Article

Detecting clinical medication errors with AI enabled wearable cameras

Check for updates

Justin Chan [1,2,6], Solomon Nsumba [3,6], Mitchell Wortsman[1], Achal Dave[4], Ludwig Schmidt[1],
Shyamnath Gollakota [1] ✉ & Kelly Michaelsen[5] ✉

Drug-related errors are a leading cause of preventable patient harm in the clinical setting. We present the first wearable camera system to automatically detect potential errors, prior to medication delivery. We demonstrate that using deep learning algorithms, our system can detect and classify drug labels on syringes and vials in drug preparation events recorded in real-world operating rooms. We created a first-of-its-kind large-scale video dataset from head-mounted cameras comprising 4K footage across 13 anesthesiology providers, 2 hospitals and 17 operating rooms over 55 days. The system was evaluated on 418 drug draw events in routine patient care and a controlled environment and achieved 99.6% sensitivity and 98.8% specificity at detecting vial swap errors. These results suggest that our wearable camera system has the potential to provide a secondary check when a medication is selected for a patient, and a chance to intervene before a potential medical error.

At least 1 in 20 patients are affected by preventable patient harm in a clinical setting based on a meta-analysis of studies involving patients of all ages from a range of specialities including surgery, intensive care, emergency department, obstetrics, and primary care[1]. Drug-related errors are a leading cause of these incidents, with up to 12% of these errors resulting in serious harm or death[1]. Studies estimate that between 140,000 to 440,000 deaths annually in the United States can be attributed to medical errors[2–4] with 80% of adverse events occurring in the hospital, and 41% occurring in the operating room[5]. Error rates for drug delivery events in hospitals are estimated at about 5–10% of all drugs given[6–8]. Drug administration errors are the most frequently reported critical incidents in anesthesia[9] and the most common cause of serious medical errors in the intensive care unit[10]. Adverse events associated with injectable medications are estimated to impact 1.2 million hospitalizations annually with 5.1 billion dollars in associated costs[11].

Syringe and vial swaps are drug errors that can result in patient harm[12–19]. For intravenous medication injections, clinicians often must remove the medication from a vial and transfer it to a syringe[19]. Vial swap errors occur at this step when the syringe is labeled incorrectly, or the wrong vial is selected for use[19]. These errors are also referred to as substitution errors, account for 20% of drug errors and lead to the wrong drug being given to a patient[19]. Another 20% of drug errors are due to syringe swaps, where the drug is labeled correctly but given in error[19,20]. A recent, highly publicized criminal trial of a nurse's vial swap error that resulted in a

patient's death has highlighted this threat to patient safety[21,22]. While syringe and vial swaps in particular have been documented most prominently in anesthesiology[12,19,20], medication administration errors have also been shown to occur across various medical settings including pediatrics[23], the emergency department[16], rehabilitation unit[24], radiology[22], medical and surgical wards[25], and the prehospital setting[18]. In the intensive care unit, 44% of medication errors occurred during drug administration[15]. In the regular hospital ward, medication administration errors accounted for 34% of preventable events, with only 2% of drug administration errors detected before they occurred[13]. Medication administration errors occur across the spectrum of clinical care settings and can cause patient harm or death[22].

Prior efforts to minimize medication errors have involved color coding medications[26], using tall man style lettering on labels[27], standardized safety protocols[28,29], additional observers[30], prefilled syringes[31] and barcode scanning[32,33]. While some of these mitigation strategies are passive, methods like barcode scanning require active participation by clinical providers prior to medication administration, and compliance with these safety mechanisms can be problematic[34–36]. Busy clinicians often develop 'workaround' techniques to decrease their workload up to 62% of the time during medication administration tasks, often administering a drug first before entering the drug name manually into the clinical record[37]. Such workarounds can lead to a threefold increased chance of a medication error[38]. Thus, there exists a need for automating medication checks in real-time prior to administration that fits into the current clinical environment.

[1]Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, USA. [2]School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA. [3]Department of Computer Science, Makerere University, Kampala, Uganda. [4]Toyota Research Institute, Los Altos, CA, USA. [5]Department of Anesthesiology & Pain Medicine, University of Washington, Seattle, WA, USA. [6]These authors contributed equally: Justin Chan, Solomon Nsumba. ✉e-mail: gshyam@cs.washington.edu; kellyem@uw.edu
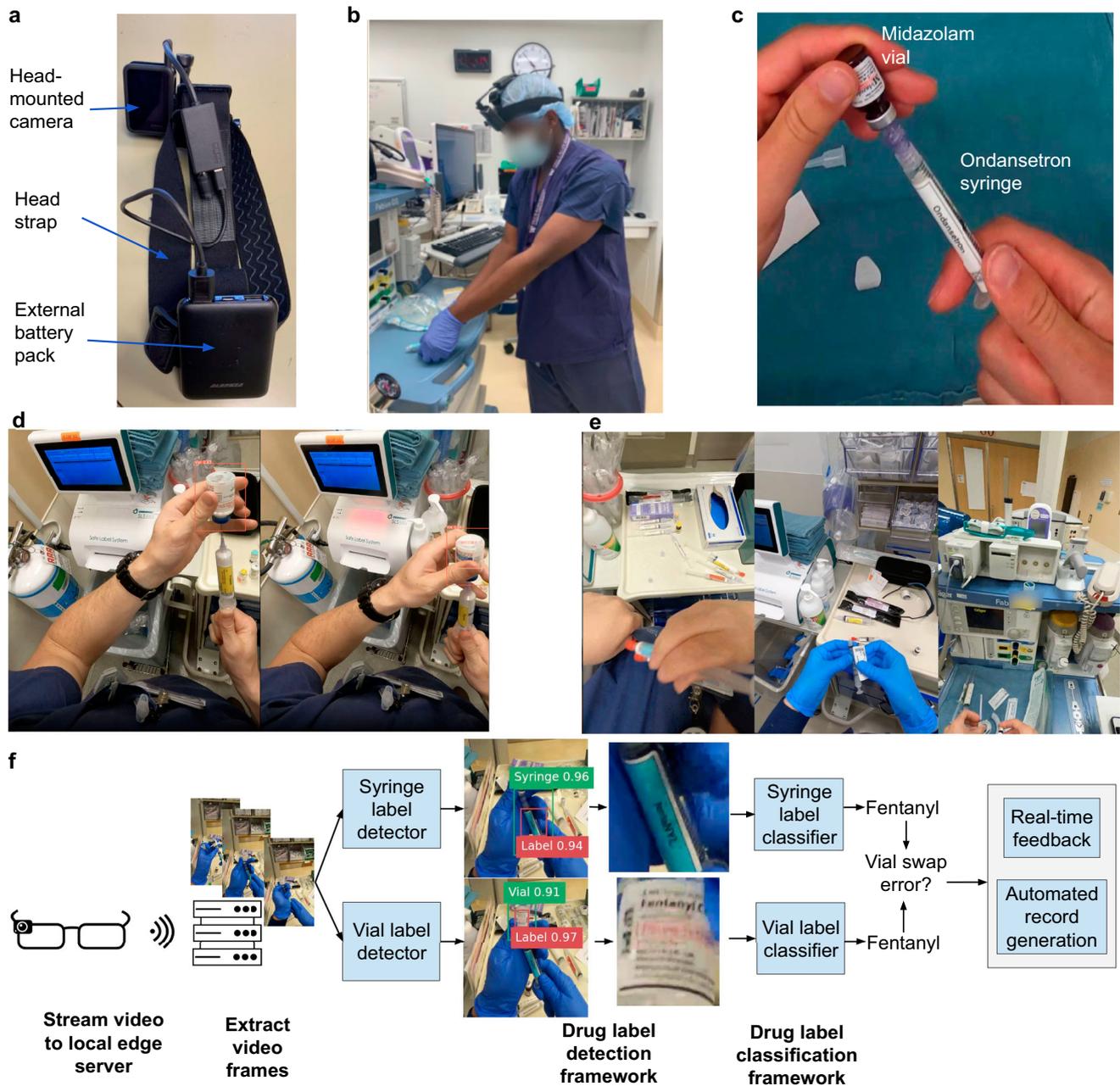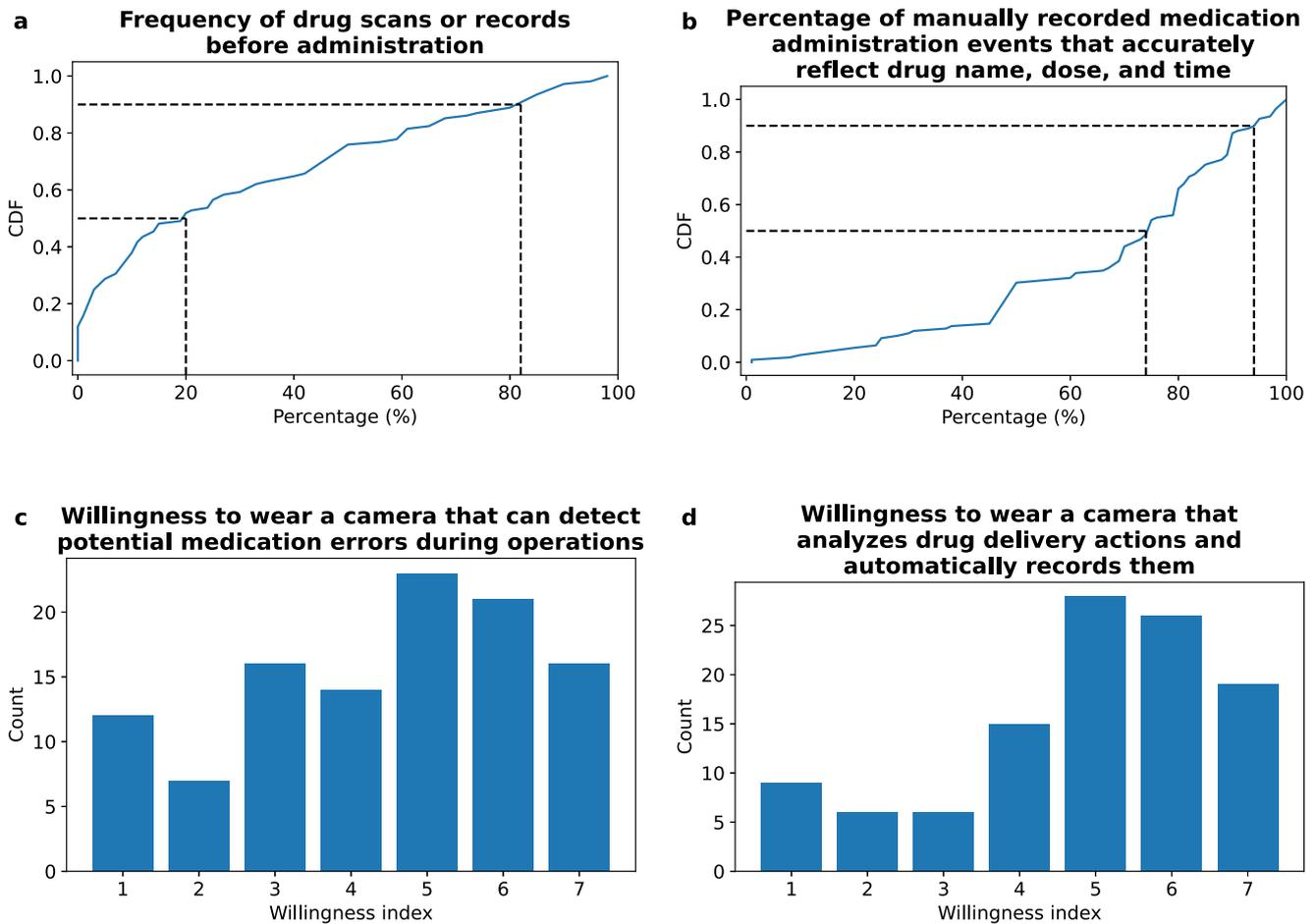
Fig. 1 | Overview of wearable camera system for detecting medication errors.
a Head-mounted camera (GoPro Hero8) attached to a head strap and powered by an external battery pack. **b** Recording drug preparation events using wearable camera system. **c** Vial swap error from midazolam vial into syringe with ondansetron label. **d** Real-world operating room use showing that syringe and vial labels are not visible in every frame and can be obscured by the provider's hand or tilted away. **e** Real-world environments containing other syringes and vials in the background unrelated to the drug drawup event. **f** Workflow demonstrating how drug drawup events captured through our system are detected and classified. Medication errors can be communicated to the provider through real-time auditory or visual feedback and automatically recorded.

Automated medication checks could serve as second set of 'eyes' verifying the work of an anesthesia provider to ensure that medication errors have not occurred, or if they had, provide a warning before the unintended medication reaches the patient. As anesthesia providers are already required to wear protective eyewear while working in the OR as a barrier against fluids, smart eyewear with built-in cameras[39–42] present themselves as a potential platform to visually detect medication errors in the operating room before they occur.

Here, we introduce a wearable camera system to automatically detect vial swap errors that occur when a provider incorrectly fills a syringe from a mismatched drug vial (Fig. 1). We demonstrate the use of

deep learning to detect syringes and vials in a provider's hand, classify the drug type on the label, and automatically check if they match in order to detect vial swap errors. Our system is trained on a large-scale drug event dataset captured from head-mounted cameras worn by anesthesiologists or certified registered nurse anesthetists performing their usual clinical workflows to prepare medications for surgery in an operating room environment. Our algorithms can detect vial swaps in real-time from videos wirelessly streamed to a local edge server with a GPU, which could enable real-time auditory or visual feedback by alerting providers to medication errors prior to drug administration, providing an opportunity to intervene.

a Frequency of drug scans or records before administration

b Percentage of manually recorded medication administration events that accurately reflect drug name, dose, and time

c Willingness to wear a camera that can detect potential medication errors during operations

d Willingness to wear a camera that analyzes drug delivery actions and automatically records them

Fig. 2 | Survey results across anesthesiology providers ($n = 109$). Cumulative distribution function (CDF) of providers' self-assessment: **a** percentage of time they scan the barcode of a drug or manually record it prior to administration and **b** percentage of accurately recorded drug administration events, including drug name, dose, and administration time. **c** Histogram displaying providers' willingness (1 lowest, 7 highest) to wear a camera capable of detecting potential medication errors, such as when a syringe label doesn't match the drug vial during syringe preparation or when a medication is selected to which the patient is allergic. **d** Histogram displaying providers' willingness (1 lowest, 7 highest) to wear a camera that analyzes drug delivery actions and automatically records them in the patient's chart.

Results

Acceptability of wearable cameras by anesthesia providers

To motivate the need for a wearable system that can automate medication checks and minimize errors, we conducted a survey across 109 anesthesiology providers spanning a range of sub-specialties including intensive care, pediatrics, pain, cardiothoracic, generalist, obstetric, and neurologic across four hospital sites. We show in Fig. 2a that the median percentage of time that providers scan the barcode of a drug or manually record its contents prior to administration (a current technique to help prevent errors) is only 20% (IQR: 46.25). When providers manually record medication administration events, they believe that a mean and SD of 68 ± 24% records accurately reflect the drug name, dose, and administration time (Fig. 2b). The survey also assessed the willingness of participants to wear a camera system that could automate these medication checks. On a scale of 1 (lowest) to 7 (highest), participants expressed a median willingness score of 5 for a wearable system that could automatically detect medication errors (Fig. 2c). Furthermore, 88% of participants indicated that they would use a lightweight, accurate, and FDA-cleared camera system if it was shown to decrease medication errors, made charting easier or if it was required by their employers (Fig. 2d).

Clinical dataset

Large publicly accessible databases such as ImageNet[43] do not have point-of-view operating room data of drug preparation events. While ImageNet contains data for syringes, they are often seen in the foreground of stock photos with a relatively uncluttered background or depicted in the context of drug abuse or recreational use rather than authorized medical care. In contrast, the visual environment of the operating room is relatively chaotic, with many tubes and wires in the background. To address this problem, we create a unique dataset that can provide insights into preparatory tasks preceding the twenty-two million surgeries performed annually in the United States alone[44], and serve as a benchmark for developing deep learning algorithms that can detect medication errors and understand the operating room environment.

Creating such a dataset requires addressing two challenges: First, the labels on syringes and vials are small, and can be difficult to read or classify if not captured on a high-resolution camera (Supplementary Fig. 1). Second, the wearable camera's field of view should be large enough to capture the drawup event, and directly face the drug label without any obstructions. To address these challenges, we collected our dataset using a high-resolution head-mounted 4K resolution camera located by the forehead that was tilted downward to detect drawup events (Supplementary Fig. 2). Other sites for a camera were considered such as the chest or anesthesia machine, but views from these sites were often blocked by people, equipment or surgical drapes and did not provide an unobstructed view of the drug delivery event.

Our dataset was collected from February 2021 to July 2023 across two clinical sites at the University of Washington. Our dataset contains video footage of 13 anesthesiology providers, and 17 operating locations over

Table 1 | Dataset statistics

Statistic	*n*
Days recorded in operating room	55
Number of operating rooms	17
Number of providers	13
Number of clinical sites	2
Drug drawup events	621
Syringe-in-hand videos	1587
Vial-in-hand videos	819
Training data	# frames
Syringe detector	1596
Vial detector	262
Syringe classifier	9332
Vial classifier	62821
Evaluation data	# frames
Syringes in operating room	609816
Vials in operating room	254117
Syringes in controlled environment	424835
Vials in controlled environment	304326

Number of different events and videos in the dataset, and number of frames used to train and evaluate the syringe and vial detector and classifier.

55 days (Table 1). It is comprised of real-world operating room videos of anesthesiologists performing drug preparation events as recorded from head-mounted cameras to capture their point-of-view (Fig. 1a, b). The dataset includes a variety of different drug preparation styles, workstation setups, lighting conditions, syringe labeling machines and operating room environments. The video streams collected from the head-mounted cameras were carefully segmented into clips containing drug draw up events. Drug draw-ups involved a provider actively drawing medication from a vial into a syringe. Drug delivery events involved a provider actively injecting a syringe containing a drug into a patient's intravenous fluid line. This dataset is further augmented with vial swap errors performed by a trained researcher in a controlled environment (Fig. 1c). Videos are annotated to indicate if a syringe and vial label can be seen (Fig. 1d), if syringes and vials are in the background (Fig. 1e), ambient lighting condition, and provider key. A subset of the videos are annotated at the frame-level with bounding boxes indicating the location of the syringe, vial, and drug label held in the provider's hand. These videos are used to generate training and validation data for our object detection framework. Videos with a drug label are annotated at the video-level with the drug name, and are used to train and evaluate our drug label classification system (Fig. 1f). All videos in the dataset are de-identified, and all protected health information present in videos are cropped or blurred. No audio recordings are present in the dataset.

Drug label detection

Accurate detection of syringe and vial drug labels in the hands of a clinical provider in a timely fashion is the first step towards recognizing potential medication errors. Achieving this requires addressing three challenges: First, the system should only detect syringes and vials in the provider's hand, not those seen in the background (Fig. 1e, Supplementary Fig. 3a). Second, the system should recognize objects regardless of whether the provider is wearing gloves (Supplementary Fig. 3b). Third, the system should be able to recognize the drug label, not just the syringe or vial object, for a variety of different drug labels and styles (Supplementary Figs. 4, 5, 6). To address these challenges, we created an annotated dataset that marks the location of syringes, vials, and drug labels only in the hands of providers, for both gloved and non-gloved providers. Further, to increase the accuracy of our system, we augment the dataset using label-preserving image transformations[45,46] (see Methods) and train two separate detectors one to detect syringes and syringe labels, and one to detect vials and vial labels. In order for the detector to run in realtime on a GPU, the video frames were rescaled from a resolution of 2160 × 3840 to 384 × 640 to obtain the drug label bounding box which is then upscaled back to the original video resolution.

We evaluated the performance of our detector at identifying syringes and vials in a medical provider's hand. We used the pretrained YOLOv5x object detection model[47], and fine-tuned it on a custom training dataset of syringes and vials from real-world operating room medication events. Our training, validation, and held-out test sets contain 1596, 308 and 111 images respectively with annotations for syringes, vials, and their labels. We note that there were no overlaps in medication events used for the training, validation, and test sets. Figure 3 shows the precision-recall curve of the object detection model for each of the four objects of interest with and without image augmentation. The AP_{50} (average precision when the intersection over union is set to 50) is 0.934 and 0.953 for the syringe and syringe label respectively when image augmentation is applied (Fig. 3a), and is 0.935 and 0.903 for the drug vial and drug vial label respectively when image augmentation is applied (Fig. 3c). In contrast, the AP_{50} is lower without the use of image augmentations and decreases to 0.897 and 0.924 for the syringe and syringe label respectively (Fig. 3b) and 0.906 and 0.801, respectively for the vial and vial label (Fig. 3d). We note that training a single object detector to recognize all four objects at once results in decreased performance with a mAP_{50} across all objects of 0.745. In contrast, with the use of two separate object detectors, the mAP_{50} across all four classes is 0.931. The inference time of the model with image augmentation is 23.9 ms on a NVIDIA A40 GPU.

Drug label classification

Once a drug label is detected, the next step is to classify the medication on the label. There are three challenges to achieving this goal: First, creating a dataset of drug labels using real-world medication events in the operating rooms alone would result in an imbalanced dataset as some medications are handled frequently while others are rarely handled. Second, drug labels can be obscured from the camera's view by the provider's fingers (Fig. 1d). Third, spurious frames of background syringes and vials may occasionally be mistaken as being in the provider's hand and can cause the classifier to produce an incorrect result. To address these challenges, we augmented the classifier training dataset with drug events that occurred less frequently in the operating room using additional drug events collected in the controlled environment (Supplementary Tables 3, 4). We curated the dataset to only contain clear and readable depictions of either a syringe label, a vial label or both (Table 2) that were not obscured by the provider's hand. To reduce the likelihood of classifying background objects incorrectly produced by the object detection model as a drug label, we created a class consisting of background object frames. To do this, we pass videos of medication events through our object detection model and use frames with a detection probability less than 0.2. Finally, to improve system accuracy we apply image transformation augmentations to our training dataset (see Methods).

We first evaluate the performance of our classifier on the set of medication events collected in a controlled setting. Here, the events consisted of a trained researcher recording correct drug drawups, where the label on the syringe and vial matched, and incorrect drug drawups, where the two labels differed. Events were recorded across different operating rooms with and without gloves on. We note that our classifiers were trained with more drug classes than what appeared in the test sets, as such the confusion matrices only display the subset of drugs that were in the evaluation data and any drug classes that were a misprediction by the classifier. Supplementary Fig. 7a, b shows that our syringe drug label classifier correctly classified syringe labels for 319 of 323 (98.8%) events with image augmentation and a background image class and in 316 of 323 (97.8%) events without these additions. We observe that the syringe labels for vecuronium was misclassified as the rocuronium drug in several events. This is likely because the background color of the label for both these drug types are identical, and furthermore

Fig. 3 | Drug label detection performance in real-world operating room environment. Precision-recall curves for **a**, **b** syringe and syringe label detector with and without image augmentation and **c**, **d** vial and vial label detector with and without image augmentation.
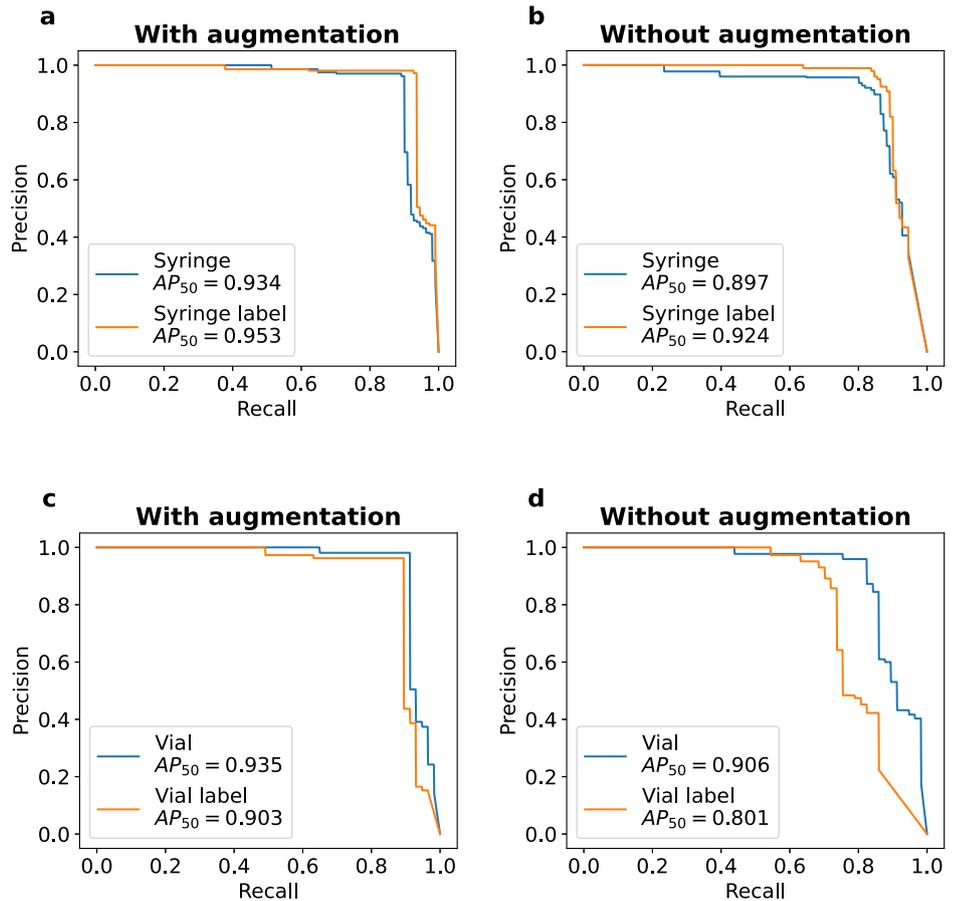
both drugs have similar spellings, sharing a common suffix. We note that drugs that have the same background color share similar medication effects[48,49], and these misclassifications are less likely to cause injury compared to misclassifications to a different color. In Supplementary Fig. 7c, d, we show that our vial drug label classifier correctly classified vial labels in 294 of 298 (98.7%) events with image augmentation and a background image class and in 309 of 328 (94.2%) events without these additions.

Next, we evaluated the performance of our system at classifying syringe drug labels and vial drug labels in real-world operating room medication events captured at the University of Washington Medical Center (UWMC) Montlake. Figure 4a, b shows the confusion matrix of our syringe classifier on medication events with a syringe label captured in the operating room at our main clinical site. The syringe classifier correctly classified syringe labels in 227 of 230 (98.7%) events with image augmentation and a background image class and in 228 of 231 (98.7%) events without these additions. Figure 4c, d shows the performance of the vial classifier on medication events containing vial labels in the operating room. The vial classifier correctly classified vial labels in 122 of 123 (99.2%) events with image augmentation and a background image class and in 122 of 143 (85.3%) events without these additions.

Predicting vial swap errors
We evaluate the feasibility of our system to detect vial swap errors. To do this we pool together the classification results for all medication events collected from the real-world operating rooms and the controlled setting, and the classifier produced a drug classification for both labels in the video. When image augmentation and a background image class were applied, there were 253 valid drug drawup events where the syringe and vial label matched, and 165 events containing vial swap errors, and the system achieved a sensitivity and specificity of 99.6% (95% CI 98.8 to 100.0%) and 98.8% (95% CI 97.1 to

100.0%) respectively at detecting vial swap errors (Supplementary Fig. 8a). Without image augmentation or a background image class, there were 277 valid drug drawup events, and 193 vial swap errors, and the sensitivity and specificity of the system was 99.3% (95% CI 98.3 to 100.0%) and 85.0% (95% CI 79.9 to 90.0%) respectively (Supplementary Fig. 8b).

Evaluating generalization
We also evaluated how performance generalizes across different environmental conditions: First, we measured performance on a dataset of medication events collected from a held-out clinical site. Second, we performed a subgroup analysis for 13 different providers, 5 of which were not used to train the system. Finally, we performed benchmark testing in different lighting conditions and distances.

Generalization to held-out clinical site. We evaluated generalization to a held-out dataset of real-world operating room medication events collected from a different hospital that is affiliated with our institution, Harborview Medical Center. This dataset is comprised of 72 and 32 events containing syringe and vial drug labels respectively. Our system correctly classifies syringe drug labels in 71 of 72 (98.6%) events with image augmentation and a background image class and 70 of 72 (97.2%) events without these additions (Supplementary Fig. 9a, b). Our vial label classifier correctly classifies vial labels in 23 of 24 (95.8%) events with image augmentation and a background image class and in 25 of 32 (78.1%) events without these additions (Supplementary Fig. 9c, d).

Subgroup analysis. We perform a subgroup analysis of medication event classification performance for drug labels after image augmentation and a background image class on the dataset of real-world events collected in the operating room setting at UWMC Montlake. Figure 5a

Table 2 | Summary of evaluation dataset for drug label classifier

Drug	Operating room		Controlled environment	
	Syringes (n = 303)	Vials (n = 175)	Syringes (n = 323)	Vials (n = 328)
Amiodarone	0	0	0	3
Cefazolin	4	1	0	0
Cisatracurium	0	0	1	0
Dexamethasone	31	15	28	25
Etomidate	0	0	2	12
Fentanyl	52	34	33	21
Glycopyrrolate	3	4	3	15
Hydromorphone	8	1	40	20
Ketamine	2	1	2	0
Ketorolac	0	0	22	26
Lidocaine	44	22	33	44
Magnesium sulfate	0	0	0	2
Metoprolol tartrate	0	0	0	1
Midazolam	20	13	33	24
Neostigmine	0	0	30	22
Ondansetron	34	26	29	36
Phenylephrine	11	0	0	0
Propofol	58	34	27	18
Rocuronium	26	17	23	33
Sugammadex	7	6	11	22
Vasopressin	3	1	1	4
Vecuronium	0	0	6	0

Number of drug preparation video events captured in the operating rooms at two clinical sites and a controlled environment.

shows the classification accuracy of both classifiers for 13 anesthesiology providers. Drug events captured by providers with ID 3 to 10 were used to train the drug label detector and classifier. The mean and standard deviation accuracy of the syringe and vial label classifier was 96.0 ± 7.6% and 99.5 ± 1.7% respectively across all 13 providers. Figure 5b shows the classification accuracy for events with a clear background, and events with syringes and vials in the background. When the background was clear of objects, the syringe and vial classifier correctly classified all drug labels within these events. When syringes are present in the background, the accuracy of the syringe and vial classifier is 98.2 and 99.2%, respectively. When vials are in the background the syringe and vial classifier obtain accuracies of 98.6 and 99.1% respectively. Figure 5c shows the performance of the classifiers for providers with and without gloves. The syringe classifier obtained accuracies of 98.4 and 98.2% with and without gloves respectively, while the vial classifier had accuracies of 97.3 and 100.0% with and without gloves respectively.

Benchmark testing. We perform benchmark testing across different environmental conditions on a held-out dataset of medication events collected by a researcher in a controlled environment. We evaluate the performance of the syringe classifier on a set of simulated drug drawup and delivery events collected in different ambient lighting conditions in the operating room. Specifically, we recorded events with five different syringe drug labels in bright, dim and dark lighting conditions respectively (Fig. 6a). These events were recorded using the head-mounted camera's default and low-light camera settings. Figure 6b shows the confusion matrix when the default camera settings were used showing a correct classification of 10 of 10 events, 9 of 9 events and 2 of

10 events for the bright, dim, and dark ambient lighting conditions respectively. When the low-light camera settings is used, Fig. 6c shows a correct classification as 10 of 10 events, 10 of 10 events and 6 of 10 events for the bright, dim, and dark conditions respectively. These results suggest the low-light camera settings are slightly better than the default camera settings under dark lighting conditions. An end-to-end implementation of the system could use a light sensor to turn on the low-light camera mode in response to darkened ambient lighting conditions. Further, the head-mounted camera in our study supports an additional light attachment[50] that could provide increased illumination under dark lighting conditions.

Next, we evaluate our system's classification performance of syringe drug labels as a function of distance. To do this, we assessed syringe classification with three different drug labels at increasing distances from both a top view as well as a side view. Figure 6d shows that when the camera is head-mounted and recording the syringe from a top view, our system was on average able to correctly classify 100.0 ± 0.0% of the frames across all the videos at a close distance of 13 cm, and 91.4 ± 9.1% of the frames at the furthest tested distance of 56 cm (about as far as a hand holding a syringe can be from a head mounted camera). Figure 6e shows that for videos where the side view of the syringe was presented to the camera, 99.5 ± 0.7% of frames were correctly classified when it is at a close distance of 15 cm and 55.6 ± 41.6% at 56 cm. This shows that although our classifier was trained on footage recorded from the top view, it is able to classify footage recorded from a side view as well. We note that to classify a drug drawup event as containing a particular syringe label, our algorithm only uses the most common drug class prediction, and does not impose a threshold on the fraction of frames containing a syringe label, as a syringe label may only be visible for a small fraction of the event.
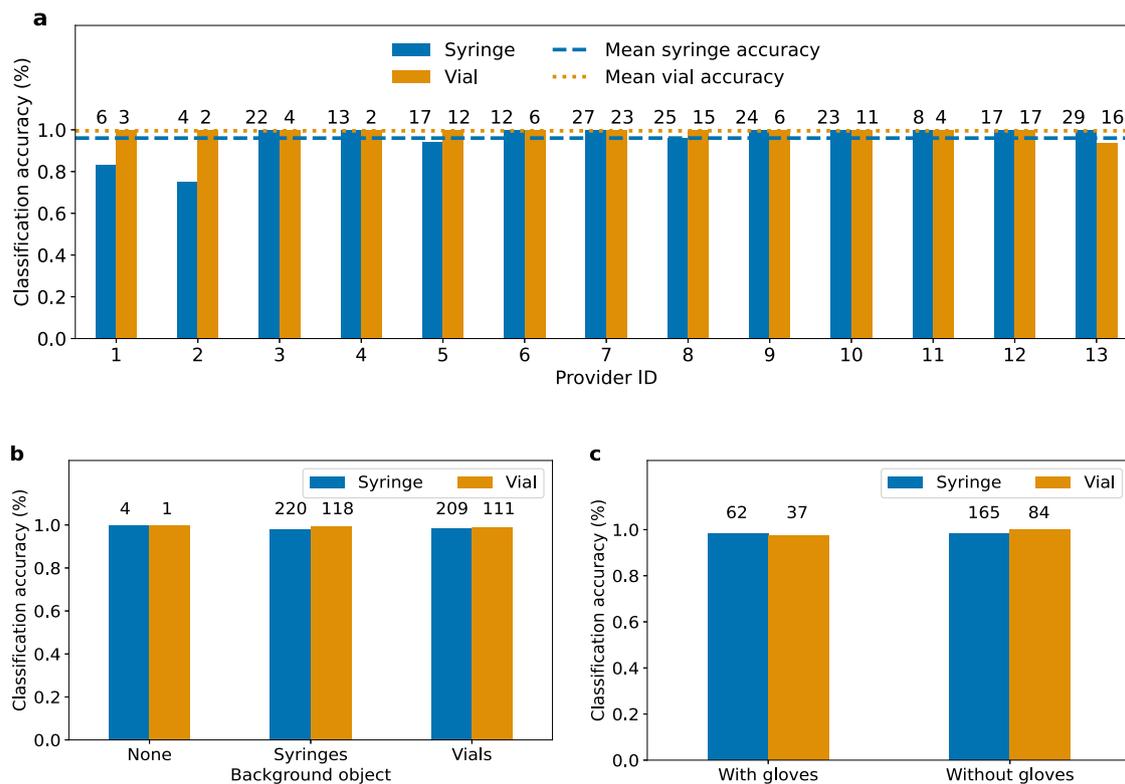
Discussion

This study demonstrates a wearable camera system for automatically detecting and classifying drug labels on syringes and vials, and flagging potential vial swap errors in the operating room. Additionally, this system also has the potential for detecting syringe swap errors, which occurs when a clinician injects an unintended medication into a patient. Syringe swap errors were not quantified in this study as they would require the clinician to point out when a mistake had occurred.

In reviewing a subset of our real world medication event dataset that was collected across a seven month period totaling 212 drug delivery events, the mean time between when a syringe was selected by a provider and when the drug was given was 9.9 ± 7.2 s (Supplementary Fig. 10). The system described in this paper took less than 25 ms on an NVIDIA A40 GPU and could be used identify syringe swaps prior to drug injection in drug delivery events recorded in our patient dataset.

The American Society of Anesthesiologists supports the use of the color-coding scheme for medication labels established by the American Society for Testing and Materials, which groups drugs into nine classes based on similar medication effects[48,49]. In our study, after image augmentation and a background image class, 6 of 7 syringe label misclassifications that occurred during drug drawup events in the real-world OR and controlled environment datasets were misclassified to drug labels of the same color. We note that if misclassifications of this type were to occur at the drug delivery stage they are likely to be less injurious compared to misclassifications to a different color category.

In its current iteration, our system has achieved the expected accuracy for clinical use as determined during our provider survey. Participants were asked about the minimum performance of the drug classifier to be considered for use in daily practice. 56% of survey participants indicated that the performance of the drug classifier should at minimum be 50–95% for use in daily practice, with the remaining 44% expressing a minimum performance of 99%. With image augmentation and a background image class, our syringe and vial classifier achieves a performance greater than 95% in the real-world deployment at UWMC Montlake and in the controlled environment.

a Syringe classifier in the OR with augmentation and background image class ($n = 230$)

b Syringe classifier in the OR without augmentation or background image class ($n = 231$)

c Vial classifier in the OR with augmentation and background image class ($n = 123$)

d Vial classifier in the OR without augmentation or background image class ($n = 143$)

Fig. 4 | Drug label classification performance in real-world operating room environments. Confusion matrices for **a**, **b** syringe labels with and without image augmentation and a background image class and **c**, **d** vial labels with and without image augmentation and a background image class.

Our design has four key limitations. First, we note that during the real-world deployment of our system at UWMC Montlake, providers were not given any instructions on how to use the system as these data were primarily used for algorithm development. In this population, an independent annotator marked that a drug label was visible in 231 of 265 (87.2%) events with syringes, and in 143 of 265 (54.0%) events with vials. In comparison, when we aimed to test the generalizability of the system at a second site, Harborview Medical Center, clinicians were given a short 2 min introduction to how the system worked was provided, including a request that the clinician try not to cover the syringe and vial labels with their hands. Here, the annotator noted that a drug label was visible in 72 of 72 (100.0%) events with syringes, and in 32 of 37 (86.5%) events with vials. This demonstrates that providing clinicians with instructions improves the visibility of drug labels and enables the wearable camera system to more reliably detect vial swaps. Additionally, we note that the drug drawup events analyzed begin only when the syringe and vial are in the provider's hand, excluding the initial pickup and inspection period when the label would also be visible to the camera. Including this initial period could further improve the system's reliability in detecting vial swaps. Furthermore, continuous learning algorithms can be developed to infer and learn from the accuracy of its

predictions, without explicit user feedback. Specifically, if the wearable system alerts the wearer of a medication error, but the wearer does not correct it, the system can infer that its predictions were incorrect and use this implicit feedback to improve its subsequent predictions.

Secondly, the data for this study was collected from a single hospital system with a specific policy for using color-coded, Codonics drug labels for syringes which is not universally adopted by all hospitals. While standardized syringe labels exist for all drugs in our training set, this is not the case for hospitals that do not use an automatic label generator[51]. We note that vial label styles vary between manufacturers, and our training and evaluation sets contain multiple vial label styles. To adapt the system to recognize a new label style, the classifier can be finetuned with examples of drug preparation events containing the new style. Future designs could generate realistic synthetic training data that would allow the system to generalize to unseen drug label styles.

Thirdly, an end-to-end deployment of the wearable camera system would require wirelessly streaming video data to a nearby edge server for further processing, which could be achieved using a combined Wi-Fi and 60 GHz WiGig link supporting multi-gigabit data rates[52], or over a terahertz link[53]. Low-power streaming algorithms can also be designed to work within

Fig. 5 | Subgroup analysis of drug label classification accuracy in real-world operating room environments. Classification performance of syringe and vial label classifiers across **a** different anesthesiology providers **b** different objects in the background and **c** whether the providers wore gloves. Numbers above each bar indicate the number of videos that were annotated in that scenario.

the power constraints of the wearable system which has a limited battery size and has to be small and comfortable for long term use. To reduce energy consumption, a heavily duty cycled high-resolution color camera that produces key frames at infrequent intervals, and a low-resolution, grayscale camera could stream video data to the edge server which combines the video streams using super-resolution and colorization algorithms[54]

Lastly, despite recording in 4K resolution, more vials than syringes were not classified by our system due to videos not containing enough clear frames for our system to produce a drug classification result. This is likely due to the smaller text of drug labels on vials. The use of higher resolution wearable cameras, such as those supporting a 5.3K resolution[55,56] may lead to improved results.

Outside of the operating room, injectable medications are most frequently administered by nurses. 44% of American nurses report giving at least five intravenous medications during a shift[57]. With 5.2 million registered nurses in the United States, there are likely tens of millions of syringe medication administrations every day[58]. Moreover, medication errors and syringe safety are nurses' top concerns according to a survey, with 97% reporting fearing potential medication administration errors[59]. The work presented here focuses on errors in the operating room due to their higher frequency in that location, however, a much larger proportion of medication administrations are performed by nurses, who could also benefit from the development of this technology.

In summary, we present a wearable camera system to detect vial swap errors prior to medication delivery by anesthesia providers. We show how deep learning algorithms trained on drug preparation events captured on head-mounted cameras can be used to detect and classify drug labels on syringes and vials in real-world operating room settings. Our proof-of-concept system may provide an automated means of detecting medication errors that can fit within the clinical setting. Further efforts include building mechanisms to provide feedback to the clinicians and creating methods for

measuring syringe volume to calculate the dose of medication delivered to a patient. Finally, when integrated with an electronic medical record system, our system opens opportunities for automatically recording drug information and reducing the overhead of manual record-keeping.

Methods

This study was approved by the University of Washington Institutional Review Board (STUDY00010313). Written consent was obtained for anesthesiology providers who recorded medication events with the wearable camera system. The rest of the operating room staff was informed of the recordings during a morning meeting and in posters on all the entrances and exits. Video recordings of drug draw up events obtained prior to patients entering the operating room was considered exempt by our IRB. Patients provided written consent for video recording during surgery during the surgical consenting process. Video footage when patients were present were de-identified prior to analysis, all reasonable measures have been taken to preserve patient anonymity and video footage will not be made available outside study personnel. Written consent was obtained to publish any identifiable images, details, or videos. All study personnel completed our institution's HIPAA and Compliance training for the handling of patient information. All studies complied with relevant ethical regulations. Randomization was not applicable and investigators were not blinded.

Dataset collection

Anesthesia providers were approached for consent before participating in the study. Providers were fitted with a head-mounted camera (GoPro Hero8) that was configured to record drug preparation events at 4K resolution and 60 frames per second. A research assistant would swap out the external battery pack every 2–3 h to ensure continuous recording was possible. Medication events in the operating room were recorded by an anesthesia provider during the preparation period between surgery cases.
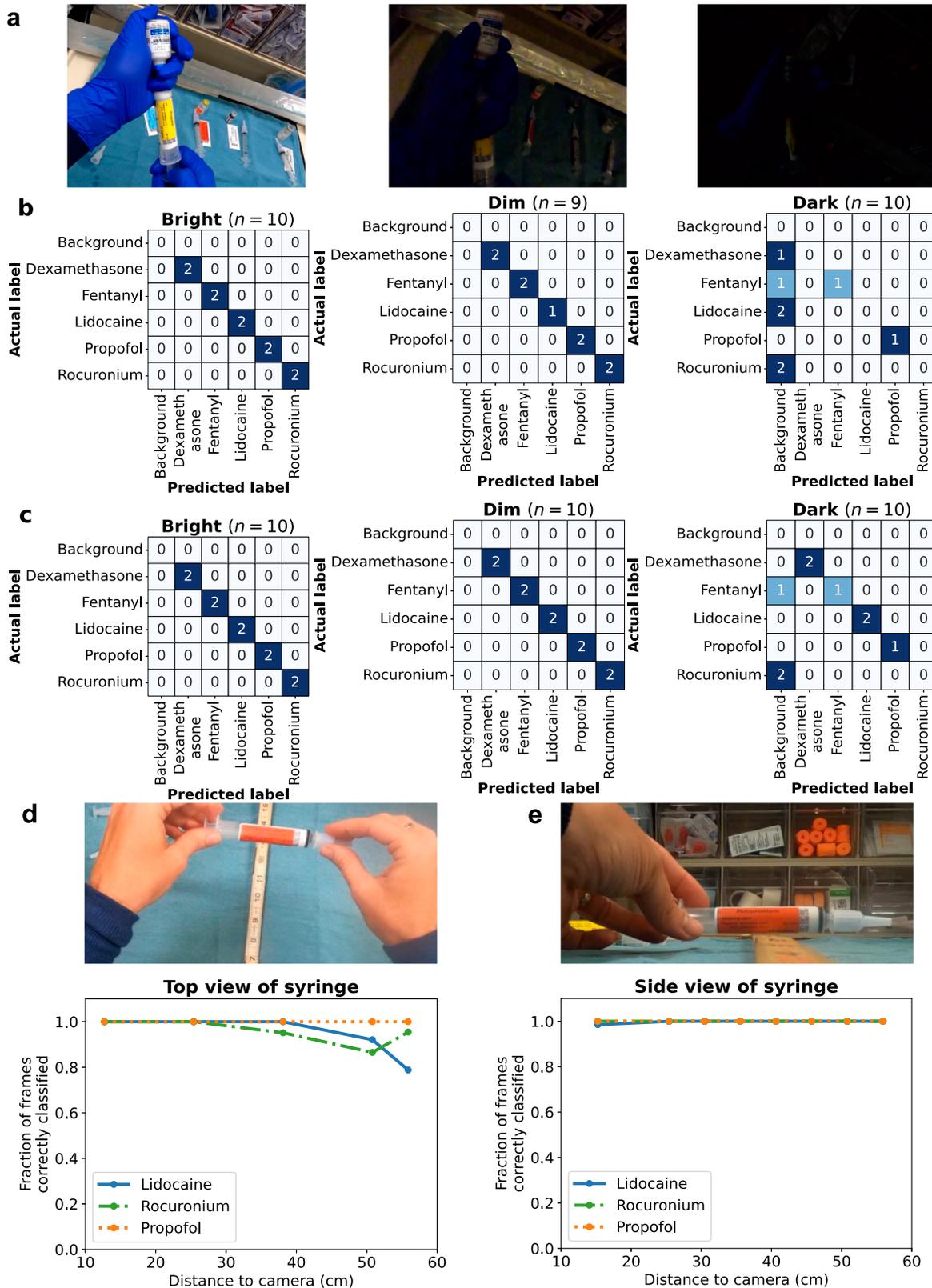
Fig. 6 | Benchmark testing. Classification performance of syringe drug labels under **a** bright, dim, and dark ambient lighting conditions in the operating room using the head-mounted camera's **b** default camera light settings and **c** low-light settings.

Fraction of syringe label video frames correctly classified at different distances when recorded from **d** top view and **e** side view.

Data was collected across multiple providers, multiple operating rooms, different types of syringes, and different lighting conditions in a single large academic hospital. Medication events in the controlled clinical environment was recorded by a researcher in an otherwise unoccupied OR.

After data collection, video footage was transferred to a HIPAA-compliant server (OneDrive) that was password-protected and deleted from the camera's memory card. Footage that did not involve drug preparation was discarded. To de-identify video footage of protected health information (PHI), all faces, ID badges, electronic medical records, and written patient and provider information were cropped or blurred using video editing software (Adobe Premiere Pro). The stream of video footage was segmented into clips capturing drug preparation events each with a duration of 2–120 s. These segmented clips were exported using matched source video settings to maintain the original quality of the footage. No audio recordings were obtained in this study and no images of patients were included in this dataset.

Dataset annotation

The segmented drug preparation clips were uploaded to a local Computer Vision Annotating Toolkit (CVAT)[60] server at the highest video quality settings for object annotation. A team of twenty volunteers were trained to annotate the location of syringes, vials, and drug labels in the dataset using bounding boxes. They were also trained to record drug label information including drug name, drug concentration, and drug label color. The videos were annotated with the drug preparation event recorded as either drug draw-up and drug delivery. We also annotated syringe-in-hand events where the provider holding the syringe without the drawing or delivery of a drug. Videos were marked to indicate if either a syringe, a vial or both objects were in the provider's hand. The ambient brightness level, whether the provider's hands were gloved, and the presence of syringes and vials in the background were also recorded as part of the annotation process. Drugs that were not being handled by the provider (i.e., syringes on the anesthesia cart) were not annotated. Volunteers were tasked with annotating keyframes at every tenth or twentieth frame. Annotations of syringes and vials for in-between frames were interpolated by CVAT.

Annotation training consisted of three major components: preparation, training, and feedback. As part of preparation, volunteers spent approximately two hours reviewing an instructional manual and a compilation of images taken from previously reviewed annotations. Volunteers then spent approximately one hour with a trainer to review training material and annotate a test clip. Lastly, volunteers were provided with written feedback on the quality of their annotations. Volunteers spent approximately 45–60 min annotating a batch of 1000 frames. Across all volunteers 300–400 h was spent annotating clips. To ensure a high level of quality in the annotations, all annotations were inspected and revised for accuracy by a senior reviewer. Senior reviewer training consisted of two months of self-training. This process took approximately 15–30 min for each batch of 1000 frames. Total review time totaled roughly 100–200 h.

The test dataset of medication events from the operating room and controlled environment used to evaluate the classifier were only annotated at the video-level with the drug name on syringes and vials handled by the provider. A senior reviewer inspected all clips in the test set and marked if the drug label on the syringe or vial was oriented towards the camera and could be read by eye.

Drug label detector training algorithm

The object detection framework for drug labels uses pretrained YOLOv5x[47] weights which are finetuned on a dataset of drug drawup events. The finetuning was performed for a maximum of 1000 epochs with early stopping if there was no improvement in validation performance for 50 epochs. Two detectors are trained, one for the syringe and syringe label, and one for the vial and vial label. The dataset was partitioned into training, validation, and test sets such that frames from the same video clip never straddled the partitioned boundaries. During model validation, it was observed that detector performance increased with video resolution of the

model. We selected the highest video resolution possible that would still allow inference to be performed in real-time on a NVIDIA A40 GPU. The validation step was also used to observe the effect of image augmentation on model performance prior to evaluating the model on the test set. The Albumentations library[61] was used for image augmentation. The following augmentations were applied: a blur with a random-sized kernel with probability 0.01, a median blur with random aperture linear size with probability 0.01, a conversion from RGB to grayscale with probability 0.01, and Contrast Limited Adaptive Histogram Equalization (CLAHE) with probability 0.01. At inference time, images were scaled to ratios of 1.0, 0.83, and 0.67. When the ratio is 0.83, horizontal flipping is performed. These images were then passed through the model, and the inverse scaling and flipping operation was applied to the coordinates of the predicted bounding boxes. The bounding box with the highest confidence score is selected as the model output.

Drug label classifier training algorithm

Our classifier leverages the pre-trained CLIP ViT-L/14 transformer model[62], fine-tuned on a training dataset of drug labels. We trained a classifier for 22 syringe drug labels and another for 20 vial drug labels (Supplementary Tables 1, 2). Some drugs come as pre-filled syringes so there is no vial equivalent at the hospitals included in the study which is why the number of syringes and vials are different. The set of drugs used to train the syringe label classifier cover all the drugs used in the drug drawup events and in 97.7% of drug delivery events in our real-world dataset (Supplementary Table 1), while the drugs used to train the vial label classifier cover all of the drug drawup events in the dataset (Supplementary Table 2). To generate the training dataset, we passed videos containing medication events through our object detection models, manually inspected the bounding boxes of objects which were detected with a probability threshold exceeding 0.8, and only included clear depictions of the drug label that could be read by a human annotator. Each video in the dataset contains a single drawup event of a single drug, and is labeled with the drug label of the syringe and vial. For each video, we extract all the frames and pass them through our object detection model and only consider frames exceeding a preset detection probability threshold. The classifier scales all images to a resolution of 224×224. The inference time of our classifier is 0.37 ms on a NVIDIA A40 GPU.

We evaluate the performance of the classifier on the drug draw-up events collected in the controlled clinical environment to determine the image augmentations to be applied to the classifiers. The image augmentations applied to the syringe classifier used the TorchVision library[63]: Image length and width is randomly resized to be within the normalized bounds [0.9, 1.0], random horizontal flip with probability 0.5 and random rotation within range $[-20, 20]°$. The image augmentations applied to the vial classifier were the same as the syringe classifier with the addition of random jitter of brightness, contrast, and saturation by a factor within the range [0.5, 1.5], and image histogram equalization with probability 0.5.

Vial swap error detection algorithm

To detect if a vial swap error has occurred, the following three key steps are performed to determine the drug label on the syringe and vial in the drug preparation event (Supplementary Algorithm 1):

Step 1: Extract frames containing a drug label. We first extract all frames in a video and pass them through the drug label detection frameworks for syringe labels and vial labels. The framework returns a bounding box for all drug labels observed in the frame, and an associated detection probability for each frame. If the highest probability exceeds *detection_probability_threshold* it is passed to the classifier for further processing. *detection_probability_threshold* is set to 0.85 for syringe labels and 0.8 for vial labels in our implementation.

Step 2: Classify drug labels in each frame. Next, we pass the extracted video frames into the corresponding drug classifier for syringe or vial labels. The logits predicted by the classifier are converted into softmax probabilities. The class with the maximum probability is marked as the predicted

drug label for the frame. If the predicted class is not a background object class, and the softmax probability is greater than *minimum_prediction_probability*, it is recorded in a counter. *minimum_prediction_probability* is set to 0.999 in our implementation.

Step 3: Compute drug label at video-level. Finally, the most frequently occurring drug label is retrieved from the counter. If that drug occurred in at least *minimum_drug_label_frames* frames, we mark that as the corresponding syringe or vial drug label for the video. *minimum_drug_label_frames* is set to 1 for syringe drug labels and 6 for vial drug labels in our implementation.

We check if the syringe and vial labels computed in the above steps match. If they do not match we mark that a vial swap error has occurred, else we mark that the drug drawup was valid (Supplementary Algorithm 2).

The thresholds in the vial swap error detection algorithm were determined by measuring their effect on the classifier's performance on drug drawup events collected in the controlled environment. These thresholds were then held constant and used to evaluate the remaining medication events in the real-world dataset and benchmark testing.

If the thresholds of *detection_probability_threshold* in step 1, *minimum_prediction_probability* in step 2 and *minimum_drug_label_frames* in step 3 are removed, the performance of the syringe label classifier on the real-world operating room dataset decreases from 227 of 230 (98.7%) events to 221 of 231 (95.7%) events. The performance of the vial label classifier decreases from 122 of 123 (99.2%) events to 132 of 138 (95.7%) events.

Statistical analysis

Algorithms to train, validate, and test computer vision algorithms were performed using PyTorch. AP_{50}, sensitivity, specificity, and 95% confidence interval analysis was performed using numpy. Figures were created using matplotlib and seaborn.

Data availability

All data necessary for interpreting the manuscript have been included. The datasets used in the current study are not publicly available but may be available from the corresponding authors on reasonable request and with permission of the University of Washington.

Code availability

Code is available at https://github.com/uw-x/mederrors.

References

1. Panagioti, M. et al. Prevalence, severity, and nature of preventable patient harm across medical care settings: systematic review and meta-analysis. *bmj* **366**, 1–11 (2019).
2. Makary, M. A. & Daniel, M. Medical error-the third leading cause of death in the US. *BMJ* **353**, i2139 (2016).
3. Kavanagh, K. T., Saman, D. M., Bartel, R. & Westerman, K. Estimating hospital-related deaths due to medical error: a perspective from patient advocates. *J. Patient Saf.* **13**, 1–5 (2017).
4. James, J. T. A new, evidence-based estimate of patient harms associated with hospital care. *J. Patient Saf.* **9**, 122–128 (2013).
5. de Vries, E. N., Ramrattan, M. A., Smorenburg, S. M., Gouma, D. J. & Boermeester, M. A. The incidence and nature of in-hospital adverse events: a systematic review. *Qual. Saf. Health Care* **17**, 216–223 (2008).
6. Berdot, S. et al. Drug administration errors in hospital inpatients: a systematic review. *PLoS One* **8**, e68856 (2013).
7. Bates, D. W., Boyle, D. L., Vliet, M. B. V., Schneider, J. & Leape, L. Relationship between medication errors and adverse drug events. *J. Gen. Intern. Med.* **10**, 199–205 (1995).
8. Belén Jiménez Muñoz, A. et al. Medication error prevalence. *Int. J. Health Care Qual. Assur.* **23**, 328–338 (2010).
9. Cooper, J., Newbower, R. & Kitz, R. An analysis of major errors and equipment failures in anesthesia management: considerations for prevention and detection. *Anesthesiology* **60**, 34–42 (1984).
10. Rothschild, J. M. et al. The Critical Care Safety Study: The incidence and nature of adverse events and serious medical errors in intensive care. *Crit. Care Med.* **33**, 1694–1700 (2005).
11. Lahue, B. J. et al. National burden of preventable adverse drug events associated with inpatient injectable medications: healthcare and medical professional liability costs. *Am. Health Drug Benefits* **5**, 1–10 (2012).
12. Fasting, S. & Gisvold, S. E. Adverse drug errors in anesthesia, and the impact of coloured syringe labels. *Can. J. Anaesth.* **47**, 1060–1067 (2000).
13. Greengold, N. L. et al. The impact of dedicated medication nurses on the medication administration error rate: a randomized controlled trial. *Arch. Intern. Med.* **163**, 2359–2367 (2003).
14. Mohanna, Z., Kusljic, S. & Jarden, R. Investigation of interventions to reduce nurses' medication errors in adult intensive care units: a systematic review. *Aust. Crit. Care* **35**, 466–479 (2022).
15. Latif, A., Rawat, N., Pustavoitau, A., Pronovost, P. J. & Pham, J. C. National study on the distribution, causes, and consequences of voluntarily reported medication errors between the ICU and non-ICU settings. *Crit. Care Med.* **41**, 389–398 (2013).
16. Patanwala, A. E., Warholak, T. L., Sanders, A. B. & Erstad, B. L. A prospective observational study of medication errors in a tertiary care emergency department. *Ann. Emerg. Med.* **55**, 522–526 (2010).
17. Moyen, E., Camiré, E. & Stelfox, H. T. Clinical review: medication errors in critical care. *Crit. Care* **12**, 1–7 (2008).
18. Vilke, G. M. et al. Paramedic self-reported medication errors. *Prehosp. Emerg. Care* **11**, 80–84 (2007).
19. Abeysekera, A., Bergman, I., Kluger, M. & Short, T. Drug error in anaesthetic practice: a review of 896 reports from the Australian incident monitoring study database. *Anaesthesia* **60**, 220–227 (2005).
20. Webster, C. S., Merry, A. F., Larsson, L., McGrath, K. A. & Weller, J. The frequency and nature of drug administration error during anaesthesia. *Anaesth. Intensive Care* **29**, 494–500 (2001).
21. Ex-nurse convicted in fatal medication error gets probation. *The New York Times* https://www.nytimes.com/2022/05/15/us/tennessee-nurse-sentencing.html (2022).
22. Bowdle, T. A., Jelacic, S., Webster, C. S. & Merry, A. F. Take action now to prevent medication errors: lessons from a fatal error involving an automated dispensing cabinet. *Br. J. Anaesth.* **130**, 14–16 (2023).
23. Kaushal, R. et al. Medication errors and adverse drug events in pediatric inpatients. *Jama* **285**, 2114–2120 (2001).
24. Trimble, A. N., Bishop, B. & Rampe, N. Medication errors associated with transition from insulin pens to insulin vials. *Am. J. Health-Syst. Pharm.* **74**, 70–75 (2017).
25. Härkänen, M., Ahonen, J., Kervinen, M., Turunen, H. & Vehviläinen-Julkunen, K. The factors associated with medication errors in adult medical and surgical inpatients: a direct observation approach with medication record reviews. *Scand. J. Caring Sci.* **29**, 297–306 (2015).
26. Webster, C. et al. Clinical assessment of a new anaesthetic drug administration system: a prospective, controlled, longitudinal incident monitoring study. *Anaesthesia* **65**, 490–499 (2010).
27. Lohmeyer, Q. et al. Effects of tall man lettering on the visual behaviour of critical care nurses while identifying syringe drug labels: a randomised in situ simulation. *BMJ Qual. Saf.* **32**, 26–33 (2023).
28. Grissinger, M. The Five Rights: a destination without a map. *Pharm. Ther* **35**, 542 (2010).
29. Martin, L. D. et al. Outcomes of a Failure Mode and Effects Analysis for medication errors in pediatric anesthesia. *Pediatr. Anesth.* **27**, 571–580 (2017).
30. Evley, R. et al. Confirming the drugs administered during anaesthesia: a feasibility study in the pilot National Health Service sites, UK. *Br. J. Anaesth.* **105**, 289–296 (2010).

31. Moreira, M. E. et al. Color-coded prefilled medication syringes decrease time to delivery and dosing error in simulated emergency department pediatric resuscitations. *Ann. Emerg. Med.* **66**, 97–106.e3 (2015).

32. Truitt, E., Thompson, R., Blazey-Martin, D., Nisai, D. & Salem, D. Effect of the implementation of barcode technology and an electronic medication administration record on adverse drug events. *Hosp. Pharm.* **51**, 474–483 (2016).

33. Bowdle, T. et al. Facilitated self-reported anaesthetic medication errors before and after implementation of a safety bundle and barcode-based safety system. *Br. J. Anaesth.* **121**, 1338–1345 (2018).

34. Bowdle, T. A. et al. Electronic audit and feedback with positive rewards improve anesthesia provider compliance with a barcode-based drug safety system. *Anesth. Analg.* **129**, 418–425 (2019).

35. Bates, D. W. & Landman, A. B. Use of medical scribes to reduce documentation burden: are they where we need to go with clinical documentation? *JAMA Intern. Med.* **178**, 1472–1473 (2018).

36. Gawande, A. *Why doctors hate their computers.* The New Yorker **12** (2018).

37. van der Veen, W. et al. Factors associated with workarounds in barcode-assisted medication administration in hospitals. *J. Clin. Nurs.* **29**, 2239–2250 (2020).

38. Van Der Veen, W. et al. Association between workarounds and medication administration errors in bar-code-assisted medication administration in hospitals. *J. Am. Med. Inform. Assoc.* **25**, 385–392 (2018).

39. Microsoft HoloLens 2 https://www.microsoft.com/en-us/hololens/ (2023).

40. Magic Leap 2 | the most immersive enterprise AR device https://www.magicleap.com/en-us/ (2023).

41. Ray-Ban Meta Smart Glasses https://www.meta.com/smart-glasses/ (2023).

42. Spectacles by Snap Inc. · The Next Generation of Spectacles https://www.spectacles.com/ (2023).

43. Deng, J. et al. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255 (2009).

44. Steiner, C. A., Karaca, Z., Moore, B. J., Imshaug, M. C. & Pickens, G. Statistical brief# 223 (2017).

45. Ciregan, D., Meier, U. & Schmidhuber, J. Multi-column deep neural networks for image classification. In *2012 IEEE conference on computer vision and pattern recognition (CVPR)*, 3642–3649 (IEEE, 2012).

46. Simard, P. Y., Steinkraus, D., Platt, J. C. et al. Best practices for convolutional neural networks applied to visual document analysis. In *Icdar*, vol. 3 (Edinburgh, 2003).

47. Jocher, G. YOLOv5 by Ultralytics https://github.com/ultralytics/yolov5 (2020).

48. Statement on labeling of pharmaceuticals for use in anesthesiology https://www.asahq.org/standards-and-practice-parameters/statement-on-labeling-of-pharmaceuticals-for-use-in-anesthesiology (2015).

49. Standard specification for user applied drug labels in anesthesiology https://www.astm.org/Standards/D4774.htm (2017).

50. GoPro Light Mod - Camera LED Light Accessory https://gopro.com/en/us/shop/mounts-accessories/light-mod/ALTSC-001-master.html (2023).

51. Janik, L. S. & Vender, J. S. Pro/con debate: color-coded medication labels. *The SAFE-T Summit and the International Standards for a Safe Practice of Anesthesia* 72 (2019).

52. Baig, G. et al. Jigsaw: Robust live 4k video streaming. In *The 25th Annual International Conference on Mobile Computing and Networking*, 1–16 (2019).

53. Nallappan, K., Guerboukha, H., Nerguizian, C. & Skorobogatiy, M. Live streaming of uncompressed HD and 4k videos using terahertz wireless links. *IEEE Access* **6**, 58030–58042 (2018).

54. Veluri, B. et al. Neuricam: Key-frame video super-resolution and colorization for IoT cameras. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*, 1–17 (2023).

55. GoPro HERO12 Black https://gopro.com/en/us/shop/cameras/hero12-black/CHDHX-121-master.html (2023).

56. GoPro HERO11 Black https://gopro.com/en/hk/shop/cameras/hero11-black/CHDHX-111-master.html (2023).

57. Hertig, J. B. et al. A comparison of error rates between intravenous push methods: a prospective, multisite, observational study. *J. Patient Saf.* **14**, 60 (2018).

58. Smiley, R. A. et al. The 2022 National Nursing Workforce Survey. *J. Nurs. Regul.* **14**, S1–S90 (2023).

59. Grissinger, M. Reducing errors with injectable medications. *Pharm. Ther.* **35**, 428–451 (2010).

60. Sekachev, B. et al. OpenCV/CVAT: v1.1.0 https://doi.org/10.5281/zenodo.4009388 (2020).

61. Buslaev, A. et al. Albumentations: fast and flexible image augmentations. *Information* **11**, 125 (2020).

62. Radford, A. et al. Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning (ICML)*, 8748–8763 (PMLR, 2021).

63. Paszke, A. et al. Automatic differentiation in PyTorch. In *31st Conference on Neural Information Processing Systems (NIPS)*, 1–4 (2017).

Acknowledgements

Author contributions

K.M. designed and performed the clinical testing; J.C., S.N., and M.W. conducted the analysis with technical supervision by L.S., A.D, S.G., K.M.; S.G., K.M., J.C., and S.N. wrote and edited the manuscript.

Competing interests

The authors declare the following competing interests: S.G. and J.C. are co-founders of Wavely Diagnostics, Inc. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41746-024-01295-2.

Correspondence and requests for materials should be addressed to Shyamnath Gollakota or Kelly Michaelsen.

Reprints and permissions information is available at http://www.nature.com/reprints

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