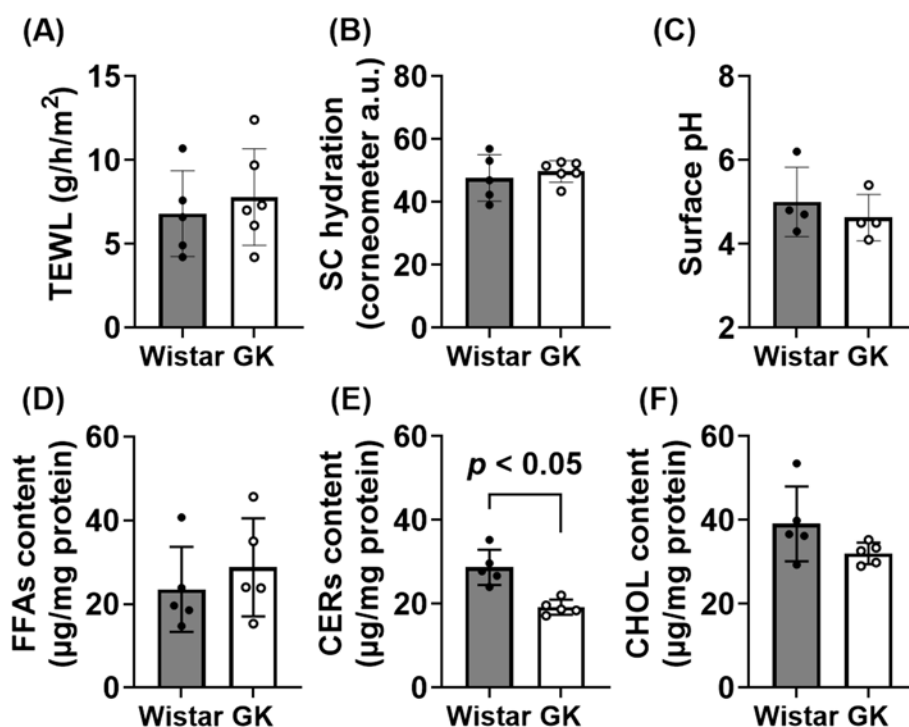
**Table 2** Logistic regression analysis related to the titration period of 6 days or more

Factors	Univariate				Multivariate			
	OR	95% CI		p value	AOR	95% CI		p value
Female	0.852	0.548	1.324	0.477	N/A			N/A
Age	1.007	0.989	1.025	0.460	N/A			N/A
BMI	0.955	0.897	1.016	0.146	N/A			N/A
Serum albumin <sup>a</sup>	0.977	0.723	1.320	0.879	N/A			N/A
Initial dose	1.029	0.771	1.375	0.845	N/A			N/A
Comorbidities <sup>b</sup> (ICD-10 code)								
Malignant neoplasm of bladder (C67)	5.059	1.286	19.905	0.020	3.831	0.854	17.176	0.079
Secondary malignant neoplasm of other and unspecified digestive organs (C78.8)	3.554	0.836	15.113	0.086	2.009	0.401	10.056	0.396
Secondary malignant neoplasm of bone and bone marrow (C79.5)	1.584	0.938	2.677	0.085	1.563	0.890	2.745	0.120
Secondary malignant neoplasm of adrenal gland (C79.7)	3.779	1.085	13.159	0.037	2.475	0.610	10.047	0.205
Myeloid leukaemia (C92)	8.525	0.943	77.076	0.056	6.771	0.595	77.025	0.123
Other hypothyroidism (E03)	3.012	0.937	9.685	0.064	2.635	0.766	9.061	0.124
Unspecified diabetes mellitus (E14)	0.451	0.236	0.863	0.016	0.438	0.217	0.884	0.021
Unspecified dementia (F03)	8.525	0.943	77.076	0.056	8.415	0.882	80.257	0.064
Aortic aneurysm and dissection (I71)	4.300	1.057	17.485	0.042	4.590	1.008	20.905	0.049
Other diseases of digestive system (K92)	4.246	0.767	23.498	0.098	4.040	0.649	25.147	0.134
Other spondylopathies (M48)	4.300	1.057	17.485	0.042	3.055	0.652	14.322	0.157
Hyperplasia of prostate (N40)	4.246	0.767	23.498	0.098	4.241	0.672	26.765	0.124

OR Odds ratio, AOR Adjusted odds ratio, CI Confidence interval, N/A Not applicable

<sup>a</sup> Data for 353 patients

<sup>b</sup> Those showing  $p < 0.1$  in univariate analysis



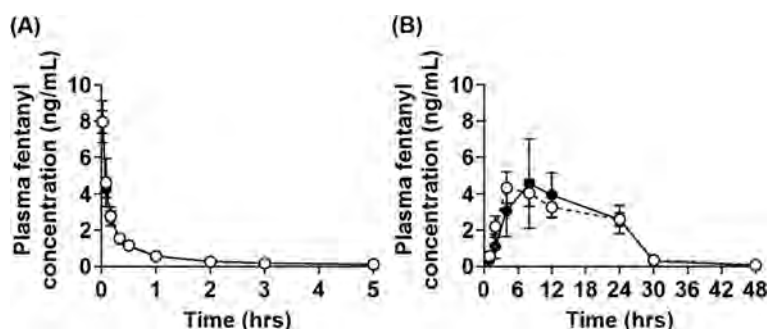
**Fig. 3** Skin parameters and intercellular lipids of SC in GK and Wistar rats. Transepidermal water loss (TEWL) (A), stratum corneum (SC) hydration (B), and surface pH (C) in the lateral abdominal skin were analyzed as skin parameters. As intercellular lipids of SC, the contents of free fatty acids (FFAs) (D), ceramides (CERs) (E), and cholesterol (CHOL) (F) were analyzed. Data are expressed as dots for individual values and bars for mean  $\pm$  S.D. ( $n=4-6$ )

#### Pharmacokinetic parameters of fentanyl after intravenous and transdermal administration in GK and Wistar rats

Plasma concentration–time profile and pharmacokinetic parameters after intravenous administration of fentanyl (10  $\mu\text{g}/\text{kg}$ ) in GK and Wistar rats are shown in Fig. 4A and Table 3. After intravenous administration, no significant difference was observed in  $AUC_{0-\infty}$ , systemic clearance ( $CL_{\text{tot}}$ ),  $k_e$ , or  $Vd_{ss}$  between GK and Wistar rats.

Plasma concentration–time profiles and pharmacokinetic parameters after transdermal administration of

fentanyl (160  $\mu\text{g}$ ) in GK and Wistar rats are shown in Fig. 4B and Table 3. After transdermal administration, no significant difference was observed in peak concentration ( $C_{\text{max}}$ ),  $AUC_{0-\infty}$ , transdermal clearance ( $CL_{\text{tot}}/F$ ),  $F$ , or  $F_{\text{skin}}$  between GK and Wistar rats. However, in GK rats, the peak time ( $T_{\text{max}}$ ) was significantly shortened by approximately 50%, and  $F_a$  was significantly increased by approximately 10% compared to Wistar rats. The  $k_a$  in GK rats was significantly higher by approximately 1.4-fold than that in Wistar rats.



**Fig. 4** Plasma fentanyl concentration after intravenous and transdermal administration in GK and Wistar rats. Fentanyl citrate injection (10  $\mu\text{g}/\text{kg}$  body weight as fentanyl) was administered intravenously to GK (open circle) and Wistar (closed circle) rats (A). Fentanyl citrate formulation (160  $\mu\text{g}$  as fentanyl) was applied to the lateral abdominal skin and the formulation was removed at 24 h (B). Data are expressed as mean  $\pm$  S.D. ( $n=5$ )

**Table 3** Pharmacokinetic parameters of fentanyl after intravenous and transdermal administration in GK and Wistar rats

	Wistar rats	GK rats
Intravenous administration		
$AUC_{0-\infty}$ (ng·hr/mL)	3.12 ± 0.16	3.22 ± 0.44
$CL_{tot}$ (L/hr/kg)	3.21 ± 0.17	3.15 ± 0.42
$k_e \times 10^{-1}$ (/hr)	2.79 ± 0.26	2.55 ± 0.34
$Vd_{ss}$ (L/kg)	6.49 ± 0.32	6.94 ± 1.77
Transdermal administration		
$C_{max}$ (ng/mL)	4.74 ± 2.32	4.40 ± 0.86
$T_{max}$ (hr)	9.60 ± 2.19	4.80 ± 1.79*
$AUC_{0-\infty}$ (ng·hr/mL)	88.6 ± 27.5	87.9 ± 12.5
$CL_{tot}/F$ (L/hr/kg)	6.51 ± 1.90	5.68 ± 0.81
$k_a \times 10^{-2}$ (/hr)	2.28 ± 0.67	3.17 ± 0.30*
$F$ (%)	53.0 ± 16.3	55.5 ± 7.1
$F_d$ (%)	69.0 ± 6.3	78.0 ± 5.8*
$F_{skin}$ (%)	76.3 ± 20.1	71.4 ± 10.5

Fentanyl citrate injection (10 µg/kg body weight as fentanyl) was administered intravenously to GK and Wistar rats. Fentanyl citrate formulation (160 µg as fentanyl) was applied to the lateral abdominal skin and the formulation was removed at 24 h. Data are expressed as mean ± S.D. (n = 5). \* Significantly different from Wistar rats at  $p < 0.05$

Discussion

To prevent unrelieved pain and improve quality of life, it is important to consider not only the dose, but also the titration period in the dose titration of transdermal fentanyl. We hypothesized that skin changes due to comorbidities would affect the titration period of transdermal fentanyl. Many patients diagnosed with cancer also have DM [22, 23], and DM can alter skin properties [24, 26, 27, 33]. In this study, therefore, we especially focused on DM as a comorbidity and conducted a retrospective study and a diabetic animal model study.

In the retrospective study, the titration period of transdermal fentanyl was defined in terms of “dose changes” and “number of rescue opioids”, not pain intensity, and showed a large inter-individual variability (range: 5–39 days). This is consistent with a report that defined the time to achieve stable pain control with various opioids in terms of “pain intensity” and “number of rescue opioids”, obtaining a median time ranging from 4–22 days for mild to severe pain [34]. On the other hand, logistic regression analysis in the retrospective study indicated that the titration period was shorter in patients with unspecified DM. Possible reasons include effects on transdermal absorption due to changes of skin properties in DM [24] and on pain perception due to diabetic neuropathy [35], but it has not been established whether there is a causal relationship.

The retrospective study has a number of limitations, including lack of validation of the titration period

evaluation criteria, the absence of plasma fentanyl concentration–time data, the entry of patients from a single center, the lack of adjustment for patient background factors, unspecified severity of comorbidities, and unknown effects of cancer chemotherapy on the skin. Additionally, since the time to reach steady state was estimated to be 5 days, the minimum period for follow-up, the effects on a shorter time scale could not be evaluated.

To test the hypothesis that skin changes due to DM would affect the titration period of transdermal fentanyl, GK rats were used as a diabetic animal model. The skin barrier function, measured in terms of TEWL, SC hydration and surface pH, differs in diabetic patients and diabetic models depending on the duration of DM onset and the severity of DM [24, 26, 27, 33]. In the GK rats used in this study, no changes in these parameters in the lateral abdominal skin were observed compared to Wistar rats, suggesting that the model in this study was an early-stage diabetic model.

The intercellular lipids of SC have a role as a skin barrier, and FFAs, CERs, and CHOL are present in approximately equimolar ratio [19]. We found that the CERs content was decreased in the lateral abdominal skin of GK rats compared to Wistar rats, whereas the levels of FFAs and CHOL were unchanged. In other diabetic animal models, OLETF rats and db/db mice, a decrease in relative mRNA expression of serine palmitoyl transferase, an enzyme related to CERs biosynthesis, was reported [26, 27]. On the other hand, a decrease of approximately 30% in CERs content has also been observed in atopic dermatitis (AD) [36]. Studies in model membranes that mimic AD indicate that the alteration in composition affects the lateral packing of lipids, the structure of drug permeation pathways and the permeability of ethyl *p*-aminobenzoate [37, 38]. Therefore, alterations in intercellular lipids of SC could affect the transdermal absorption of fentanyl.

In the pharmacokinetic experiments in rats, optimized  $k_a$  after transdermal administration was lower than  $k_e$  after intravenous administration, which is consistent with the observation of flip-flop kinetics in humans [13–16]. Generally, the steady state is reached in about five times the half-life, which is dependent on the elimination half-life of the drug and is inversely proportional to  $k_e$ . However, in the case of very slow absorption from the administration site, so-called flip-flop kinetics may be observed in which the apparent terminal half-life is dependent on  $k_a$ . Under this condition ( $k_a < k_e$ ), the half-life is inversely proportional to  $k_a$ . In GK rats,  $k_a$  was significantly higher and  $T_{max}$  was significantly shorter after transdermal administration compared to Wistar rats, although no difference was observed in  $CL_{tot}$  after

intravenous administration. These results suggest that fentanyl rapidly reaches the steady state during repeated transdermal administration in DM and support the findings in the retrospective study in humans.

To investigate the steps that affect the transdermal absorption of fentanyl in GK rats,  $F_a$  and  $F_{skin}$  were calculated.  $F_a$  is commonly used in the assessment of transdermal fentanyl, reflecting distribution from the formulation to the SC [4, 5, 39–41]. On the other hand,  $F_{skin}$ , the fraction absorbed through the skin, could be affected by reservoir effects, binding, and metabolism in the skin [42–45]. The higher value of  $F_a$  in GK rats compared to Wistar rats, despite the lack of difference in  $F_{skin}$ , suggests that the decrease of CERs content in GK rats was one of factors contributing to an increase in the distribution of fentanyl from the formulation to the SC.

A further possible issue is that DM patients may be using skincare products containing CERs to improve their skin condition. This might increase the content of CERs in the SC, and at least partially reverse the increase of  $k_a$  and the shortened titration period of fentanyl after transdermal administration. However, the effects of supplementing CERs on transdermal absorption of fentanyl are currently unknown, and further investigation is needed.

This study focused on type 2 DM, because no patient with type 1 DM was identified in this retrospective study and because type 2 DM accounts for approximately 90% of all DM [24, 46]. However, the prevalence of type 1 DM is increasing [46], and skin disorders are observed in type 1 DM as well as type 2 DM [47]. Thus, type 1 DM may also affect the transdermal absorption of fentanyl, though further investigation will be needed to confirm this.

A shorter titration period of transdermal fentanyl would contribute to improving patients' quality of life. However, in clinical practice the dose is titrated after an interval of several days, because a long time is required to reach the steady state. Our findings indicate that the steady state is reached earlier in DM, so that it may be feasible to carry out pain assessment and dose adjustment of fentanyl at an earlier time point in cancer patients with DM as a comorbidity.

## Conclusion

Our results indicate that transdermally administered fentanyl reaches the steady state earlier in DM. For rapid pain relief in cancer patients with DM, transdermal fentanyl dose adjustment of fentanyl may be feasible at an earlier time point than in cancer patients without DM.

## Abbreviations

AD	Atopic dermatitis
AUC	Area under the curve
AOR	Adjusted odds ratio

BMI	Body mass index
CERs	Ceramides
CHOL	Cholesterol
CI	Confidence interval
CYP	Cytochrome P450
$C_{max}$	Peak concentration
$CL_{tot}$	Systemic clearance
$CL_{tot}/F$	Transdermal clearance
DM	Diabetes mellitus
ESMO	European Society for Medical Oncology
$F$	Transdermal bioavailability
$F_a$	Fentanyl release ratio from the formulation
FFAs	Free fatty acids
$F_{skin}$	Skin availability
GK rats	Goto-Kakizaki rats
HPLC	High-performance liquid chromatography
ICD-10	International Classification of Diseases 10th Revision
IS	Internal standard
$k_a$	Absorption rate constant
$k_e$	Elimination rate constant
OLETF rats	Otsuka Long-Evans Tokushima Fatty rats
OR	Odds ratio
SC	Stratum corneum
TEWL	Transepidermal water loss
$T_{max}$	Peak time
$Vd_{ss}$	Volume of distribution at the steady state

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40780-024-00402-5>.

Supplementary Material 1.

## Acknowledgements

The authors wish to thank Institute for Experimental Animals, Kanazawa University.

## Authors' contributions

SM, AF, TS, and YS designed the overall study and drafted the manuscript. SM, MT, YI, TK, and YH collected and analyzed the data in the retrospective study. SM, HM, KO, and TA collected and analyzed the data on skin parameters. SM, KT and YT collected and analyzed the data on intercellular lipids of SC. SM, AF, TS, and YS collected and analyzed the data in the pharmacokinetic study. YK contributed analytic tools. All authors contributed to the final version of the manuscript for submission.

## Funding

This study was supported in part by JSPS KAKENHI, Grant Numbers JP15K08092, JP20K07195, JP23K06209, JP23K06253 and JST SPRING, Grant Number JPMJSP2135.

## Data availability

The data analyzed in this study are available from the corresponding author on reasonable request.

## Declarations

### Ethics approval and consent to participate

This retrospective study was approved by the Medical Ethics Committee of Kanazawa University (2019–007 (3045)). All animal experiments were approved by the Committee on Animal Experimentation of Kanazawa University (AP-204140).

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.



**Author details**

<sup>1</sup>Department of Clinical Pharmacokinetics, Graduate School of Medical Sciences, Kanazawa University, 13-1 Takara-Machi, Kanazawa 920-8641, Japan. <sup>2</sup>Department of Hospital Pharmacy, University Hospital, Kanazawa University, 13-1 Takara-Machi, Kanazawa 920-8641, Japan. <sup>3</sup>Department of Biochemistry, School of Medicine, Tokyo Women's Medical University, Tokyo, Japan. <sup>4</sup>Department of Biological Science and Nursing, Graduate School of Medicine, Yokohama City University, Yokohama, Japan. <sup>5</sup>Department of Bio-Engineering Nursing, Graduate School of Nursing, Ishikawa Prefectural Nursing University, Kahoku, Japan. <sup>6</sup>The United Graduate School of Agricultural Sciences, Kagoshima University, Kagoshima, Japan. <sup>7</sup>Laboratory of Cosmetic Sciences, Institute of Ocean Energy, Saga University, Saga, Japan. <sup>8</sup>Department of Molecular Pharmacotherapeutics, Faculty of Pharmacy, Kanazawa University, Kanazawa, Japan. <sup>9</sup>Department of Clinical Pharmacy and Healthcare Science, Faculty of Pharmacy, Institute of Medical, Pharmaceutical and Health Science, Kanazawa University, Kanazawa, Japan. <sup>10</sup>AI Hospital/Macro Signal Dynamics Research and Development Center, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Kanazawa, Japan.

Received: 18 September 2024 Accepted: 2 December 2024

Published online: 18 December 2024

**References**

- Mercadante S, Bruera E. Opioid switching: a systematic and critical review. *Cancer Treat Rev*. 2006;32:304–15.
- Fallon M, Giusti R, Aielli F, Hoskin P, Rolke R, Sharma M, et al. Management of cancer pain in adult patients: ESMO Clinical Practice Guidelines. *Ann Oncol*. 2018;29:iv166–91.
- Paice JA, Ferrell B. The management of cancer pain. *CA Cancer J Clin*. 2011;61:157–82.
- Fiset P, Cohane C, Browne S, Brand SC, Shafer SL. Biopharmaceutics of a new transdermal fentanyl device. *Anesthesiology*. 1995;83:459–69.
- Solassol I, Caumette L, Bressolle F, Garcia F, Thézenas S, Astre C, et al. Inter- and intra-individual variability in transdermal fentanyl absorption in cancer pain patients. *Oncol Rep*. 2005;14:1029–36.
- Holdsworth MT, Forman WB, Killilea TA, Nystrom KM, Paul R, Brand SC, et al. Transdermal fentanyl disposition in elderly subjects. *Gerontology*. 1994;40:32–7.
- Heiskanen T, Mätzke S, Haakana S, Gergov M, Vuori E, Kalso E. Transdermal fentanyl in cachectic cancer patients. *Pain*. 2009;144:218–22.
- Nomura M, Inoue K, Matsushita S, Takahari D, Kondoh C, Shitara K, et al. Serum concentration of fentanyl during conversion from intravenous to transdermal administration to patients with chronic cancer pain. *Clin J Pain*. 2013;29:487–91.
- Barratt DT, Bandak B, Klepstad P, Dale O, Kaasa S, Christrup LL, et al. Genetic, pathological and physiological determinants of transdermal fentanyl pharmacokinetics in 620 cancer patients of the EPOS study. *Pharmacogenet Genomics*. 2014;24:185–94.
- Kuip EJM, Zandvliet ML, Koolen SLW, Mathijssen RHJ, van der Rijt CCD. A review of factors explaining variability in fentanyl pharmacokinetics; focus on implications for cancer patients. *Br J Clin Pharmacol*. 2017;83:294–313.
- Gourlay GK, Kowalski SR, Plummer JL, Cousins MJ, Armstrong PJ. Fentanyl blood concentration-analgesic response relationship in the treatment of postoperative pain. *Anesth Analg*. 1988;67:329–37.
- Nakamura A, Hasegawa M, Ito H, Minami K, Koike K, Habu-Tomita N, et al. Distinct relations among plasma concentrations required for different pharmacological effects in oxycodone, morphine, and fentanyl. *J Pain Palliat Care Pharmacother*. 2011;25:318–34.
- McClain DA, Hug CC. Intravenous fentanyl kinetics. *Clin Pharmacol Ther*. 1980;28:106–14.
- Kokubun H, Ebinuma K, Matoba M, Takayanagi R, Yamada Y, Yago K. Population pharmacokinetics of transdermal fentanyl in patients with cancer-related pain. *J Pain Palliat Care Pharmacother*. 2012;26:98–104.
- Lane ME. The transdermal delivery of fentanyl. *Eur J Pharm Biopharm*. 2013;84:449–55.
- Bista SR, Haywood A, Hardy J, Norris R, Hennig S. Exposure to Fentanyl After Transdermal Patch Administration for Cancer Pain Management. *J Clin Pharmacol*. 2016;56:705–13.
- Peng PWH, Sandler AN. A Review of the Use of Fentanyl Analgesia in the Management of Acute Pain in Adults. *Anesthesiology*. 1999;90:576–99.
- Kalia YN, Guy RH. Modeling transdermal drug release. *Adv Drug Deliv Rev*. 2001;48:159–72.
- van Smeden J, Janssens M, Gooris GS, Bouwstra JA. The important role of stratum corneum lipids for the cutaneous barrier function. *Biochim Biophys Acta Mol Cell Biol Lipids*. 2014;1841:295–313.
- Koroukian SM, Murray P, Madigan E. Comorbidity, disability, and geriatric syndromes in elderly cancer patients receiving home health care. *J Clin Oncol*. 2006;24:2304–10.
- Williams GR, Mackenzie A, Magnuson A, Olin R, Chapman A, Mohile S, et al. Comorbidity in older adults with cancer. *J Geriatr Oncol*. 2016;7:249–57.
- Jalving M, Gietema JA, Lefrandt JD, de Jong S, Reyners AKL, Gans ROB, et al. Metformin: Taking away the candy for cancer? *Eur J Cancer*. 2010;46:2369–80.
- Giovannucci E, Harlan DM, Archer MC, Bergenstal RM, Gapstur SM, Habel LA, et al. Diabetes and Cancer: A Consensus Report. *CA Cancer J Clin*. 2010;60:207–21.
- Quondamatteo F. Skin and diabetes mellitus: What do we know? *Cell Tissue Res*. 2014;355:1–21.
- Falanga V. Wound healing and its impairment in the diabetic foot. *Lancet*. 2005;366:1736–43.
- Park H-Y, Kim J-H, Jung M, Chung CH, Hasham R, Park CS, et al. A long-standing hyperglycaemic condition impairs skin barrier by accelerating skin ageing process. *Exp Dermatol*. 2011;20:969–74.
- Kim J, Yoon NY, Kim DH, Jung M, Jun M, Park H, et al. Impaired permeability and antimicrobial barriers in type 2 diabetes skin are linked to increased serum levels of advanced glycation end-product. *Exp Dermatol*. 2018;27:815–23.
- Portenoy RK, Southam MA, Gupta SK, Lapin J, Layman M, Inturrisi CE, et al. Transdermal fentanyl for cancer pain. Repeated dose pharmacokinetics. *Anesthesiology*. 1993;78:36–43.
- Goto Y, Kakizaki M, Masaki N. Production of spontaneous diabetic rats by repetition of selective breeding. *Tohoku J Exp Med*. 1976;119:85–90.
- Tsukui K, Suzuki M, Amma M, Tokudome Y. Ionic composition of Shotokuseki extract alters cell differentiation and lipid metabolism in three-dimensional cultured human epidermis. *Cytotechnology*. 2024;76:279–90.
- Ishida T, Naito T, Sato H, Kawakami J. Relationship between the plasma fentanyl and serum 4 $\beta$ -hydroxycholesterol based on CYP3A5 genotype and gender in patients with cancer pain. *Drug Metab Pharmacokinet*. 2016;31:242–8.
- Hisaka A, Sugiyama Y. Analysis of nonlinear and nonsteady state hepatic extraction with the dispersion model using the finite difference method. *J Pharmacokinet Biopharm*. 1998;26:495–519.
- Man MQ, Wakefield JS, Mauro TM, Elias PM. Alterations in epidermal function in type 2 diabetes: Implications for the management of this disease. *J Diabetes*. 2022;14:586–95.
- Fainsinger RL, Fairchild A, Nekolaichuk C, Lawlor P, Lowe S, Hanson J. Is Pain Intensity a Predictor of the Complexity of Cancer Pain Management? *J Clin Oncol*. 2009;27:585–90.
- Feldman EL, Callaghan BC, Pop-Busui R, Zochodne DW, Wright DE, Bennett DL, et al. Diabetic neuropathy. *Nat Rev Dis Primers*. 2019;5:41.
- Imokawa G, Abe A, Jin K, Higaki Y, Kawashima M, Hidano A. Decreased level of ceramides in stratum corneum of atopic dermatitis: an etiologic factor in atopic dry skin? *J Invest Dermatol*. 1991;96:523–6.
- Uche LE, Gooris GS, Beddoes CM, Bouwstra JA. New insight into phase behavior and permeability of skin lipid models based on sphingosine and phytosphingosine ceramides. *Biochim Biophys Acta Biomembr*. 2019;1861:1317–28.
- Uche LE, Gooris GS, Bouwstra JA, Beddoes CM. Barrier Capability of Skin Lipid Models: Effect of Ceramides and Free Fatty Acid Composition. *Langmuir*. 2019;35:15376–88.
- Marier J-F, Lor M, Potvin D, Dimarco M, Morelli G, Saedder EA. Pharmacokinetics, tolerability, and performance of a novel matrix transdermal delivery system of fentanyl relative to the commercially available reservoir formulation in healthy subjects. *J Clin Pharmacol*. 2006;46:642–53.
- Marier J-F, Lor M, Morin J, Roux L, Di Marco M, Morelli G, et al. Comparative bioequivalence study between a novel matrix transdermal delivery

system of fentanyl and a commercially available reservoir formulation. *Br J Clin Pharmacol*. 2007;63:121–4.

41. Kress HG, Boss H, Delvin T, Lahu G, Lophaven S, Marx M, et al. Transdermal fentanyl matrix patches Matrifen and Durogesic DTrans are bioequivalent. *Eur J Pharm Biopharm*. 2010;75:225–31.
42. Walter K, Kurz H. Binding of drugs to human skin: influencing factors and the role of tissue lipids. *J Pharm Pharmacol*. 1988;40:689–93.
43. Yourick JJ, Koenig ML, Yourick DL, Bronaugh RL. Fate of chemicals in skin after dermal application: does the in vitro skin reservoir affect the estimate of systemic absorption? *Toxicol Appl Pharmacol*. 2004;195:309–20.
44. van Eijl S, Zhu Z, Cupitt J, Gierula M, Götz C, Fritsche E, et al. Elucidation of xenobiotic metabolism pathways in human skin and human skin models by proteomic profiling. *PLoS One*. 2012;7:e41721.
45. Dumont C, Prieto P, Asturiol D, Worth A. Review of the availability of in vitro and in silico methods for assessing dermal bioavailability. *Appl In Vitro Toxicol*. 2015;1:147–64. Available from: <http://www.liebertpub.com/doi/10.1089/aivt.2015.0003>.
46. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res Clin Pract*. 2019;157:107843.
47. de Macedo GMC, Nunes S, Barreto T. Skin disorders in diabetes mellitus: an epidemiology and physiopathology review. *Diabetol Metab Syndr*. 2016;8:63.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

# Virtual imaging trials in medicine: A brief takeaway of the lessons from the first international summit

Ehsan Samei<sup>1</sup> | Ehsan Abadi<sup>1</sup> | Predrag Bakic<sup>2</sup> | Kristina Bliznakova<sup>3</sup> |  
 Hilde Bosmans<sup>4</sup> | Ann-Katherine Carton<sup>5</sup> | Alejandro F. Frangi<sup>6</sup> | Stephen Glick<sup>7</sup> |  
 Joseph Y. Lo<sup>1</sup> | Paul Kinahan<sup>8</sup> | Andrew Maidment<sup>9</sup> | Francesco Ria<sup>1</sup> |  
 Ioannis Sechopoulos<sup>10</sup> | William Paul Segars<sup>1</sup> | Rie Tanaka<sup>11</sup> |  
 Liesbeth Vancoillie<sup>1</sup>

<sup>1</sup>CVIT Radiology, Duke University, Durham, North Carolina, USA

<sup>2</sup>Radiology, Lund University, Malmö, Sweden

<sup>3</sup>Biomedical Engineering, Medical University of Bulgaria, Varna, Bulgaria

<sup>4</sup>Radiology, KU Leuven, Leuven, Belgium

<sup>5</sup>GE Healthcare, Buc, France

<sup>6</sup>Computational Medicine, University of Manchester, Manchester, UK

<sup>7</sup>FDA, Silver Spring, Maryland, USA

<sup>8</sup>Bioengineering, University of Washington, Seattle, Washington, USA

<sup>9</sup>Radiology, University of Pennsylvania, Philadelphia, Pennsylvania, USA

<sup>10</sup>Medical Imaging, Radboud University, Nijmegen, The Netherlands

<sup>11</sup>Radiological Technology, Kanazawa University, Kanazawa, Ishikawa, Japan

## Correspondence

Liesbeth Vancoillie, CVIT Radiology, Duke University, Durham, NC 27705, USA.  
 Email: [liesbeth.vancoillie@duke.edu](mailto:liesbeth.vancoillie@duke.edu)

## Abstract

**Background:** The rapid advancement of medical technologies presents significant challenges for researchers and practitioners. While traditional clinical trials remain the gold standard, they are often limited by high costs, lengthy durations, and ethical constraints. In contrast, in-silico trials and digital twins have emerged not only as efficient and ethical alternatives but also as a complementary technology that can extend beyond classical trials to predict and design new strategies. The successful application of digital twins in industries like nuclear energy, automotive engineering, and aviation underscores their potential in human health.

**Methods:** In April 2024, Duke University hosted the first international summit on Virtual Imaging Trials in Medicine (VITM). The summit brought together over 130 experts from academia, industry, and regulatory bodies to discuss the latest developments, challenges, and future directions in this field. The event featured plenary speakers, presentations, and panel discussions, emphasizing the integration of clinical and in-silico methods to enhance medical evaluations.

**Results:** Key takeaways included the necessity of diverse and realistic digital patient representations, the integration of physics and biology in simulations, and the development of robust validation frameworks. The summit also highlighted the importance of regulatory science and the establishment of Good Simulation Practices to ensure the credibility and reliability of virtual trials.

**Conclusion:** The key discussions and insights from the VITM summit underscore the potential of in-silico trials to revolutionize medical research and patient care through personalized, efficient, and ethical evaluation methods. The collaborative efforts and recommendations from this summit aim to drive future advancements in virtual imaging trials in medicine.

## KEYWORDS

digital twins, in-silico trials, virtual imaging

## 1 | INTRODUCTION

Medical imaging has become a cornerstone of medical practice and research. In a 2001 survey, leading general internists ranked computed tomography (CT)

and magnetic resonance imaging (MRI) as the top two technology innovations and advances that had the greatest impact on the practice of medicine. Mammography alone was ranked top 5.<sup>1</sup> These successes have led to further applications and technology innovations

**TABLE 1** Terminology of in-silico, virtual, and digital twin trials.

Term	Definition
Virtual Clinical Trials	Representational methods that emulate clinical reality
Virtual Imaging Trials	Representational methods that emulate imaging reality
<i>In-Silico</i> Clinical Trials	Representational computational methods that emulate clinical reality
Medical Digital Twin	Representational computational methods with a personalized focus

providing an even greater value to healthcare.<sup>2</sup> A natural consequence of this rapid advancement is the increasing variability and complexity of technologies, which has created a challenge in evaluating and optimizing their design and clinical utility. This challenge spans scientific research, product development, and clinical practice. The ultimate evaluation of new innovations is through clinical trials, yet such trials are not practical in most scenarios given their ethical constraints, high costs, lengthy durations, difficulties in recruiting sufficient subjects, and the absence of a definitive ground truth. The rapid advancement of Artificial Intelligence (AI) in healthcare has further challenged our medical evaluation processes to ensure transparency, generalizability, and interpretability.

In recent years, in-silico trials and digital twins have emerged as an alternative approach for the evaluation and optimization of technological innovations.<sup>3</sup> This approach has been highlighted in major reports from the National Academies and the Food and Drug Administration (FDA),<sup>4,5</sup> which point out to its successful application in industries such as nuclear energy, automotive engineering, and aviation—underscoring their potential for human health. The pioneering and expected applications in medicine have indeed been subject of much interest, reflected in public periodicals as well.<sup>6</sup> Among those applications, the in-silico approach further offers solutions to AI-related challenges of access to unambiguous ground truth, especially for rare diseases, and interpretability.

The terms in-silico, virtual, and digital twin are often interchangeably used. Table 1 offers a summary of the terms. But invariably they all include realistic representation of individuals or populations of varying attributes, detailed models of medical technologies and applications, and computational representations of outcome assessments. In-silico trials address key challenges in medical evaluation such as prohibitive cost, time, and the ethics of experimentation, offering an efficient mechanism to test interventions before their clinical application. They further provide innovative ways to personalize, quantify, and optimize healthcare. Yet, as a new methodological approach, the prospect of in-silico trials requires the scientific community to collegially

achieve consensus for their responsible, qualified, and effectual implementation.

To that end, in April 2024, Duke University hosted the first international summit of Virtual Imaging Trials in Medicine (VITM). Organized by a committee of leading scientists of the discipline around the globe, the conference brought together experts, researchers, and practitioners in medical imaging and therapy utilizing in-silico trials and digital twins. This first-of-its-kind event aimed to consolidate the latest developments, summarize the current status, and envision future prospects for in-silico virtual trials and digital twins in medicine. The 2 and a half days of the summit featured plenary speakers, proffered presentations, and perspectives from industry and regulatory bodies. Well-attended by over 130 participants, comprising 74.8% from academia, 17.6% from industry, and 7.6% from regulatory and governmental bodies with attendees from North America, Europe, Asia, and Oceania, the summit served as a forum to discuss methods, opportunities, challenges, limitations, and future directions of virtual trials in medicine.

This white paper provides a summary of the discussion points and consensus outcome of the VITM summit. The paper is organized into four themes that were a major focus of the event, aiming to inform the current status and potential trajectory of virtual imaging trials as an emerging discipline in medicine. The paper also includes an Appendix offering few broad-category use cases featured at the summit.

### 1.1 | Theme 1—real plus virtual: maximizing the complementary role of clinical and in-silico methods in medical evaluation

Proficient and efficient evaluation of medical products is a crucial need in medicine. The requisite proficiency and efficiency can be enabled by integration of diverse clinical trial methodologies, focusing on optimizing the complementary aspects of real and virtual methods. Common among all types of trials is the need and challenge of ensuring the diversity of studied population, addressing various diseases including those that are rare, and modeling the complexity of the research design and avoidance of bias. Within this landscape, real, virtual, and hybrid clinical trials offer comparable advantages and disadvantages summarized in Table 2.

A key consideration for incorporating a virtual approach as a trial tool is its level of reliability and trustworthiness. These can be enhanced by combining the virtual approach with complementary physics experiments. This aligns with the emphasis of one of the plenary speakers of this event, Dr. Mitchel Schnall of the University of Pennsylvania, who highlighted the goal of having quality metrics that closely reflect clinical outcomes, and in the same vein, clinical outcomes



**TABLE 2** Advantages and disadvantages of real, virtual, and real-virtual hybrid trials.

Type	Advantages	Disadvantages
Clinical	<ul style="list-style-type: none"> <li>• Closely (but not exactly) reflecting clinical practice</li> <li>• Providing definitive clinical endpoints</li> </ul>	<ul style="list-style-type: none"> <li>• Slow, costly nature, large sample size requirement (impossible for rare pathologies)</li> <li>• Ethics of repeated, serial exams, or potentially harmful interventions</li> <li>• Confluence of confounding factors</li> </ul>
Virtual	<ul style="list-style-type: none"> <li>• Subject safety (e.g., no radiation exposure)</li> <li>• Longitudinal monitoring capabilities</li> <li>• Rare disease study capabilities</li> <li>• “Clinical” testing of new concepts</li> <li>• Reduction in animal testing</li> <li>• Providing quantitative endpoints</li> </ul>	<ul style="list-style-type: none"> <li>• Needed validation for realism and credibility</li> <li>• Complexity of implementation by users</li> <li>• Current lack of practical robust offerings</li> </ul>
Hybrid	<ul style="list-style-type: none"> <li>• Potential balance between efficiency (of virtual) and accuracy (of clinical) methods</li> </ul>	<ul style="list-style-type: none"> <li>• Needed validation for realism and credibility</li> <li>• Limited experience and integration strategy for virtual and clinical trials</li> </ul>

that should be expected and meaningful for virtual trials. Such a goal, while worthwhile, is not absolute: striving for perfection can be counterproductive, as *perfect is often the enemy of good*.

Recognizing virtual assessment as a component of technology evaluation, the FDA has instituted Medical Device Development Tools (MDDTs) as non-clinical evaluations of medical devices.<sup>7</sup> Future strategies should facilitate uniform testing across imaging modalities and realistically optimized associated parameters through simulations followed by clinical validation. In that effort, the focus should extend from specific parameters to quantitative biomarkers to understand how simulation processes and accuracy impact imaging endpoints. Imaging and clinical endpoints are certainly not necessarily identical either. They are distinct measures in medical research, each providing different information about treatment efficacy and patient outcomes. Understanding these distinctions is crucial for ensuring that research findings are relevant and applicable to broad patient populations.

## 1.2 | Theme 2—Physics plus biology: the intersection of physical and biological simulations impacts our understanding of metabolism, treatment, and disease progression

Real trials involve the entire complexity of reality, but they can only reveal a part of it due to practical limitations. Simulations focus on a limited aspect of reality, but within that scope, they can offer more detailed insights. These distinctions are important for understanding the strengths and limitations of different research approaches and methodologies.

Virtual trials have historically had separate development of biological and physical aspects, and the challenges of translating and combining the two. Physics and biology take different approaches to simulation, each with their own assets and limits. Physical mod-

els often involve well-defined laws and equations that can predict outcomes with high precision. In contrast, modeling biological systems is more complex due to high variability, numerous interacting components, and the influence of countless internal and external factors. In imaging trials, it is more straightforward to model the physical aspects of imaging trials, but not so the biological aspects. This complexity makes it challenging to create simple models in biology that are accurate and predictive. However, this task can be a bit more manageable when the variabilities are captured for individual patients, as opposed to aggregates across patients which introduce additional sources of variability.

This perspective was highlighted by another plenary speaker of the summit, David Hormuth from the University of Texas at Austin, in utilizing digital twins to tailor treatments to the unique characteristics and requirements of an individual. While clinical trials typically address the general population, evaluating utility across a population might disregard individual experiences. An “average” patient represents no single individual within a cohort of individuals. Consequently, although trials provide insights into broader trends for a given technology or intervention across a population, they naturally fall short of fully capturing what might be optimum for an individual patient.

Looking ahead, therapy and healthcare are becoming more personalized and responsive. Concepts like prospective adaptive therapy and response prediction are driving this shift. Additionally, individualized surveillance decisions are seeking to ensure that monitoring strategies are customized for each person. These advancements promise more effective and precise care, better meeting the diverse needs of patients in the future. Towards that goal, there is need for the integration of physics- and biology-based models, as they not only invoke complementary mindsets, but together, they cover a broader spectrum of clinical reality and offer different dimensions of phenotypes that are instrumental to personalized care.

Of course, as in any modeling effort, isolated or integrated, there is a need for a structured approach to validation of the models and processes. As integrated models are particularly complex, validation should begin with simpler, more fundamental aspects, and progressively move to more complex and advanced components. This progressive approach ensures that each level of complexity is thoroughly tested and verified before moving on to the next, more advanced level. This approach helps ensure that each stage is solidly built upon a verified foundation, reducing the risk of overlooking critical issues as complexity increases.

### 1.3 | Theme 3—virtual meets diversity: strategies to generate digital patient representations with enhanced diversity

Representing diversity is a key requirement of any clinical trial—recognizing that evaluation, claims, or solutions are only valid if generalized across a population or a specific subgroup. This need for diversity is also applicable to practitioners of medical interventions including image interpreters. While current imaging practice often assumes an “average” interpreter or observer, no single observer represents such a hypothetical average interpreter. Thus, there is a need to include diverse observers and incorporating their variability, which can be dataset-dependent, to assure generalizability.

This highlights the significance of virtual trials incorporating diversity in their constructs. In representing diversity, virtual trials provide unique opportunities. Virtual trials offer the capacity to represent a diverse patient population with fewer virtual patients compared to physical clinical trials. Instead of replicating the sampling of the data of clinical trials, the virtual trial can deploy differing distribution of configurations, including both uniform and skewed sampling towards “less frequent, but different in characteristics” cases (so-called edge cases), thus enabling a simpler and targeted trial design. An illustrative application of diversity in virtual trials involves simulating rare diseases across diverse patient cohorts to clarify disease responses and evaluate simulation boundaries. In the trial design for rare diseases where large-scale trials are impractical, Bayesian approaches are particularly beneficial. Such designs can provide evidence of efficacy and safety by integrating virtual and real data.

Recognizing the importance of diversity in trials highlights the need to create a “science of diversity” to understand the complex factors that influence imaging examinations and trial outcomes. Often diversity is targeted by using large datasets, assuming that higher numbers translate to higher diversity. Typically, diversity is characterized by generic attributes such as age, sex, race, etc. However, these broad representations rarely capture all attributes that influence the outcome.

There is a need for a systematic of diversity at multiple scales. Virtual trials offer a controlled way to do so, with its inherent unique advantage of defining and controlling for the ground truth of the subjects and the interventions. Such a control also offers an opportunity to consider matching the data and interpretations so as to optimize or fine-tune the intervention for specific subpopulations. This is a unique asset to balance generalizability (ensured by population statistics) with personalized care, recognizing that while models must account for diverse populations, individual patient needs should not be overlooked.

Addressing diversity gaps in clinical trials requires a multi-layered approach. Options include generating numerous digital phantoms based on real data, incorporating aging models, manual modifications using 3D modeling, leveraging deep learning methods such as generative AI, developing a platform to create diversity and exploring other strategies. While digital phantoms offer flexibility in simulating diverse scenarios, they may still follow a distribution, limiting their effectiveness in truly addressing diversity. However, incorporating interpolation techniques can mitigate this limitation. Diversity in virtual trials is also crucial when dealing with under-represented groups or when real data collection is challenging. It is essential to recognize biases and actively pursue inclusivity in trial design and data representation, ensuring that virtual trials reflect the diverse global population accurately.

### 1.4 | Theme 4—virtual meets reality: overcoming barriers to accessibility and widespread implementation

To take advantage of what virtual trials can offer, we need to have reliable and easy to use tools and resources. These resources are currently not where they need to be. There are various challenges and opportunities towards such readiness and confidence. Those include factors related to the validation and realism of simulation methods, transparency associated with “black box” imaging tools such as vendor-specific post processing algorithms including AI systems (which incidentally in-silico models can tackle), managing the open-source versus proprietary nature of some tools, and overcoming the lack of broad standards for trial design and execution. There is a significant need for openness, transparency, and collaboration to address these challenges and fully capitalize on the opportunities within medical imaging research to synergistically integrate experimental and modeling aspects to enhance healthcare outcomes.

This level of confidence is essential, especially when virtual trials are planned to address the needs in regulatory science. As Alejandro Frangi of the University of Manchester, another plenary speaker emphasized, the

regulatory evidence has a pivotal role as the cornerstone of scientific validation, going towards controlled trials based on credible scientific foundations, prioritizing simulation to adhere to the principle of “first do no harm.” This point was also emphasized by Tina Morrison of the FDA: the FDA, as a large multi-component entity devoted to regulatory science, plays a crucial role in shaping the landscape of virtual and in-silico trials. She emphasized the distinctions and interconnections between basic science, regulatory science, and translational science. This framework supports the development and evaluation of products, post-market surveillance, and public health preparedness through multiple programs, including the Centers of Excellence in Regulatory Science and Innovation (CERSI) program.<sup>8</sup> Pertinent to virtual trials is the establishment and the integration with the current regulatory scenario of Good Simulation Practices (GSP), which build upon other established Good Practices, setting the level of needed details and credibility assessment guidelines.<sup>9</sup>

In terms of availability of tools and resources, current offerings range from independently developed tools to FDA-provided and FDA-approved programs.<sup>10–13</sup> Those include applications with Monte Carlo packages like PENELOPE<sup>14</sup> and GEANT4,<sup>15</sup> demonstrating a balance between open-source software and self-developed technologies. Key factors to consider include accessibility, quality, and regulatory compliance. There is a dynamic tension between making open-source software widely accessible, ensuring high quality, and meeting regulatory requirements. Balancing these factors is challenging because increasing one aspect (e.g., accessibility) might impact the others (e.g., quality or regulatory compliance). There is also a risk of overregulation.

Currently, there exists a prevalent uncertainty within the community regarding the reliability of virtual imaging trials which underscores the need for improvement in robust validation of simulation technologies. Further, there is a deficiency in crucial biological data which hinders the accuracy and reliability of virtual trials, contributing to widespread skepticism regarding their predictive capabilities. To build trustworthiness encompassing credibility, reproducibility, and accessibility, it is essential to address these data gaps, thereby enhancing the validity of virtual trials and fostering greater confidence in their outcomes. Towards that goal, several actions may be considered:

1. Dedicate efforts in the use of FDA guidelines for Medical Device Development Tools.
2. Develop a quality standard for simulation studies, similar to the AAPM Task Group 268 report.<sup>16</sup>
3. Consider establishing certification standards or good simulation practices to increase trust in in-silico results, for example, when submitted for regulatory review.

4. Explore a neutral third-party validation mechanism to validate proprietary simulation aspects from manufacturers, ensuring protection of intellectual property while encouraging collaboration.
5. Require manufacturers to provide research access to reconstruction algorithms and data through contractual agreements or incentives from regulators.
6. Continuously monitor clinical reports for changes in disease detection correlated with imaging protocols to help validate virtual results.
7. Establish a community of stakeholders (e.g., The Regulatory Science In-Silico X Collaborator Community—ISXCC) to coordinate and facilitate efforts across stakeholders.
8. Clarify designations and terminology within this space, which can offer consistent and clear definition of terms such as digital twins, in silico, and virtual.
9. Develop stages for granularity and quality of digital data.
10. Ensure the context of use for an in-silico tool is always noted to ensure relevance to regulatory practices and beyond.

By implementing these action items, the medical imaging community can improve the credibility, reproducibility, and accessibility of virtual trials, thereby enhancing their overall trustworthiness and effectiveness. Equally important is the need to effectively incorporate the perspectives of the patient and the clinical providers into the prospects of using virtual trials for evidence generation, including AI-driven medical research. Imaging clinicians can provide insights on image interpretation processes that are important to simulate, and current limitations with imaging methods that motivate further optimization. Patients are the ultimate recipient of the “good” that virtual trials would offer and thus their voice and agency should be respected. They deserve a better trial, a better product, and better care.

## 2 | CONCLUSIONS

The Virtual Imaging Trails in Medicine summit concluded with a powerful call to action to address emerging challenges and fully harness the potential of in-silico virtual trials and digital twins in healthcare. Opening remarks underscored the limitations of traditional clinical trials and advocated for the adoption of innovative methodologies. Contributions from plenary speakers, industry experts, and regulatory bodies highlighted the crucial role of regulatory science in steering this transformation. Collaboration emerged as a central theme for achieving progress. The summit underscored the necessity of transparency, collaboration, and inclusivity to unlock the full potential of in-silico trials and digital twins, aiming to revolutionize patient care on a global scale.



## ACKNOWLEDGMENTS

The summit was sponsored by Center for Virtual Imaging Trials (CVIT), Duke University, GE Healthcare US, GE Mammo France, Medical Physics Graduate Program of Duke University, Metis Health Analytics, LLC, Siemens Healthineers US, Triangle Center of Excellence in Regulatory Science and Innovation, and University of Manchester. Affiliates included American Association of Physicists in Medicine, European Federation of Organisations for Medical Physics, In Silico UK, and International Organization for Medical Physics.

## CONFLICT OF INTEREST STATEMENT

This report was prepared by the VITM summit directors and the organizing committee as a factual general consensus of the discussions that took place at the summit. The statements do not necessarily represent the views of all workshop participants, the summit directors, or the organizing committee.

## REFERENCES

- Fuchs VR, Sox HCJ. Physicians' views of the relative importance of thirty medical innovations. *Health Affairs*. 2001;20(5):30-42.
- Bercovich E, Javitt MC. Medical Imaging: From Roentgen to the Digital Revolution, and Beyond. *Rambam Maimonides Med J*. 2018;9(4):e0034. <https://doi.org/10.5041/RMMJ.10355>
- Abadi E, Segars WP, Tsui BMW, et al. Virtual clinical trials in medical imaging: a review. *J Med Imaging (Bellingham)*. 2020;7(4):042805. <https://doi.org/10.1117/1.JMI.7.4.042805>. Epub 2020 Apr 11.
- Food US and Administration D. Credibility of Computational Models Program: Research on Computational Models and Simulation Associated with Medical Devices; 2024. <https://www.fda.gov/medical-devices/medical-device-regulatory-science-research-programs-conducted-osel/credibility-computational-models-program-research-computational-models-and-simulation-associated>
- U.S. Food and Drug Administration. Modeling & Simulation at FDA; 2024. <https://www.fda.gov/science-research/about-science-research-fda/modeling-simulation-fda>
- The Guardian. How digital twins may enable personalised health treatment; 2024. <https://www.theguardian.com/science/2023/nov/12/digital-twin-personalised-medical-treatment>
- U.S. Food and Drug Administration. Medical Device Development Tools (MDDT); 2004. [https://www.fda.gov/medical-devices/medical-device-development-tools-mddt#:~:text=MDDT%20Categories,-The%20FDA%20recognizes&text=Non%2Dclinical%20Assessment%20Model%20\(NAM,%2C%20effectiveness%2C%20or%20device%20performance](https://www.fda.gov/medical-devices/medical-device-development-tools-mddt#:~:text=MDDT%20Categories,-The%20FDA%20recognizes&text=Non%2Dclinical%20Assessment%20Model%20(NAM,%2C%20effectiveness%2C%20or%20device%20performance)
- U.S. Food and Drug Administration. Centers of Excellence in Regulatory Science and Innovation (CERSIs); 2024. <https://www.fda.gov/science-research/advancing-regulatory-science/centers-excellence-regulatory-science-and-innovation-cersis#:~:text=The%20CERSIs%20promote%20innovation%20in,awards%2C%20of fellowships%2C%20and%20competitions>
- U.S. Food and Drug Administration. Good Clinical Practice; 2024. <https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/good-clinical-practice#:~:text=FDA%20regulates%20scientific%20studies%20that,biological%20products%2C%20and%20medical%20devices>
- U.S. Food and Drug Administration. VICTRE: In Silico Breast Imaging Pipeline; 2024. <https://cdhr-rst.fda.gov/victre-silico-breast-imaging-pipeline>
- University D. Center for Virtual Imaging Trials; 2024. <https://cvit.duke.edu/>
- Barufaldi B, Maidment ADA, Dustler M, et al. Virtual clinical trials in medical imaging system evaluation and optimisation. *Radiat Prot Dosimetry*. 2021;195(3-4):363-371.
- Wu M, FitzGerald P, Zhang J, Segars WP, Yu H, Xu Y, De Man B. XCIST—an open access x-ray/CT simulation toolkit. *Phys Med Biol*. 2022;67(19). <https://doi.org/10.1088/1361-6560/ac9174>
- Sempau J, Fernández-Varea JM, Acosta E, Salvat F. Experimental benchmarks of the Monte Carlo code penelope. *Nucl Instrum Methods Phys Res B*. 2003;207(2):107-123. [https://doi.org/10.1016/S0168-583X\(03\)00453-1](https://doi.org/10.1016/S0168-583X(03)00453-1)
- Agostinelli S, Allison J, Amako K, et al. Geant4—a simulation toolkit. *Nucl Instrum Methods Phys Res A*. 2003;506(3):250-303. [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8)
- Sechopoulos I, Rogers DWO, Bazalova-Carter M, et al. RECORDS: improved Reporting of monte Carlo RaDiation transport Studies: report of the AAPM Research Committee Task Group 268. *Med Phys*. 2018;45(1):e1-e5.
- Barufaldi B, Higginbotham D, Bakic PR, Maidment ADA. Open-VCT: a GPU-accelerated virtual clinical trial pipeline for mammography and digital breast tomosynthesis. *Proceedings Volume 10573, Medical Imaging 2018: Physics of Medical Imaging*; 1057358 (2018). <https://doi.org/10.1117/12.2294935>
- Rafferty EA, Park JM, Philpotts LE, et al. Assessing radiologist performance using combined digital mammography and breast tomosynthesis compared with digital mammography alone: results of a multicenter, multireader trial. *Radiology*. 2013;266(1):104-113.
- Tomic H, Costa AC, Bjerken A, et al. Simulation of breast lesions based upon fractal Perlin noise. *Phys Med*. 2023;114:102681.
- Marshall NW, Bosmans H. Performance evaluation of digital breast tomosynthesis systems: comparison of current virtual clinical trial methods. *Phys Med Biol*. 2022;67(22):1-31.
- Sauer TJ, Buckler AJ, Abadi E, et al. Development of physiologically-informed computational coronary artery plaques for use in virtual imaging trials. *Med Phys*. 2024;51(3):1583-1596.
- Segars WP, Sturgeon G, Mendonca S, Grimes J, Tsui BMW. 4D XCAT phantom for multimodality imaging research. *Med Phys*. 2010;37(9):4902-4915.
- Abadi E, Harrawood B, Sharma S, Kapadia A, Segars WP, Samei E. DukeSim: A Realistic, Rapid, and Scanner-Specific Simulation Framework in Computed Tomography. *IEEE Trans Med Imaging*. 2019;38(6):1457-1465. <https://doi.org/10.1109/TMI.2018.2886530>. Epub 2018 Dec 12.
- Baraghoshi D, Strand M, Humphries SM, Estepar SJR, Sanchez-Ferrero VG, Charbonnier JP. Quantitative CT evaluation of emphysema progression over 10 years in the COPD Gene study. *Radiology*. 2023;307(4):e222786.
- Humphries SM, Mackintosh JA, Jo HE, Walsh SL, Silva M, Calandriello L. Quantitative computed tomography predicts outcomes in idiopathic pulmonary fibrosis. *Respirology*. 2022;27(12):1045-1053.
- Abadi E, Jadick G, Lynch DA, Segars WP, Samei E. Emphysema Quantifications With CT Scan: Assessing the Effects of Acquisition Protocols and Imaging Parameters Using Virtual Imaging Trials. *Chest*. 2023;163(5):1084-1100. <https://doi.org/10.1016/j.chest.2022.11.033>. Epub 2022 Dec 1.
- Abadi E, Sturgeon GM, Agasthya G, et al. Airways, vasculature, and interstitial tissue: anatomically informed computational modeling of human lungs for virtual clinical trials. *Proceedings Volume 10132, Biomedical Optics and Imaging 2017:1605-7422* (2017). <https://doi.org/10.1117/12.2254739>
- Ho FC, Sotoudeh-Paima S, Segars WP, Samei E, Abadi E. Development and Application of a Virtual Imaging Trial framework for Airway Quantifications via CT. *Proc SPIE Int Soc Opt Eng*. 2023;12463:124631B. <https://doi.org/10.1117/12.2654263>. Epub 2023 Apr 7.



29. McCabe C, Solomon J, Segars WP, Abadi E, Samei E. Synthesizing heterogeneous lung lesions for virtual imaging trials. Proceedings Volume 12925, Medical Imaging 2024: Physics of Medical Imaging; 129251N (2024). <https://doi.org/10.1117/12.3006199>
30. Sotoudeh-Paima S, Segars WP, Ghosh D, Luo S, Samei E, Abadi E. A systematic assessment and optimization of photon-counting CT for lung density quantifications. *Med Phys*. 2024;51(4):2893-2904.
31. Sotoudeh-Paima S, Samei E, Abadi E. CT-HARMONICA: physics-based CT harmonization for reliable lung density quantification. Proceedings Volume 12465, Medical Imaging 2023: Computer-Aided Diagnosis; 124651J (2023). <https://doi.org/10.1117/12.2654346>
32. Harrison R, Elston B, Doot RK, Lewellen T, Mankoff DA, Kinahan PE. A virtual clinical trial of FDG-PET imaging of breast cancer: effect of variability on response assessment. *Transl Oncol*. 2014;7:138-146.
33. Wangerin KA, Muzi M, Peterson LM, et al. A virtual clinical trial comparing static versus dynamic PET imaging in measuring response to breast cancer therapy. *Phys Med Biol*. 2017;62:3639-3655.
34. Zhang R, Alessio AM, Pierce LAI, et al. Improved attenuation correction for respiratory gated PET/CT with extended-duration cine CT: a simulation study. Proceedings of the SPIE, Volume 10132, id. 101321I 7 pp. 2017.
35. Li Z, Benabdallah N, Luo J, Wahi RL, Thorek DLI, Jha AK. ISIT-QA: in silico imaging trial to evaluate a low-count quantitative SPECT method across multiple scanner-collimator configurations for 223Ra-based radiopharmaceutical therapies. *J Nucl Med*. 2024;65:810-817.
36. Wangerin KA, Muzi M, Peterson LM, Linden HM, Novakova A, O'Sullivan F. Effect of 18F-FDG uptake time on lesion detectability in PET imaging of early stage breast cancer. *Tomography*. 2015;1:53-60.
37. Sarraimi-Foroushani A, Lassila T, MacRaid M, et al. In-silico trial of intracranial flow diverters replicates and expands insights from conventional clinical trials. *Nat Commun*. 2021;12(1):3561.
38. Liu Q, Sarraimi-Foroushani A, Wang Y, MacRaid M, Kelly C, Lin F. Hemodynamics of thrombus formation in intracranial aneurysms: an in silico observational study. *APL Bioeng*. 2023;7(7):036102.
39. Frangi AF, Denison T, Myles P, Ordish J, Brown P, Turpin R. *Unlocking the Power of Computational Modelling and Simulation Across the Product Lifecycle in Life Sciences: a UK Landscape Report*. InSilicoUK Pro-Innovation Regulations Network; 2023.

**How to cite this article:** Samei E, Abadi E, Bakic P, et al. Virtual imaging trials in medicine: A brief takeaway of the lessons from the first international summit. *Med Phys*. 2025;52:1950–1959. <https://doi.org/10.1002/mp.17587>

## APPENDIX: USE CASES OF IN-SILICO TRIALS

Many current and potential application of in-silico trials were highlighted at the summit. For brevity, here we highlight few such applications.

### Optimized use of imaging for breast cancer diagnosis

VCTs have been used to investigate and optimize DBT systems. Barufaldi et al.<sup>12</sup> used open software OpenVCT,

developed at the University of Pennsylvania<sup>17</sup> to evaluate detectability of breast lesions in DBT and DM. By simulating a clinical DM/DBT system, and using virtual ellipsoidal masses and single microcalcifications, VCTs closely reproduced (< 4% of the area under the ROC curve, see Figure A1) clinical results.<sup>18</sup> More recently, OpenVCT has been used to optimize DBT acquisition for improved detectability of microcalcifications, and in detailed modeling of lesions based upon the Perlin noise.<sup>19</sup> A comparison for virtual clinical trial methods for performance evaluation of DBT systems can be found in the paper of Marshall et al.<sup>20</sup>

### Specific evaluation of coronary artery disease (CAD)

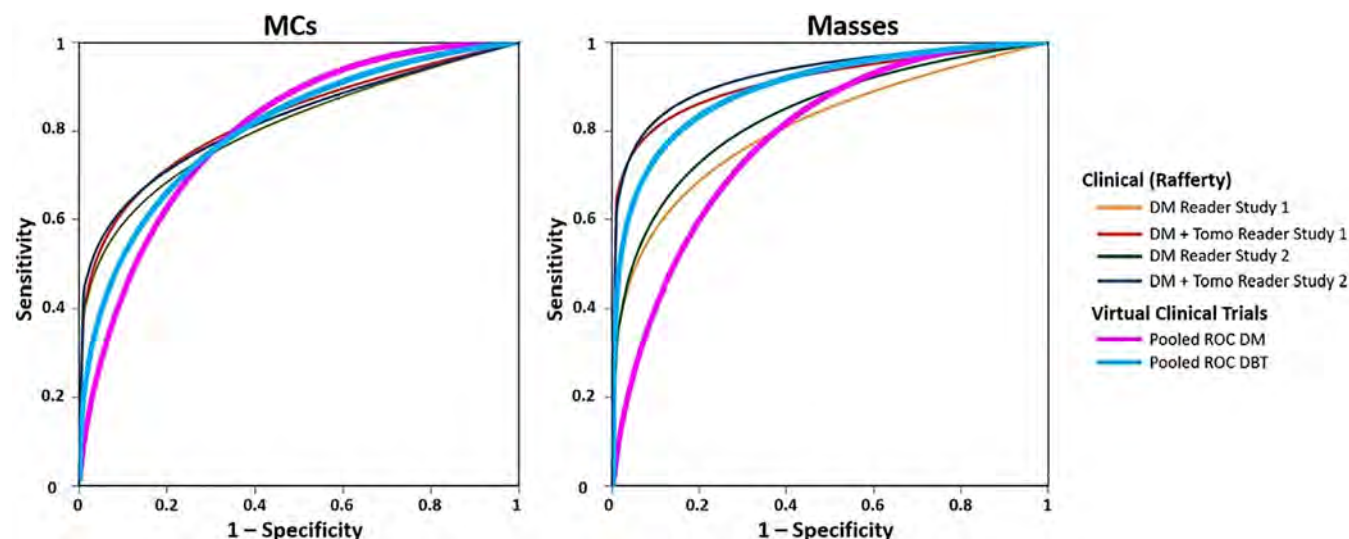
Heart disease is the leading cause of death worldwide with coronary artery disease (CAD) responsible for 10% of all deaths annually. VITs provide a safer, faster, less expensive means to evaluate, optimize, and compare cardiac imaging technologies and methods. In the example below, Sauer et al.<sup>21</sup> developed realistic, physiologically-informed models for coronary plaques based on histology data. The plaque models were combined with the whole-body XCAT computational human phantom<sup>22</sup> and DukeSim CT simulator<sup>23</sup> to perform an initial simulation comparing standard energy-integrating detector (EID) CT with the emerging technology of photon-counting detector (PCD) CT. The results, shown in Figure A2 for a plaque modeled within the right coronary artery (RCA), demonstrate the potential of PCD-CT to better resolve cardiac plaques and their individual components.

### Precise lung disease quantification

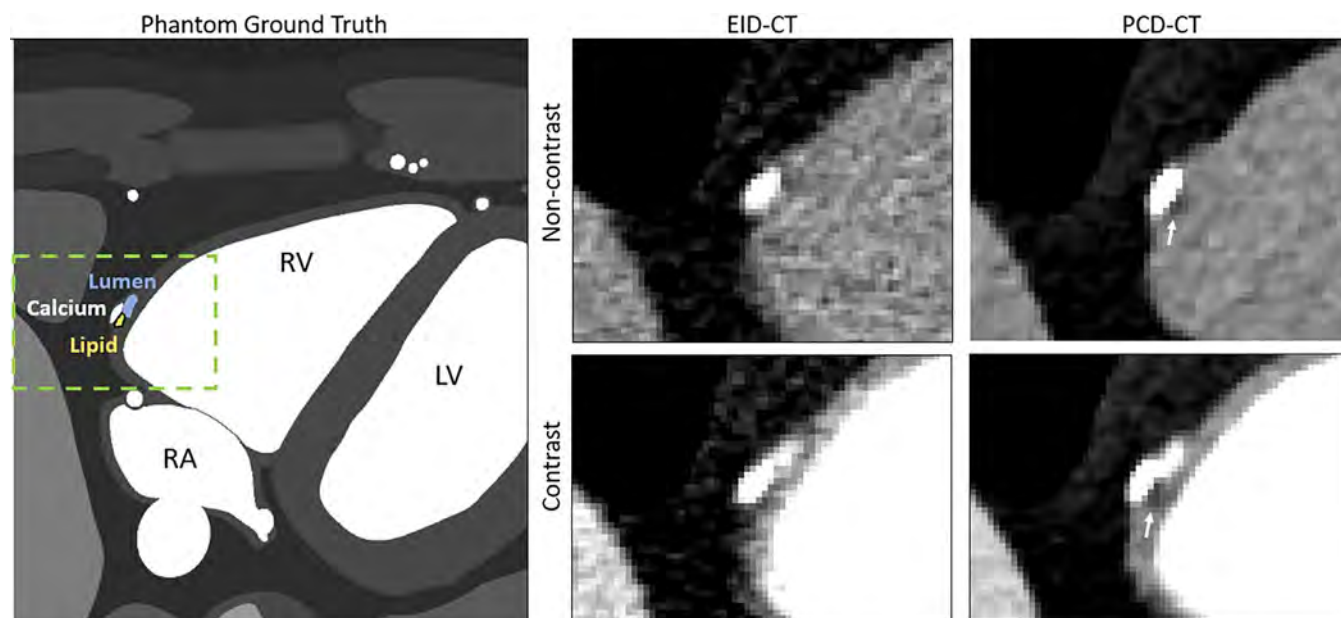
Virtual imaging trials have also been instrumental in advancing lung quantification via CT. CT is a crucial tool for objectively characterizing the severity of lung diseases.<sup>24,25</sup> However, the consistency and accuracy of these quantifications are often compromised by variations in acquisition (e.g., scanner makes and models, and acquisition protocols) and post-processing techniques (e.g., image reconstruction and quantification methods). These limitations have been effectively mitigated by harnessing the unique capabilities of virtual imaging trials (Figure A3). On the developmental front, numerous models have been introduced to synthesize lungs with varying degrees of disease severity and diverse patient attributes.<sup>26–29</sup> On the application side, virtual tools have been employed to improve accuracy and reduce quantification variability by optimizing acquisitions<sup>30</sup> and developing image harmonization techniques.<sup>31</sup>

### Longitudinal management cancer via PET-CT

Virtual imaging trials have been in use in nuclear medicine research for a decade, albeit initially as a niche research area<sup>32</sup> but with increasing use in



**FIGURE A1** ROC curves of clinical and virtual lesion detectability in Digital Mammography (DM) and Digital Breast Tomosynthesis (DBT), fitted for the MRM analysis using the ROC + KIT software (University of Chicago).



**FIGURE A2** (Left) Transaxial image slice of the XCAT heart with a plaque inserted in the RCA. The calcium and lipid components of the plaque are labeled as well as the lumen of the RCA. (Middle and right) DukeSim generated non-contrast and contrast enhanced EID and PCD CT simulations showing the corresponding CT-imaged renditions of the plaque zoomed in on the region indicated by the green box shown on the left. The calcium component of the plaque appears white in the images. The lipid component appears more visible in the PCD-CT images, appearing as a dark spot next to the calcium as indicated by the white arrows.

PET imaging,<sup>33</sup> PET/CT imaging,<sup>34</sup> and more recently SPECT imaging.<sup>35</sup> The primary use of PET/CT imaging is the use of 18-fluorodeoxyglucose (FDG) in detection, diagnosis, staging, and evaluating response to therapy in cancer, and these areas have received the most attention to date in Virtual imaging trials. An example is the ability to optimize the uptake time (typically 30–60 min) between FDG injection and PET scanning for the detection of breast cancer. Wangerin et al.<sup>36</sup> modeled the patient biodistribution, tumor biology, scanner acquisi-

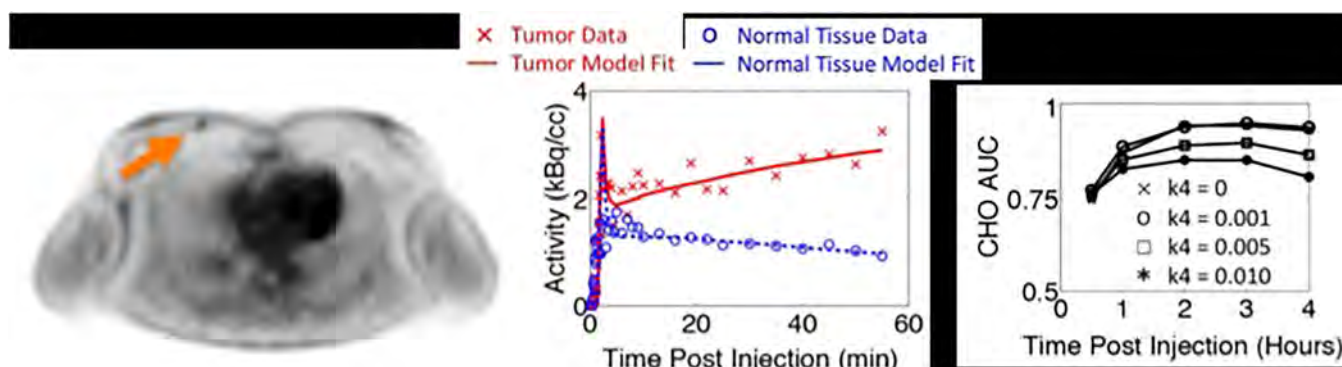
tion with image reconstruction, and human observers to estimate the optimal FDG uptake time that maximizes human observer detectability (Figure A4). This study would be infeasible to conduct with patients in practice.

### Actualization and evaluation of dynamic radiography

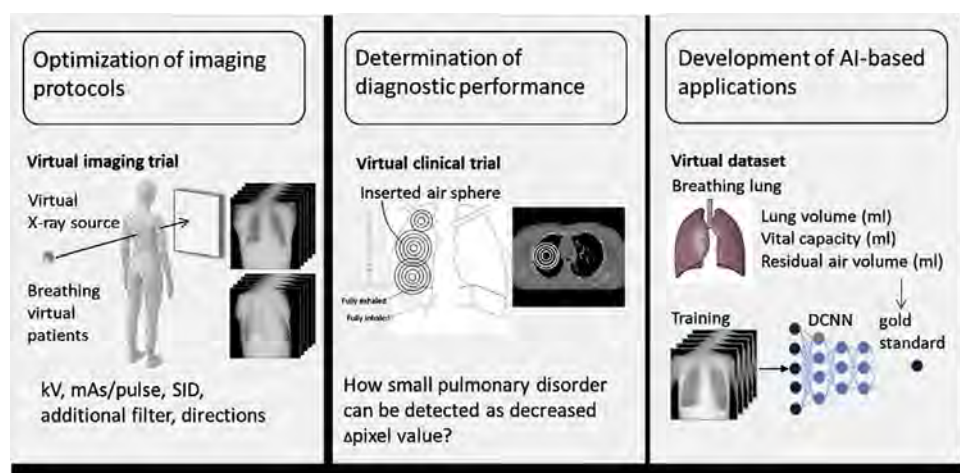
Virtual imaging trials have also contributed to the actualization and evaluation of dynamic radiography using a flat-panel detector. The imaging system is available in



**FIGURE A3** A demonstration of the utility of virtual imaging trials for robust CT quantifications of lungs. A patient is imaged using a low and a high dose scan. The overlaid red mask inside the lung denotes the emphysema regions quantified for each scan, exemplifying a major variability in the emphysema quantifications due to the differences in the acquisition dose levels. The emphysema quantification variability is reduced by developing a harmonization algorithm<sup>8</sup> developed taking advantage of virtual imaging trials developed for chest CT imaging.



**FIGURE A4** Illustration of a virtual imaging trial estimating the optimal FDG uptake time that maximizes the human observer detectability (as measured by the area under the curve [AUC] of the receiver operating characteristic curve [ROC] for a channelized Hotelling Observer [CHO]) for a tumor of a digital twin of a breast cancer patient. Left: FDG-PET image of a breast cancer patient with the tumor indicated. Center: Modeled decay-corrected FDG uptake curves for the patient. Right: Result of 9000 i.i.d simulations showing that showing that an uptake time of 120–180 min improved breast cancer detectability compared to typical times of 30–60 min.



**FIGURE A5** Overview of virtual imaging trials to accelerate real clinical research and technical development in dynamic radiography.

Japan, the US, the UK, and Asian countries, and several researchers have reported its clinical usefulness in chest and orthopedics examinations. However, there are still many challenges to its widespread use, such as determining optimal imaging conditions and diagnos-

tic performance, and providing AI-based applications. Recent studies have reported that real clinical studies can be conducted efficiently based on VIT results and that AI models trained on virtual dataset work well in real patients (Figure A5).



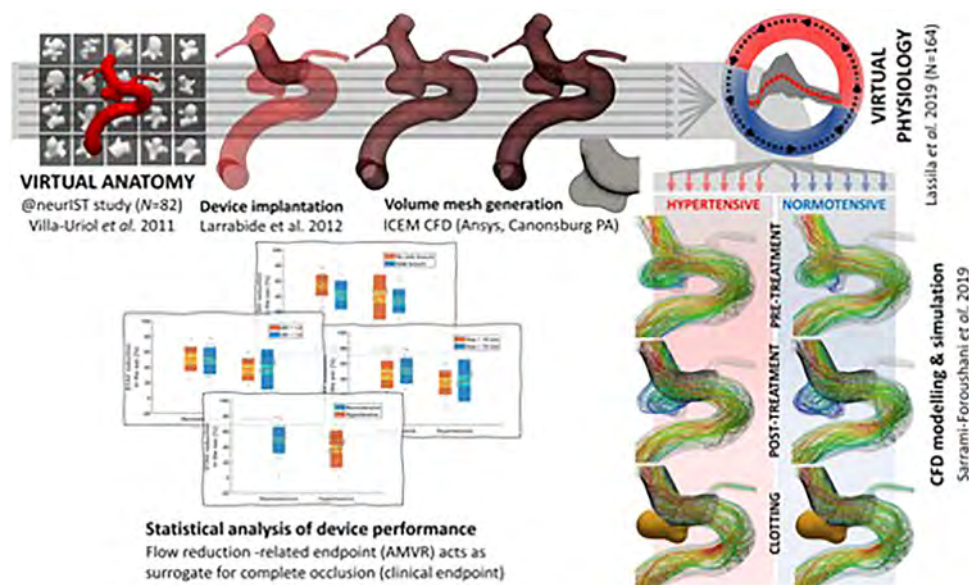


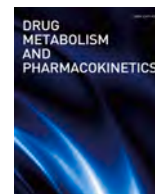
FIGURE A6 The FD-PASS study.

### In-silico trials of endovascular devices and interventions

The FD-PASS study, shown in Figure A6,<sup>37</sup> illustrated the power of In Silico Trials. The flow diverter performance assessment (FD-PASS) in-silico trial modeled the treatment of intracranial aneurysms in 164 virtual patients with 82 distinct anatomies simulating both normotensive and hypertensive patient phenotypes using a flow-diverting stent. Computational fluid dynamics were employed to quantify post-treatment flow reduction. The predicted FD-PASS flow-diversion success rates mirrored those previously reported in three clinical trials or registries (PUFS, PREMIER and ASPIRE). This in-silico

approach enabled a broader investigation of factors associated with insufficient flow reduction than what was feasible in a conventional trial. The findings demonstrated that in-silico trials of endovascular medical devices could replicate the results of traditional clinical trials to a 5% accuracy and perform virtual experiments and subgroup analyses that are challenging or impossible in conventional trials, thereby uncovering new insights into treatment failure, such as in the presence of side branches or hypertension. Several other related case studies are emerging in the literature<sup>38</sup> and recent reports have shown the potential to similar applications of in-silico trials in other device and pharm domains.<sup>39</sup>





## Regular Article

## Drug-drug interaction between trastuzumab emtansine (T-DM1) and orally administered tacrolimus in a patient and in rats

Tomohiro Nakamura<sup>a</sup>, Moeko Shimada<sup>b</sup>, Makiko Takabayashi<sup>b</sup>, Arimi Fujita<sup>a,b</sup>,  
 Takahiro Kawakami<sup>a,b</sup>, Hiroyuki Maruyama<sup>c</sup>, Hirofumi Terakawa<sup>d</sup>,  
 Takahiro Ishimoto<sup>a,b</sup>, Tsutomu Shimada<sup>a,b,e</sup>, Toshihiro Miyamoto<sup>c</sup>, Yoshimichi Sai<sup>a,b,\*</sup>

<sup>a</sup> Department of Clinical Pharmacokinetics, Graduate School of Medical Sciences, Kanazawa University, 13-1 Takara-machi, Kanazawa, 920-8641, Japan

<sup>b</sup> Department of Hospital Pharmacy, University Hospital, Kanazawa University, 13-1 Takara-machi, Kanazawa, 920-8641, Japan

<sup>c</sup> Department of Hematology, Faculty of Medicine, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, 13-1 Takara-machi, Kanazawa, 920-8640, Japan

<sup>d</sup> Department of Breast Surgery, Kanazawa University Hospital, 13-1 Takara-machi, Kanazawa, 920-8640, Japan

<sup>e</sup> Department of Clinical Pharmacy and Healthcare Science, Faculty of Pharmacy, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Kakuma-machi, Kanazawa, 920-1192, Japan



## ARTICLE INFO

## Keywords:

Antibody-drug conjugates  
 Trastuzumab emtansine  
 Tacrolimus  
 Drug-drug interactions  
 Gastrointestinal absorption

## ABSTRACT

We have experienced a 2.3-fold increase in the trough concentration of orally administered tacrolimus in a patient for 7 days after intravenous administration of trastuzumab emtansine (T-DM1), an antibody-drug conjugate. Retrospective clinical study revealed no change in factors known to alter the pharmacokinetics of tacrolimus, and the change in the trough concentration of tacrolimus was reversed after switching T-DM1 to trastuzumab (Tmab). Thus, interaction with T-DM1 was suspected as the cause of the increased trough concentration of tacrolimus. An animal study in rats showed that T-DM1 significantly increased the  $AUC_{0-\infty}$  of orally administered tacrolimus by more than 2-fold on day 0, day 3, and day 7, whereas no change was observed in the case of intravenous administration of tacrolimus. T-DM1 also significantly increased  $F_a \cdot F_g$  of tacrolimus by more than 2-fold on day 7. In contrast, Tmab itself had no effect on the blood concentration of tacrolimus. These results suggest that T-DM1 increased the blood concentration of orally administered tacrolimus, and the effect persisted for 1 week after T-DM1 administration. Metabolites of T-DM1 excreted via the biliary route may contribute to the increase in the gastrointestinal absorption of tacrolimus.

## 1. Introduction

Tacrolimus is an immunosuppressive agent widely used to suppress rejection after organ transplantation, to suppress graft-versus-host disease (GVHD) after bone marrow transplantation, and to treat autoimmune diseases [1–8]. Tacrolimus shows high inter- and intra-individual variability in pharmacokinetics, which may be influenced by multiple factors, including concomitant use with drugs that induce or inhibit CYP3A, hematocrit, albumin concentration, diarrhea, and obesity [9–24]. Such effects are important, because tacrolimus has a narrow therapeutic index and can cause diverse side effects, including severe renal impairment [25]. Therefore, strict control of the blood concentration of tacrolimus is essential to achieve efficacy and minimize side

effects. However, it is difficult to predict the blood concentration of tacrolimus, and unexpected fluctuations in blood concentration of tacrolimus often occur in clinical situations due to the addition or discontinuation of concomitant drugs or changes in the condition of patients.

Trastuzumab emtansine (T-DM1) is an antibody-drug conjugate (ADC) composed of trastuzumab (Tmab), a monoclonal antibody against the human epidermal growth factor receptor 2 (HER2); *N*-succinimidyl 4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate (MCC), a non-cleavable thioether linker, and *N*<sup>2</sup>-deacetyl-*N*<sup>2</sup>′-(3-mercapto-1-oxopropyl)-maytansine (DM1), a cytotoxic drug (payload) [26]. After intravenous administration, T-DM1 is metabolized and the metabolites predominantly undergo biliary excretion [27]. T-DM1 is mainly used for

\* Corresponding author. Department of Clinical Pharmacokinetics, Graduate School of Medical Sciences, Kanazawa University, 13-1 Takara-machi, Kanazawa, 920-8641, Japan.

E-mail address: [sai-ys@staff.kanazawa-u.ac.jp](mailto:sai-ys@staff.kanazawa-u.ac.jp) (Y. Sai).

<https://doi.org/10.1016/j.dmpk.2025.101486>

Received 20 December 2024; Received in revised form 18 March 2025; Accepted 28 April 2025

Available online 3 May 2025

1347-4367/© 2025 The Japanese Society for the Study of Xenobiotics. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

the treatment of HER2-positive metastatic breast cancer and is administered once every three weeks due to its relatively long elimination half-life ( $t_{1/2}$ ) of approximately 4 days in humans [28,29]. Compared to small molecule drugs, ADCs are considered less likely to interact with metabolic enzymes such as CYPs, as well as transporters such as P-glycoprotein and breast cancer resistance protein [30–32]. Consequently, drug-drug interactions are not usually considered when T-DM1 is administered in clinical situations.

We experienced a case of a 2.3-fold increase in the trough concentration of tacrolimus in a patient for 7 days after the first intravenous administration of T-DM1. We considered the possibility of an interaction between T-DM1 and tacrolimus in this case, but we could find no report of a relevant drug-drug interaction. Therefore, we conducted a detailed retrospective clinical study of the interaction between T-DM1 and tacrolimus in this case, and also conducted animal experiments to examine the putative pharmacokinetic interaction.

## 2. Materials and methods

### 2.1. Retrospective study of the case

Patient information was acquired from the electronic medical record system of Kanazawa University Hospital (MegaOak HR, NEC, Tokyo, Japan). This study was approved by the ethics committee of Kanazawa University Hospital (No. 2021-086).

### 2.2. Reagents

Prograf Injection 5 mg (tacrolimus) was purchased from Astellas Pharma (Tokyo, Japan). Kadcyla for Intravenous Infusion 100 mg (T-DM1) and Herceptin for Intravenous Infusion 60 mg (Tmab) were purchased from Chugai Pharmaceutical (Tokyo, Japan). All other reagents were commercial products of guaranteed grade.

### 2.3. Animals

Seven-week-old male Sprague-Dawley rats (Japan SLC Corporation, Hamamatsu, Japan) were acclimatized for 1 week with free access to water and feed. Prior to the start of the experiment, rats were fasted for 16 h with free access to water. Animal experiments were conducted in accordance with the standards set forth in the Guidelines for the Care and Use of Laboratory Animals at the Takara-machi campus of Kanazawa University (No. AP-214274).

### 2.4. Preparation of tacrolimus, T-DM1, and Tmab solutions

Tacrolimus solution was prepared from Prograf Injection at 2.5 mg/mL in saline. T-DM1 solution was prepared from Kadcyla for Intravenous Infusion at 20 mg/mL in sterile water for injection. Tmab solution was prepared from Herceptin for Intravenous Infusion at 21 mg/mL in sterile water for injection.

### 2.5. Tacrolimus administration to rats

T-DM1 (20 mg/kg), Tmab (19.6 mg/kg), or saline (1 mL/kg) as a control group was rapidly injected intravenously via a tail vein under isoflurane anesthesia. Tacrolimus was administered orally (3.2 mg/kg) at 1 h (day 0), or 3, 7, 14, or 21 days after administration of T-DM1, Tmab, or saline. In other groups, tacrolimus was administered intravenously (1.0 mg/kg) at 1 h (day 0), or 3 or 7 days after administration of T-DM1, Tmab, or saline. The dosing intervals were chosen to evaluate the effect of concurrent administration (day 0), administration at the time when the blood concentration of tacrolimus was elevated in the patient (day 7), and administration based on the T-DM1 dosage regimen (day 21), as well as day 3 and day 14 as intermediate dosing intervals. Heparin treated whole blood samples were collected from the jugular

vein under isoflurane anesthesia at 0.5, 1, 2, 4, 8, 12, and 24 h after administration of tacrolimus, and all samples were stored at  $-80^{\circ}\text{C}$  until analysis.

### 2.6. Quantification of tacrolimus using LC-MS

The determination of tacrolimus concentration was optimized based on a previous report [33] with some modifications. Whole blood samples were supplemented with 25 ng/mL ascomycin as an internal standard, and after sonication, methyl tert-butyl ether was added. The mixture was vortexed for 30 s. After centrifugation, the supernatant was collected and evaporated to dryness. The residue was taken up in a mixture of mobile phase A (2 mM ammonium acetate and 0.1 % formic acid in water), and mobile phase B (2 mM ammonium acetate and 0.1 % formic acid in acetonitrile) (A:B = 50:50, v/v), and filtered to obtain the assay sample. Tacrolimus solutions of 5, 30, and 50 ng/mL were also prepared as quality controls for low, medium, and high concentrations.

Blood concentrations of tacrolimus were determined by LC-MS using an LCMS-2020 system (Shimadzu, Kyoto, Japan) equipped with LC-20AD, SIL-20AC, CTO-20A, SPD-20A, and CBM-20A units (Shimadzu), and a CAPCELL PAK C8DD(S-5) column (2.0  $\times$  150 mm, 5  $\mu\text{m}$  particle size, Osaka Soda, Osaka, Japan). Gradient conditions were as follows: 0 min (10 % mobile phase B), 0.1–15 min (linear increase to 100 % mobile phase B), 15.1–20 min (100 % mobile phase B), 20.1–30 min (10 % mobile phase B). Flow rate was 0.4 mL/min, measurement time was 30 min, and column temperature was  $40^{\circ}\text{C}$ . Mass spectrometry was performed using cation-mode electrospray ionization, with tacrolimus and ascomycin monitored at  $m/z$  826.5 and 809.1. Labsolutions software was used for data analysis.

### 2.7. Pharmacokinetic analysis

The area under the blood concentration-time curve ( $AUC_{0-24\text{h}}$ ,  $AUC_{0-\infty}$ ), maximum blood concentration ( $C_{\text{max}}$ ), time to reach maximum blood concentration ( $T_{\text{max}}$ ), elimination half-life ( $t_{1/2}$ ), and mean residence time ( $MRT$ ) were calculated by model-independent moment analysis using Napp version 2.31 software [34]. Total clearance ( $CL_{\text{tot}}$ ), oral clearance ( $CL_{\text{tot}}/F$ ), bioavailability ( $F$ ), the fraction of dose absorbed in the portal vein ( $F_a \cdot F_g$ ), hepatic availability ( $F_h$ ), and mean absorption time ( $MAT$ ) were calculated using the following equations. The value of hepatic blood flow ( $Q_h$ ) was assumed to be 80 mL/min/kg [35].

$$CL_{\text{tot}} / F = D_{\text{po}} / AUC_{0-\infty, \text{po}}$$

$$CL_{\text{tot}} = D_{\text{iv}} / AUC_{0-\infty, \text{iv}}$$

$$F = (AUC_{0-\infty, \text{po}} / D_{\text{po}}) / (AUC_{0-\infty, \text{iv}} / D_{\text{iv}})$$

$$F_h = 1 - CL_{\text{tot}} / Q_h$$

$$F_a \cdot F_g = F / F_h$$

$$MAT = MRT_{\text{po}} - MRT_{\text{iv}}$$

### 2.8. Statistical analysis

GraphPad Prism 10 (GraphPad Software Inc., San Diego, CA, USA) was used for statistical processing of the data. The Student's unpaired  $t$ -test or one-way ANOVA with Tukey-Kramer method were used to test for significant differences ( $p$  value less than 5 %).

### 3. Results

#### 3.1. The case

The patient was 50s-year-old female with a history of acute myeloid leukemia, peripheral blood stem cell transplant, chronic GVHD enteritis, chronic kidney disease (post-transplant nephropathy), and atrophic gastritis. She was under follow-up after 4 courses of docetaxel + Tmab therapy and 4 courses of epirubicin + cyclophosphamide combination therapy post-surgery for right breast cancer. In addition to oral administration of tacrolimus 1.5 mg/day, she was taking amlodipine, bonoprazan, budesonide, fluconazole, ibandronic acid, levothyroxine, mesalazine, sulfamethoxazole trimethoprim (ST combination), and ursodeoxycholic acid. She was admitted to the hospital for initiation of chemotherapy due to lung metastasis of recurrent HER2-positive breast cancer 3 years after surgery. T-DM1, an antibody-drug conjugate, was selected as the first-line treatment considering the patient's medical history, and T-DM1 was started at a two-step dose reduction (2.4 mg/kg, 120 mg) after consultation with a hematologist.

The first day of T-DM1 administration was day 0. The doses of tacrolimus, T-DM1 and Tmab, trough concentration, and concentration/dose (C/D) ratio of tacrolimus are summarized in Fig. 1, and laboratory values are shown in Table 1. The trough concentration of tacrolimus (1.5 mg/day) was stable at 3.0–4.0 ng/mL (C/D ratio 2.0–2.7) before T-DM1 initiation. On day 5, grade 3 elevations in aspartate aminotransferase (AST) and alanine aminotransferase (ALT) were observed, and on day 7, the trough concentration of tacrolimus reached 8.2 ng/mL (C/D ratio 5.5). During this period, there were no changes in concomitant medications other than T-DM1, no changes in diet or intake that could affect absorption or excretion of tacrolimus, and no serious diarrhea or bleeding symptoms. Although a transient increase in the trough concentration of tacrolimus due to hepatic impairment was considered, after consultation with the hematologist, the dose of tacrolimus was reduced to 1.0 mg/day. On day 23, the trough concentration of tacrolimus decreased to 4.5 ng/mL, but the C/D ratio was 4.5, which was still 1.9-fold greater than the C/D ratio before T-DM1 was initiated. After 3 courses of T-DM1, the case was switched to Tmab at the physician's discretion. By day 175, the trough concentration of tacrolimus had decreased to 2.2 ng/mL (C/D ratio 2.2). There were no major changes in concomitant drugs, diet, defecation habit, or laboratory values after the change to Tmab. No tacrolimus-related adverse events, such as renal dysfunction, hyperkalemia, or GVHD enteritis flare-ups, were seen from day 0 to day 175.

#### 3.2. Effect of T-DM1 or Tmab on the pharmacokinetics of orally administered tacrolimus in rats

The blood concentration-time profiles of tacrolimus after oral administration in the T-DM1 and Tmab groups are shown in Fig. 2. The pharmacokinetic parameters are shown in Table 2, and the  $AUC_{0-\infty}$  values from day 0 to day 21 are shown in Fig. 3. In the T-DM1 group,  $AUC_{0-\infty}$  was significantly increased compared to the saline group by 2.2-fold, 3.8-fold, and 1.8-fold on day 0, day 3, and day 7, respectively ( $p < 0.05$ ). However, by day 14 and day 21, there was no significant difference between the two groups in any of the parameters. In the Tmab group, there was no significant difference in blood concentration or pharmacokinetic parameters compared with the saline group at any time point from day 0 to day 21.

#### 3.3. Effect of T-DM1 or Tmab on the pharmacokinetics of intravenously administered tacrolimus in rats

The pharmacokinetic parameters of tacrolimus after intravenous administration to rats are shown in Table 3. There was no significant difference in any of the parameters between the saline and T-DM1 groups or between the saline and Tmab groups from day 0 to day 7.

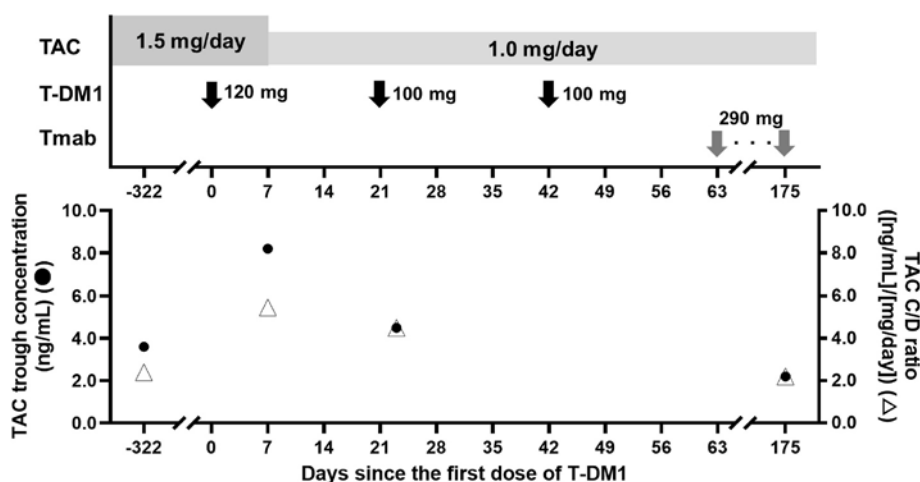
#### 3.4. Effect of T-DM1 or Tmab on $F$ , $F_a \cdot F_g$ , $F_h$ and $MAT$ of tacrolimus after oral administration to rats

The calculated values of  $F$ ,  $F_a \cdot F_g$ ,  $F_h$ , and  $MAT$  after oral and intravenous administration of tacrolimus are shown in Table 4. In the T-DM1 group,  $F$  and  $F_a \cdot F_g$  tended to increase on day 0 and day 3, and were significantly increased by 2.5-fold and 2.3-fold, respectively, on day 7 compared to the saline control. There was no significant difference in  $F_h$ .  $MAT$  was significantly prolonged on day 3 and day 7 ( $p < 0.05$ ). There was no significant difference between the Tmab and saline groups in any of these parameters.

### 4. Discussion

Our retrospective clinical study suggested that there is a drug-drug interaction between T-DM1 and tacrolimus, and we conducted animal experiments in rats to confirm this and examine the mechanism.

The patient was a woman in her 50s who had been taking tacrolimus for long-term treatment of chronic GVHD. Seven days after the first intravenous dose of T-DM1, initiated for the treatment of HER2-positive metastatic breast cancer, the trough concentration of tacrolimus and C/

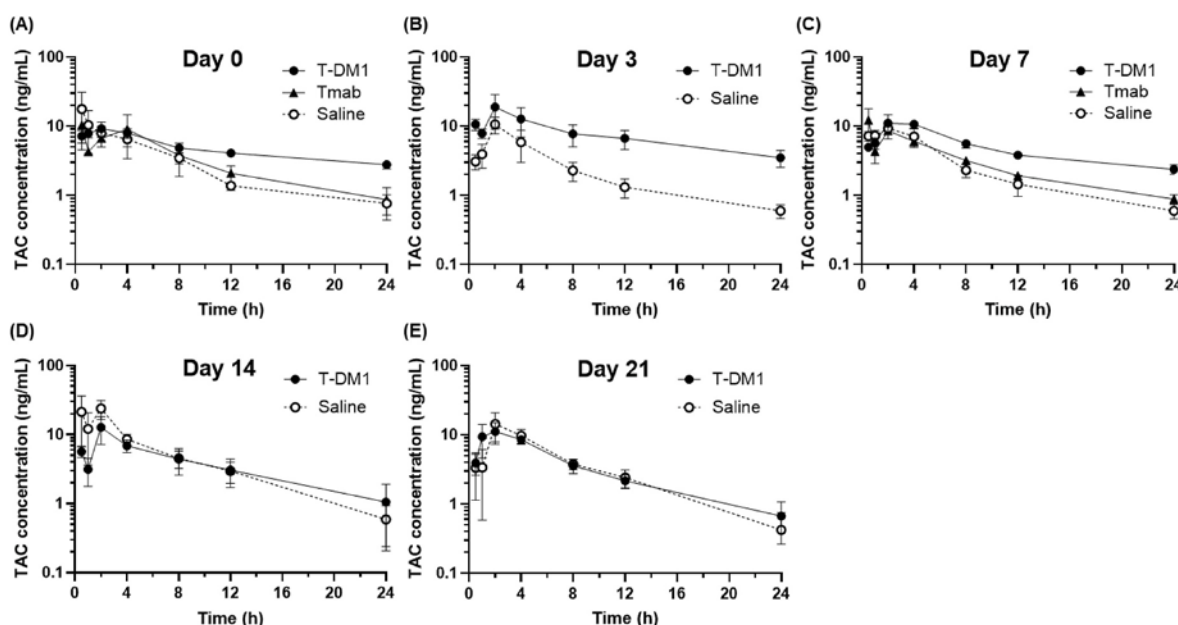


**Fig. 1.** Clinically administered doses of tacrolimus (TAC), T-DM1 and trastuzumab (Tmab), trough concentration of tacrolimus (●), and concentration/dose (C/D) ratio of tacrolimus (△)

The patient was taking tacrolimus (1.5 mg/day or 1.0 mg/day) orally and then T-DM1 (120 mg or 100 mg) or Tmab (290 mg).

**Table 1**  
The patient's laboratory values.

Days since the first dose of T-DM1	−322	0	5	6	7	12	21	23	42	45	63	65	175
Hb (g/dL)	10.9	12.2	13.0	13.4	12.9	12.9	12.5	13.5	12.6	13.6	12.7	13.0	12.6
Ht (%)	34.3	36.6	38.7	39.7	39.3	39.6	38.9	40.8	38.8	40.6	39.1	40.1	39.5
Plt ( $\times 10^3$ )	501	428	174	164	172	353	382	354	374	307	404	439	402
Scr (mg/dL)	1.15	1.10	0.90	1.02	1.07	0.86	1.13	0.96	1.10	1.02	1.25	1.01	0.89
eGFR (mL/min/1.73m <sup>2</sup> )	38.6	40.3	50.2	43.8	41.5	52.7	39.1	46.8	40.3	43.8	35.0	44.2	50.5
AST (IU/L)	14	13	181	76	39	23	42	50	54	69	46	44	27
ALT (IU/L)	16	14	279	199	137	48	57	59	83	83	41	49	29
T-Bil (mg/dL)	0.3	0.4	0.8	0.7	0.6	0.6	0.4	0.5	0.6	0.7	0.4	0.5	0.6
Alb (g/dL)	—	3.1	3.0	3.2	3.1	3.1	3.5	3.3	3.6	3.5	3.3	3.2	3.8



**Fig. 2.** Blood concentration-time profiles of tacrolimus (TAC) in rats after oral administration on day 0 (A), day 3 (B), day 7 (C), day 14 (D), or day 21 (E) after T-DM1 or trastuzumab (Tmab)

T-DM1 (20 mg/kg), Tmab (19.6 mg/kg), or saline (1 mL/kg) in the control group was rapidly injected intravenously in rats. Tacrolimus (3.2 mg/kg) was administered orally at 1 h, 3, 7, 14, or 21 days after administration of T-DM1, Tmab, or saline. Values are shown as the mean  $\pm$  S.E.M. ( $n = 4-6$ ).

D ratio were increased 2.3-fold compared with those before T-DM1 initiation. These changes in the trough concentration of tacrolimus and C/D ratio were reversed after switching from T-DM1 to Tmab. During this period, there were no changes in diet, other medications or general condition, including serious diarrhea or bleeding symptoms, that might have affected the absorption or excretion of tacrolimus [17–24]. Although we cannot completely rule out the possibility that the hepatotoxicity influenced tacrolimus pharmacokinetics in our case, it has been reported that mild hepatic impairment (Child-Pugh score 6–7) does not affect the pharmacokinetics of tacrolimus [36], and our patient showed a Child-Pugh score of 6 (serum albumin 3.2 g/dL, <3.5 g/dL which increased the score by 1; hepatic encephalopathy was absent; no ascites; serum bilirubin level 0.7 mg/dL, <2.0 mg/dL; prothrombin time activity  $122 > 70$  %; prothrombin time-international normalized ratio  $0.9 < 1.7$  on day 6). Therefore, it seems unlikely that the temporary hepatotoxicity had any marked effect on the blood concentration of tacrolimus. These data support the existence of an interaction between T-DM1 and tacrolimus. The  $t_{1/2}$  of a single intravenous infusion of T-DM1 is approximately 4 days, and the blood concentration of T-DM1 on day 7, when the trough concentration of tacrolimus was still increased, would have been approximately 30 % of  $C_{max}$  [37]. Trastuzumab (antibody moiety of T-DM1) is mainly composed of the Fab and Fc regions, and even if taken up nonspecifically by cells, the Fc region binds to the neonatal Fc receptor and is recycled out of the cell by endosomes [38], resulting in a long residence time of T-DM1. Thus, the

increase in the trough concentration of tacrolimus in our patient for 7 days after T-DM1 administration is consistent with a relatively long  $t_{1/2}$  of T-DM1 (approximately 4 days). On the other hand, the trough concentration of tacrolimus reverted to the pre-administration level after the change of medication from T-DM1 to Tmab, suggesting that T-DM1 and/or its metabolites containing DM1, rather than Tmab itself, caused the change in the trough concentration of tacrolimus.

Since an effect of factors not measured in the retrospective study could not be ruled out, we conducted animal experiments to confirm the putative interaction. The  $AUC_{0-\infty}$  of tacrolimus after oral administration to rats was increased significantly on day 0, day 3, and day 7 after T-DM1 administration, reproducing the clinical phenomenon. The effect of T-DM1 persisted for 1 week after T-DM1 administration and was no longer observed after the second week. Based on the results of non-clinical and clinical studies, the pharmacokinetic profile, including the  $t_{1/2}$  of T-DM1, is considered to be similar in rats and humans in the clinically relevant dose range [27,39]. This suggests that after 14 days of T-DM1 administration, the majority of T-DM1 would have been eliminated, which is consistent with the absence of a significant change in  $AUC_{0-\infty}$  of orally administered tacrolimus by this time.

Next, to investigate which phase of tacrolimus absorption, distribution, metabolism, and excretion was affected by administration of T-DM1, tacrolimus was administered intravenously after T-DM1 administration. There was no significant difference in the pharmacokinetics of intravenously administered tacrolimus between the T-DM1 group and



Table 2

Pharmacokinetic parameters of tacrolimus (3.2 mg/kg) orally administered at various time points after T-DM1 or trastuzumab (Tmab) treatment in rats.

		Day 0	Day 3	Day 7	Day 14	Day 21
$AUC_{0-24h}$ (ng · h/mL)	Saline	77.5 ± 16.0	61.2 ± 17.6	68.9 ± 10.3	126.8 ± 37.8	92.0 ± 22.2
	T-DM1	115.7 ± 12.3	182.8 ± 65.0	122.5 ± 6.1 <sup>a,b</sup>	93.5 ± 29.9	87.1 ± 14.7
	Tmab	82.5 ± 33.9		73.8 ± 7.5		
$AUC_{0-∞}$ (ng · h/mL)	Saline	93.4 ± 15.6	68.5 ± 18.4	76.4 ± 11.8	133.0 ± 39.1	96.7 ± 25.8
	T-DM1	201.5 ± 31.0 <sup>a,b</sup>	256.9 ± 79.1*	196.1 ± 23.8 <sup>a,b</sup>	147.4 ± 73.5	101.5 ± 20.5
	Tmab	92.1 ± 38.7		84.5 ± 7.4		
$C_{max}$ (ng/mL)	Saline	21.0 ± 12.6	11.6 ± 2.5	12.1 ± 1.3	30.0 ± 12.8	17.5 ± 5.6
	T-DM1	11.7 ± 1.5	20.3 ± 9.4	15.0 ± 2.7	14.9 ± 4.7	14.8 ± 3.5
	Tmab	13.7 ± 4.1		14.5 ± 5.0		
$T_{max}$ (h)	Saline	2.4 ± 0.6	2.2 ± 0.5	1.7 ± 0.5	2.1 ± 0.6	2.8 ± 0.5
	T-DM1	1.9 ± 0.6	1.4 ± 0.4	3.3 ± 0.4 <sup>b</sup>	3.3 ± 1.3	2.3 ± 0.6
	Tmab	1.4 ± 0.9		1.3 ± 0.4		
$t_{1/2}$ (h)	Saline	10.7 ± 3.8	9.5 ± 0.9	10.2 ± 1.8	5.1 ± 0.9	4.1 ± 1.2
	T-DM1	20.6 ± 2.8 <sup>b</sup>	15.2 ± 0.9*	18.8 ± 4.8	9.9 ± 6.6	7.4 ± 3.1
	Tmab	7.4 ± 0.9		8.7 ± 0.5		
$CL_{tot}/F$ (L/h/kg)	Saline	40.5 ± 8.0	62.0 ± 16.3	47.7 ± 8.1	35.3 ± 10.2	43.4 ± 11.5
	T-DM1	17.5 ± 2.6	16.7 ± 3.7*	17.7 ± 2.3 <sup>a</sup>	82.1 ± 48.2	40.6 ± 10.3
	Tmab	51.5 ± 14.9		38.9 ± 3.9		

Values are shown as the mean ± S.E.M. (n = 4–6). \* $p < 0.05$ ; Student's unpaired  $t$ -test; T-DM1 vs saline group. <sup>a</sup> $p < 0.05$ ; one-way ANOVA with Tukey-Kramer test; T-DM1 vs saline group. <sup>b</sup> $p < 0.05$ ; one-way ANOVA with Tukey-Kramer test; T-DM1 vs trastuzumab (Tmab) group.

the saline group from day 0 to day 7. However,  $F_a \cdot F_g$  was significantly higher in the T-DM1 group than in the saline group, though  $F_h$  was unchanged, suggesting that administration of T-DM1 affects the absorption process of tacrolimus in the gastrointestinal tract.  $MAT$  was significantly prolonged on day 3 and day 7, suggesting that T-DM1 may have prolonged tacrolimus absorption.

Tacrolimus is absorbed over a wide range of sites from the duodenum to the colon in rats [40], as well as in humans [41]. A rat in vivo study using Tmab-MCC-[<sup>3</sup>H]DM1 indicated that T-DM1 is mainly metabolized to DM1, MCC-DM1, and Lys-MCC-DM1 [27]. It has been reported that DM1 inhibits CYP3A but not P-glycoprotein [42–44], whereas Lys-MCC-DM1 does not inhibit CYP3A or P-glycoprotein [42,45]. As 50 % of the radioactive metabolites underwent biliary excretion within 7 days after administration of radiolabeled T-DM1 [27], these known and/or unknown metabolites might affect the absorption process of orally administered tacrolimus, possibly by inhibiting CYP3A and/or P-glycoprotein, resulting in delayed gastrointestinal absorption. It has been reported that T-DM1 did not affect the pharmacokinetics of

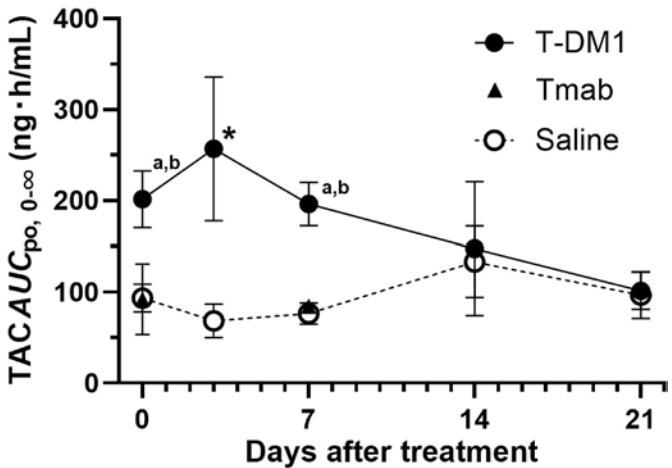


Fig. 3. The  $AUC_{0-∞}$  of tacrolimus (TAC, 3.2 mg/kg) orally administered at various intervals (day 0 to day 21) after T-DM1 or trastuzumab (Tmab) treatment in rats

Values are shown as the mean ± S.E.M. (n = 4–6). \* $p < 0.05$ ; Student's unpaired  $t$ -test; T-DM1 vs saline group. <sup>a</sup> $p < 0.05$ ; one-way ANOVA with Tukey-Kramer test; T-DM1 vs saline group. <sup>b</sup> $p < 0.05$ ; one-way ANOVA with Tukey-Kramer test; T-DM1 vs trastuzumab (Tmab) group.

Table 3

Pharmacokinetic parameters of tacrolimus (1.0 mg/kg) intravenously administered at various time points after T-DM1 or trastuzumab (Tmab) treatment in rats.

		Day 0	Day 3	Day 7
$AUC_{0-24h}$ (ng · h/mL)	Saline	299.7 ± 25.9	278.9 ± 12.9	283.5 ± 23.3
	T-DM1	365.4 ± 50.6	319.4 ± 28.7	297.9 ± 23.1
	Tmab	276.5 ± 22.5		306.1 ± 13.9
$AUC_{0-∞}$ (ng · h/mL)	Saline	327.2 ± 31.3	297.4 ± 16.1	308.6 ± 22.2
	T-DM1	399.2 ± 52.8	342.3 ± 26.0	317.3 ± 25.2
	Tmab	289.5 ± 23.3		340.3 ± 18.6
$C_{max}$ (ng/mL)	Saline	66.2 ± 11.6	73.7 ± 5.7	69.6 ± 5.7
	T-DM1	89.2 ± 7.4	72.5 ± 9.2	69.5 ± 7.4
	Tmab	88.8 ± 13.5		72.5 ± 7.2
$t_{1/2}$ (h)	Saline	7.4 ± 0.8	6.4 ± 0.5	7.5 ± 1.1
	T-DM1	7.8 ± 0.7	6.7 ± 0.7	6.5 ± 0.6
	Tmab	6.0 ± 0.2		7.8 ± 0.1
$CL_{tot}$ (L/h/kg)	Saline	3.2 ± 0.4	3.4 ± 0.2	3.3 ± 0.3
	T-DM1	2.7 ± 0.3	3.0 ± 0.2	3.2 ± 0.2
	Tmab	3.5 ± 0.3		3.0 ± 0.2

Values are shown as the mean ± S.E.M. (n = 4–6).

Table 4

Calculated pharmacokinetic parameters of tacrolimus after oral or intravenous administration at various time points after T-DM1 or trastuzumab (Tmab) treatment in rats.

		Day 0	Day 3	Day 7
$F$ (%)	Saline	8.9 ± 1.7	7.2 ± 2.0	7.7 ± 1.3
	T-DM1	15.8 ± 3.2	23.5 ± 7.4	19.3 ± 2.7 <sup>a,b</sup>
	Tmab	9.9 ± 4.3		7.8 ± 0.8
$F_a \cdot F_g$ (%)	Saline	27.3 ± 8.2	24.8 ± 7.7	25.2 ± 6.0
	T-DM1	35.5 ± 8.9	62.2 ± 21.2	58.9 ± 11.7 <sup>a,b</sup>
	Tmab	37.4 ± 18.1		20.3 ± 2.7
$F_h$ (%)	Saline	32.7 ± 7.6	29.0 ± 4.2	30.7 ± 5.1
	T-DM1	44.5 ± 6.5	37.7 ± 4.7	32.8 ± 5.0
	Tmab	26.6 ± 6.0		38.2 ± 3.3
$MAT$ (h)	Saline	5.7 ± 5.9	2.8 ± 1.2	0.3 ± 1.1
	T-DM1	19.5 ± 3.4	12.7 ± 1.7*	17.8 ± 6.1 <sup>a,b</sup>
	Tmab	3.1 ± 1.1		1.6 ± 1.4

Values are shown as the mean ± S.E.M. (n = 4–6). \* $p < 0.05$ ; Student's unpaired  $t$ -test; T-DM1 vs saline group. <sup>a</sup> $p < 0.05$ ; one-way ANOVA with Tukey-Kramer test; T-DM1 vs saline group. <sup>b</sup> $p < 0.05$ ; one-way ANOVA with Tukey-Kramer test; T-DM1 vs trastuzumab (Tmab) group.

intravenously administered paclitaxel or intravenously administered pertuzumab [46,47], and that brentuximab vedotin, an ADC, did not affect the pharmacokinetics of intravenously administered midazolam [48]. These reports are consistent with the results of intravenous administration in the present study.

The idea that the factor responsible for the increase in the trough concentration of tacrolimus is not the antibody moiety of T-DM1 but the ADC and/or its metabolites is supported by our finding that there was no significant change in the blood concentration of tacrolimus in the Tmb group after oral or intravenous administration of tacrolimus. The clinical data and animal study indicate that T-DM1 increases the absorption of tacrolimus for approximately 1 week. This is the first report of an interaction between an ADC and an orally administered drug, and raises the possibility that drug-drug interactions may occur with other ADCs, which have previously been considered to have little effect on concomitant drugs. Therefore, the potential for interaction of ADCs with other orally administered drugs should be further examined. Since tacrolimus is an immunosuppressive drug that shows large individual variations in pharmacokinetics and a narrow therapeutic index, it will be particularly important for clinicians to consider potential drug-drug interactions with T-DM1 or other ADCs, and to monitor blood concentration regularly. Further studies are needed to elucidate in detail the mechanism of the interaction between T-DM1 and tacrolimus, as well as to investigate interactions with other ADCs.

## 5. Conclusion

An increase in the blood concentration of orally administered tacrolimus was observed after intravenous T-DM1 administration in a patient and the interaction was confirmed in an animal study. The effect was shown to persist for 1 week after T-DM1 administration, and the pharmacokinetic parameters suggest that metabolites of T-DM1, rather than Tmb itself, affected the gastrointestinal absorption of tacrolimus.

## CRediT authorship contribution statement

**Tomohiro Nakamura:** Writing – original draft, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Moeko Shimada:** Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization. **Makiko Takabayashi:** Writing – review & editing, Conceptualization. **Arimi Fujita:** Writing – review & editing, Supervision. **Takahiro Kawakami:** Writing – review & editing. **Hiroyuki Maruyama:** Investigation. **Hirofumi Terakawa:** Investigation. **Takahiro Ishimoto:** Writing – review & editing. **Tsutomu Shimada:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Toshihiro Miyamoto:** Investigation. **Yoshimichi Sai:** Writing – review & editing, Project administration, Methodology, Funding acquisition, Conceptualization.

## Declaration of competing interest

The authors declare that they have no competing interests.

## Acknowledgements

This work was supported in part by JST SPRING [Grant Number JPMJSP2135] and JSPS KAKENHI [Grant Number JP23K06209]. We thank Satoshi Mizuno and Mayu Kajino, PhD students at the Graduate School of Medical Sciences, Kanazawa University, for their technical supports in the LC/MS measurement of tacrolimus blood concentration.

## References

- [1] Starzl TE, Todo S, Fung J, Demetris AJ, Venkataraman R, Jain A. FK 506 for liver, kidney, and pancreas transplantation. *Lancet* 1989;2:1000–4. [https://doi.org/10.1016/s0140-6736\(89\)91014-3](https://doi.org/10.1016/s0140-6736(89)91014-3).
- [2] European FK506 Multicentre Liver Study Group. Randomised trial comparing tacrolimus (FK506) and cyclosporin in prevention of liver allograft rejection. European FK506 Multicentre Liver Study Group. *Lancet* 1994;344:423–8. [https://doi.org/10.1016/S0140-6736\(94\)91766-3](https://doi.org/10.1016/S0140-6736(94)91766-3).
- [3] U.S. Multicenter FK506 Liver Study Group. A comparison of tacrolimus (FK 506) and cyclosporine for immunosuppression in liver transplantation. *N Engl J Med* 1994;331:1110–5. <https://doi.org/10.1056/NEJM199410273311702>.
- [4] Hiraoka A, Ohashi Y, Okamoto S, Moriyama Y, Nagao T, Kodera Y, et al. Phase III study comparing tacrolimus (FK506) with cyclosporine for graft-versus-host disease prophylaxis after allogeneic bone marrow transplantation. *Bone Marrow Transplant* 2001;28:181–5. <https://doi.org/10.1038/sj.bmt.1703097>.
- [5] Ratanatharathorn V, Nash RA, Przepiorka D, Devine SM, Klein JL, Weisdorf D, et al. Phase III study comparing methotrexate and tacrolimus (prograf, FK506) with methotrexate and cyclosporine for graft-versus-host disease prophylaxis after HLA-identical sibling bone marrow transplantation. *Blood* 1998;92:2303–14. <https://doi.org/10.1182/blood.V92.7.2303>.
- [6] Nash RA, Pineiro LA, Storb R, Deeg HJ, Fitzsimmons WE, Furlong T, et al. FK506 in combination with methotrexate for the prevention of graft-versus-host disease after marrow transplantation from matched unrelated donors. *Blood* 1996;88:3634–41. <https://doi.org/10.1182/blood.V88.9.3634.bloodjournal8893634>.
- [7] Furst DE, Saag K, Fleischmann MR, Sherrer Y, Block JA, Schnitzer T, et al. Efficacy of tacrolimus in rheumatoid arthritis patients who have been treated unsuccessfully with methotrexate: a six-month, double-blind, randomized, dose-ranging study. *Arthritis Rheum* 2002;46:2020–8. <https://doi.org/10.1002/art.10427>.
- [8] Yocum DE, Furst DE, Kaine JL, Baldassare AR, Stevenson JT, Borton MA, et al. Efficacy and safety of tacrolimus in patients with rheumatoid arthritis: a double-blind trial. *Arthritis Rheum* 2003;48:3328–37. <https://doi.org/10.1002/art.11363>.
- [9] Mañez R, Martin M, Raman D, Silverman D, Jain A, Warty V, et al. Fluconazole therapy in transplant recipients receiving FK506. *Transplantation* 1994;57:1521–3. <https://doi.org/10.1097/00007890-199405270-00022>.
- [10] Jensen C, Jordan M, Shapiro R, Scantlebury V, Hakala T, Fung J, et al. Interaction between tacrolimus and erythromycin. *Lancet* 1994;344:825. [https://doi.org/10.1016/S0140-6736\(94\)92383-3](https://doi.org/10.1016/S0140-6736(94)92383-3).
- [11] Iwasaki K, Matsuda H, Nagase K, Shiraga T, Tokuma Y, Uchida K. Effects of twenty-three drugs on the metabolism of FK506 by human liver microsomes. *Res Commun Chem Pathol Pharmacol* 1993;82:209–16.
- [12] Venkataramanan R, Zang S, Gayowski T, Singh N. Voriconazole inhibition of the metabolism of tacrolimus in a liver transplant recipient and in human liver microsomes. *Antimicrob Agents Chemother* 2002;46:3091–3. <https://doi.org/10.1128/AAC.46.9.3091-3093.2002>.
- [13] Christians U, Schmidt G, Bader A, Lampen A, Schottmann R, Linck A, et al. Identification of drugs inhibiting the in vitro metabolism of tacrolimus by human liver microsomes. *Br J Clin Pharmacol* 1996;41:187–90. <https://doi.org/10.1111/j.1365-2125.1996.tb00181.x>.
- [14] Furlan V, Perello L, Jacquemin E, Debray D, Taburet AM. Interactions between FK506 and rifampicin or erythromycin in pediatric liver recipients. *Transplantation* 1995;59:1217–8.
- [15] Takahashi K, Yano I, Fukuhara Y, Katsura T, Takahashi T, Ito N, et al. Distinct effects of omeprazole and rabeprazole on the tacrolimus blood concentration in a kidney transplant recipient. *Drug Metab Pharmacokinet* 2007;22:441–4. <https://doi.org/10.2133/dmpk.22.441>.
- [16] Shapiro R, Venkataramanan R, Warty VS, Scantlebury VP, Rybka W, McCauley J, et al. FK 506 interaction with danazol. *Lancet* 1993;341:1344–5. [https://doi.org/10.1016/0140-6736\(93\)90852-8](https://doi.org/10.1016/0140-6736(93)90852-8).
- [17] Rodrigo E, de Cos MA, Sánchez B, Ruiz JC, Piñera C, Fernández-Fresnedo G, et al. High initial blood levels of tacrolimus in overweight renal transplant recipients. *Transplant Proc* 2005;37:1453–4. <https://doi.org/10.1016/j.transproceed.2005.02.055>.
- [18] Sawamoto K, Huang TT, Sugimoto N, Mizutani Y, Sai Y, Miyamoto K. Mechanisms of lower maintenance dose of tacrolimus in obese patients. *Drug Metab Pharmacokinet* 2014;29:341–7. <https://doi.org/10.2133/dmpk.dmpk-13-rg-110>.
- [19] Hochleitner BW, Bösmüller C, Nehoda H, Frühwirth M, Simma B, Ellemunter H, et al. Increased tacrolimus levels during diarrhea. *Transpl Int* 2001;14:230–3. <https://doi.org/10.1007/s001470100331>.
- [20] Bekersky I, Dressler D, Mekki QA. Effect of low- and high-fat meals on tacrolimus absorption following 5 mg single oral doses to healthy human subjects. *J Clin Pharmacol* 2001;41:176–82. <https://doi.org/10.1177/00912700122009999>.
- [21] Undre NA, Schäfer A, European Tacrolimus Multicentre Renal Study Group. Factors affecting the pharmacokinetics of tacrolimus in the first year after renal transplantation. *Transplant Proc* 1998;30:1261–3. [https://doi.org/10.1016/s0041-1345\(98\)00234-6](https://doi.org/10.1016/s0041-1345(98)00234-6).
- [22] Jain AB, Fung JJ, Tzakis AG, Venkataramanan R, Abu-Elmagd K, Alessiani M, et al. Comparative study of cyclosporine and FK 506 dosage requirements in adult and pediatric orthotopic liver transplant patients. *Transplant Proc* 1991;23:2763–6.
- [23] Abu-Elmagd KM, Fung JJ, Alessiani M, Jain A, Takaya S, Venkataramanan R, et al. Strategy of FK 506 therapy in liver transplant patients: effect of graft function. *Transplant Proc* 1991;23:2771–4.
- [24] Jain AB, Venkataramanan R, Cadoff E, Fung JJ, Todo S, Krajack A, et al. Effect of hepatic dysfunction and T tube clamping on FK 506 pharmacokinetics and trough concentrations. *Transplant Proc* 1990;22:57–9.
- [25] Japanese FK 506 Study Group. Japanese study of FK 506 on kidney transplantation: the benefit of monitoring the whole blood FK 506 concentration. *Transplant Proc* 1991;23:3085–8.
- [26] Junttila TT, Li G, Parsons K, Phillips GL, Sliwowski MX. Trastuzumab-DM1 (T-DM1) retains all the mechanisms of action of trastuzumab and efficiently inhibits

- growth of lapatinib insensitive breast cancer. *Breast Cancer Res Treat* 2011;128: 347–56. <https://doi.org/10.1007/s10549-010-1090-x>.
- [27] Shen B-Q, Bumbaca D, Saad O, Yue Q, Pastuskovas CV, Khojasteh SC, et al. Catabolic fate and pharmacokinetic characterization of trastuzumab emtansine (T-DM1): an emphasis on preclinical and clinical catabolism. *Curr Drug Metab* 2012; 13:901–10. <https://doi.org/10.2174/138920012802138598>.
- [28] Verma S, Miles D, Gianni L, Krop IE, Welslau M, Baselga J, et al. Trastuzumab emtansine for HER2-positive advanced breast cancer. *N Engl J Med* 2012;367: 1783–91. <https://doi.org/10.1056/NEJMoa1209124>.
- [29] Krop IE, Beeram M, Modi S, Jones SF, Holden SN, Yu W, et al. Phase I study of trastuzumab-DM1, an HER2 antibody-drug conjugate, given every 3 weeks to patients with HER2-positive metastatic breast cancer. *J Clin Oncol* 2010;28: 2698–704. <https://doi.org/10.1200/JCO.2009.26.2071>.
- [30] Seitz K, Zhou H. Pharmacokinetic drug-drug interaction potentials for therapeutic monoclonal antibodies: reality check. *J Clin Pharmacol* 2007;47:1104–18. <https://doi.org/10.1177/0091270007306958>.
- [31] Huang S-M, Zhao H, Lee J-I, Reynolds K, Zhang L, Temple R, et al. Therapeutic protein-drug interactions and implications for drug development. *Clin Pharmacol Ther* 2010;87:497–503. <https://doi.org/10.1038/clpt.2009.308>.
- [32] Han TH, Zhao B. Absorption, distribution, metabolism, and excretion considerations for the development of antibody-drug conjugates. *Drug Metab Dispos* 2014;42:1914–20. <https://doi.org/10.1124/dmd.114.058586>.
- [33] Park JS, Cho HR, Kang MJ, Choi YS. A rapid and sensitive method to determine tacrolimus in rat whole blood using liquid-liquid extraction with mild temperature ultrasonication and LC-MS/MS. *Arch Pharm Res (Seoul)* 2016;39:73–82. <https://doi.org/10.1007/s12272-015-0681-9>.
- [34] Hisaka A, Sugiyama Y. Analysis of nonlinear and nonsteady state hepatic extraction with the dispersion model using the finite difference method. *J Pharmacokinet Biopharm* 1998;26:495–519. <https://doi.org/10.1023/a:1023294632129>.
- [35] Lebrec D, Blanchet L. Effect of two models of portal hypertension on splanchnic organ blood flow in the rat. *Clin Sci (Lond)* 1985;68:23–8. <https://doi.org/10.1042/cs0680023>.
- [36] Bekersky I, Dressler D, Alak A, Boswell GW, Mekki QA. Comparative tacrolimus pharmacokinetics: normal versus mildly hepatically impaired subjects. *J Clin Pharmacol* 2001;41:628–35. <https://doi.org/10.1177/00912700122010519>.
- [37] Girish S, Gupta M, Wang B, Lu D, Krop IE, Vogel CL, et al. Clinical pharmacology of trastuzumab emtansine (T-DM1): an antibody-drug conjugate in development for the treatment of HER2-positive cancer. *Cancer Chemother Pharmacol* 2012;69: 1229–40. <https://doi.org/10.1007/s00280-011-1817-3>.
- [38] Roopenian DC, Akilesh S. FcRn: the neonatal Fc receptor comes of age. *Nat Rev Immunol* 2007;7:715–25. <https://doi.org/10.1038/nri2155>.
- [39] Poon KA, Flagella K, Beyer J, Tibbitts J, Kaur S, Saad O, et al. Preclinical safety profile of trastuzumab emtansine (T-DM1): mechanism of action of its cytotoxic component retained with improved tolerability. *Toxicol Appl Pharmacol* 2013;273: 298–313. <https://doi.org/10.1016/j.taap.2013.09.003>.
- [40] Kagayama A, Tanimoto S, Fujisaki J, Kaibara A, Ohara K, Iwasaki K, et al. Oral absorption of FK506 in rats. *Pharm Res* 1993;10:1446–50. <https://doi.org/10.1023/a:1018967107612>.
- [41] Tsunashima D, Kawamura A, Murakami M, Sawamoto T, Undre N, Brown M, et al. Assessment of tacrolimus absorption from the human intestinal tract: open-label, randomized, 4-way crossover study. *Clin Ther* 2014;36:748–59. <https://doi.org/10.1016/j.clinthera.2014.02.021>.
- [42] Davis JA, Rock DA, Wienkers LC, Pearson JT. In vitro characterization of the drug-drug interaction potential of catabolites of antibody-maytansinoid conjugates. *Drug Metab Dispos* 2012;40:1927–34. <https://doi.org/10.1124/dmd.112.046169>.
- [43] Choi W-G, Park R, Kim DK, Shin Y, Cho Y-Y, Lee HS. Mertansine inhibits mRNA expression and enzyme activities of cytochrome P450s and uridine 5'-diphosphoglucuronosyltransferases in human hepatocytes and liver microsomes. *Pharmaceutics* 2020;12:220. <https://doi.org/10.3390/pharmaceutics12030220>.
- [44] Wong S, Bumbaca D, Yue Q, Halladay J, Kenny JR, Salphati L, et al. Abstract A136: nonclinical disposition, metabolism, and in vitro drug-drug interaction assessment of DM1, a component of trastuzumab emtansine (T-DM1). *Mol Cancer Ther* 2011; 10:A136–7. <https://doi.org/10.1158/1535-7163.targ-11-a136>.
- [45] Li C, Menon R, Walles M, Singh R, Upreti VV, Brackman D, et al. Risk-based pharmacokinetic and drug-drug interaction characterization of antibody-drug conjugates in oncology clinical development: an international consortium for innovation and quality in pharmaceutical development perspective. *Clin Pharmacol Ther* 2022;112:754–69. <https://doi.org/10.1002/cpt.2448>.
- [46] Krop IE, Modi S, LoRusso PM, Pegram M, Guardino E, Althaus B, et al. Phase 1b/2a study of trastuzumab emtansine (T-DM1), paclitaxel, and pertuzumab in HER2-positive metastatic breast cancer. *Breast Cancer Res* 2016;18:34. <https://doi.org/10.1186/s13058-016-0691-7>.
- [47] Noguchi E, Tamura K, Hattori M, Horiguchi J, Sato N, Kanatani K, et al. Trastuzumab emtansine plus pertuzumab in Japanese patients with HER2-positive metastatic breast cancer: a phase Ib study. *Breast Cancer* 2019;26:39–46. <https://doi.org/10.1007/s12282-018-0887-z>.
- [48] Han TH, Gopal AK, Ramchandren R, Goy A, Chen R, Matous JV, et al. CYP3A-mediated drug-drug interaction potential and excretion of brentuximab vedotin, an antibody-drug conjugate, in patients with CD30-positive hematologic malignancies. *J Clin Pharmacol* 2013;53:866–77. <https://doi.org/10.1002/jcph.116>.

## Regular Article

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

Naoko Yoshida,<sup>\*,a,b,c</sup> and Shoei Maeda<sup>c</sup>

<sup>a</sup>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; <sup>b</sup>Society for Medicines Security Research, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan; <sup>c</sup>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.<sup>1)</sup> 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.<sup>2)</sup> 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.<sup>3-7)</sup> 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.<sup>8)</sup> 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.<sup>9)</sup> 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



© 2025 Author(s) BPB Reports applies the Creative Commons Attribution (CC BY) license to works we published. The license was developed to facilitate open access - namely, free immediate access to, and unrestricted reuse of, original works to all types. Under this license, authors agree to make articles legally available for reuse, without permissions or fees, for virtually any purpose. Anyone may copy, distribute, or reuse these articles, as long as the author and original source are properly cited. <https://creativecommons.org/licenses/by/4.0/>



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.<sup>2,5-7,10-13</sup> 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.<sup>3,5-7</sup> 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;<sup>14</sup> reference to the Act on Quality, Efficacy and Safety Assurance of Medical Devices (Pharmaceutical and Medical Device Act);<sup>15</sup> 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	89.58	0.906	0.812
	Falsified	12			
Test	Genuine	31	82.35	0.889	0.779
	Falsified	12			

AUC, area under the receiver operating characteristic curve.

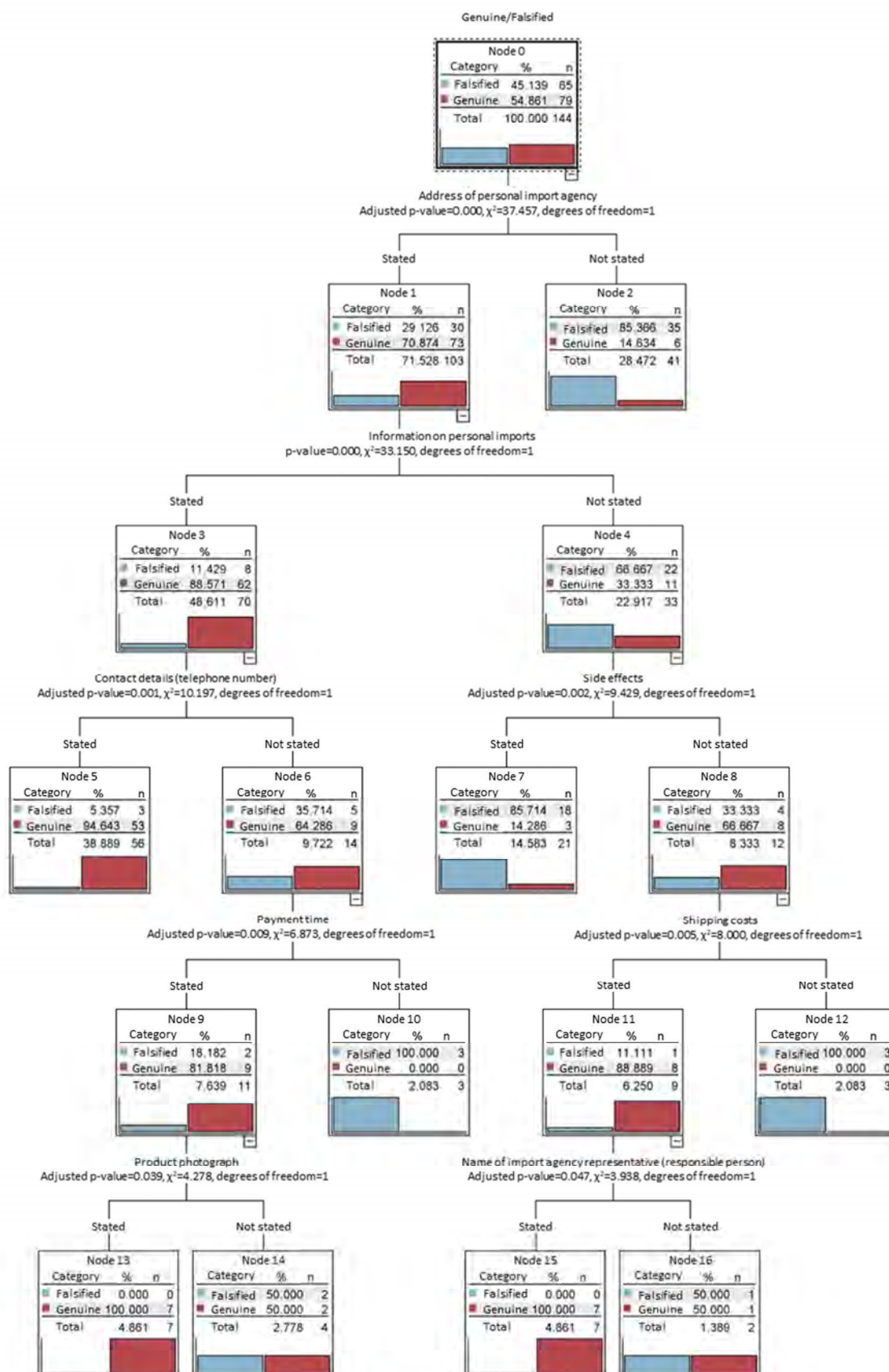
29.0% for products with the manufacturer's name on the outer package and 71.6% for products without the manufacturer's name, which was significantly higher ( $p < 0.01$ ). The rate of falsified medicines was 39.6% for products with a manufacturer's address and 56.2% for products without a manufacturer's 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 sig-

nificant 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.

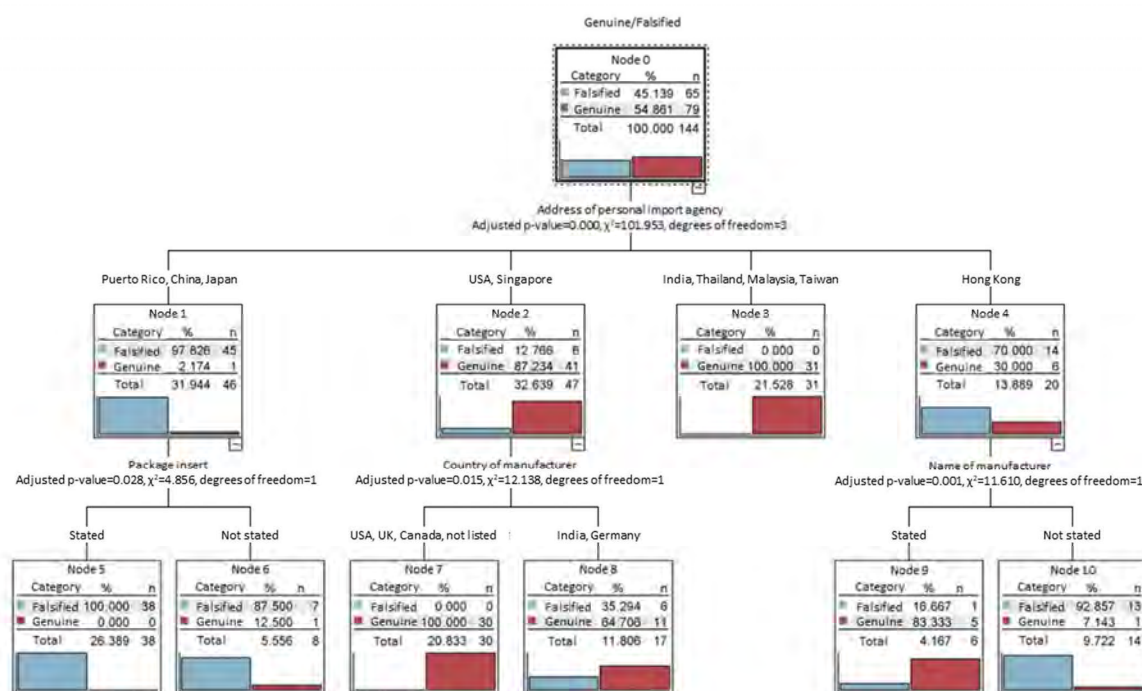


Fig. 2. Tree Diagrams Created from the Information on Each Product Obtained

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.

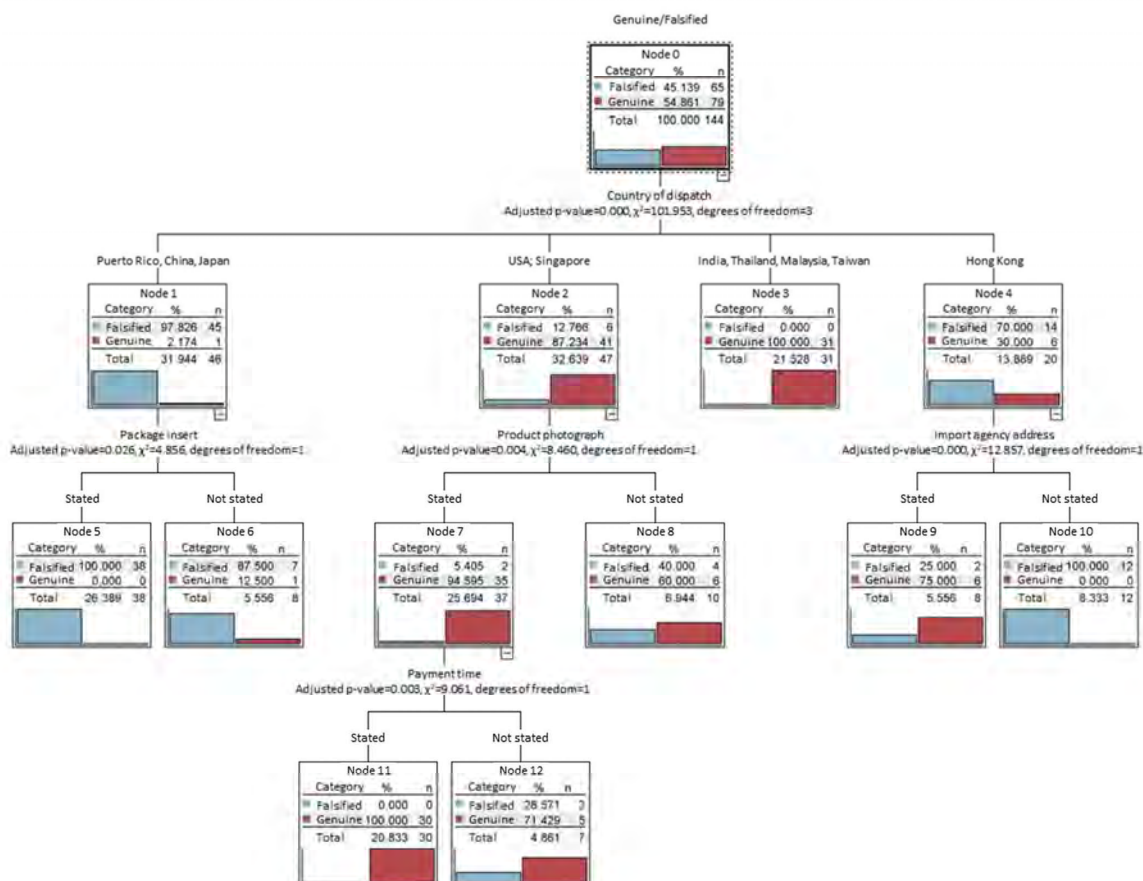
The conditions for a 90% or higher probability of being a genuine medicine were that the product was shipped from the United States, Singapore, Thailand, Malaysia, or Taiwan, or that the product was shipped from India or Hong Kong and the address of the personal import agency was listed. The conditions 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	93.75	0.985	0.97
	Falsified	2			
Test	Genuine	29	94.12	0.925	0.821
	Falsified	2			

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	93.75	0.987	0.973
	Falsified	1			
Test	Genuine	29	94.12	0.925	0.851
	Falsified	2			

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 purchas-

es, 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.<sup>16)</sup>

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,<sup>17)</sup> 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 package are 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.<sup>18,19)</sup> Additionally, non-destructive methods using the latest miniaturized devices is also being developed.<sup>20,21)</sup> 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

- 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](http://apps.who.int/gb/ebwha/pdf_files/WHA70/A70_23-en.pdf), accessed 18 February, 2025.
- 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](http://who.int/entity/medicines/regulation/ssffc/publications/GSMS_Report.pdf), accessed 18 February, 2025.
- 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)
- 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).
- Zhu S, Yoshida N, Kimura K, Matsushita R, Tsuboi H. Falsified vardenafil tablets available online. *J. Pharm. Biomed. Anal.*, **177**, 112872 (2020).
- 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).
- 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).
- 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.
- 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).
- 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](https://mhlw-grants.niph.go.jp/system/files/2019/193041/201925012B_upload/201925012B0011.pdf), accessed 27 November, 2024.
- 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).
- 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).
- 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).
- Act on Specified Commercial Transactions. <https://laws.e-gov.go.jp/law/351AC0000000057>, accessed 4 December 2024.
- Act on Quality, Efficacy and Safety Assurance of Medical Devices. <https://laws.e-gov.go.jp/law/335AC0000000145>, accessed 4 December 2024.
- 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.
- 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.
- 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).
- 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).
- 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).
- 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).

## Regular Article

Distribution Survey of Individual Transactions for Pharmaceutical Products *via* X (Formerly Twitter)Ruoyu Zhang<sup>a</sup> and Naoko Yoshida<sup>\*,a,b,c</sup>

<sup>a</sup>Clinical Pharmacy and Healthcare Sciences, Faculty of Pharmacy, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa 920–1192, Japan; <sup>b</sup>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; and <sup>c</sup>Society for Medicines Security Research, Kakuma-machi, Kanazawa, Ishikawa 920–1192, Japan.

\*Correspondence: [naoko@p.kanazawa-u.ac.jp](mailto:naoko@p.kanazawa-u.ac.jp)

Received February 20, 2025; accepted May 23, 2025

Individual transactions for pharmaceutical products *via* social networking sites are deemed inappropriate distribution routes for pharmaceuticals and related products and may be a guise for business-to-consumer transactions. To prevent risks to public health, we investigated the status of individual transactions involving pharmaceutical products on the main social networking site used in Japan, X (formerly Twitter), to elucidate the actual status regarding individual transactions of pharmaceutical products. To detect posts involving such individual transactions, we used text mining to extract terms from 1389 text posts intercepted by the Counterfeit and Illegal Drugs Information Center that were suspected of being individual pharmaceutical transactions. We found that the Japanese hashtag #Okusuri mogumogu was often used in posts related to transactions involving psychotropic medicines. To gain further insight into the actual transaction process, we purchased two types of medicines among those being distributed *via* X. The pharmaceutical products received included gifts of psychotropic medicines that were not ordered and goods such as snacks and commodities. To evaluate the quality of products received, we conducted visual observation, authentication with the manufacturer, and Raman scattering analysis to distinguish the samples received from legally distributed pharmaceuticals in Japan; we also conducted qualitative and quantitative analyses of the active pharmaceutical ingredients in these products using HPLC. According to the results, the products obtained were unlikely to be falsified. The inappropriate distribution of pharmaceuticals should be continuously and closely monitored to prevent health hazards caused by inappropriate use of these products.

**Key words** individual transaction, social networking site, Raman scattering analysis, inappropriate distribution

## INTRODUCTION

Social networking sites (SNS) are communication tools used to interact with others *via* the Internet. SNS such as X (formerly Twitter) are now widely used to communicate information to others, obtain the latest information, and even purchase products. Because of its convenience, over the past two decades, Internet users have increasingly used SNS to find and share health information.<sup>1)</sup> However, the vast amount of information on the Internet has become difficult to regulate,<sup>2)</sup> and SNS have become a fountainhead for some inappropriate transactions,<sup>3–5)</sup> including the purchase of pharmaceutical products that should be sold through formal distribution channels with license management.

Individual transactions involving pharmaceuticals and other products purchased through the Internet are recognized as an inappropriate distribution route. Many in the general public believe that online pharmacies and social media platforms are safe and feel confident that they can obtain licensed medications online. Research has shown that counterfeit as

well as substandard and falsified medicines are prevalent on these platforms.<sup>6)</sup> If the quality of pharmaceutical products purchased *via* individual transactions is poor or the products are counterfeit, they may cause harm to the health of the purchaser.<sup>7)</sup> To prevent the inappropriate use of pharmaceutical products, active surveillance and guidance are considered necessary from the perspective of pharmaceutical security and quality assurance.

Few cases of individual transactions involving pharmaceutical products on SNS have been comprehensively investigated and appropriate countermeasures have not yet been initiated, despite this being a global problem.<sup>4)</sup> In Japan, individual sales of prescription medicines through free markets or online auction websites or applications have been identified.<sup>8,9)</sup> However, intervention or control measures to deal with such websites make it impossible for sellers to promote these profitable pharmaceutical products so as to attract buyers. Sellers then make greater use of SNS, such as X, to post information about their products, which are mainly related to prescription medicines but also include traditional Chinese herbal medicines,



OTC medicines, as well as medicines not licensed for use in Japan.<sup>5)</sup> The unauthorized trade of pharmaceuticals has the potential to harm the health of purchasers and cause economic losses, which has a negative impact on society. To address this problem, we must understand the risk of individual pharmaceutical transactions on SNS such as X so that immediate measures can be initiated to prevent the occurrence of health hazards. Some posts on SNS have been identified in which psychotropic drugs are being offered. To evaluate the quality of pharmaceuticals obtained *via* X in the surveillance of traded pharmaceuticals, we conducted visual observation, authenticity checks of manufacturers and distributors, quantification of the content of the main pharmaceutical ingredients using HPLC, and Raman scattering analysis to distinguish these from products that are officially distributed in Japan.

In this research, we investigated the actual situation of individual transactions involving pharmaceutical products purchased *via* SNS, an inappropriate distribution channel of pharmaceutical products. We aimed to contribute to improving the monitoring methods of pharmaceutical transactions to prevent public health hazards.

## MATERIALS AND METHODS

**Pharmaceutical Product Selection: Text Mining** To target the frequency of pharmaceutical products offered on X, after excluding overlapping information, we selected 1389 posts from 1790 suspected pharmaceutical trafficking posts from December 2012 to June 2022 on X, using information provided by the Counterfeit and Illegal Drugs Information Center. We extracted individual terms from all postings in the database, and the number of appearances was counted *via* text mining. Text mining, text data mining, or text analytics is the process of deriving high-quality information from text.<sup>10)</sup> Individual word cloud extraction was performed using MATLAB Text Analytics Toolbox (MathWorks Inc., Natick, MA, U.S.A.). We included no displaying words with fewer than four duplicates to eliminate words with a low duplicate frequency (Supplementary Table S1). We also extracted keywords from posts on X that were suspected of involving individual trafficking in pharmaceutical products, according to a manual search, and that were accessible as of September 8, 2022. This process yielded the names of pharmaceutical products that had not appeared in the word cloud (Supplementary Table S2).

### Sample Collection

**Trial Purchase** Trial purchases were carried out to obtain detailed information about individual pharmaceutical product transactions *via* SNS. We contacted accounts on X, playing the role of an anonymous buyer posting about the transfer of two kinds of pharmaceuticals (other than psychotropic pharmaceuticals, products A and B) that were frequently traded, using the words we extracted for the names of pharmaceutical products (Supplementary Table S1) as targets; both of these pharmaceuticals were original products. The trial purchases were conducted with the permission of the Ministry of Health, Labour and Welfare. The information regarding these two pharmaceutical products is not available to the public to prevent the transferee from knowing that

the transaction was carried out for research purposes. The content of our investigation included the specific methods of communication between the buyer and the seller, the purchase price of the product, the shipping method, and the risks involved in the transaction (*i.e.*, problems that may exist in the transaction). To prevent any infringement of the Japanese Pharmaceutical Affairs Law, we checked the product sold, which comprised asking for an image before the product was sent. The trial purchase period was from December 13, 2022 to January 20, 2023 and communication was *via* direct message (DM) or user comments expressing a desire to obtain the product. After negotiating with the holders of accounts that responded to our post, we obtained the target pharmaceuticals through a private transaction. We confirmed the stock, type, and photo of the target pharmaceutical products, inquired regarding the quantity and price, and confirmed the payment and delivery method. The delivery address was a post office box that we had engaged. We forwarded the packages received at the post office box to our university, then collected, sorted, and recorded the acquired products.

**Sample Code Definition** As for the definition of sample codes for each product obtained, a sample code was assigned as follows: transaction number–product code–serial number for each product code. If there was only one sample for the same product code, a serial number for each product code was not assigned. The product code also indicated the difference in dosage formulation, with the first sample of product A in dosage form pattern 1 obtained in the first transaction being indicated as Twi1-A1-1. Product codes for products obtained as a gift contained the letter G.

### Quality Evaluation of Received Products

**Observation of Product Appearance** The appearance of the obtained products was carefully observed and visually compared with the officially distributed products in Japan. The main points checked were whether the (press through packaging) (PTP) was damaged, as well as the printing and identification codes on the packaging and the product itself.

**Authenticity Investigation** To clarify the origin of each product, we prepared a questionnaire including the obtained product information and the product images to send to each manufacturer or to collect through verbal inquiry by contacting the manufacturer.

**Identification: Raman Scattering Spectroscopic Analysis** To identify differences between the obtained products and Japanese commercial products, we conducted Raman scattering spectroscopic analysis.<sup>11)</sup> We used a handheld Raman scattering spectrometer, which nondestructively focuses a laser on the surface of a pharmaceutical product (tablet or capsule) without removing it from the PTP. After fixing with sufficient fixation, the Raman spectra were measured. Except for two official Japanese pharmaceutical products that could not be collected owing to purchasing controls during our study period, the collected samples and the corresponding Japanese commercial products were measured 10 times at different locations on the product itself. An integral count was set for 5 times at the same spot on different sample tablets/capsules, obtaining a total of 17 pharmaceutical product spectra; the average value was then calculated.



**Scattering Conditions** The measurement wavelength range was 200–3000 cm<sup>-1</sup>, and the maximum laser power was 50 mW (compact Raman spectrophotometer PR-1W; JASCO Corporation, Tokyo, Japan). The obtained samples and the Japanese commercial products were assessed, with Raman spectra measured under the following conditions. Laser: high; exposure time: 2 s; integration times: 3; measurement times: 10 for Japanese commercial products to reduce handling errors and 5 times for samples.

**Analysis Tools** From the obtained spectra, characteristic peaks and agreement ratios were compared using JASCO Spectrum Manager Ver. 2.0. Spectrum matching was performed using Panorama Soft ver. 2.0 (S.T. Japan Inc., Tokyo, Japan) to create a library of average spectra values for Japanese commercial products. The average values of sample spectra were used to calculate the percentage of agreement. The search algorithm parameters were 90 minimum evaluation values and scalar products as the algorithm for comparison.

**Active Pharmaceutical Ingredient (API) Testing: HPLC-Photodiode Array (PDA)** To evaluate the quality of the two purchased target pharmaceutical products, we determined the API content using HPLC-PDA. However, we excluded gifts obtained in the trial purchases, which included psychotropic agents. We did not analyze any pharmaceutical products that were not targets of our trial purchase. The analysis methods are not available to the public.

Except for samples that contained fewer than four tablets or capsules, we analyzed a total of 34 samples of target product A and 7 samples of target product B. The API content testing method was based on the United States Pharmacopeia (2018),<sup>12)</sup> as it was not listed in the Japanese Pharmacopeia.

## RESULTS AND DISCUSSION

**Current Situation of Individual Pharmaceutical Transactions via X** We first identified 568 unduplicated posts from among 1389 posts suspected of being illicit pharmaceutical transactions, based on the specific content of the posts and characteristics of the sending account. We ultimately determined that there was a very high likelihood that the posts were sent for the purpose of making an illicit transaction. We then conducted text mining to obtain the unique characteristics of posts suspected of being pharmaceutical product transactions. Many of the posts where illicit pharmaceutical trafficking was suspected used the hashtags “#Okusuri mogumogu” (663 posts, 47.7%) and “#Okusuriyuzurimasu” (91 posts, 6.6%). The number of words indicating the name of the pharmaceutical product was 186. There were 121 (65.1%) posts indicating psychotropic agents; among the psychotropic agents, the terms DEPAS, SILENCE, and MYSLEE appeared most frequently. This is consistent with the characteristics identified in previous study.<sup>5)</sup> Other than psychotropic agents, pharmaceuticals with a risk of abuse were found, including insomnia medications, analgesics, antitussives, antihistamines, and antiparkinsonian agents (Supplementary Tables S1, S2).

We filtered other hashtags or search terms. #Okusuri mogumogu and #Okusuriyuzurimasu were frequently used in individual transactions involving pharmaceutical products,

based on text mining analysis of posts about pharmaceutical transactions provided by the Information Center for Counterfeit Medicine. Using those keywords and hashtags, the results of a manual search revealed 23 new words/terms in posts suspected of involving individual transactions (Supplementary Table S2). The type of pharmaceutical that appeared most frequently in posts on X involving suspected individual pharmaceutical transactions was psychotropic agents. In such transactions, if the product is shipped from overseas, this is considered as personal import. In the case of psychotropic transactions, the seller may be in violation of the Narcotics and Psychotropic Substances Control Law. Inappropriate distribution of these pharmaceutical products carries a high risk of health hazards and criminal activity.

A list of target products to be purchased was compiled by combining the results obtained through text mining and manual searches on X. During the purchase period in this study, 34 posts were found regarding the transfer of target pharmaceutical products in the trial purchase. All sellers in trial purchases were contacted; of these, 17 (50%) transactions were completed.

**Transaction Details** All target pharmaceutical products and gifts (which were not requested for purchase) that we received as part of the trial purchases were all distributed in Japan. The main flow of transactions was as follows: we played the role of a buyer in these transactions and contacted accounts that posted information about individual pharmaceutical transactions on X *via* user comments or DMs. In these transactions, the transferor received a confirmation of remittance from the transferee; then, the transferor contacted the transferee regarding shipment. None of the transactions required a prescription to be presented, nor was the identity of the buyer verified through personal identification or other means. In all cases, the transaction price was negotiated to be higher than the market price; it is possible that the rate of completed transactions would be lower if a lower price had been offered by the buyer (in this case, the authors). Additionally, the transferor may have obtained the medicine at an even lower price, which would be more profitable for them. In all cases, the transferee was able to choose the method of payment during the negotiations. In Fig. 1, trackable postal service (14 cases) was the most commonly used shipping method, followed by use of a Japanese transport company (2 cases) linked to the free market site, and Japan regular mail (1 case). In most cases, the sender's information was inadequately described: no name, no address or only the name of the prefecture, and no telephone number. In some cases, the product was sent as a shipment from the transferee. However, we confirmed with the logistics providers or services used that all the shipments were sent domestically. In no case were the actual contents (pharmaceuticals) listed in the “Contents” column of the shipping invoices. In some cases, the names of the goods were not listed. The specific product names listed and the number of cases are shown in Fig. 2. Confidentiality of all information is considered to be a possible motivator for obtaining pharmaceuticals through individual transactions.

Target pharmaceutical products were obtained *via* 17 completed transaction cases. In one case, a generic product

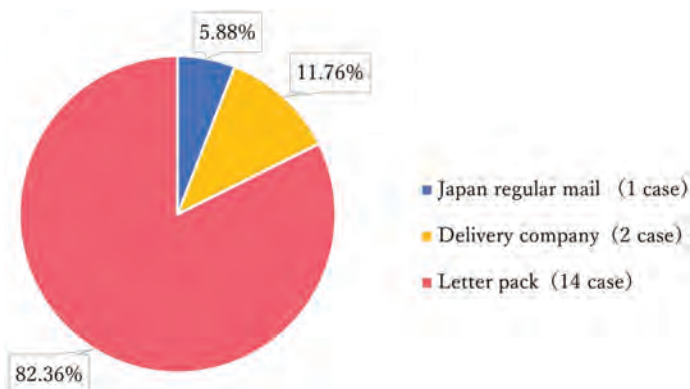


Fig. 1. Shipping Methods Used in All Purchases

Letter pack is a service in which packages of A4 size up to 4kg, including correspondence, can be sent throughout Japan at a flat rate ([https://www.post.japanpost.jp/service/letterpack/index\\_en.html](https://www.post.japanpost.jp/service/letterpack/index_en.html)).

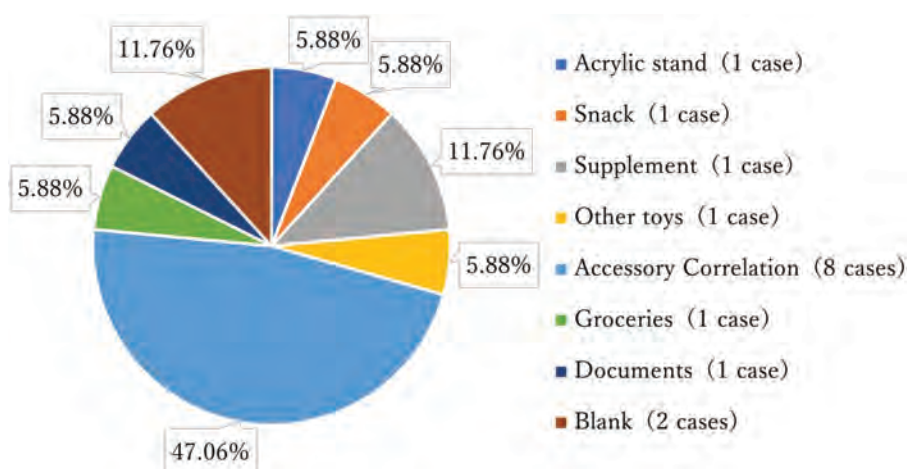


Fig. 2. Item Names on Shipping Invoices

was delivered, although we had requested for the original product; this highlights the risk of fraud during such online transactions. All products were for the Japanese market. Some products that had not been requested were delivered as a gift, which included snacks and daily necessities in some cases. These gifts also included eight different types of psychotropic agents and two types of cough suppressants. In two cases, the number of items ordered did not match the number received. In one sample, some of the packaging was damaged, and two samples had a strong tobacco smell. In one case, although the picture we confirmed in advance showed pillow packaging, the product was delivered without pillow packaging, only PTP sheets. In no case were instructions for usage and dosage included in the package. In one case, the product was not received because the shipping address was not correctly written on the shipping invoice. Although we contacted the transferor, we lost contact with them and were therefore unable to obtain the products. This highlights the risk in these types of transactions that money is paid but the goods are never received. In the case of person-to-person transactions, especially when the transferor uses a disposable X account, there is no way to address any problems if the transferor does not respond.

During negotiations, in many cases, the transferor wished to change to another SNS such as LINE or Telegram. After changing the contact method from DMs to another SNS, there were cases in which the seller requested us to delete the previous DMs. In some cases, our X account was deemed by the seller to be a disposable account and the transaction was refused. In these negotiations, we found that transferors were often vigilant and may have shared transaction information about transferors and other parties with each other. Upon rechecking the accounts *via* which transactions had been completed, we found that five accounts had been deleted or suspended either by the transferee or by X (the administrator or auto patrol), and posts related to previous transactions had been deleted. However, we also confirmed that some accounts remained active and continued to post their items (Fig. 3).

In some cases, transferees were encouraged to carry out these transactions on a regular basis. Some transferees attempt to purchase large quantities of pharmaceutical products. After completing the purchase, we were asked by some sellers to exchange contact information with them, such as *via* Line, so as to become a long-term buyer of their products. We declined these invitations but understood that long-term buyers might maintain long-term deals with sellers in this way even if the

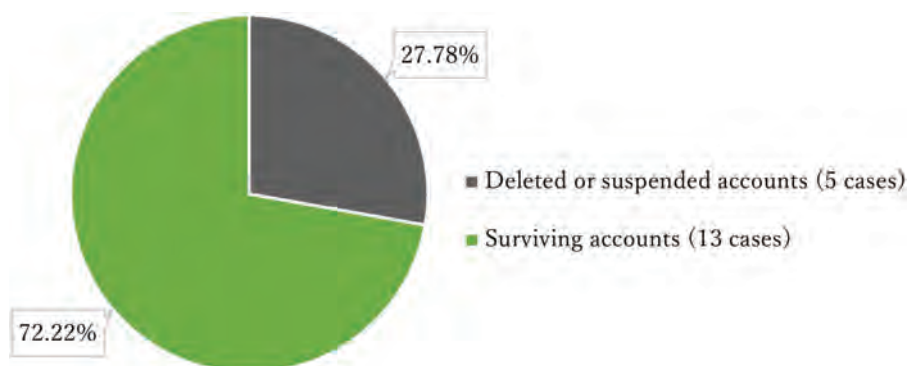


Fig. 3. Account Status for Completed Transactions

seller's account on X is canceled owing to being reported or self-deletion. The sale of prescription drugs to long-term buyers yields benefits for them, such as income stability even if they do not post new products for sale on X. Additionally, such long-term relationships and transactions will not be detected by SNS security systems or administrators. These sellers may even increase their income *via* referrals from buyers to new customers. It was impossible to determine how the transferor obtained the pharmaceuticals. Some sellers with whom transactions were concluded continued to post content offering pharmaceutical products, and others offered to trade in large quantities or on a regular basis, suggesting that these sellers may have been fronts for business transactions. Additionally, psychotropic agents were received as gifts in some cases, highlighting the risk of unintentional acquisition of these controlled substances.

To minimize the inappropriate distribution of pharmaceuticals and other products and prevent harm to the health of the public, it is extremely important to monitor and provide guidance to buyers regarding these online transactions by raising awareness about the risks involved.

### Quality Evaluation

**Visual Observation** In a visual comparison of each of the obtained products with the officially distributed products in Japan, no obvious differences were found, except that some of the products received *via* X as free gifts had damaged PTP sheets. The lot or serial number and expiration date were verified on the obtained pharmaceutical product packages (purchased as well as gifted products). None of the products that had expiration dates stamped on them were found to be past their expiration date. Because these sellers are not licensed to sell pharmaceutical products, only a small number of pharmaceutical products were received as gifts. Among these gift products, the API may deteriorate under inappropriate storage conditions, affecting the overall quality of the product and its dissolution properties. It is also possible that an unsealed PTP package indicates that the original tablet or capsule had been replaced with an unknown product. The API content of these products could not be accurately quantified owing to an insufficient number of tablets or capsules obtained as gift samples; however, it is reasonable to assume that quality problems may be present owing to damaged packaging and for other reasons. Obtaining and consuming such pharmaceutical products is

very dangerous from a health and hygiene standpoint because products may be received that were not ordered, are expired, or have quality problems. The quality of inappropriately distributed pharmaceutical products that deviate from regular distribution routes is not guaranteed, presenting a health hazard owing to poor-quality pharmaceutical products.

**Authentication** For all products obtained, we investigated their authenticity with the manufacturers. Product serial numbers and expiration dates were confirmed with each manufacturer. One of these companies replied verbally; all others responded by returning our questionnaire or *via* e-mail. All serial numbers were substantiated, and the expiration date for one sample was confirmed to be prior to the order date.

In the authenticity investigation, we asked the manufacturers to analyze the samples we obtained; however, all of the manufacturers declined our request. Instead, they compared photographs of our pharmaceutical product samples with the lot numbers we provided to check whether the medicines we received were indeed their products. If someone falsifies the lot number or replaces the package contents but uses genuine product packaging, questionnaires and emails cannot be used to correctly identify problematic products. Thus, as a result of our investigation with the manufacturing companies, we could not accurately determine the authenticity of pharmaceutical products obtained and whether these were products officially distributed in Japan.

**Identification Analysis** We analyzed and compared the data collected, with the greatest possible control of manual measurement errors. No differences were found between samples that could be measured *via* dissimilarity identification using Raman scattering analysis and the genuine Japanese marketed product, in terms of the spectra and intensities, such as Twi12-G5 and Twi24-G14. The spectra generally had high coincidence rates (Supplementary Fig. S1). The software calculated that the coincidence rates of the Raman spectra for measured samples in comparison with those of the Japanese commercial products were all 99% or higher (Table 1). The high rate of concordance obtained indicated a low possibility of poor quality among the pharmaceutical products received in the trial purchase, with most within the expiration date.

Raman scattering analysis has long been used in pharmaceutical monitoring and analysis. There is a high possibility that the Raman intensities detected may be different if

Table 1. Quality Assessment Results

#	Sample code	Coincidence rate		HPLC
		Measuring location 1	Measuring location 2	Results (AVE)
1	Tw12-A7	99.22	—	100.00
2	Tw15-B	99.98	—	99.84
3	Tw15-A5	100.00	—	100.00
4	Tw15-A1	99.80	—	100.00
5	Tw15-G1	—	—	—
6	Tw15-G2	99.98	—	—
7	Tw16-B	99.98	—	100.01
8	Tw17-A64	99.89	—	99.99
9	Tw18-A1-1	99.61	—	100.05
10	Tw18-A1-2	99.69	—	100.04
11	Tw18-A1-3	99.69	—	100.04
12	Tw18-A1-4	99.27	—	100.04
13	Tw18-A1-5	99.75	—	100.04
14	Tw19-A2	99.93	99.89	100.00
15	Tw19-A1	99.64	—	100.04
16	Tw19-G3	99.98	—	—
17	Tw19-G4	99.97	—	—
18	Tw112-A1	99.81	—	100.04
19	Tw112-G5	99.39	—	—
20	Tw114-B	99.99	—	100.21
21	Tw114-G6	99.97	—	—
22	Tw114-A3	100.00	—	100.00
23	Tw114-G7	—	—	—
24	Tw119-A2-1	99.95	99.80	100.00
25	Tw119-A2-2	99.91	99.99	100.00
26	Tw119-A2-3	99.98	99.98	100.00
27	Tw122-A2-1	99.65	99.88	100.00
28	Tw122-A1-1	99.76	—	100.04
29	Tw122-A1-2-1	99.83	—	100.04
30	Tw122-A3-1	99.44	—	100.00
31	Tw122 A3-2	99.53	—	100.00
32	Tw122-A2-2	99.97	99.92	100.00
33	Tw122-A1-3	99.77	—	100.04
34	Tw122-A4	100.00	—	100.01
35	Tw124-A3	99.83	—	100.00
36	Tw124-G8-1	99.88	—	—
37	Tw124-G8-2	99.90	—	—
38	Tw124-G9	99.93	—	—
39	Tw124-G10	99.87	—	—
40	Tw124-G11	99.84	—	—
41	Tw124-G12	99.44	—	—
42	Tw124-G13	99.38	—	—
43	Tw124-G14	99.01	—	—
44	Tw125-A18-1	99.63	—	100.04
45	Tw125-A1-2	99.13	—	100.03
46	Tw128-A1	99.74	—	100.04
47	Tw129-A20-1	99.66	99.98	100.00
48	Tw129-A2-2	99.79	99.95	100.01
49	Tw130-A1-1	99.43	—	100.00
50	Tw130-A1-2	99.78	—	100.00
51	Tw134-A1-1	99.54	—	100.04
52	Tw134-A1-2	99.43	—	100.04
53	Tw134-B-1	99.91	—	97.27
54	Tw134-B-2	99.89	—	99.88
55	Tw134-B-3	99.96	—	99.39
56	Tw134-B-4	99.96	—	101.76

“—” Not measured; tablets and single-color capsules were measured at one location. In the case of two-color capsules, the coincidence rate was measured at two locations for matching. Only two target pharmaceutical products for trial purchase in this study were measured using HPLC-PDA. \*Raman coincidence rate evaluation reference range: 99.00–100.00. \*Reference range of the pharmacopoeia for active pharmaceutical ingredients in acceptable products: 98.00–102.00. Because the actual operation has been modified based on the pharmacopoeia, this range is for reference only. HPLC-PDA: HPLC-photodiode array.

they were not properly stored prior to sale on the SNS. For example, expired medicines lose quality in their crystal or powder components or undergo changes in their properties (e.g., density) owing to incorrect storage, which may lead to different spectra,<sup>11)</sup> even if they are genuine Japanese commercial products. Therefore, results using this method

can only serve as a preliminary reference, and analysis with greater precision is necessary if more detailed information is needed.

**API Testing** In HPLC-PDA determination, we confirmed that the content of the main active ingredient was 100%; all reached the standard in all measured samples of target



products A and B. In this study, the number of samples was fewer than the 10–20 tablets/capsules needed for reliable determination of the API content by HPLC, in accordance with the Japanese Pharmacopoeia. Therefore, a minimum of three tablets was used for content measurement, which yielded acceptable results in API analysis; however, these results may still be substandard owing to the limited number of tablets/capsules. Although the quality of products obtained in our trial purchase did not raise suspicions of counterfeit or substandard pharmaceutical products, in the worst case, there is a high possibility that these products may pose a health hazard to purchasers who receive products for which there are no safety measures in place. The quality of such inappropriately distributed pharmaceutical products that deviate from regular marketing channels is not guaranteed, and there is a risk of harm to health owing to poor-quality pharmaceuticals and other products.

## CONCLUSION

In this study, the current status regarding transactions between individuals *via* X involving pharmaceutical products in Japan was revealed to mainly include psychotropic agents, analgesics, antitussives, and medicines for allergic diseases. Through our trial purchase survey, we found some situations in which the target product was not received, despite being paid for, and some items received were pharmaceutical products for which the expiration date had passed. The inappropriate distribution of these pharmaceutical products carries a high risk of causing harm to human health as well as involving criminal activity. It is necessary to continue to monitor posts on SNS like X regarding individual transactions of such pharmaceutical products and to alert the public about the risks of engaging in such transactions.

**Acknowledgments** This work was supported by JST SPRING [Grant No. JPMJSP2135] and the Ministry of Health, Labour and Welfare of the Japan Pharmaceuticals and Medical Devices Regulatory Science Policy Research Program [Grant No. JPMHLW21KC1007]. We would like to express our gratitude to the Counterfeit and Illegal Drugs Information Center for providing us with the submitted information, and to JASCO Corporation for conducting the Raman scattering analysis.

**Conflict of Interest** The authors declare no conflict of interest.

**Supplementary Materials** This article contains supplementary materials.

## REFERENCES

- 1) Chou WS, Oh A, Klein WMP. Addressing health-related misinformation on social media. *JAMA*, **320**, 2417–2418 (2018).
- 2) Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. *J. Med. Internet Res.*, **23**, e17187 (2021).
- 3) Alrogy W, Jawdat D, Alsemari M, Alharbi A, Alasaad A, Hajeer AH. Organ trade using social networks. *Saudi J. Kidney Dis. Transpl.*, **27**, 971–976 (2016).
- 4) Mackey TK, Liang BA. Global reach of direct-to-consumer advertising using social media for illicit online drug sales. *J. Med. Internet Res.*, **15**, e105 (2013).
- 5) Hakariya H, Ikejiri T, Yokoyama N, Saito Y. A survey of illegal medication trading through Twitter in Japan. *Yakugaku Zasshi*, **142**, 901–904 (2022).
- 6) Moureaud C, Hertig J, Dong Y, Muraro IS, Alhabash S. Purchase of prescription medicines *via* social media: a survey-based study of prevalence, risk perceptions, and motivations. *Health Policy*, **125**, 1421–1429 (2021).
- 7) Fincham JE. Negative consequences of the widespread and inappropriate easy access to purchasing prescription medications on the Internet. *Am. Health Drug Benefits*, **14**, 22–28 (2021).
- 8) Kishimoto K, Takeuchi T, Fukushima N. Present status of displaying pharmaceutical products for sale on flea market applications for smartphones and the responses to illicit selling by service providers. *Yakugaku Zasshi*, **137**, 1533–1541 (2017).
- 9) Ohtani H, Imaoka A, Akiyoshi T. Current status of illegal trade in pharmaceutical products on Internet auction sites in Japan and responses of site administrators to such transactions. *Yakugaku Zasshi*, **135**, 529–534 (2015).
- 10) Hearst MA. “What is text mining?”: <<https://www.ibm.com/topics/text-mining>>, accessed 20 June, 2024.
- 11) Buckley K, Matousek P. Recent advances in the application of transmission Raman spectroscopy to pharmaceutical analysis. *J. Pharm. Biomed. Anal.*, **55**, 645–652 (2011).
- 12) Non-publication of pharmaceutical ingredients, *United States Pharmacopeia. USP 41-NF 36*, United States Pharmacopeial Convention, Rockville, MD, U.S.A. (2018).



# OPEN Medicine quality assessment in Nepal using semi randomised sampling and evaluation of a small scale dissolution test and portable Raman spectrometers

Robin Schreiber<sup>1</sup>, Md. Ahsanul Haque<sup>2,3</sup>, Mohammad Sofiqur Rahman<sup>2,4</sup>,  
Bhupendra Kumar Poudel<sup>5</sup>, Balmukunda Regmi<sup>6</sup>, Kazuko Kimura<sup>2,7</sup> & Naoko Yoshida<sup>7,8</sup>✉

Substandard and falsified medicines threaten global health and require reliable data and screening technologies to combat their spread. This study examined the quality of 241 samples containing azithromycin, cefixime, esomeprazole and losartan collected from licenced private vendors in the Saptari (121 samples; convenience sampling) and Kathmandu (120 samples; randomised sampling) districts of Nepal. Nearly 10% (24 samples; 95% CI 6.5–14.5) of samples failed pharmacopoeial quality analysis and were classified as 'substandard' or 'probably substandard'. No falsified medicines were identified. Small-scale dissolution acceptance criteria were applied to all 20 three-unit combinations of 213 samples tested in the first stage of the United States Pharmacopoeia dissolution test. Approximately 1% of these results were false positives when compared with the final United States Pharmacopoeia dissolution test results, suggesting the test's usefulness in encouraging dissolution testing in resource-limited contexts. In the narrow sense of presence/absence, two portable Raman spectrometers reliably detected azithromycin, cefixime and losartan in most samples based on effective methods for detecting falsified medicines; however, none of the substandard samples were identified. The findings suggest that falsified medicines are less prevalent in Nepal and the surrounding region than suggested by regional concerns about Nepal and global concerns about low- and middle-income countries. Nevertheless, the Nepalese government should continue to ensure the quality of all distributed medicines.

**Keywords** Quality, Substandard and falsified medicines, Nepal, Screening technology, Dissolution test, Portable Raman scattering analysis

Nearly one-third of the global population still lacks access to essential medicines, especially in low- and middle-income countries (LMICs)<sup>1–3</sup>. The full realisation of the human right to health remains a major global challenge despite the United Nations member states' adoption of target 3.8 of the Sustainable Development Goals in 2015<sup>4–6</sup>. The target links access to safe, effective, high-quality and affordable essential medicines for all people with the realisation of universal health coverage<sup>4,5,7</sup>. However, improvements in health service coverage have stagnated since the Sustainable Development Goals officially came into force<sup>5</sup>.

<sup>1</sup>Clinical Pharmacy and Healthcare Sciences, Division of Pharmaceutical Sciences, Graduate School of Medical Sciences, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan. <sup>2</sup>Graduate School of Medical Sciences, Medi-Quality Security Institute, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan. <sup>3</sup>Stark Neuroscience Research Institute, Indiana University, Indianapolis, IN 46202, USA. <sup>4</sup>Department of Clinical Pharmacy, School of Pharmacy, University of California, 521 Parnassus Ave, San Francisco, CA 94143, USA. <sup>5</sup>Western Regional Hospital, Pokhara Academy of Health Sciences, Pokhara, Nepal. <sup>6</sup>Department of Pharmacy, Maharajgunj Medical Campus, Institute of Medicine, Tribhuvan University, Kathmandu, Nepal. <sup>7</sup>Society for Medicines Security Research, 4F Venture Business Laboratory, Kanazawa University Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan. <sup>8</sup>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. ✉email: naoko@p.kanazawa-u.ac.jp

The World Health Organisation (WHO) estimates that about one in ten (10.5%) of medicines in LMICs are substandard and falsified (SF) medicines<sup>8</sup>. The World Health Assembly 2017 agreed on the following definition of SF medicines as mutually exclusive categories: 1. 'Substandard (or 'out-of-specification') medicines are authorised medical products that fail to meet either their quality standards, specifications, or both' and 2. 'Falsified medicines are medical products that deliberately or fraudulently misrepresent their identity, composition, or source'<sup>9,10</sup>. SF medicines can lead to harm, treatment failure and even death of the consumer<sup>10–15</sup>. These medicines further cause a loss of trust in medicines and healthcare stakeholders and pose a significant economic burden<sup>9,10,12–16</sup>. The cost of SF medicines is estimated at USD 10–200 billion annually<sup>14,16</sup>. The available data suggest that substandard medicines may be six times more prevalent as a global average than falsified medicines in LMICs<sup>17</sup>. However, international media focus more on the reporting of falsified high-priced products<sup>18</sup>. Nevertheless, the spread of falsified medicines has been described as a pandemic and has been reported more frequently in recent years<sup>19</sup>.

The South Asian country of Nepal, with a population of roughly 30 million people, is landlocked in the Himalayas<sup>20</sup>. Despite a significant and rapid reduction in general poverty in recent years, Nepal remains one of the slowest-growing economies in Asia and is classified as a lower-middle-income country<sup>3,21</sup>. The country is bordered by China and India, both of which were named 'global primary producers of falsified medicines' by the Organisation for Economic Co-operation and Development (OECD) in 2020<sup>22</sup>. While Nepal maintains limited, monitored trade with China, its open, visa-free border with India allows unrestricted movement of people and goods<sup>23</sup>. Nepal relies on imports of medical products, raw materials and packaging materials from both neighbour countries, according to the Department of Drug Administration (DDA; March 2023).

The spread of SF medicines in a country is known to be strongly inversely linked to the strength of its healthcare system<sup>18,24,25</sup>. Nepal's medicine regulatory system is currently classified at WHO maturity level 1, i.e. as having limited capacity<sup>26,27</sup>. Additionally, Nepal's hot and humid climate, which is categorised as climatic zone IVb by international pharmaceutical standards<sup>28,29</sup>, requires suitable storage conditions to preserve the quality of medicines. Reliable epidemiological data are required to effectively combat SF medicines and their dissemination, but such data remain limited worldwide<sup>10,14,17,30–32</sup>. To date, no study has assessed the pharmaceutical quality of medicines in Nepal's private sector using both pharmacopoeial tests and field-based screening tools.

In LMICs, including Nepal, quality tests such as the dissolution test are often neglected given capacity limitations. However, they play a critical role in preventing adverse effects from the use of poor-quality medicines such as treatment failure and increased antimicrobial resistance<sup>9,10,12–16,33</sup>. Dissolution and drug-release tests are expected to play a further important role in the future regulation of medicine quality<sup>34</sup>, necessitating even more worldwide testing capacities. A novel small-scale dissolution test recently developed by Rahman and Yoshida et al. reduces required testing capacity compared to the USP dissolution test<sup>35</sup>. To assess the dissolution performance of a sample, it applies modified compliance criteria of the USP dissolution test to three tested units: the average dissolution rate (Q) of the three units is not less than or equal to Q + 6%, and no individual Q is less than or equal to Q + 2%, while maintaining the USP test conditions<sup>35</sup>. However, the small-scale dissolution test has not yet been field tested, and no data are available on its comparability to the USP dissolution test.

Screening technologies to identify SF medicines on-site are urgently needed to safeguard pharmaceutical supply chains, especially in resource-limited contexts. They can also help reduce the number of samples requiring costly laboratory analysis. Portable Raman spectrometers are analytical devices that use laser-based spectroscopy to identify the molecular composition of substances based on their Raman scattering signatures. These instruments are relatively inexpensive, potentially non-destructive and have previously been shown to detect falsified medicines and unlicensed products<sup>36–41</sup>, making them potentially useful screening technologies. However, they have so far been unable to detect substandard medicines, limiting their broader applicability.

This semi-randomised, cross-sectional study in the Saptari (convenience sampling) and Kathmandu (randomised sampling) districts investigated the quality of medicines sold by Nepalese pharmacies and legal retail shops of the private sector (hereinafter referred to as 'licensed vendors'). The study focused on pharmaceutical products that contained one of four selected active pharmaceutical ingredients (APIs): azithromycin (AZM), cefixime (CFIX), esomeprazole (ESM) and losartan (LST). The results were intended to identify improvements for manufacturers, the Nepalese regulatory authorities and the Nepalese government. The findings could also support the development of effective strategies to combat the problem of SF medicines and improve and ensure medicine quality in Nepal. In addition, the study assessed the applicability of a small-scale dissolution test and two portable Raman spectrometers as potential screening technologies for the effective detection of SF medicines.

## Methods

### Chemicals and materials

The AZM dihydrate standard (#PHR1088, LOT: LRAC6480, purity: 94.6% as dihydrate) and the LST potassium standard (#PHR1602, LOT: LRAC5655, purity: 99.9%) were purchased from Sigma-Aldrich Chemie GmbH (Steinheim, Germany). The omeprazole standard (#217087-09-7, LOT: J0H412, purity: 99.9%) was purchased from The United States Pharmacopoeial Convention (USP; North Bethesda, MD, USA). The CFIX trihydrate standard (#034-25261, LOT: SKE6566, purity: 99.0%), lansoprazole (#129-05863, LOT: CAE1292, purity: 99.8%), metronidazole (#139-14931, LOT: DLG2869, purity: minimum 98.0%), benzophenone (#023-01072, LOT: CTP1333, purity: minimum 98.0%), diclofenac sodium (#043-22851, LOT: LEM5673, purity: minimum 98.0%), K<sub>2</sub>HPO<sub>4</sub>, H<sub>3</sub>PO<sub>4</sub>, MeCN and MeOH were purchased from FUJIFILM Wako Pure Chemical Corporation (Osaka, Japan). All of the solvents were high-performance liquid chromatography (HPLC)-grade. NaH<sub>2</sub>PO<sub>4</sub> × 2 H<sub>2</sub>O, Na<sub>2</sub>HPO<sub>4</sub>, Na<sub>3</sub>PO<sub>4</sub> × 12 H<sub>2</sub>O, 5M HCl, and 5M and 10M NaOH solutions were purchased from Nacalai Tesque, Inc. (Kyoto, Japan). KH<sub>2</sub>PO<sub>4</sub> was purchased from FUJIFILM Wako Pure Chemical Corporation (Osaka, Japan) and Nacalai Tesque, Inc. (Kyoto, Japan).

The AZM standard product Zithromac 250 mg (tablets) was purchased from Pfizer Inc. (New York, NY, USA), the CFIX standard product Cefspan 100 mg (capsules) was purchased from Choseido Pharmaceutical Co., Ltd. (Tokushima, Japan), the ESM standard product Nexium 20 mg (capsules) was provided by AstraZeneca PLC (Cambridge, UK) and the LST standard product NU-LOTAN 50 mg (tablets) was purchased from Organon & Co. (Jersey City, NJ, USA).

### Devices and instrumentation

HPLC analysis was performed at Kanazawa University in Kanazawa, Ishikawa, Japan using HPLC systems from three manufacturers. A summary of the chromatographic conditions is given in Table 1. For the analysis of AZM and LST, a system by Hitachi High-Tech Science Corporation (Tokyo, Japan) that was equipped with a Hitachi Elite LaChrom Organizer, an Elite LaChrom L-2130 pump, an Elite LaChrom L-2200 autosampler, an Elite LaChrom L-2300 column oven and an Elite LaChrom L-2455 diode array detector was used. AZM was analysed using GL Sciences Inc.'s (Tokyo, Japan) InertSustain C18 4.6 mm I.D. × 150 mm (5-μm particle) packing L1 column. The system for analysing LST was equipped with the Mightysil RP-18 GP 4.6 mm I.D. × 150 mm (5-μm particle) Cica reagent column by Kanto Chemical Co., Inc. (Tokyo, Japan). ChromAssist Data Station Version 3.0 by Hitachi High-Tech Science Corporation (Tokyo, Japan) was used as a software/digital workstation for data acquisition. For the analysis of CFIX, a system by Shimadzu Corporation (Kyoto, Japan) was equipped with a CBM-20A prominence communications bus module, an LC-40D pump (pump A), an LC-10AD pump (pump B), a SIL-10A XL auto-injector, a CTO-20AC prominence column oven and an SPD-M20A prominence photodiode array detector. CFIX was analysed using the Shim-pack CLC-ODS (M) 4.6 mm I.D. × 150 mm (5-μm particle) RP18 column by Shimadzu Corporation (Kyoto, Japan). Shimadzu Corporation's (Kyoto, Japan) Shimadzu LabSolutions was used as a software/digital workstation for data acquisition. For the analysis of ESM, a system by Jasco, Inc. (Tokyo, Japan) equipped with an LC-Net II/ADC instrument, a PU-4180 pump, an AS-4550 autosampler, a CO-1560 column oven, a UV2075 Plus UV/VIS detector, a MD-2018Plus photodiode array detector, and a CD-2095 chiral detector was used. ESM was analysed using the Phenomenex 4.6 mm I.D. × 150 mm (5-μm particle) NX-C18 column by Phenomenex Inc. (Torrance, CA, USA). ChromNAV 2 by Jasco, Inc. (Tokyo, Japan) was used as a software/digital workstation for data acquisition.

Dissolution tests were performed using two Toyama NTR-VS6P dissolution testers and one Toyama NTR-6100 dissolution tester by Toyama Sangyo Co., Ltd. (Osaka, Japan) that were each equipped with six 1000 mL dissolution vessels and either six USP dissolution apparatus type 1 baskets (for CFIX) or six USP dissolution apparatus type 2 paddles (for AZM, ESM and LST) in conformance with USP 41.

Raman scattering analysis was conducted using a C13560 ultra-compact, portable Raman spectrometer (96 mm × 14.5 mm × 60 mm; 90 g) manufactured by Hamamatsu Photonics K.K. (Shizuoka, Japan) with a silicon substrate provided by the manufacturer<sup>38</sup> and an Inspector500 portable Raman spectrometer manufactured by SciAps Inc. (Laramie, WY, USA) with a polystyrene standard provided by the manufacturer. C13560 operation software and NuSpec Pro software were supplied by Hamamatsu Photonics K.K. (Shizuoka, Japan) and SciAps Inc. (Laramie, WY, USA), respectively, and corresponding drivers were installed on a personal computer before the analysis. The Unscrambler X 10.5 by CAMO Software AS (Oslo, Norway) was used for data analysis and visualisation.

The calculation of HPLC results, statistical analysis, and data visualisation were performed using Excel MSO 365 (Microsoft Corp., Redmond, WA, USA).

API	Column	Injection volume	Mobile phase	Flow rate	Oven temperature	Detection wavelength
AZM	InertSustain C18 4.6 mm I.D. × 150 mm (5-μm particles) L1 C18 column GL Sciences Inc (Tokyo, Japan)	20 μL	35% of A: pH 7.5 buffer 65% of B: acetonitrile (v/v)	0.9 mL/ min	45 °C	210 nm
CFIX	Shim-pack CLC-ODS (M) 4.6 mm I.D. × 150 mm (5-μm particles) C18 column Shimadzu Corporation (Kyoto, Japan)	10 μL	72.5% of A: pH 6.86 buffer 27.5% of B: acetonitrile (v/v)	1.0 mL/ min	40 °C	288 nm: Cefixime 254 nm: internal standard
ESM	Phenomenex NX-C18 4.6 mm I.D. × 150 mm (5-μm particles) C18 column Phenomenex Inc (Torrance, CA, USA)	10 μL	40% of A: pH 7.3 buffer 60% of B: acetonitrile (v/v)	0.8 mL/ min	40 °C	302 nm
LST	Mightysil RP-18 GP 4.6 mm I.D. × 150 mm (5-μm particles) C18 column Kanto Chemical Co., Inc. (Tokyo, Japan)	10 μL	50% of A: pH 4.0 buffer 50% of B: acetonitrile (v/v)	0.9 mL/ min	35 °C	250 nm

**Table 1.** Overview of the chromatographic conditions.



## Sampling design

### *Active pharmaceutical ingredient selection criteria*

AZM, CFIX, ESM and LST were selected because these medications are commonly sold by Nepalese licenced vendors, they can harm patients if drug quality does not meet pharmaceutical standards, and, except for ESM, they are included in the National List of Essential Medicines Nepal<sup>42</sup>. ESM was selected for its higher market price compared with omeprazole (which is included in the National List of Essential Medicines Nepal) and the resulting anticipated high criminal activity associated with this active ingredient.

### *Sample size selection criteria*

Based on Cochran's formula, with parameters set to  $\pi=0.105$  (the estimated population proportion of SF medicines),  $e=0.05$  (the desired margin of error),  $\kappa=1.96$  (corresponding to a 95% confidence level), the minimum required number of collected samples was 145; however, the target was set at 240 samples—120 samples from each district, including 60 samples of each API. Samplers were permitted to make slight adjustments if necessary.

### *District selection criteria*

The Saptari district was selected given its geographical location near India, its distance to the Nepalese capital city of Kathmandu and as it has a relatively high population density<sup>43</sup>.

The Kathmandu district was chosen because it is Nepal's capital city, is centrally located geographically, is easily accessible for sampling and represents the most populous and economically important district in Nepal.

### *Sampling site selection criteria*

All licenced vendors including wholesalers with affiliated licenced vendors within the selected districts were eligible for sampling.

### *Selection of sampling sites*

For the randomisation of sampling sites within a district, an official registration list of licenced vendors that comprised shop names, addresses and other shop-related information was requested from responsible regional offices in the Saptari and Kathmandu districts. However, no official registration list was available from the national or regional Nepalese government institutions responsible for the Saptari district. Therefore, conducting a list-randomisation of sampling sites in the Saptari district was impossible.

For the Kathmandu district, the official registration list contained more than 3,000 licenced vendors. This list was transferred to Microsoft Excel, formatted, and arranged as a top-to-bottom list. Subsequently, the Microsoft Excel ' $=\text{RAND}()$ ' function was used to generate random numbers in the columns adjacent to the information for each sampling site. A filter was used to sort the rows of random numbers from 1 to  $n$  with the sampling site information, resulting in a randomised table. This randomised table of sampling sites was used as a top-to-bottom priority list.

The priority list, which was used to select the first 360 sampling sites (priority numbers 1 to 360) as the study's target sampling sites, was communicated to the samplers. Samplers were instructed to strictly visit the sampling sites according to the priority list from lower numbers to higher numbers starting from 1. The samplers were divided into two groups: Team C and Team D. Teams C and D were instructed to collect samples from odd- and even-numbered sampling sites, respectively. If a licenced vendor was closed or non-operational, samplers were allowed to visit a nearby licenced vendor, even if the shop was not included in the priority list. However, the next sampling site visited was required to continue following the priority list. Written instructions containing this information were distributed with the priority list to the samplers before sampling started in the Kathmandu district.

### *Sampling preparation*

All the samplers were healthcare workers and pharmacy students who underwent two training sessions. Held before sampling began in the two districts, the sessions covered the terms, criteria and details of the sampling procedure. Physical copies of the training materials were distributed to all the samplers. Researchers supervised the samplers during the sampling process.

In the Saptari district, the samplers reported a low density of licenced vendors. To achieve a sufficient geographical distribution of sampling sites, samplers were split into two groups—Team A and Team B—that travelled to several towns within the Saptari district to collect samples from all licenced vendors that were visually identified along the road. During this process, no licenced vendor that was identified along the road was skipped. During sampling in the Kathmandu district, samplers who were split into Teams C and D reported that some licenced vendors on the priority list were either temporarily or permanently closed or had been renamed, and others were difficult, inefficient or impossible to access given heavy traffic volume; these factors made strict adherence to the priority list difficult. Therefore, the teams collected samples only from the sampling sites on the priority list that were accessible.

### *Sampling terms and criteria*

The sampling terms included the requirements that the same API not be collected twice from the same sampling site and that each sample consists of a minimum of 30 dosage units. In cases in which multiple products containing the same API were available, the cheapest product was acquired. However, samplers were advised to avoid repeated purchases of the same product—particularly those from the same batch—although this practice was not explicitly prohibited.

In the Saptari district, samplers were allowed to collect four samples per sampling site using convenience sampling given the unavailability of an official list of licenced vendors and the low density of licenced vendors. In the Kathmandu district, the collection of two samples per sampling site following a randomised approach was required.

Sampling was conducted using a ‘mystery shopper’ method<sup>44,45</sup> in which samplers were asked to enter the sampling sites in pairs, purchase the samples and collect packaging and leaflets, if available while pretending to be casual customers.

Temperature and humidity were secretly recorded inside and outside the sampling sites using a hygro-thermometer, and the number of pharmacists and workers was counted after the sample was received. Additionally, the samplers were required to query vendors about the number of units of each of the sampled medicines sold monthly at the sampling site. The collected information was documented with other sample-specific documentation in a sampling form (Annex 1) after the sampling site was left. The samples and the corresponding sampling forms were stored together in separate plastic zip lock bags and transported to an air-conditioned car to Tribhuvan University in Kathmandu, Nepal, where they were stored under appropriate storage conditions until transportation to Kanazawa University, Japan. After sample acquisition, the samples were stored under suitable conditions at Tribhuvan University in Kathmandu and Kanazawa University in Kanazawa.

#### *Definition of the sample identification code*

The sample identification codes (IDs) expressed as ‘X-YZ(-2)’ were assigned according to the sampling team, API and a consecutive API count. ‘X’ represented the sampling teams A, B, C or D, with Teams A and B sampling in the Saptari district and Teams C and D sampling in the Kathmandu district. ‘Y’ indicated the labelled API (1: ESM, 2: LST, 3: CFX and 4: AZM), and the API count ‘Z’ was consecutively numbered from 1 to 15 to provide an overview for the sampling team of the number of samples that were collected per API. In addition, Sample B-115-2 was collected from the same sampling site as sample B-115 and was the 16<sup>th</sup> sample collected by sampling Team B in the Saptari district.

#### *Mapping the sampling sites*

The Google Maps (Mountain View, CA, USA) search engine was used to locate individual sampling sites which were indicated on the sampling forms. The identified sampling site locations and areas were mapped in Figs. 1 and 2 using Microsoft Snipping Tool version 2022.2410.21.0 (Microsoft Corp., Redmond, WA, USA) to extract the base map from Google Maps; Microsoft Paint version 11.2410.28.0 (Microsoft Corp., Redmond, WA, USA) to indicate sampling locations and areas; and Microsoft Word MSO 365 version 2411 (Microsoft Corp., Redmond, WA, USA) to add lines, arrows, and text fields.

If samplers documented areas or neighbourhoods which included sampling sites without specifying addresses, the sampling sites were only mapped under specific conditions: the shop identified during the Google search had to have the same shop name in Nepali as indicated on the priority list, its location had to be within the documented neighbourhoods or areas provided by the samplers and the Google Maps representation was required to show an image with the correct shop name of the sampling site for confirmation.

For sampling site locations which could not be located using this method, the documented neighbourhoods or areas were indicated as encircled areas of the Saptari district and hatched areas of the Kathmandu district.

#### *Authentic product acquisition*

Authentic products which corresponded to the collected sample products were purchased directly from manufacturers after sampling and were received before March 14, 2023. Next, the products were transported to Kanazawa University, Japan and stored under suitable conditions. Manufacturers were informed about the study purpose before the purchase or collection of the authentic products.

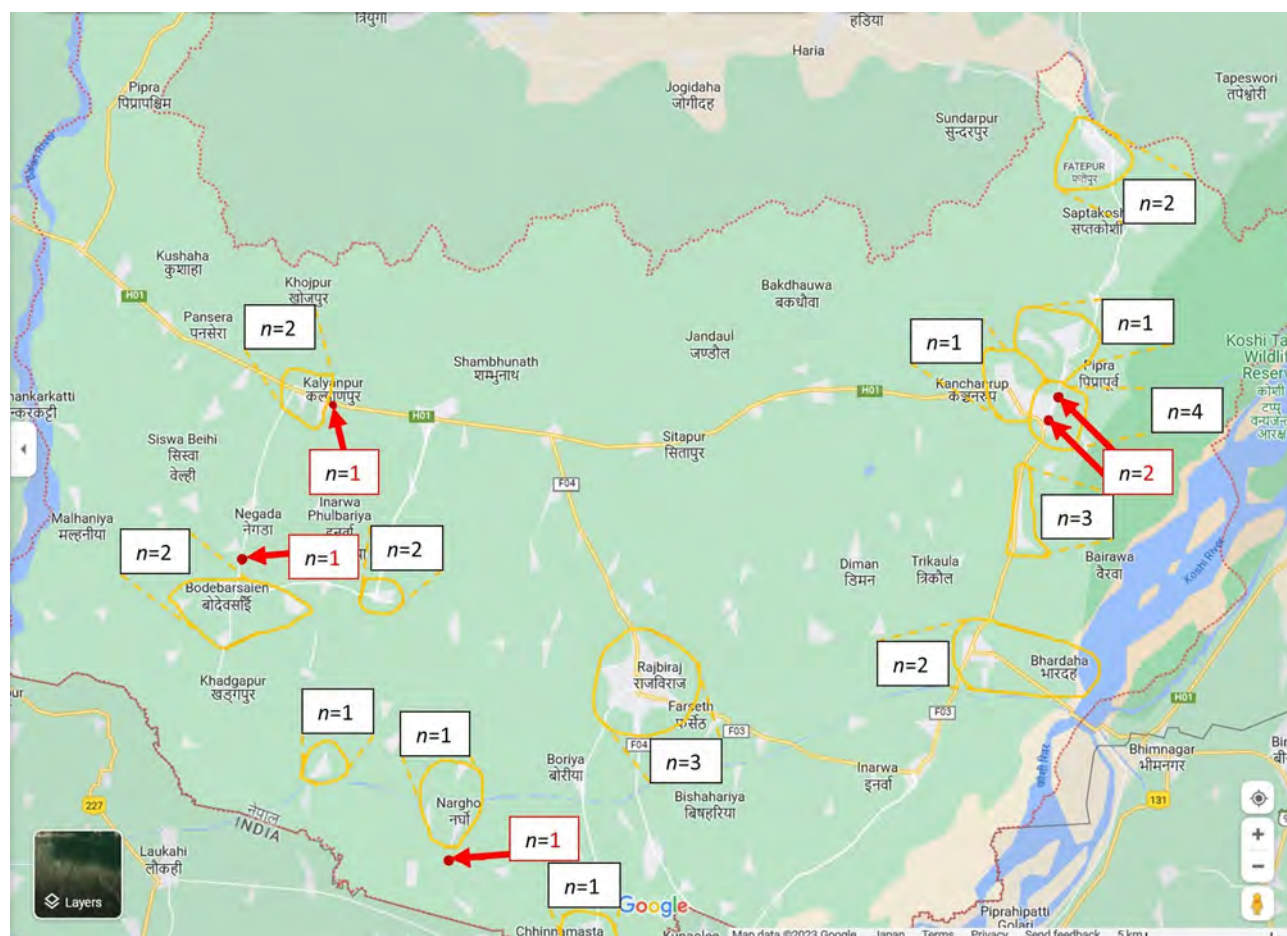
#### *Visual observation test*

All of the 241 samples were visually examined using the ‘tool for visual inspection of medicines’ produced by the International Council of Nurses in partnership with the USP and modified by the International Pharmaceutical Federation<sup>46</sup>. Detailed information from the labelling, packaging, blisters, dosage units and leaflets, if available, was carefully recorded and documented, and images were taken of each sample. Although visual defects may suggest violations of Good Manufacturing Practices, they were not included in the definition of ‘substandard’ in this study, as doing so could be misleading and inconsistent with pharmacopoeial standards, which define substandard medicines based on chemical and performance criteria.

#### *Authenticity investigation*

All available email and physical addresses of the labelled manufacturers were collected from the sample blisters and packages, the manufacturers’ websites, social media appearances on Facebook (Meta Platforms, Inc., Cambridge, MA, USA), LinkedIn (Sunnyvale, CA, USA) and X (formerly Twitter, X Corp., San Francisco, CA, USA) and advertisements and job offers published by the manufacturers using the Google (Mountain View, CA, USA) search engine’s text and image search feature.

Authenticity investigation forms were compiled by recording the samples’ brand or trade name, stated manufacturer and distributor, dosage form and strength, quality specifications as per the labelled pharmacopoeia, batch number and manufacturing and expiry dates. All relevant images of the sample, including the packaging, blisters and any problematic dosage units were attached to the form. Emails with the corresponding authenticity investigation form, a cover letter and the request to return completed authenticity forms were sent to all of the 36 manufacturers.



**Fig. 1.** Map of the Saptari district showing the sampling sites and approximate areas in which sampling sites were visited. Shows the sampling map of the Saptari district. The red arrows indicate the five localised sampling sites and the orange circles and nearby  $n = x$  markings indicate sampling areas that include  $n$  sampling sites. These circles contain a total of 25 sampling sites as documented in the sampling form, which could not be localised. Samplers did not document the address nor the area of one sampling site, which consequently could not be mapped. The source of the map data was Google Maps ©2023 (Mountain View, CA, USA). The base map was extracted from this using Microsoft Snipping Tool version 2022.2410.21.0 (Microsoft Corp., Redmond, WA, USA). Sampling locations and areas were marked using Microsoft Paint version 11.2410.28.0 (Microsoft Corp., Redmond, WA, USA), and lines and text fields were added using Microsoft Word MSO 365 version 2411 (Microsoft Corp., Redmond, WA, USA).

Manufacturers who did not respond to the initial email for 14 days were classified as non-responder type 1 and a reminder email was sent. A manufacturer was classified as non-responder type 2 if a completed authenticity form was not received within 6 months of sending the reminder email. Subsequently, the authenticity forms of all manufacturers classified as non-responder type 2 were printed and dispatched together with the cover letter and the request to return the completed authenticity form by Japan Post's Express Mail Service. Manufacturers who did not answer the postal authenticity enquiry within 6 months were classified as non-responder type 3.

The dispatch dates of all authenticity enquiry emails and the posting and receipt dates of the completed authenticity forms were recorded to investigate the authenticity response time. The response time was defined as the income difference between the dispatch date and the response receipt date. The minimum, median and maximum response times were determined for email and postal transmission.

#### *Legitimacy status investigation*

Relevant information regarding the samples, including the manufacturer names and their licence numbers, trade/product names, batch numbers, APIs, strengths and dosage formulations, was sent to the DDA, which checked the legitimacy status of the samples using the latest available product and manufacturer authorisation lists.

#### **Chemical analysis**

The chemical analysis was conducted in the analytical laboratory of Kanazawa University in Japan. All applicable samples underwent the dissolution test '<711>', the uniformity of dosage units test, the assessment of content

uniformity ' $<905>$ ', and the assay test as per the USP 41 instructions and acceptance criteria<sup>47</sup>. Assay test values were calculated as an average of 10 individual unit contents which were tested in the first stage of the uniformity of dosage units test ' $<905>$ '.<sup>47</sup> However, for samples which required second-stage testing, the assay value was calculated from the total 30 individual contents tested (first- and second stage).

A summary of the HPLC methods is provided in Table 1, a summary of the dissolution test conditions is given in Table 2. However, the detailed HPLC and chemical analysis methods are fully described in Supplementary Material 1. The individual acceptance criteria which corresponded to USP 41 criteria for each of the four APIs<sup>47</sup> are provided in Supplementary Material 2. In this study, samples which failed quality testing in an intermediate stage but did not have sufficient units to finalise quality testing (hereinafter referred to as 'interim failed' or 'interim failing' samples) were judged as 'passed'.

Calibration curve and quality control (QC) solutions were freshly prepared for each HPLC run, and all other solutions used in analysis preparation, such as buffer solutions, were freshly prepared for each analysis day. QC solutions of three concentrations were prepared for the analysis of AZM, ESM and LST. For CFIX, QC solutions were prepared only during validation.

The method validation followed general Good Laboratory Practice guidelines<sup>48</sup> and the 'The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use' Q2(R2) guideline<sup>49</sup>. AZM, CFIX, ESM and LST standard products approved in Japan (see Chemicals and Materials section) were analysed before sample analysis. A detailed overview of the validation results is given in Supplementary Material 4.

The peak data for API and internal standards were calculated by the respective software programmes specified in the Devices and Instrumentation section, and the data were exported to Excel MSO 365 (Microsoft Corp., Redmond, WA, USA).

Images of failing units were taken after the acid stage (only for ESM samples) and after the buffer stage of the dissolution tests.

### Price analysis

During the sampling process, the individual sample prices which were paid by samplers were documented in the sampling form. However, the samplers did not collect the individual payment receipts. Prices which were indicated in Indian rupees (INR) were converted into Nepalese rupees (NPR) using the rounded INR-to-NPR exchange rate of April 27, 2022 (INR1 = NPR1.6)<sup>50</sup>. Subsequently, all prices were converted into United States dollars (USD) using the NPR-to-USD exchange rate of April 27, 2022 (USD1 = NPR122.5964)<sup>51</sup>, the date when sampling started in the Saptari district. The unit prices of the samples were plotted according to the results of the quality analysis. The median price per unit was determined for each API of the collected samples. The median medicine price ratios (MPRs), 25th percentile MPR, 75th percentile MPR, minimum MPR and maximum MPR in the absolute median and the 25th and 75th percentile, minimum, and maximum prices per unit were divided by the international reference unit price (IRP). The IRPs were taken from the reference unit price list of the supplier median prices published by the Management Sciences for Health and the WHO in 2016<sup>52,53</sup>.

### Small-scale dissolution test evaluation

To evaluate the applicability of the small-scale dissolution test previously proposed by Rahman and Yoshida et al.<sup>35</sup> as a screening technology, the small-scale dissolution test criteria were applied to this study. The screening technology uses various compliance criteria to evaluate dissolution test results using  $n = 3$  units which are tested for dissolution (instead of the 6–24 units required for the USP dissolution test). The following compliance criteria resulted for the four APIs: the AZM and CFIX dissolution test results averaged  $Q \geq 86\%$ , with all individual  $Qs \geq 82\%$ , and the ESM and LST dissolution test results averaged  $Q \geq 81\%$ , with all individual  $Qs \geq 77\%$ .

These small-scale dissolution test compliance criteria were then applied to all.

$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{6}{3} = 20$  possible  $n = 3$  combinations of the  $n = 6$  individual Q results obtained in the first stage of the dissolution test using the methods described in the Chemical Analysis section.

The evaluation was expressed as the compliance of the small-scale dissolution test result of each  $n = 3$  combination with the final test results of the USP 41 dissolution test. The results were compliant if test results were either both 'pass' or both 'fail', they were false negative (= falsely non-compliant) if the small-scale dissolution test result was 'fail' but the USP 41 dissolution test was 'pass', and they were false positive (= true non-compliant) if the small-scale dissolution test was 'pass' and the USP 41 dissolution test was 'fail'. Importantly, the small-scale dissolution test results were not compared with the USP 41 first-stage test result but to the last (= final) tested stage result if further test stages were conducted.

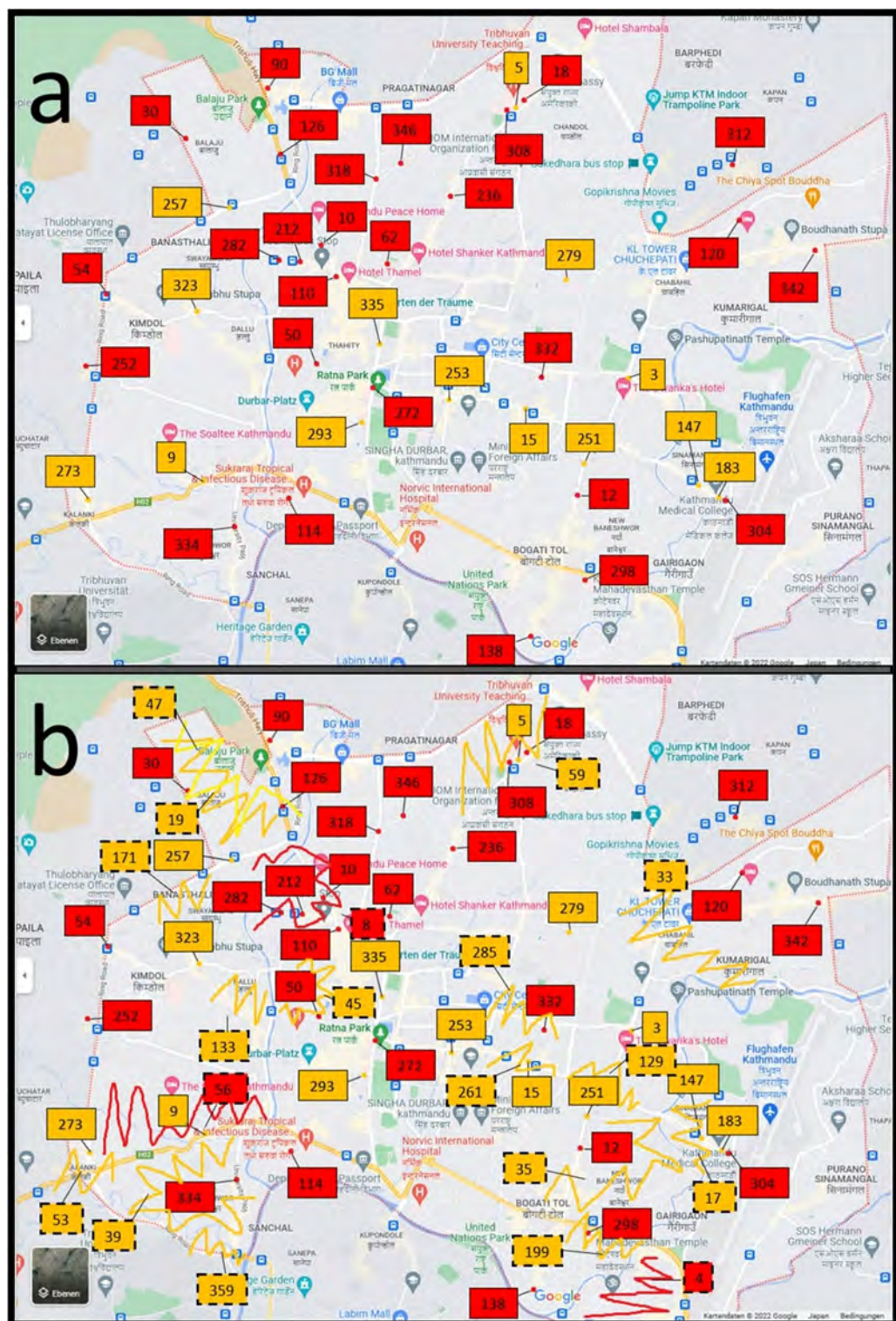
The compliance of both test results was expressed as the overall percentage of samples in agreement. The rate of false-negative judgement was expressed as the overall percentage of samples which required retesting and the individual rate for each of the 20 combinations per  $n = 6$  Q results which required retesting despite an expected 'pass' for the sample. The rate of false-positive acceptance of the samples was expressed as an overall percentage, a total percentage of all individual combinations per  $n = 6$  Q results and the individual rate for each of the 20 combinations per  $n = 6$  Q results of samples which were false-positively accepted in the small-scale dissolution test.

### Raman scattering analysis

Before analysis, the supplied software was installed on a personal computer and the C13560 and Inspector500 portable Raman spectrometers were set up in accordance with manufacturer instructions.

Raman scattering analysis was conducted in accordance with the methods applied by Schreiber et al.<sup>41</sup>. A summary of the Raman spectroscopy conditions is given in Table 3. For the C13560 device, the output was set at





15 mW ('High'), the excitation laser wavelength was set at 785 nm, the scanning time was 1000 ms per scan and the spectral X-axis wavenumber interval was set between 98 and 1852  $\text{cm}^{-1}$ . Before taking the measurements, the dark signal was measured by inserting the silicon substrate provided by the manufacturer. Subsequent calibration of the software was performed using the Raman shift peak of the substrate near 521  $\text{cm}^{-1}$ . As per the instructions, the emitted laser wavelength was reset manually to 785 nm immediately after calibration.

For the Inspector500, the output was set at 300 mW ('High'), the excitation laser wavelength was set at 1030 nm, the scanning time was set to the default 'automatic' setup with a maximum of 8 s and the spectral X-axis wavenumber interval was between 150 and 2450  $\text{cm}^{-1}$ . Before measurements were taken, a calibration test was conducted in accordance with the instructions. If the test failed, the device was recalibrated repeatedly until it passed the test.

◀**Fig. 2.** Map of the Kathmandu district showing the individual sampling sites and approximate areas in which sampling sites were visited. Shows a map of the Kathmandu district in which the locations of sample acquisition are indicated. Sampling sites that could be localised are shown as dots with connected boxes (continuous frame). Hatched areas with connected boxes (dashed frame) each represent one sampling site for which only the area could be localised. The boxes contain the corresponding priority number of the sampling site. The orange colour indicates that the sampling was done by Team C, whereas the red colour indicates sites that were visited by Team D. The source of the map data was Google Maps ©2023 (Mountain View, CA, USA). The base map was extracted from this using Microsoft Snipping Tool version 2022.2410.21.0 (Microsoft Corp., Redmond, WA, USA). Sampling locations and areas were marked using Microsoft Paint version 11.2410.28.0 (Microsoft Corp., Redmond, WA, USA), and lines and text fields were added using Microsoft Word MSO 365 version 2411 (Microsoft Corp., Redmond, WA, USA). (a): localised sampling sites and (b) localised sampling sites and hatched areas corresponding to one sampling site each, which were documented as area only.

API	Apparatus type	Medium	Rotation speed	Temperature	Testing time
AZM	Apparatus 2: Paddle	pH 6.0 buffer	75 rpm	37.0 °C ± 0.5 °C	30 min
CFIX	Apparatus 2: Paddle	pH 7.2 buffer	100 rpm	37.0 °C ± 0.5 °C	45 min
ESM	Apparatus 2: Paddle	Acid stage: 0.1 M HCl Buffer stage: pH 6.8 buffer	100 rpm	37.0 °C ± 0.5 °C	Acid stage: 120 min Buffer stage: 30 min
LST	Apparatus 2: Paddle	Distilled, deaerated water	50 rpm	37.0 °C ± 0.5 °C	30 min

**Table 2.** Overview of the dissolution test conditions.

Raman spectrometer	Laser wavelength	Spectral resolution	Scan time	Spectral data points per spectrum	Output
C13560 device	785 nm	10 cm <sup>-1</sup>	1000 ms	50 spectral data points averaged	15 mW
Inspector500	1030 nm	8–10 cm <sup>-1</sup>	Automatic (max. 8000 ms)	50 spectral data points averaged	300 mW

**Table 3.** Overview of the Raman spectroscopy testing conditions. N/A = not available.

The measurement procedure was identical for both devices: the tablets were placed in front of the laser source. Measurements were taken by placing the spectrometer’s attachment for analysis in direct contact with each tablet to completely cover the tablet surface to ensure that the laser directly hit the tablet. During analysis, a black opaque fabric was used to cover the set-up and prevent light contamination. One measurement was obtained as an average of five spectral data points from the Raman scattering analysis. A total of 10 measurements per tablet were performed on randomly chosen spots of the sample surface. Five measurements were taken from each of the top and bottom of the tablet. Therefore, a total of 50 spectral data points were averaged, resulting in one spectrum for each sample for each device and creating a spectral library for each of the products. Capsules were opened and the granules were placed in a small plastic film bag before being analysed in the same way as for the tablets.

For the C13560, the Spectral X-axis wavenumber interval was set between 403 and 1852 cm<sup>-1</sup>. Raman spectra were line plotted and colourised using the Unscrambler X 10.5 (CAMO Software AS, Oslo, Norway).

**Statistical analysis**

Descriptive statistical analysis including the calculation of means, standard deviations (SDs) and relative standard deviations (%RSDs) was performed using Excel MSO 365 Versions 2307–2404 and 2506 (Microsoft Corp., Redmond, WA, USA). Fisher’s exact test (any cell value of the contingency table < 5) and Pearson’s chi-squared test (all cell values of the contingency table ≥ 5) were used where appropriate; these tests were performed using SPSS version 29.0.2.0 (IBM Corporation, Armonk, NY, USA). Statistical significance was assumed at the *p* < 5% level for all tests.

**Results**

**Sampling**

*Sample collection*

A total of 241 samples were collected from 91 sampling sites in Nepal. Of the total, 121 samples were acquired from April 27 to May 6, 2022, in the Saptari district (31 sampling sites) and 120 samples were collected from May 14 to June 8, 2022, in the Kathmandu district (60 sampling sites). Although the samplers did not present prescriptions, none of the samples were refused to them during sampling.

The 241 samples consisted of a total of 59 products and trade names from 113 different batches produced by 36 manufacturers based in two countries. According to the sample labelling, 173 samples originated in Nepal

(71.8%) and 68 in India (28.2%). A total of 138 samples (57.3%) were sampled as blisters only and had no secondary packaging.

A comprehensive, anonymised overview of samples is presented in Supplementary Material 3 (Supplementary Tables S1a–b, S2a–d, S3 and S4).

#### *Sampling sites*

Out of the 91 sampling sites, 88 sites (96.7%) were confirmed as licenced vendors: 87 sampling sites (95.6%) were categorised as legal pharmacies and one site (1.1%) was a legal wholesaler. The shop category of the three sampling sites was not specified by the samplers, but the eight samples acquired from these sampling sites were included in the study. Sampling sites IDs 1–31 were located in the Saptari district, whereas sampling sites IDs 32–91 were located in the Kathmandu district. Of the 31 sampling sites in the Saptari district, a total of 30 site locations ( $n=5$ ) or areas ( $n=25$ ) could be identified and were mapped (Fig. 1). In the Kathmandu district, all 60 sampling sites were identified and mapped either as locations if available or as areas; among the sites, 41 (68.3%) were documented or identified as sampling site locations and 19 (31.7%) were documented as sampling site areas (Fig. 2).

Sixty-five of the 91 sampling sites (71.4%; 95% CI 61.0–80.4) had indoor temperatures above the 2022 recommended 25 °C, with a maximum temperature reaching 30 °C in five sampling sites. Additionally, the relative humidity level of one sampling site (1.1%; 95% CI 0.0–6.0), reached 75%. No sampling site exceeded the stability testing conditions of pharmaceutical climatic zone IVb.

A comprehensive, anonymised overview of sampling sites is presented in Supplementary Material 3 (Supplementary Tables S1b and S5a–c).

### **Visual observation, authenticity investigation and legitimacy status investigation**

#### *Visual observation*

Of the 241 collected samples, a total of 16 samples (6.6%; 95% CI 3.8–10.6) failed the visual observation test. Specifically, the dosage units of all of the 16 failing samples (AZM: 0, CFIX: 0, ESM: 12, and LST: 4) showed cracks (A-103, A-106, A-108, B-101, B-105, B-110, C-207 and D-211), broken additional tablet pieces within the blisters (C-211), or inhomogeneous coatings (B-115-2, C-103, C-110, C-113, D-103, D-113 and D-205). One tablet of sample B-110 (shown in Fig. 3) had severe cracks and was separated before chemical analysis; the tablet was used for further investigations, which are not discussed in this study. The separation of one unit of sample B-110 likely biased the SF medicine occurrence estimation of this study. Overall, only three samples (1.2%; C-106, D-314 and D-404) contained a leaflet, which was available only in English.

#### *Authenticity investigation*

A total of 16 of the 36 manufacturers (44.4%; 95% CI 27.9–61.9) replied with completed authenticity forms. Twelve manufacturers (75%) were based in Nepal and four (25%) in India (Supplementary Tables 1a–b), but the difference was not significant: Fisher's exact test; 2-Tail (1,  $n=36$ ) = 0.9,  $p=0.343$ ;  $p>0.05$ . Emailed forms led to six completed authenticity forms, and posted forms led to an additional ten replies, an increase from 16.7% to 44.4%. Thirty manufacturers (83.3%) were classified as both non-responder type 1 and non-responder type 2 before the authenticity investigation forms were sent by post, and 20 manufacturers (55.6%) were classified as non-responder type 3 after forms were sent by post. In the authenticity investigation, the 16 manufacturers who responded declared 118 samples (49.0%) as genuine and no sample was suspected by the manufacturers of being falsified in their responses, whereas no response was received for the remaining 123 samples (51.0%).

#### *Authenticity investigation response time analysis*

The median response time for six replies from manufacturers following emailing using the sending date as the dispatch date was 7.5 days (mean = 7.2 days, SD = 4.0 days), with minimum and maximum response times of two and 14 days, respectively.

For forms sent by post, the median response time for 10 replies from manufacturers was 23.5 days (arithmetic mean = 25.6 days, SD = 10.2 days), with minimum and maximum response times of 14 and 47 days, respectively, including an unknown travel time for each sent package. Determining the exact arrival date for forms sent by post and mail delivery times was not possible; therefore, all response times for postal delivery included individual mail delivery times.

#### *Legitimacy status investigation*

The DDA stated that the indicated information regarding manufacturer name and licence number, product name, batch number, API, strength, and dosage formulation of all 241 collected samples were legitimate and that the medications were authorised for distribution in Nepal.

### **Chemical analysis**

#### *Method validation*

The validation parameters for each API (AZM, CFIX, ESM and LST) and the test results for the four standard products are summarised in Supplementary Material 4 (Supplementary Tables S6a–d and S7).

#### *Summary of the chemical analysis results*

A total of 24 of the 241 samples (9.96%; 95% CI 6.5–14.5) failed one or more of the pharmacopoeial quality tests (AZM: 1, CFIX: 3, ESM: 7 and LST: 13). One AZM sample (0.41%) interim failed quality testing without sufficient sample material to continue. The API identity was confirmed in all 241 samples, and no additional APIs were detected with the analytical methods used. However, 28 samples (CFIX: 22 and ESM: 6) could not





**Fig. 3.** Side view of the cracked and damaged tablet of sample B-110 which was separated before analysis of the sample. Shows the image of a unit of sample B-110. The unit exhibits severe cracks and damages in the enteric coating and the tablet core. The markings on the ruler indicate the distance in cm. Intermediate markings are in mm.

API	Samples tested	Pass	Fail	
			Interim	Final
Azithromycin	60	58	1	1
Cefixime	60	57	0	3
Esomeprazole	61	54	0	7
Losartan	60	47	0	13
Total	241	216	1	24

**Table 4.** Overview of the quality analysis results.

undergo dissolution testing because USP 41 evaluation criteria were not defined for their formulations (uncoated dispersible tablets and film-coated tablets). Therefore, 213 samples (88.4%; AZM: 60, CFIX: 38, ESM: 55 and LST: 60) were tested for dissolution.

Table 4 provides an overview of the quality analysis test results. The individual chemical analysis results of the pharmacopoeial tests are shown and described in detail in Supplementary Material 5 (Supplementary Tables S8-S10, S11a-d, S12a-b, S13a-b, S14a-b and S15a-b).

#### Uniformity of Dosage Units Test:

Three of the 241 tested samples (1.2%; 95% CI 0.3–3.6) failed the uniformity of dosage units test, and another three samples (1.2%) interim failed the first stage of this test. However, insufficient units were left to continue the test.

#### Assay Test:

Overall, six of the 241 samples (2.5%; 95% CI 0.9–5.3) failed the assay test; however, none of the samples had an average API content below 80%, a value which can be considered a ‘moderate deviation’.



**Dissolution Test:**

Eighteen of the 213 tested samples (8.5%; 95% CI 5.1–13.0) failed the dissolution test. Figures 4, 5 and 6 provide images of representative failing units per sample that were taken immediately after drawing the sample solution.

**Overview of test-failing samples**

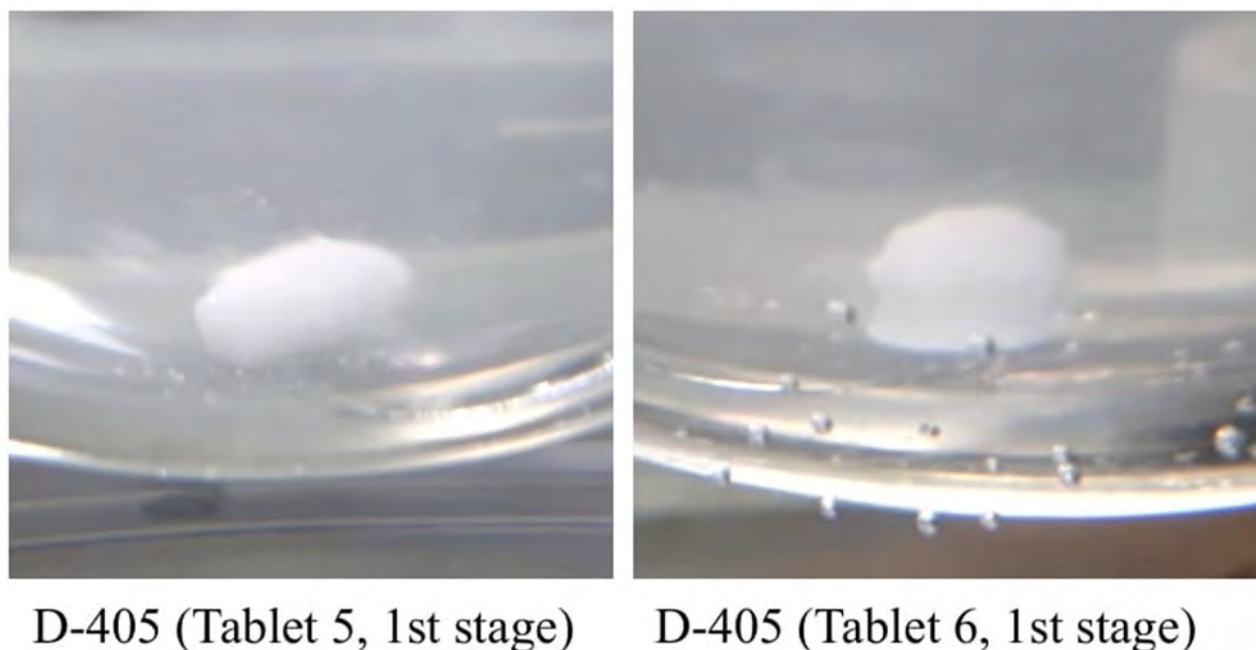
Twenty-four samples of the four APIs (AZM, CFIX, ESM and LST) which represented nine of the 59 products (15.3%; 95% CI 7.2–27.0) and 13 of the 113 batches (11.5%; 95% CI 6.3–18.9) manufactured by 9 of the 36 manufacturers (25.0%) failed at least one test, as summarised in Supplementary Tables S11a–d. Representative HPLC chromatograms are presented in Supplementary Material 6 (Supplementary Figs. S1–S4). Twenty (83.3% of the total) and four test-failing samples (16.7% of the total) were labelled as having been manufactured in Nepal (out of 173) and India (out of 68), respectively. However, the difference between the origin of manufacture was insignificant: Fisher's exact test; 2-tail (1,  $n = 241$ ) = 1.76,  $p = 0.185$ ;  $p > 0.05$ . Among the 118 authenticated samples (see Visual Observation, Authenticity Investigation and Legitimacy Status Investigation section), 16 samples (13.6%) of three products failed quality testing, whereas eight samples (6.5%) of the 123 samples for which no reply was received failed quality testing. However, the observed difference between the groups was insignificant: Pearson's chi-squared test; (1,  $n = 241$ ) = 3.34,  $p = 0.343$ ;  $p > 0.05$ .

The average proportion of test-failing antibiotics (3.33%; 95% CI 0.9–8.3; 4 of 120 AZM and CFIX samples) was lower than that of medicines against non-communicable disease (NCDs; 16.5%; 95% CI 10.4–24.4; 20 of 121 ESM and LST samples) and the difference was significant: Fisher's exact test; 2-tail (1,  $n = 241$ ) = 11.7,  $p = 0.001$ ;  $p < 0.05$ . Furthermore, the average proportion of test-failing samples collected in the Saptari district (13.22%; 95% CI 7.8–20.6; 16 of 121 samples) was higher than that of those collected in the Kathmandu district (6.67%; 95% CI 2.9–12.7; 8 out of 120 samples). However, no significant difference was observed: Pearson's chi-squared test; (1,  $n = 241$ ) = 2.89,  $p = 0.09$ ;  $p > 0.05$ .

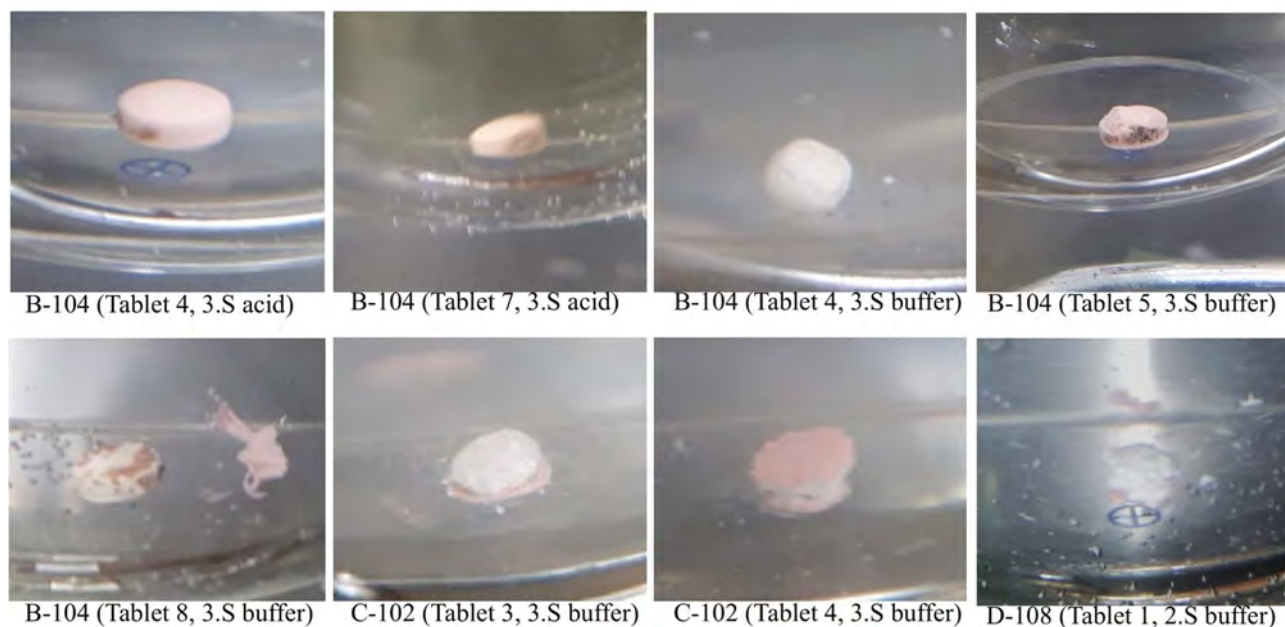
**Price analysis**

Price documentation was missing for one CFIX sample (B-313) and three LST samples (B-201, C-211, C-213), all of which passed the performed quality tests.

The unit prices of the samples were plotted according to the results of the quality analysis (see Fig. 7a–d). The rounded median prices per unit for each of the four APIs were USD 0.245 (arithmetic mean = USD 0.245; %RSD = 0.0%) for AZM, USD 0.163 (arithmetic mean = USD 0.159; %RSD = 9.4%) for CFIX, USD 0.086 (mean = USD 0.095; %RSD = 22.1%) for ESM and USD 0.062 (mean = USD 0.063; %RSD = 4.3%) for LST. IRPs were only available for omeprazole 20 mg tablets and capsules and 40 mg injection vials but not for ESM. Thus, the MPRs of ESM could not be assessed. According to the documentation, all of the 60 AZM samples were sold at the same MPR of 1.08 (SD = 0.00). The median MPR of CFIX was 0.98 (SD = 0.09), with minimum and maximum MPRs of 0.61 and 1.14, respectively. The median MPR of LST was 0.54 (SD = 0.02), with minimum



**Fig. 4.** Images of representative units of AZM sample D-405 taken immediately after the dissolution test. Shows images of representative units of AZM sample D-405 (units 5 and 6), which did not pass the USP dissolution test. The tablet core remains which did not sufficiently dissolve are distinctly visible. The images were taken immediately after the sample solution was drawn from the dissolution tester.



**Fig. 5.** Images of representative ESM units of failing samples B-104, C-102 and D-108 taken immediately after the dissolution test acid or buffer stages. Shows images of representative units of ESM samples B-104, C-102 and D-108, which did not pass the USP dissolution test. The images were taken immediately after the sample solution was drawn from the dissolution tester for both the acid and buffer stages. Image taking for sample A-105 was missed. 3.S acid = third stage of the acid stage. 2.S buffer = second stage of the buffer stage. 3.S buffer = third stage of the buffer stage.

and maximum MPRs of 0.53 and 0.64, respectively. The MPRs for the individual APIs are shown in Fig. 8. A comprehensive overview of the individual price data is provided in Supplementary Material 7 (Supplementary Figs. S5–S7).

For AZM, the median and average prices per unit were identical between failed and passed samples (USD 0.245). Similarly, for CFIX, the median unit price was the same for both groups (USD 0.163), although the average price was slightly higher among failed samples (USD 0.163 vs. USD 0.159). For ESM, failed samples had a lower median unit price (USD 0.082 vs. USD 0.087) but a slightly higher average unit price (USD 0.099 vs. USD 0.094) compared to passed samples. In the case of LST, the median unit price was equal between both groups (USD 0.062), while the average unit price was marginally higher for failed samples (USD 0.064 vs. 0.062).

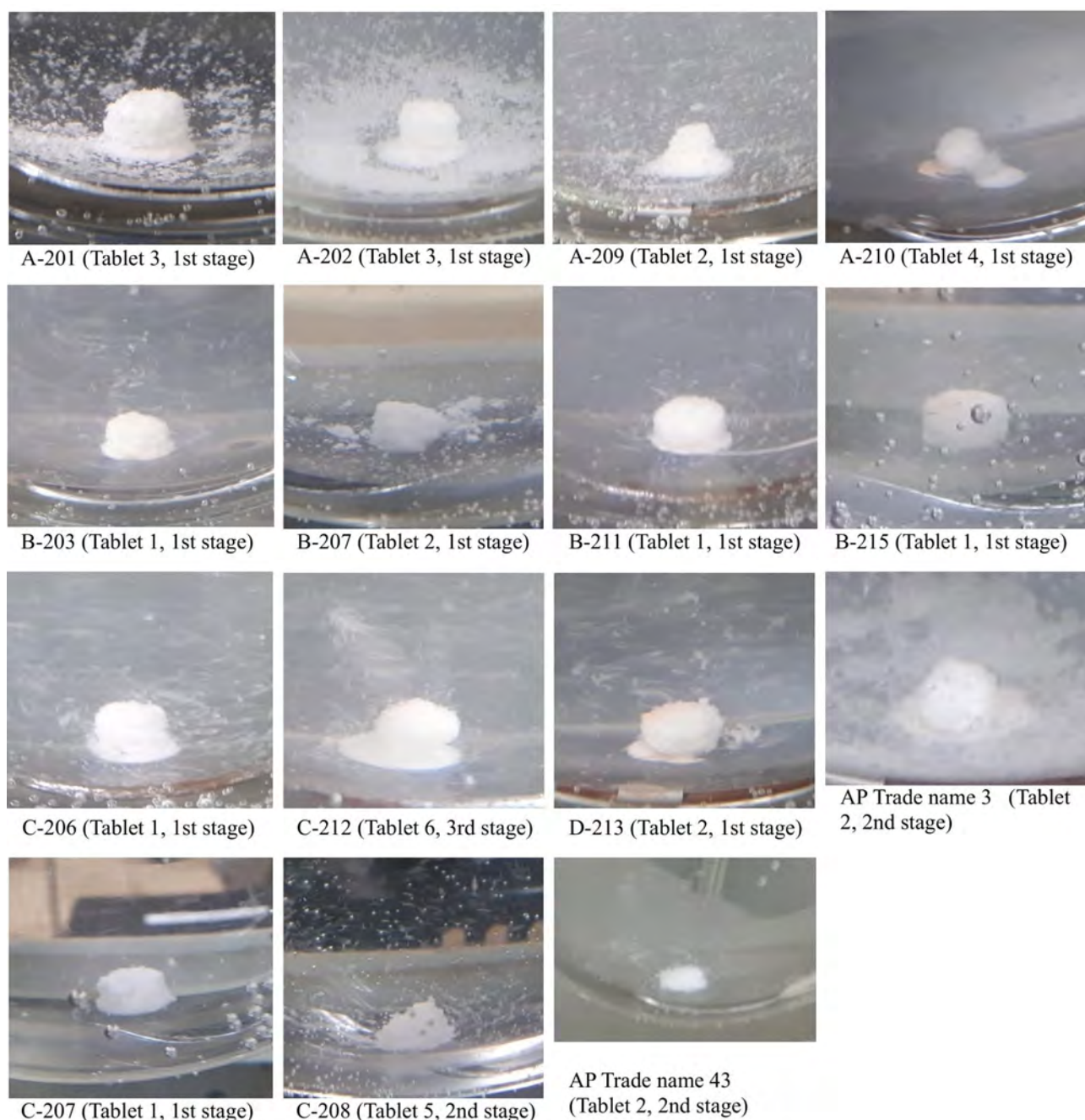
### Small-scale dissolution test evaluation

The small-scale dissolution test criteria ( $\bar{Q}$  dissolution rate of the sample  $\geq Q_{USP} + 6\%$ ; the three individual dissolution rates  $\geq Q_{USP} + 2\%$ ) were applied to all of the 20 combinations of  $n=3$  units for each of the  $n=6$  dissolution rates per sample measured in the first stage of the USP dissolution test, resulting in 4,260 results for all possible  $n=3$  combinations.

Of the 213 tested samples, agreement was observed between small-scale and USP dissolution test results in all of the 20  $n=3$  combinations of 192 samples (90.1%), which were in all four APIs. Overall, 245  $n=3$  combinations (AZM: 183 CFIX: 0 ESM: 21, LST: 41) of the total 4,260  $n=3$  combinations (5.75%; 95% CI 4.8–6.5) were false negatives and 39  $n=3$  combinations (AZM: 1 CFIX: 0 ESM: 38, LST: 0) of the total 4,260  $n=3$  combinations (0.92%; 95% CI 0.7–1.3) were false positives when compared with the last-tested stage USP dissolution test results (AZM: 0.1% false-positive [0.08%; 95% CI 0.0–0.5; 1 of 60 samples; 1 of 1,200 three-unit combination], 15% false-negative [15.3%; 95% CI 13.3–17.4; 10 of 60 samples, 183 of 1,200 three-unit combinations]; CFIX: 0 of 38 samples false-positive or false-negative; ESM: 3% false-positive [3.5%; 95% CI 2.5–4.7; 4 of 55 samples; 38 of 1,100 three-unit combinations], 2% false-negative [1.9%; 95% CI 1.2–2.9; 3 of 55 samples, 21 of 1,100 three-unit combinations]; LST: 0 false-positive, 3% false-negative [3.4%; 95% CI 2.5–4.6; 3 of 60 samples, 41 of 1,200 three-unit combinations]).

### Raman scattering analysis

Raman spectra of all 241 samples and the reference standards for AZM, CFIX, omeprazole and LST) were obtained using both portable Raman spectrometers. For AZM, CFIX and LST samples, 172 of 180 (95.6%; 95% CI 91.4–98.1) spectra collected with the C13560 and all 180 of the spectra collected with the Inspector500 revealed characteristic peaks that corresponded to the respective reference standard. However, the visible AZM Raman peaks observed with the C13560 were small and thus difficult to identify. No characteristic peak was observed for the ESM samples when these were visually compared with the omeprazole reference standard. The spectra are presented in Supplementary Material 8 (Supplementary Figs. S8–S15).



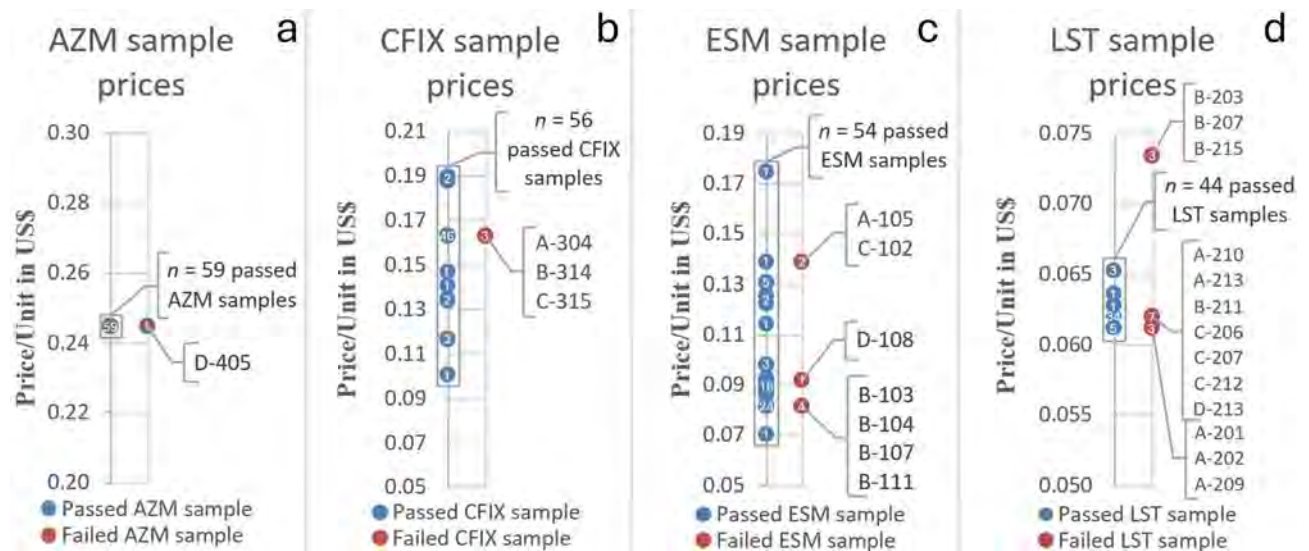
**Fig. 6.** Images of representative LST units of failing samples immediately after the dissolution test. Shows images of representative units of LST samples A-201, A-210, B-203, B-207, B-211, B-215, C-206, D-213, AP Trade name 3, C-207, C-208, and AP Trade name 43, which did not pass the USP dissolution test. The images were taken immediately after the sample solution was drawn from the dissolution tester, but image-taking for sample A-213 was missed. Sample C-208 and AP Trade name 43 passed the USP dissolution test and were compared with C-207, which was labelled the same as Trade name 43 but presented a different batch number.

Notably, the detection limits for API quantities across different formulations were not assessed, and spectral comparisons of passing and failing samples were not successful in predicting sample quality given that their spectra were often too similar. Variations in spectral profiles appeared to be influenced more by surface colour differences than by differences in sample quality, e.g. LST product Trade name 3.

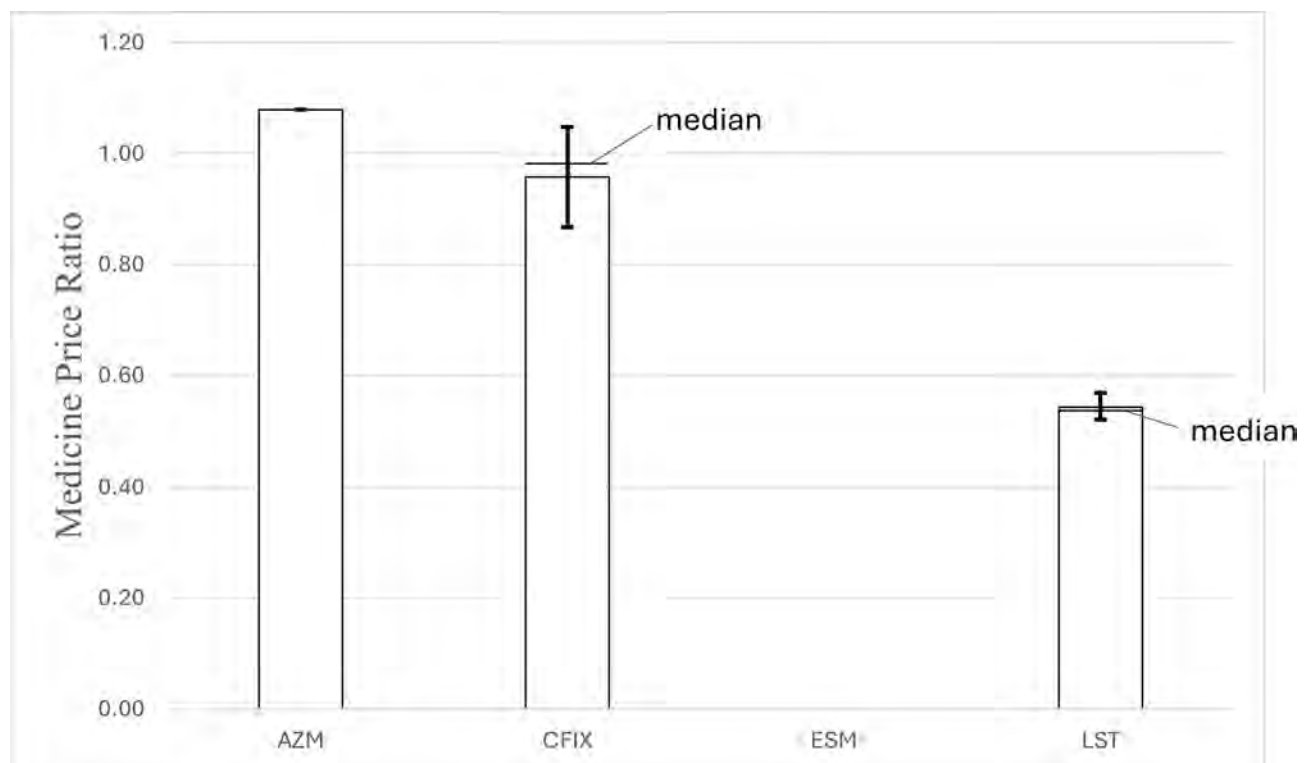
## Discussion

In this study, 241 samples of four APIs—AZM, CFI, ESM and LST – were collected from 91 licenced vendors in the Saptari and Kathmandu districts of Nepal using randomly conducted convenience sampling (Saptari district) and the randomised selection of sampling sites (Kathmandu district; Figs. 1 and 2; Supplementary Table





**Fig. 7.** a–d: Unit prices of AZM, CFIX, ESM and LST samples in US dollars, categorised by quality analysis results. (a)–(d) shows the available unit prices of the AZM (a), CFIX (b), ESM (c) and LST (d) samples in US dollars per unit. Blue dots represent samples that passed quality tests, whereas red circles indicate samples that failed at least one pharmacopoeial test. Each subplot includes a white number indicating the count of samples with a similar or identical unit price. Unit prices for four samples (B-313, B-201, C-211 and C-213) were unavailable and are therefore not included in the figure.



**Fig. 8.** Medicine Price Ratios (MPRs) of AZM, CFIX and LST samples. Illustrates the median MPRs of AZM, CFIX and LST including the standard deviation (SD). All 60 AZM samples were documented to be sold at an MPR of 1.08 (mean: 1.08, SD: 0.00), the median MPR of CFIX was 0.98 (mean: 0.96, SD: 0.09), and the median MPR of LST was 0.54 (mean: 0.54, SD: 0.02). The international reference unit price of ESM was not available, therefore the MPR could not be calculated.



S1a–b; Supplementary Figs. S1 and S2). A total of 24 samples (10.0%; 95% CI 6.5–14.5) failed at least one of the three pharmacopoeial quality tests conducted (assay, uniformity of dosage units and dissolution; Table 4; Figs. 4, 5 and 6; Supplementary Tables S11a–d, S12a–b, S13a–b, S14a–b and S15a–b). This result was in line with the WHO's global estimate of 10.5% SF medicines in LMICs<sup>8</sup> and a recent systematic review by Ozawa et al. which estimated a 10.2% prevalence of SF medicines in Asian LMICs<sup>17</sup>; the 95% confidence interval of the current study (6.5–14.5%) overlaps with both reference estimates.

### Classification of the pharmacopoeial test-failing samples

Of the 24 samples that failed pharmacopoeial testing (= 'test-failing' or 'SF' samples), 16 authenticated samples were classified as 'substandard' based on chemical analysis and confirmation of authenticity by the manufacturers (B-103, B-104, B-107, B-111, A-201, A-202, A-209, A-210, A-213, B-203, B-207, B-211, B-215, C-206, C-212, D-213). The eight samples for which no response was received from the respective manufacturers (A-105, C-102, D-108, C-207, A-304, B-314, C-315, D-405) were classified as 'probably substandard' according to the approach proposed by Ozawa et al.<sup>17</sup>, as they had failed pharmacopoeial testing but had average API contents above 80% of the declared amount, and no evidence of falsification and/or criminal intent was identified in this study. For simplicity, 'probably substandard' samples are referred to as 'substandard' hereinafter. However, although no indications of falsification were found for any sample, the classifications of the eight 'probably substandard' samples remain uncertain—especially given the 56% non-response rate—, and this limitation must be acknowledged.

### Implications of the occurrence of SF samples in Nepal

Among the 24 substandard samples, medicines against NCDs, i.e. ESM and LST, included a significantly higher proportion of substandard medicines than antibiotics, i.e. AZM and CFIX, (20 samples [8.3%] versus 4 samples [1.7%]; Fisher's exact test; 2-Tail;  $p < 0.05$ ). Given that antibiotics have been commonly reported as SF medicines<sup>8,11,17,54</sup> and NCDs cause a high burden of disease in Nepal and globally<sup>55,56</sup>, this finding is noteworthy and suggests that future medicine quality studies should continue to elucidate the quality of medicines against NCDs. The result further indicates that poor-quality medicines against NCDs may present a potential contributing factor to the burden of disease in Nepal. Nonetheless, substandard antibiotics were also identified in this study. SF antibiotics can impact treatment and potentially contribute to antimicrobial resistance given their poor dissolution (D-405; Fig. 4, Supplementary Table S12a) and slightly lower average content or non-uniform content in dosage units (A-304, B-314 and C-315; Supplementary Table S13b)<sup>33,57–59</sup>.

Overall, the findings of this study are unexpectedly moderate, considering that Nepal's regulatory system is classified at WHO maturity level 1. This classification reflects limited regulatory capacity, which may include, for example, inadequate oversight of manufacturing standards due to infrequent inspections of industry and laboratories, limited post-market surveillance, and insufficient or delayed mechanisms for identification of SF medicines or preventing their dissemination in practice<sup>26,27</sup>. Despite concerns expressed by Nepalese healthcare professionals who recently described Nepal as vulnerable to an influx of falsified medicines<sup>60</sup>, no evidence of falsification was found for any sample in this study. This finding is consistent with previous reports from Nepal<sup>61–64</sup> and the DDA's statement that no official or unofficial reports on falsified medicines have been issued in Nepal. However, the absence of detected falsified medicines in this study may be attributable to methodological and sampling limitations. Accordingly, the possibility of their presence in Nepal cannot be conclusively dismissed.

Previously, four studies reported that nine of 214 samples (4.2%; 95% CI 1.9–7.8)<sup>62</sup>, nine of 172 samples (5.2%; 2.4–9.7)<sup>63</sup>, 13 of 40 samples (32.5%; 95% CI 18.6–49.1)<sup>64</sup> and 37 of 244 samples (15.2%; 95% CI 10.9–20.3)<sup>61</sup> were SF medicines. All reported test-failing samples had specifically been classified as 'substandard', and no study had reported any falsified medicines. Therefore, the data may suggest that falsified medicines could be less prevalent in Nepal—and potentially in the entire geographical region—than the global average data for LMICs or the OECD data would indicate<sup>17,22</sup>. It is important to note that the study reporting the alarmingly high failure rate of 32.5% was based on a small, non-representative sample size and reported discrepancies in analytical test results between two laboratories. The studies reporting 4.2% and 5.2% 'substandard medicines' did not provide information on randomisation or sampling methodologies; in the latter case, the specific medicines sampled were not reported. Therefore, their comparability to Dhakal et al.'s study, which employed randomised multi-district sampling in public health care facilities, and the present study is limited. Although the average test-failing samples were different between Dhakal et al.'s study and the present study (15%; 95% CI 10.9–20.3 vs. 10%; 95% CI 6.5–14.4), this difference was not significant: Pearson's chi-squared test; (1,  $n = 485$ ) = 2.99,  $p = 0.084$ ;  $p > 0.05$ . This may imply that the quality of the distributed medicines could be comparable throughout the country, but the underlying data for this, such as analysed APIs and sample sizes, remain limited.

### Assessment of the visual observation test results

Only one of the 16 samples (6.25%) that failed the visual observation test also failed the pharmacopoeial analysis result (C-207), which exhibited noticeable porosity. This underscores the critical importance of full laboratory analysis for assessing medicine quality, while also demonstrating the benefits of visual observation in identifying poorly manufactured or mechanically compromised units; and, by extension, suggesting potential manufacturing flaws. Furthermore, the low number of collected leaflets indicates that many medicines from licenced vendors in Nepal are distributed to consumers without accompanying written information, i.e. instructions for use or health information, thus potentially leading to usage errors such as unintentional misapplication or dosing by the consumer (Supplementary Table S1b). Therefore, the increased distribution of leaflets with medicines (ideally in Nepali and/or English) should be encouraged.

### Assessment of the authenticity investigation

In the authenticity investigation, the use of the postal delivery method in addition to email delivery led to an increase of nearly 28% (from 6 to 16 responses) in manufacturers' responses. This approach led to valuable data, although no email response had been received for 6 months or longer before the documents were sent by post. This study's findings may encourage researchers to send authenticity investigation documents by post in addition to email delivery in future studies. The median response times for authenticity responses in this study were 7.5 days for email and 23.5 days for post, with a maximum response time of 47 days. This finding adds to data provided by Hauk et al.<sup>65</sup> which suggest that comprehensive data collection during authenticity investigation should allow for at least 2 months for manufacturer response.

### Assessment of price–quality relationships

Although the MPRs of some collected CFIX samples and all LST samples were relatively low (Fig. 8), a low unit price was not a key parameter for a simple prediction of quality issues (Fig. 7a–d). On the other hand, three LST samples that failed the dissolution test critically were sold at the highest unit price among all LST samples (Fig. 7d; B-203, B-207, B-215; all produced in Nepal) and exceeded the indicated—and legally binding—maximum retail price by approximately 20%. However, as these samples were visually similar from others of the same labelled product and batch numbers that were sold at standard unit prices, it could be argued that vendors were not aware of the quality issues and had 'simply' increased the prices illegally.

No clear association was found between unit price and sample quality based on descriptive statistics due to the absence of a consistent trend, but further statistical analysis would be required to confirm this. It should also be noted that the reported prices were based on documentation by the samplers; original receipts were not available. The collection of payment receipts was deliberately omitted by the samplers, as requesting them is uncommon in Nepalese pharmacies and could have revealed the mystery shoppers.

### Assessment of the small-scale dissolution test

This study marks the first application in a field survey of a novel small-scale dissolution test developed by Rahman and Yoshida et al. Compared with the final USP dissolution test results of the 213 tested samples, the small-scale dissolution test showed overall false-positive and false-negative rates of approximately 1% (0.9%; 95% CI 0.7–1.3) and 6% (5.8%; 95% CI 4.8–6.5), respectively (AZM: 0.1% false-positive [95% CI 0.0–0.5], 15% false-negative [95% CI 13.3–17.4]; CFIX: 0 samples false-positive or false-negative; ESM: 3% false-positive [95% CI 2.5–4.7], 2% false-negative [95% CI 1.2–2.9]; LST: 0 false-positive, 3% false-negative [95% CI 2.5–4.6]).

These findings suggest that the required capacity for dissolution testing may be considerably lower than the USP test capacity, whereas the risk for false positive results remains low, making it a suitable screening technology for identification of poorly dissolving medicines. The technology is particularly useful in contexts in which dissolution testing is not performed because of capacity issues, which are often observed in LMICs<sup>9,10,12–16,33</sup>.

The small-scale dissolution test was applied to the  $\binom{6}{3} = 20$  three-unit combinations derived from the first-stage USP dissolution test results for all of the 213 tested 'real-world' samples collected in Nepal. These were then compared with the final USP dissolution test judgement of the samples. For computational simplicity and to avoid dramatically overrepresenting the combinations of those that underwent the second and/or third stage. Therefore, only the 20-three unit combinations  $\binom{6}{3}$  from the first stage were used, rather than those from subsequent stages ( $\binom{12}{3}$  or  $\binom{24}{3}$ ), because the results of the analysis within a sample were expected to be relatively similar given the nature of manufacturing. However, the occurrence of false positives and false negatives can be attributed to the inconsistent dissolution performance throughout the individual units within certain samples (see 'Supplementary Material\_Chemical Analysis Results' available as Excel table).

By reducing resources such as the amount of dissolution medium and LC mobile phase and run time that are required for analysis and by reducing maintenance and the working time required of the analyst, this small-scale dissolution test offers a cost-effective solution for the large-scale dissolution testing of distributed products and may thus be particularly beneficial in resource-limited contexts. For example, it could serve as a preliminary screening stage prior to full-scale dissolution testing of the poorly dissolving samples identified. When used in this way, false negatives from the small-scale dissolution test would require additional resources for full-scale dissolution testing but would not pose a risk for consumers in terms of pharmacopoeial dissolution quality.

In addition, the test may be useful in medicine quality surveys in which only small unit quantities of a sample can be obtained or to increase the capacity of samples that undergo dissolution testing<sup>35</sup>. However, the test is not appropriate as a regulatory replacement. Given the insufficient reliability of its results, replacing full-scale dissolution testing as a batch release requirement for manufacturers may also not be favourable. The test may be applied to larger data sets in future research to elucidate its broader applicability. In addition, the acceptance criteria may be modified to avoid false-positive judgements. More stringent criteria for the optimisation of the compliance of such a small-scale dissolution test may be worth investigating.

### Assessment of the portable Raman spectrometers

Raman scattering analysis using two portable spectrometers showed promise as rapid, potentially non-destructive screening technologies for detecting the APIs in most AZM, CFIX and LST samples through characteristic peaks observed near specific Raman shift wavenumbers (Supplementary Figs. S8–S13); moreover, CFIX results aligned with previous findings<sup>66</sup>. However, despite consistent analytical conditions, no corresponding peak was observed in the ESM samples compared with the omeprazole reference standard (Supplementary Figs. S14 and S15). This mismatch is not surprising, given the low sensitivity of Raman scattering analysis for APIs in coated formulations. Due to the limited penetration depth of the laser, the technique is particularly susceptible to

interference or fluorescence caused by enteric or film coatings – especially when the API content is low relative to the total mass of excipients.

Currently, few evidence supports the application of Raman scattering analysis to identify substandard medicines. Nevertheless, this would be highly desirable, as consumers and the healthcare system must also be protected from substandard products that can potentially cause health hazards, lead to under-treatment, or—in the case of antimicrobials—contribute to the development of antimicrobial resistance<sup>9,10,13,14,33,57–59</sup>. Comparative spectral analysis could benefit from information on excipients, product composition and authentic samples from the same batches as test-failing products. The findings highlight the limitations of relying solely on API peak identification using Raman scattering analysis, particularly when the API content is correct or varies moderately, e.g. exceeds 80%, in both compliant and substandard samples. Future studies should consider formulation-specific characteristics when Raman spectra are interpreted.

### Limitations

A limitation of this study is that sampling in the Saptari district could not be conducted using randomised sampling. Therefore, the SF medicine occurrence estimate of this study may not reflect the true prevalence in both districts and a national prevalence in Nepal. Additionally, twice as many samples per sampling site were taken from the Saptari district than from the Kathmandu district (4 vs. 2 samples) and the collection of individual samples in the Saptari district was not randomised. Both factors potentially biased the prevalence estimates, for instance, the population of failing LST samples such as those labelled with Trade name 3 may have been overrepresented because of unintentional preferential selection. Nepalese licenced vendors usually dispense blisters without secondary packaging material, which lists the amount of outer packaging collected. The study design may have prevented the detection of falsified medicines given only four APIs were collected exclusively from licenced vendors in just two districts. The authenticity investigation was based on the delivery of sample information, e.g. product names, lots, and pictures, but did not include the delivery of sample material to the manufacturers, limiting the information at the manufacturers' disposal to assess the samples. Additionally, it could not be confirmed that products from non-responding companies were not falsified. Given these limitations, this investigation cannot definitively conclude the absence of falsified medicines. Notably, the SF medicine occurrence estimate of this study was likely biased given that one unit of sample B-110 (Fig. 3), which was visibly damaged and expected to fail the dissolution test, was separated before the dissolution test was conducted for analytical purposes.

### Policy implications and future research

The results of this study should encourage the Nepalese government to continue its focus on combatting poor-quality medicines and ensuring that all medicines distributed in Nepal meet high-quality standards. Additionally, it should continue strengthening its regulatory capacity, aiming to enhance the WHO maturity levels (and therefore regulatory oversight) with a focus on product quality across all approved medicines in Nepal, including imported medicines. Improving pharmaceutical quality of medicines in production, distribution, and storage is crucial to ensure that all medicines distributed in Nepal meet the quality standards. Furthermore, the distribution of correct leaflets with the medicines should be promoted to increase medicine therapy safety.

Up-to-date registration lists of licenced vendors should be readily available from the responsible regional or national authorities for all districts and regions to ensure the integrity of the pharmaceutical supply chain, enabling regulatory compliance of all stakeholders within the chain and mitigating risks with the rapid identification and removal of non-compliant or illicit vendors. This practice would ideally include functional de-registration procedures for expired and unauthorised vendor licences and strategies to ensure the accuracy and update of such lists.

Future research steps could include multi-district sampling, testing a more diverse range of medicines—including additional APIs and dosage strengths—, conducting validation studies of screening technologies under field conditions, and investigating the causes of poor dissolution performance in certain samples, such as issues related to drug integrity or manufacturing problems. Additionally, the small-scale dissolution test (and dissolution testing in general) could be applied in medicine quality studies within and outside Nepal. In some contexts, this test may encourage the adoption of dissolution testing. Portable Raman scattering analysis could serve as a screening technology to detect products with no APIs, one or more APIs that are different from the stated one, or considerably low API content labelled as containing AZM, CFIX and LST, thereby assisting in the removal of these products. Nevertheless, detecting substandard medicines remains a challenge.

### Conclusion

This semi-randomised, cross-sectional study found that nearly 10% of medicines that contained AZM, CFIX, ESM and LST collected from licenced vendors in the Saptari and Kathmandu districts in Nepal were substandard or probably substandard. Notably, no falsified medicines were detected given the limited methods and sampling scope. Medicines against NCDs were significantly more affected than antibiotics, which may contribute to the burden of disease in Nepal.

The small-scale dissolution test showed only roughly 1% false-positive results compared with the USP dissolution test, indicating its suitability as a screening technology to enhance dissolution testing capacity, which is often lacking in resource-limited settings. Portable Raman scattering spectrometers reliably detected AZM, CFIX and LST in nearly all samples in the narrow sense of API presence or absence, which does not necessarily support their broader value for quality testing. Overall, they were unable to distinguish between compliant and substandard samples. The development of more effective screening technologies for substandard medicines remains urgently needed.

Concerted efforts are required to ensure the integrity of medical products in Nepal and worldwide to protect consumers, healthcare systems and economies from harm.

## Data availability

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

Received: 22 January 2025; Accepted: 14 August 2025

Published online: 21 August 2025

## References

- Chan, M. *Access to Medicines: Making Market Forces Serve the Poor in Ten Years in Public Health, 2007–2017: Report by Dr Margaret Chan, Director-General, World Health Organisation 14–24* (World Health Organisation, 2017).
- Ozawa, S., Shankar, R., Leopold, C. & Orubu, S. Access to medicines through health systems in low- and middle-income countries. *Health Policy Plan.* <https://doi.org/10.1093/heapol/czz119> (2019).
- The World Bank Group. *Data: World bank country and lending groups.* <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups> (2024).
- United Nations. *The 17 goals: the sustainable development goals.* <https://sdgs.un.org/goals> (2016).
- World Health Organisation. *Universal health coverage (UHC) key facts.* [https://www.who.int/news-room/fact-sheets/detail/universal-health-coverage-\(uhc\)](https://www.who.int/news-room/fact-sheets/detail/universal-health-coverage-(uhc)) (2023).
- United Nations OHCHR. Access to medicines—a fundamental element of the right to health. *OHCHR and the right to development.* <https://www.ohchr.org/en/development/access-medicines-fundamental-element-right-health> (cited 2024 Jun 01).
- World Health Organisation. *SDG target 3.8|achieve universal health coverage, including financial risk protection, access to quality essential health-care services and access to safe, effective, quality and affordable essential medicines and vaccines for all.* *World Health Statistics.* [https://www.who.int/data/gho/data/themes/topics/indicator-groups/indicator-group-details/GHO/sdg-target-3.8-achieve-universal-health-coverage-\(uhc\)-including-financial-risk-protection](https://www.who.int/data/gho/data/themes/topics/indicator-groups/indicator-group-details/GHO/sdg-target-3.8-achieve-universal-health-coverage-(uhc)-including-financial-risk-protection) (cited 2024 Jun 01).
- World Health Organisation. *1 in 10 medical products in developing countries is substandard or falsified.* <https://www.who.int/news/item/28-11-2017-1-in-10-medical-products-in-developing-countries-is-substandard-or-falsified> (2017).
- World Health Organisation. *Seventieth world health assembly update, 29 May 2017.* <https://www.who.int/news/item/29-05-2017-seventieth-world-health-assembly-update-29-may-2017> (2017).
- World Health Organisation. *Substandard and Falsified Medical Products.* <https://www.who.int/news-room/fact-sheets/detail/substandard-and-falsified-medical-products> (2018).
- World Health Organisation. *A Study on the Public Health and Socioeconomic Impact of Substandard and Falsified Medical Products* (World Health Organisation, 2017).
- Rahman, M. S. et al. The health consequences of falsified medicines- a study of the published literature. *Trop. Med. Int. Health.* **23**(12), 1294–1303 (2018).
- Salami, R. K. et al. Health, economic, and social impacts of substandard and falsified medicines in low- and middle-income countries: a systematic review of methodological approaches. *Am. J. Trop. Med. Hyg.* **109**(2), 228–240 (2023).
- Buckley, G. J. & Gostin, L. O., et al.; Countering the problem of falsified and substandard drugs in Washington (DC): *National Academies Press (US)* 1–14, 55–84, 137–196 (Committee on Understanding the Global Public Health Implications of Substandard, Falsified, and Counterfeit Medical Products; Board on Global Health; Institute of Medicine, 2013).
- Ozawa, S., Higgins, C. R., Nwokike, J. I. & Phanouvong, S. Modeling the health and economic impact of substandard and falsified medicines: A review of existing models and approaches. *Am. J. Trop. Med. Hyg.* **107**(1), 14–20 (2022).
- Ozawa, S. et al. Prevalence and estimated economic burden of substandard and falsified medicines in low- and middle-income countries: A systematic review and meta-analysis. *JAMA Netw. Open.* **1**(4), e181662. <https://doi.org/10.1001/jamanetworkopen.2018.1662> (2018).
- Ozawa, S., Chen, H. H., Lee, Y. A., Higgins, C. R. & Yemeke, T. T. Characterizing medicine quality by active pharmaceutical ingredient levels: A systematic review and meta-analysis across low- and middle-income countries. *Am. J. Trop. Med. Hyg.* **106**(6), 1778–1790 (2022).
- World Health Organisation. *WHO Global Surveillance and Monitoring System for Substandard and Falsified Medical Products* (Health Organisation, 2017).
- Nayyar, G. M. L. et al. Responding to the pandemic of falsified medicines. *Am. J. Trop. Med. Hyg.* **92**(6 Suppl), 113–118 (2015).
- Central Intelligence Agency. *The world factbook – south Asia: Nepal* <https://www.cia.gov/the-world-factbook/countries/nepal/> (2024).
- The World Bank Group. *Nepal country economic memorandum—climbing higher: toward a middle-income Nepal* <https://documents1.worldbank.org/curated/en/358501495199225866/pdf/115156-CEM-PUBLIC-SAREC-70p-Country-Economic-Memorandum-19-May-2017.pdf> (2017).
- OECD Public Governance Directorate and the European Union Intellectual Property Office (EUIPO). *Trade in counterfeit pharmaceutical products* [https://www.oecd-ilibrary.org/sites/a7c7e054-en/index.html?itemId=/content/publication/a7c7e054-en&csp\\_840766a83c2cbec116466a60bc97c49d&itemIGO=oecd&itemContentType=book&execsumm-d1e306](https://www.oecd-ilibrary.org/sites/a7c7e054-en/index.html?itemId=/content/publication/a7c7e054-en&csp_840766a83c2cbec116466a60bc97c49d&itemIGO=oecd&itemContentType=book&execsumm-d1e306) (cited 2024 Jun 01).
- Cho, S. H. et al. Determination of anabolic-androgenic steroid adulterants in counterfeit drugs by UHPLC-MS/MS. *J. Pharm. Biomed. Anal.* **111**, 138–146 (2015).
- Johnston, A. & Holt, D. W. Substandard drugs: A potential crisis for public health. *Br. J. Clin. Pharmacol.* **78**(2), 218–243 (2013).
- Odiase, P. O. Recalibrating African health laws to combat substandard and falsified medical products: beyond covid-19. *Int. J. Civ. Law Legal Res.* **1**(2), 1–9 (2021).
- Bhattarai, B. Department of Drug Administration, Ministry of Health and Population, Government of Nepal. National joint annual review, 2078/79 <https://mohp.gov.np/uploads/articles/DDA.pdf> (cited 2024 Jun 01).
- World Health Organisation. *Egypt and Nigeria medicines regulators achieve high maturity level in WHO classification and WHO launches list of regulatory authorities that meet international standards* <https://www.who.int/news/item/30-03-2022-egypt-and-nigeria-medicines-regulators-achieve-high-maturity-level-in-who-classification-and-who-launches-list-of-regulatory-authorities-that-meet-international-standards> (2022).
- World Health Organisation, International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. *Q1F stability guideline annex 10* [https://database.ich.org/sites/default/files/Q1F\\_Stability\\_Guideline\\_WHO\\_2018.pdf](https://database.ich.org/sites/default/files/Q1F_Stability_Guideline_WHO_2018.pdf) (2018).
- World Health Organisation. *Stability testing of active pharmaceutical ingredients and finished pharmaceutical products: stability conditions for WHO member states by region* [https://cdn.who.int/media/docs/default-source/medicines/norms-and-standards/guidelines/regulatory-standards/trs953-annex2-appendix1-stability-conditions-table-2018.pdf?sfvrsn=74032aec\\_12&download=true](https://cdn.who.int/media/docs/default-source/medicines/norms-and-standards/guidelines/regulatory-standards/trs953-annex2-appendix1-stability-conditions-table-2018.pdf?sfvrsn=74032aec_12&download=true) (2021).



30. Newton, P. N. et al. Covid-19 and risks to the supply and quality of tests, drugs, and vaccines. *Lancet Glob. Health.* **8**(6), e754–e755. [https://doi.org/10.1016/S2214-109X\(20\)30136-4](https://doi.org/10.1016/S2214-109X(20)30136-4) (2020).
31. Mackey, T. K. Prevalence of substandard and falsified essential medicines: still an incomplete picture. *JAMA Netw. Open.* **1**(4), e181685. <https://doi.org/10.1001/jamanetworkopen.2018.1685> (2018).
32. Sweileh, W. M. Substandard and falsified medical products: bibliometric analysis and mapping of scientific research. *Global Health.* **17**(1), 114. <https://doi.org/10.1186/s12992-021-00766-5> (2021).
33. Cavany, S. et al. The uncertain role of substandard and falsified medicines in the emergence and spread of antimicrobial resistance. *Nat. Commun.* **14**(1), 6153. <https://doi.org/10.1038/s41467-023-41542-w> (2023).
34. Anand, O., Yu, L. X., Conner, D. P. & Davit, B. M. Dissolution testing for generic drugs: an FDA perspective. *AAPS J.* **13**(3), 328–335 (2011).
35. Rahman, M. S. et al. Small-scale dissolution test screening tool to select potentially substandard and falsified (SF) medicines requiring full pharmacopoeial analysis. *Sci. Rep.* **11**(1), 12145. <https://doi.org/10.1038/s41598-021-91443-5> (2021).
36. Zhu, S., Yoshida, N., Kimura, K., Matsushita, R. & Tsuboi, H. Falsified vardenafil tablets available online. *J. Pharm. Biomed. Anal.* **177**, 112872. <https://doi.org/10.1016/j.jpba.2019.112872> (2019).
37. Sanada, T., Yoshida, N., Matsushita, R., Kimura, K. & Tsuboi, H. Falsified tadalafil tablets distributed in Japan via the internet. *Forensic. Sci. Int.* **307**, 110143. <https://doi.org/10.1016/j.forsciint.2020.110143> (2020).
38. 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**(1), 3. <https://doi.org/10.3390/pharmacy9010003> (2020).
39. Hajjou, M., Qin, Y., Bradby, S., Bempong, D. & Lukulay, P. Assessment of the performance of a handheld Raman device for potential use as a screening tool in evaluating medicines quality. *J. Pharm. Biomed. Anal.* **74**, 47–55 (2012).
40. Dégardin, K., Guillemain, A. & Roggo, Y. Comprehensive study of a handheld Raman spectrometer for the analysis of counterfeits of solid-dosage form medicines. *J. Spectrosc. (Hindawi)*. **1**, 1–13. <https://doi.org/10.1155/2017/3154035> (2017).
41. Schreiber, R. et al. Falsified and problematic methandienone products available online: active pharmaceutical ingredient identification by portable Raman spectrometers and quantification by ultra-high-performance liquid chromatography–Fourier transform mass spectrometry. *AAPS Open*. <https://doi.org/10.1186/s41120-024-00093-0> (2024).
42. Department of Drug Administration, Ministry of Health and Population, Government of Nepal. *National list of essential medicines Nepal, sixth revision* <https://www.dda.gov.np/content/essential-drug-list> (2022).
43. Brinkhoff, T. Central Bureau of Statistics Nepal. *Saptari—population* [https://citypopulation.de/en/nepal/admin/madhesh/15\\_\\_saptari/](https://citypopulation.de/en/nepal/admin/madhesh/15__saptari/) (2023).
44. Newton, P. N. et al. Guidelines for field surveys of the quality of medicines: A proposal. *PLOS Med.* **6**(3), e52. <https://doi.org/10.1371/journal.pmed.1000052> (2009).
45. Madden, J. M., Quick, J. D., Ross-Degnan, D. & Kafle, K. K. Undercover careseekers: Simulated clients in the study of health provider behavior in developing countries. *Soc. Sci. Med.* **45**(10), 1465–1482 (2001).
46. 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> (cited 2024 Jun 01).
47. United States Pharmacopoeial Convention, Inc. *The United States Pharmacopoeia 41 The National Formulary 36* (USP, 2018).
48. OECD. Good laboratory practice (GLP) <https://www.oecd.org/chemicalsafety/testing/good-laboratory-practiceglp.htm> (OECD, cited 2024 Jun 01).
49. Committee for Medicinal Products for Human Use, International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. *ICH guideline Q2(R2) on validation of analytical procedures* [https://www.ema.europa.eu/en/documents/scientific-guideline/ich-guideline-q2r2-validation-analytical-procedures-step-2b\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/ich-guideline-q2r2-validation-analytical-procedures-step-2b_en.pdf) (ICH, 2022).
50. MBH Media, Inc. Indian Rupee (INR) To Nepalese Rupee (NPR) Exchange Rate History for 2022 <https://www.exchange-rates.org/exchange-rate-history/inr-npr-2022> (cited 2024 Jun 01).
51. Exchange Rates UK. *US Dollar to Nepalese Rupee Spot Exchange Rates for 2022* <https://www.exchangerates.org.uk/USD-NPR-spot-exchange-rates-history-2022.html> (cited 2024 Jun 01).
52. MSH (Management Sciences for Health). *International Medical Products Price Guide, 2015 Edition* <https://msh.org/wp-content/uploads/2020/03/msh-2015-international-medical-products-price-guide.pdf> (MSH, 2016).
53. World Health Organisation, Health Action International. *Measuring Medicine Prices, Availability, Affordability and Price Components* 2nd Edition (World Health Organisation, 2008).
54. Akpobolokemi, T., Martinez-Nunez, R. T. & Raimi-Abraham, B. T. Tackling the global impact of substandard and falsified and unregistered/unlicensed anti-tuberculosis medicines. *J. Med. Access*. <https://doi.org/10.1177/23992026211070406> (2022).
55. Nepal Health Research Council (NHRC), Ministry of Health and Population (MoHP), Institute for Health Metrics and Evaluation (IHME), Monitoring Evaluation and Operational Research (MEOR). *Nepal Burden of Disease 2019: A Country Report Based On The 2019 Global Burden Of Disease Study* [https://nhrc.gov.np/wp-content/uploads/2022/02/BoD-Report-Book-includ-Cover-ma-il-6\\_compressed.pdf](https://nhrc.gov.np/wp-content/uploads/2022/02/BoD-Report-Book-includ-Cover-ma-il-6_compressed.pdf) (2021).
56. Institute for Health Metrics and Evaluation (IHME). *Global Burden of Disease 2021: Findings From the GBD 2021 Study* (IHME, 2024).
57. Holmes, A. H. et al. Understanding the mechanisms and drivers of antimicrobial resistance. *Lancet* **387**(10014), 176–187 (2015).
58. Kelesidis, T. & Falagas, M. E. Substandard/counterfeit antimicrobial drugs. *Clin. Microbiol.* **28**(2), 443–464 (2015).
59. Zabala, G. A. et al. Substandard and falsified antibiotics: Neglected drivers of antimicrobial resistance?. *BMJ Glob. Health.* **7**(8), e008587. <https://doi.org/10.1136/bmjgh-2022-008587> (2022).
60. Bhandari, B. & Rayamajhi, G. Counterfeit healthcare products: Nepal at a vulnerable Position. *J. Nepal Med. Assoc.* **60**(256), 1070–1072 (2022).
61. Dhakal, N. et al. Assessment of quality of essential medicines in public health care facilities of Nepal: Findings of nationwide study. *PLOS Glob. Public Health.* **3**(5), e0001841. <https://doi.org/10.1371/journal.pgph.0001841> (2023).
62. Aryal, B. P. Some drugs supplied by Govt, Pvt sector are substandard: NHRC in *myRepublica* <https://myrepublica.nagariknetwork.com/news/33228/> (2017).
63. Karki, K. B. et al. Quality of drug in health facilities: a cross-sectional study. *J. Nepal Health Res. Counc.* **18**(4), 644–648 (2021).
64. Gyanwali, P. et al. Surveillance of quality of medicines available in the Nepalese market: A study from Kathmandu valley. *J. Nepal Health Res. Counc.* **13**(31), 233–240 (2016).
65. Hauk, C., Hagen, N. & Heide, L. Identification of substandard and falsified medicines: Influence of different tolerance limits and use of authenticity inquiries. *Am. J. Trop. Med. Hyg.* **104**(5), 1936–1945 (2021).
66. Rahman, M. S. et al. A comprehensive analysis of selected medicines collected from private drug outlets of Dhaka city, Bangladesh in a simple random survey. *Sci. Rep.* **12**(1), 234. <https://doi.org/10.1038/s41598-021-04309-1> (2022).

## Acknowledgements

This study was made possible by the generous support of the Japan Society for the Promotion of Science, for which the authors express their deep gratitude. The authors would like to express their appreciation to the Department of Drug Administration in Nepal, Tribhuvan University and the Nepalese government for their cooperation. The authors would like to thank AstraZeneca for providing NEXIUM 20 mg capsules. The contribution

of R.S. was supported by the ‘Monbukagakusho:MEXT’ PhD scholarship from the Japanese Ministry of Education, Culture, Sports, Science and Technology in Japan. We also thank Anahid Pinchis (MBA, BSc) from Edanz (<https://jp.edanz.com/ac>) for their assistance in editing a draft of this manuscript.

## Author contributions

R.S.: Conceptualisation, methodology, investigation, data curation, method validation and formal analysis (AZM, ESM, LST), Visualization, software, writing – original draft, writing – review & editing. Md.A.H.: Method validation and formal analysis (CFIX), writing – review & editing. M.S.R.: Conceptualisation, methodology, NHRC approval, sampler training, Writing – review & editing. B.K.P.: Resources, selection of samplers, sampling and supervision, writing – review & editing. B.R.: Principal investigator in Nepal, NHRC approval, conceptualisation, methodology, resources, selection of samplers, sampling and supervision, collection of authentic products, Writing – review & editing. K.K.: Conceptualisation, methodology, data curation, writing – review & editing, supervision. N.Y.: Principal investigator in Japan, conceptualisation, methodology, investigation, data curation, funding acquisition, project administration, writing – review & editing.

## Funding

This work was supported by JSPS KAKENHI Grant number JP19KK0236. The funding bodies had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Declarations

## Competing interests

The authors declare no competing interests.

## Ethical approval

Ethical approval for the collection of medicine sample materials on Nepalese territory was obtained from the Nepal Health Research Council on February 9, 2022 (Ref. No.: 2008) prior to the commencement of this study. This study did not involve or use any experimental animals, human participants, organs, tissues, or vulnerable groups. The role of the participating samplers was solely to physically collect medicine sample materials and data and transport both Tribhuvan University; no personal data about the samplers was collected, nor were they subject of investigation in this study.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-16340-7>.

**Correspondence** and requests for materials should be addressed to N.Y.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025