Assisted Medication Management in Elderly Care Using Miniaturised Near-Infrared Spectroscopy

SIMON KLAKEGG, University of Oulu
JORGE GONCALVES and CHU LUO, The University of Melbourne
AKU VISURI and ALEXEY POPOV, University of Oulu
NIELS VAN BERKEL, ZHANNA SARSENBAYEVA, The University of Melbourne
VASSILIS KOSTAKOS, The University of Melbourne
SIMO HOSIO, University of Oulu
SCOTT SAVAGE, Northern Health
ALEXANDER BYKOV, IGOR MEGLINSKI and DENZIL FERREIRA, University of Oulu

Near-infrared spectroscopy[1] (NIRS) measures the light reflected from objects to infer highly detailed information about their molecular composition. Traditionally, NIRS has been an instrument reserved for laboratory usage, but recently affordable and smaller devices for NIRS have proliferated. Pairing this technology with the ubiquitous smartphone opens up a plethora of new use cases. In this paper, we explore one such use case, namely medication management in a nursing home/elderly care centre. First, we conducted a qualitative user study with nurses working in an elderly care centre to examine the protocols and workflows involved in administering medication, and the nurses' perceptions on using this technology. Based on our findings, we identify the main impact areas that would benefit from introducing miniaturised NIRS. Finally, we demonstrate via a user study in a realistic scenario that miniaturised NIRS can be effectively used for medication management when leveraging appropriate machine learning techniques. Specifically, we assess the performance of multiple pre-processing and classification algorithms for a selected set of pharmaceuticals. In addition, we compare our solution with currently used methods for pharmaceutical identification in a local care centre. We hope that our reflection on the multiple aspects associated with the introduction of this device in an elderly care setting can help both academics and practitioners working on related problems.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI; Ubiquitous and mobile computing**

KEYWORDS

Near-infrared spectroscopy, medication management, elderly care, user study, preprocessing, machine learning.

ACM Reference format:

Simon Klakegg, Jorge Goncalves, Chu Luo, Aku Visuri, Alexey Popov, Niels van Berkel, Zhanna Sarsenbayeva, Vassilis Kostakos, Simo Hosio, Scott Savage, Alexander Bykov, Igor Meglinski, and Denzil Ferreira. 2018. Assisted Medication Management in Elderly Care Using Miniaturised Near-Infrared Spectroscopy. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 2, Article 69 (June 2018), 24 pages.
DOI: https://doi.org/10.1145/3214272

[1] This work is supported by the Academy of Finland (Grants 286386-CPDSS, 285459-iSCIENCE, 304925-CARE, 313224-STOP, 260321-APP, 290596-AVB), Marie Skłodowska-Curie Actions (645706-GRAGE), Government of Russian Federation (U01-074-APP, AVB) and the Russian Science Foundation (15-14-10008-IVM).
Author's addresses: S. Klakegg, A. Visuri, A. Popov, S. Hosio, A. Bykov, I. Meglinski, D. Ferreira, University of Oulu, Oulu 90014, Finland, email: {firstname.lastname}@oulu.fi; J. Goncalves, C. Luo, N. van Berkel, Z. Sarsenbayeva, V. Kostakos, The University of Melbourne, VIC-3010 Parkville, Australia, email: {firstname.lastname}@unimelb.edu.au; S. Savage, Northern Health, 3076 Epping, Australia, email: Scott.Savage@nh.org.au

2474-9567/2018/6-ART69 $15.00

https://doi.org/10.1145/3214272

1 INTRODUCTION

One of the tasks where humans excel is intuitively recognising everyday objects in their environments. Machine vision is a widely used method to extend this human ability to computers [58]. While machine vision using visible light is remarkably efficiently in detecting shape and texture, it cannot consider objects' chemical composition as an input parameter [6]. For instance, machine vision cannot be used to check if food meets certain composition criteria, or if certain drugs, or even textiles, have the desired mix of ingredients. Near-infrared spectroscopy (NIRS) uses near-infrared light capable of penetrating the surface and traverse the chemical structure of an object. This allows the user to retrieve information about the inner composition of the sample in the form of a spectrum, which acts as a fingerprint [56]. In turn, this enables accurate and detailed object identification. NIRS has been shown to be an efficient apparatus in research across many different areas, including brain research [33], fabricating biotissue-mimicking phantoms [64], pharmaceutical technology [13,26], and blood glucose sensing [28,49].

NIRS scanners have been used in research laboratories for years [7]. However, only recently has the technology matured sufficiently to allow for end-user hardware which can be small and robust enough to be carried around, and still capable of producing reliable results. The NIRS device used in our study (DLP NIRscan Nano [45]) costs under 1000 US dollars at the time of writing this article and weighs just 80 grams – a mere fraction of the price and weight of high-end NIRS hardware. This dramatic drop in cost encourages us to consider everyday scenarios for this technology, and for the first time put it in the hands of non-expert end-users. Furthermore, coupling NIRS hardware with commodity devices (tablets, smartphones) opens a range of exciting research avenues to explore [29].

In this paper, we explore the potential for using an in situ NIRS scanner placed at an elderly care centre to enable the identification of pharmaceuticals before they are administered to the patients. In such a scenario, a user can scan a pill or medicine to confirm that it is the right one to take at that moment. This is particularly important, as medication mismanagement causes substantial financial and human costs [57,61]. The contribution of our work is three-fold:

- We systematically examine the current medication procedures at a nursing home to discover problem areas related to mismanagement of medicine that could be mitigated using NIRS. In addition, we explore the usability of the technology and how it could be incorporated in the elderly care centre.
- We validate the performance of miniaturised NIRS for pharmaceutical identification through multiple preprocessing and classification algorithms. While there exists a large body of work dedicated to exploring the accuracy of laboratory benchtop NIRS, there is little reference literature on the just recently available miniaturised versions.
- We evaluate the NIRS pharmaceutical identification method through a user study with nurses in a elderly care centre. The performance of our solution is compared with identification methods currently implemented at the care centre.

2 RELATED WORK

2.1 Medication Mismanagement

Improper use of medication is an unnecessary healthcare expenditure. For instance, in the United Kingdom, The National Health Service (NHS) recently estimated that they had spent between £1 billion up to a total of £2.5 billion on preventable medication errors [15]. Furthermore, drug waste costs the NHS £300 million a year [61]. In the US, an estimated 125,000 Americans die annually due to medication errors [5], highlighting that there exist more than just economic consequences. Hartnell *et al.* analysed barriers related to medication error reporting in hospitals [20]. Their results reveal that close to 20% of all medication doses are incorrectly administered and that the costs of avoidable drug-related deaths in elderly care are extensive. The under-reporting of incidents is another identified prevalent issue associated with the logging of medication errors.

Metsälä *et al.* systematically reviewed several studies on medication mismanagement in elderly acute care to find the most common causes [41]. The results indicated that polypharmacy (*i.e.*, the use of multiple medicines), lack of knowledge, cognitive status, faulty communication and information, and incorrectly filled medication boxes are all contributing factors. The nurses in the study stated that improved systems (*e.g.*, more structured filling of medication boxes [9], implementation of technology applications), better-standardised medication sheets and patient information, and improved communication with doctors could help to mitigate the problems.

To alleviate this considerable challenge, new innovative solutions are being developed. Medisafe, an application running on smart devices and computers, has shown great potential in enabling more efficient medication management for the user through alarms and medication-taking support [40]. Chang *et al.* [9] built a smart medication dispensing system for medication reminding and recording in contexts such as elderly care. Hayakawa *et al.* [21] evaluated a smartphone-based medication self-management system with intelligent pill boxes that can determine when a pill has been consumed, and showed a positive effect on the user's medical adherence. A key difference of our work is that we rely on chemically identifying the pill that a nurse or patient is holding in their hand, rather than relying on schedules, people's memory, or patient's' ability to visually compare the pill to a reference image. Hence, our work seeks to minimise the potential for human error in medication dispensing by chemically and unambiguously determining what pill a person is currently holding.

2.2 Object Identification

Identifying everyday objects computationally using machine vision is an active area of research [49]. For instance, Hart *et al.* [19] developed a smartphone application for pill detection using computer vision techniques. The implemented probabilistic framework relies on three visual features: size, shape, and colour. The result illustrates how smartphones can be used to rapidly identify pills in-situ. A similar approach (*i.e.*, utilising the same set of features) for identification of pharmaceutical is HelpmePills [10]. By relying on the smartphone's camera and computational power, the application aims to accommodate mobile pill recognition for elderly. The work highlights how such a tool could alleviate multiple aspects related to medication mismanagement.

There is a large body of work dedicated to developing and improving machine vision accuracy. For example, Hough Forest was identified as a quick and efficient method for detection, tracking, and recognition [16]. While machine vision for object identification has seen significant improvement over the years, there are still various challenges to overcome before it can reach equivalent accuracy to that of the primate visual system [17]. To this end, hyperspectral imaging, which can be considered as an extension of machine vision [38], has been proposed as a way to improve identification accuracy. This technique has been used successfully for early detection of fungal infection in citrus fruit [35].

Furthermore, there have been several developments regarding other methods of computational identification of objects. For instance, Pérez *et al. [48]* explored the use of RFID tags in a hospital for trackability and matching between patient and their prescribed medication. The system received positive feedback during evaluation and the users' comments indicated that it could reduce the number of adverse medication events. As an alternative, multimodal tactical sensing is now being explored as an option for enabling human like touch capabilities in machines. It relies on force, vibration, and temperature collected from exploratory movements to identify objects [66]. Naturally, these approaches cannot distinguish the inner composition of samples, which is one of the main benefits behind NIRS as explored in this paper.

2.3 Miniaturised NIRS

NIRS probes samples with infrared light (wavelengths of 750 to 2,500 nm) and measures the magnitude of the absorbance across the spectrum. The absorbance across the spectrum is an outcome of the vibrations of the atoms in the sample, and can therefore be used to extract information about a sample's internal molecular composition [47]. This information can be used for advanced analytics, *e.g.*, identification of pharmaceuticals

and the level of each substance present in the tablets [14]. NIRS devices have conventionally been utilised in a laboratory setting, but given recent hardware improvements (see
Table 1), the form factor is now significantly smaller and lighter. The new generation of miniaturised NIRS devices enables the user to bring the analytic instrument to the sample of interest, enabling in situ scanning. This has enabled a range of interesting use cases in areas such as detecting food allergens or determining food quality [32] and chemical analysis of pharmaceutical excipients grinded to a powder state [2,59]. Klakegg *et al.* [29] explored how miniaturised NIRS devices could be commoditised for non-expert end users. They conducted a study where participants utilised a custom 3D printed enclosure and accompanying smartphone app to identify different types of samples. The result from a scan included information about the identified object, such as the name and composition. Their results show that novice end-users can successfully use NIRS in combination with assistive software.

Table 1. Overview of commercially available miniaturised NIRS

Brand	Spectral range (nm)	Weight (g)	Physical Dimensions (L x W x H) (mm)	OS Support
TellSpec [60]	900 - 1700	136	82.2 x 66 x 45	iOS, Android
SCiO [55]	700 - 1100	35	67.7 x 40.2 x 18.8	iOS, Android
MicroNIR [42]	950 - 1650	< 200	(L x Diameter) 146 x 44.5	Windows
DLP NIRscan Nano [45]	900 - 1700	80	62 x 58 x 36	iOS, Android, Windows

3 STUDY 1 - CARE CENTRE WORKFLOW AND POTENTIAL OF MINIATURISED NIRS IN MEDICATION MANAGEMENT

We conducted a user study at a collaborating elderly care centre to better understand the challenges that nurses encounter in regard to medication management. This would enable us to determine potential difficulties in the distribution of patients' medicine, which could be mitigated with miniaturised NIRS. Previous work has already demonstrated how untrained novice users are able to utilise miniaturised NIRS [29]. Therefore, in this work we set to explore how the technology could be integrated into the care centre and other scenarios.

3.1 Protocol

We conducted semi-structured interviews with 11 care workers at a local care centre (2 males, 9 females; ages: 17 - 56 years old, M = 35.09, SD = 12.62; work experience: 1-28 years, M = 11.27, SD = 8.19). The care workers had been at this particular institution for 1-8 years, M = 5.55, SD = 3.00 and occupied different positions (1 student, 8 practical nurses, 2 registered nurses). Their roles and responsibilities included: basic and rehabilitating care, medication, stimulating and assisting during everyday activities, housekeeping, and sampling (*i.e.*, gathering of matter for medical diagnosis). This study was conducted in the context of a broader technological intervention study our research group is conducting. We have previously sporadically observed and interviewed staff at the care centre over a period of 6 months to better understand their tasks, information needs, and job procedures [24,31]. Hence, we had already built a rapport with the staff we interviewed for our NIRS study.

We carried out the interviews with participants in groups of 2 or 3, to foster discussion and allow nurses to build on each other's experiences and insights while still allowing all nurses to speak out. The interview

session took approximately 45 minutes per group. After an introduction to the device and its potential use cases in the care centre, the nurses were shown the miniaturised NIRS (as shown in Fig. 3) before we proceeded with an open-ended interview.

The interview questions probed the participants about the structure of medical procedures (*e.g.*, pill allocation and distribution) in the care centre, and to describe any out of the ordinary past events. We were also interested in how the nurses perceived the usefulness of the device, and if they thought it could improve their work quality or the quality of their patients' lives. We concluded the interview with questions on the relevance for miniaturised NIRS in a homecare setting.

3.2 Findings

We start by outlining the current medical procedures in the elderly care facility. Following, we highlight some of the primary issues related to medication management and their causes in nursing homes. The potential impact of miniaturised NIRS in the care centre is shortly described before the applicability for home care is assessed. We conclude the section by summarizing all main impact areas of the device.

3.2.1 Medication Process. A typical medication process in this elderly care facility consists of the following steps. A patient's health issues and needs are discussed with a medical doctor, who issues a new prescription based on their diagnosis. This is communicated to the nurses who update the medication list in their PC-based system. In addition, the nurses check the system for drug interactions between two or more medications (to avoid unwanted side effects), and change the medication list (or time when given to the patient) according to the found drug interactions. Non-prescription drugs (*e.g.*, painkillers) are also assigned to the patients by doctors.

All medication is purchased and administered on a per-patient basis (*i.e.*, there is no sharing of pharmaceuticals between patients and each have their own pill case). New medication is ordered and delivered every two weeks. The prescription drugs come in cardboard boxes (one box per patient, as shown in Fig. 1b). Each prescription box contains two weeks of medication for a patient, divided into small transparent bags, one for each of the four daily doses (*i.e.*, breakfast, lunch, dinner, supper). Each individual bag has a text label with the patient's name, drug information, time, date, and a QR code that can be scanned for additional drug information. During the nightshift, the responsible nurse prepares the following day's medication by removing the pills from the bags and placing them in the appropriate patient's pill case. While the bags already come pre-sorted for each daily dose, preparation by the nightshift nurse is still a key step for various reasons:

1. Verifying that the bag contains the correct pills according to the patient's latest updated prescription.
2. Removing any medication that has been placed in the bag erroneously.
3. Checking for harmful effects between each pill and other food sources such as milk and juice. In addition, some pills need to be taken before, with, or following a meal. If this is the case, the pills are spread out according to the planned timeslot, ensuring that the pharmaceutical is correctly ingested without any unwanted interactions.

For these three reasons, the nurses often have to identify each of the pharmaceuticals found in the prescription bag. This workload comes on top of other pressing tasks, which have to be completed during the nightshift. However, this process was sometimes complicated and in the interviews the nurses stated issues such as:

> *"We've had times when the prescription bags did not have the correct medication, or had the wrong amounts."*
>
> *"We have to pay a lot of attention to make sure that the medication we receive for each patient is up-to-date, sometimes changes are not delivered to the pharmacy, or the pharmacy makes mistakes."*

Another statement was made on the usability of the tool the nurses had at disposal during this vital process:

> *"We have a system where we can scan a QR code from the distribution bag, and see images and information about the medication, but even then, the pills change often and are sometimes very [visually] similar and hard to identify from others."*

After the pharmaceuticals have been divided, the pill cases are stored in a cabinet using separate boxes, one for each of the four distribution timeslots as shown in Fig. 1a. Non-prescription pharmaceuticals, which are needed on ad-hoc basis (*e.g.*, patient feels pain or has trouble sleeping), are stored in one joint box (Fig. 1c). The name of the patient is written on the package.

(1a) (1b) (1c)

Fig. 1. (1a): Medication cabinet with four boxes (breakfast, lunch, dinner, supper) and basket of non-prescription drugs on top. (1b): Patients' prescription bags. (1c): Close up of basket containing non-prescription drugs (*e.g.*, painkillers, sleeping pills).

3.2.2 Medication Mismanagement. All nurses are required to pass a medication administration test every three years. However, the medication products change frequently, and it is challenging to constantly keep track. Pharmacies commonly purchase the cheapest brand, and pills containing a specific substance often change in both price and appearance. This, combined with the strenuous work environment, makes it challenging to accurately administer medication:

> *"We have had cases where we gave either the wrong medicine to a patient, or the wrong amount, usually due to time constraints or human error".*
> *"There have been scenarios where two patients' medicine have gotten mixed together".*
> *"Many of the mistakes are simple human errors, like dismiss giving a pill hourly or forget to remove a pill with unwanted interactions from the pill case".*

The handling of medication is reserved for the most experienced care workers. It requires in-depth knowledge of the patients, and the ability to precisely identify each patient. Students or rotating employees are not allowed to take on this task autonomously until they have had sufficient training. Despite the considerable knowledge of the experienced care workers, some tasks still impose a big problem.

> *"Sometimes we find random pills on the floor after mealtimes, which have been clearly spit out, but we have no idea from who the pill actually is."*

When the nurses are able to identify the owner of the pill, they contact the doctor to inform them of the mismanagement. This is a high priority event and the personnel will follow instructions received, such as

measuring the blood pressure or sugar balance of the affected patient. However, it is not simple to identify pharmaceuticals found in this manner simply by visual inspection.

3.2.3 Impact on Work and Patients' Life Quality. The nurses were positive about the capabilities of the device and stated that it could be a useful tool to reduce medication mismanagement, due to difficulties in identifying different medications from one another. Although they currently had a tool available (scanning of QR code on prescription bags for additional information and images), distinguishing tablets was still perceived as a tough and time-consuming task. Accurately distributing the pharmaceuticals would enable healthier patients, which is one of the care centre's main objectives.

> *"This type of device would significantly increase our work efficiency and reduce mismanagement of medication. This would lead to increased patient safety, which is our primary goal."*

While the care workers could see the gain from having a miniaturised NIRS to assist them, they were undecided on how it would directly apply in all parts of their work. One participant said that it did not necessarily have to be incorporated in every part of their workflow, but it could be accessible as needed.

> *"Even if this device was not required in every case, it would still be beneficial to have it available to validate during the medication identification process."*

3.2.4 Usability in homecare. In homecare, the medication is delivered either weekly or daily to the patient's home by a nurse. Ingestion of the tablets is the patient's responsibility and delimiting health (*e.g.*, Alzheimer) frequently causes inconsistencies in medication schedules.

> *"Homecare patients often suffer from memory problems and simply lose track of either what medication they have already taken, or what they should take and how much. Sometimes they even forget where the pills should be taken from (i.e., the pill dispenser) and take pills directly from the pill box."*

For miniaturised NIRS to work in homecare, nurses repeatedly stated that the device should come with a display, that contained information about the tablet scanned, daily dosage, and a list of the user's medication. This would allow a patient to track their pharmaceutical intake, receive a warning in case of overconsumption, and be notified if a pill has not been taken when required. Additionally, as elderly tend to mix pharmaceuticals, a miniaturised NIRS solution could assist in differentiating and identifying pills. Visiting care workers could use the logs of the device to determine how well the patient is medicating. It was emphasized that the design should be simple enough for an elder to use.

> *"This would be extremely useful for patients in homecare, as long as they were able to use the device and know how it works."*

3.2.5 Main Impact Areas of Miniaturised NIRS. The following list is intended to clarify and describes the main areas which our study indicates could benefit from utilising a miniaturised NIRS for medication management.

1. When prescription bags arrive from the pharmacy the nurses must verify the content, check for dangerous effects between each pill and other food sources and sort the tablets in the correct pill case and position. As one bag contains multiple unlabeled tablets, the nurses scan a QR code to see the names of the tablets with the corresponding image. However, if two different tablets look very similar (*e.g.*, same colour, shape, and size), it is difficult to reliably create a link between the physical pill and the image on the smartphone. In such a scenario, sorting by only visual prowess becomes challenging and potentially dangerous.

2. When pills are found after meals (*e.g.*, fallen out of a patient's mouth) or are discovered in atypical places, identifying the pill could help to narrow down the search for the patient who potentially forgot to take it. Currently, it is challenging to backtrack to the patient based on the pharmaceutical. One could either start to open all the non-prescription drugs packages or randomly scan the QR codes on prescription bags (if not disposed of), in hopes of finding a visual match. Even then it is not guaranteed to be the same tablet.
3. The nursing work environment is hectic and can lead to inhibited cognition. In stressful moments, the care workers might mix multiple pills which they are carrying. Instead of throwing them away, or distributed based on their gut feeling, an accurate identification could provide them with the desired information and comfort.
4. Patients in a homecare setting frequently find themselves mixing up pills, taking the wrong doses or tablets and may even forget to ingest their medicine. In such scenarios, it could be beneficial to have a system in place with countermeasures (*e.g.*, notifications, warnings, pharmaceutical information on scan, logs for visiting nurses, automatic verification of drug interactions).

4 STUDY 2 - MINIATURISED NIRS PERFORMANCE FOR PHARMACEUTICAL IDENTIFICATION

The user study conducted in Study 1 captured a clear need for the nurses to have a tool which could be utilised to accurately identify pills. However, before introducing miniaturised NIRS to the care environment, it is vital to measure numerous aspects of the performance. There exists substantial reference literature for pharmaceutical analysis using the benchtop versions of the device, but the miniaturised version has received far less attention in this regard. We start by building a knowledge base of pharmaceuticals before describing different data preparation techniques for noise removal. Lastly, we present the results of various machine learnings parameters and an evaluation of the proposed solution in a realistic scenario.

4.1 Creating the Knowledge Base

To conduct a thorough investigation of miniaturised NIRS for the purpose of pharmaceutical analysis, we built a knowledge base of 20 different pills (shown in Fig. 2) typically used in care centres. We chose items with different compositions, such as content, size, and degree of transparency, in addition to pharmaceuticals that are visually highly similar. The reason for this is to systematically benchmark how the miniaturised NIRS works on different and more challenging tablet form factors. For example, some pharmaceuticals come in capsules where the powder is stored inside a gelatin container which can interfere with the signal. Other tablets might be transparent, causing less light to be reflected to the lens (*i.e.*, it passes through the object). Additionally, small objects have a less surface area that covers the lens of the device and different shapes (*e.g.*, curved, flat) may both affect the quality of the reflected signal. Visually similar tablets were included to demonstrate how efficiently NIRS can distinguish them, *i.e.* top two rows in Fig. 2.

The pills are sampled in the same way an end user would scan them in a potential deployment; without any form of sample preparation (*e.g.*, grinding to powder state). The purpose of this knowledge base was to use it for benchmarking combinations of pre-processing and machine learning algorithms and to measure if the performance would be sufficiently high for implementation in a medical context.

To choose an appropriate sample size for our experiment, we investigated what has been the standard in the literature. Balabin *et al.* [3] used a training set with an average of 27 samples for each object type when investigating density of biodiesel. Another study collected on average 18 samples for their calibration set of pharmaceutical excipients [8]. Ruckebusch *et al.* [51] used a training set of 17 cotton and 14 polyester samples for quantitative analysis of textiles. An experiment analysing flour and wheat grain used 14 samples for each unique group in the training set [18]. We decided to use 20 unique samples for each object in each category, which is in line with recommendations from literature.

4.1.1 Configuration. When using NIRS there are trade-offs between scan time, signal-to-noise ratio (SNR), and scan precision. Effectively, by scanning an object for a longer period of time we can obtain more reliable data. Drawing on literature [25], we decided to scan our samples with the configurations as shown in Table 2.

Table 2. Scan configuration used during data collection

Method	Start nm	End nm	Width nm	Dig. Resolution	Num. of scans to average	Time
Hadamard	900	1700	8	225	6	2s

Our chosen configuration uses the Hadamard scan method, which multiplexes several wavelengths together and decodes individual wavelengths. Noise in the incident signal is distributed evenly over the spectrum to minimize its effect. Compared to the alternative, column scan, it collects more light and provides a greater SNR [45]. The entire spectral band of the device (*i.e.*, 900-1700 nm) is utilised when scanning, with a width of 8 nm. A digital resolution of 225 is chosen to oversample by 2.25, satisfying the Nyquist-Shannon Sampling Theorem [46]. We obtained a total of six scans each time, and then averaged these together (*i.e.*, signal averaging), which is a common way to increase the SNR while sampling [22]. In practice, signal averaging will appear to the users as one continuous scan with a 2s duration (*i.e.*, the users will only press scan once), however the scan is subsequently comprised of six individual scans. For clarity and simplicity, when discussing a 'scan', we implicitly refer to a scan that is taken with the configuration described in Table 2.

Furthermore, in some cases decreasing the scan time can cause a decrease in accuracy. Here, we chose to have a 2 second completion time for each scan, which in prior research has been categorised as suitable for a common task that needs to maintain the flow of interaction [11]. This enables a user to scan frequently without being annoyed by long waiting times, while still retaining important spectral characteristics in the scan result that can contain useful information for analysis.

4.1.2 Pharmaceuticals. Table 3 displays an overview of the 20 different types of pharmaceuticals we scanned for our knowledge base. These tablets represent a diverse set of samples, with varying characteristics and compositions (both similar and non-similar samples were included), and were purchased at a local pharmacy.

Table 3. Overview of pharmaceuticals in knowledge base.

#	Pharmaceutical	Weight (mg)	Shape	Brand
1	Retafer	100	Round	Orion
2	Probiotic	400	Round	Apteekki
3	Ibuxin Rapid	400	Oval	Recipharm
4	Zinc	550	Round	Vitabalans
5	Caffeine	200	Round	Bulk Powders
6	Pyrvin	100	Round	Orion
7	Multivitamin	635	Round	Orion
8	Kalcipos	500	Oval	Recipharm
9	C-Vitamin	500	Round	Pirkka
10	Lactase Enzyme	400	Round	Yliopiston Apteekki
11	Paracetamol	500	Round	Ratiopharm
12	Magnesium	375	Oval	Vitabalans
13	Disperin	500	Round	Orion
14	Imodium Plus	125	Oval	Imodium

15	Fexorin	120	Oval	Orion
16	B Neuro	300	Oval (Capsule)	Bethover
17	Toncils	600	Round	Vitabalans
18	Ranixal	300	Oval	Ratiopharm
19	Omega-3	1250	Oval (Capsule)	Möller
20	Kelasin	500	Round	Valioravinto

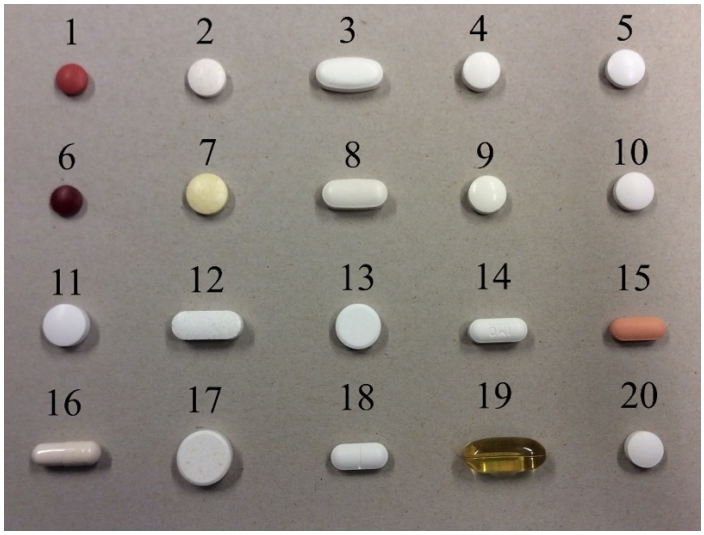


Fig. 2. Overview of the pharmaceuticals selected for the study.

4.1.3 Scan Collection. The list of items was collected from local vendors. The temperature and humidity were stable at 20° and 40% during the data collection. These are two of the main environmental conditions that could have a minor effect on the scan quality [67]. Each sample was carefully aligned on a custom 3D-printed sample holder (Fig. 3) and then scanned. The holder is designed to facilitate correct object placement, while blocking stray light with a lid. The interior of the lid is coated with insulating black tape to absorb light scattered from the sample. A total of 400 spectra were collected (20 pharmaceuticals x 20 samples).

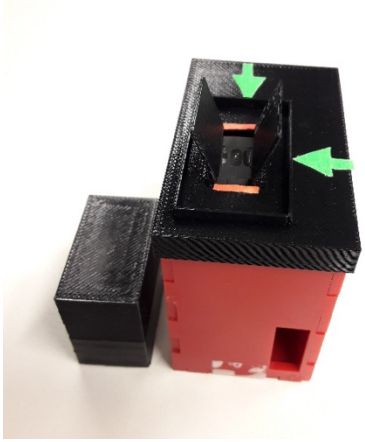


Fig. 3. Pharmaceutical sample holder with lid to block ambient light and arrows indicating where to place the sample.

4.2 Preprocessing

We applied a range of different preprocessing techniques to the spectra and investigated their effect. It is an important step in the analysis of NIR data, as the technique is prone to various sources of noise. The main purpose of preprocessing is to lessen the requirement for training data, to increase the accuracy of statistical classification, and to reduce the complexity of spectrum interpretation and the underlying prediction models [39].

4.2.1 Smoothing: Savitzky-Golay. Savitzky-Golay is a filtering technique which is used to smoothen a sequence of data points in order to increase the SNR. Through convolution, adjoining data points are fitted with a polynomial of order n using the least-squares method [54]. It is widely used to filter chemical data and mathematically it operates as follow:

$$x_j * = \frac{1}{N} \sum_{h=-k}^{k} c_h x_{j+h} \tag{1}$$

Where $x_j *$ is the new value, N is the normalising coefficient, k is the number of neighbour values at each side of j and c_h is the pre-computed coefficient that depend on the selected polynomial order and degree. A polynomial of order 3 and a window size of 11 were chosen, providing sufficient smoothing while still retaining the distinctive features of the spectra. This technique requires k values at each side of j, as a result k values are dropped at the start and end of the spectrum.

4.2.2 Background Correction: True Reflectance. In order to obtain the true reflectance measurements of the scanned object, it is important to correct for background noise (dark current, ambient light). For example, the spectrum of the fluorescent lamps (which are present at the experiment facilities) have some intensity in the NIR region [43], which can have an effect on accuracy. It is in general much more challenging to account for ambient light than dark current. This is because the reference spectra for ambient light changes based on the sample's position and thus a new reference would be needed every scan. In contrast, dark current is not affected by the samples position, but requires reference scans to be removed. In our experiment we satisfy the prerequisites for using the following equation (*i.e.*, no ambient light reaches the NIRS device's lens [1], by covering the sample with a lid when scanning.

$$Spectrum = \frac{x - x_o}{y - x_o} \tag{2}$$

x_o is the dark current, y is the reflectance from the calibrated highly reflectance standard covering the measurement window and x is the reflectance from the object covering entirely the measurement window.

4.2.3 Scatter Correction: Standard Normal Variate. Fluctuations in the effective path length (*e.g.*, different object sizes) and scattering effects (*e.g.*, object surface or composition) can cause interference in the spectrum. A common way to remove these effects is through scatter correction using standard normal variate (SNV) [4]. It simply normalises the spectrum, and should preferably be conducted after filtering [62].

4.2.4 Derivatives: Savitzky-Golay. Derivatives are primarily used to reduce baseline drift and resolve overlapping absorbance peaks. This can be done in combination with smoothing of the spectrum using the Savitzky-Golay filter. When the polynomial is fitted and the parameters are calculated, the derivatives of order n can be found and used for the derivative estimate. This is done for all the data points in the spectrum [50]. Some of the negative effects of derivatives include increased noise and difficulties interpreting the spectrum.

4.2.5 Combinations. The effect of each preprocessing algorithm varies greatly (*e.g.*, smoothing, hardware drift correction, scattering correction) and often works best in synergy with others. In this paper, we use two alternative combinations for preprocessing:

1. Preprocessing 1: Savitzky-Golay smoothing, derivatives and SNV scatter correction.
2. Preprocessing 2: True reflectance background correction, Savitzky-Golay smoothing and derivatives, and SNV scatter correction.

The first alternative applies three of the most popular preprocessing methods on the data. These are fast and efficient algorithms that moderately change the nature of the data. The second option is more complex as it includes background correction. There is no evidence in literature on which is the most appropriate way, and thus we explore these two combinations to determine which produces the best performance for miniaturised NIR spectroscopy data. In Fig. 4 we can see the effect of the alternative preprocessing approaches. The first graph shows how spread-outs the raw spectra for caffeine tablets are. There is some noise present, however it can be difficult to visualise this due to the high resolution. In the second graph, we can observe how the spectra are more condensed after SNV and how some of the noise has been filtered away. The various peaks are now more pronounced, but the spectra have also changed the form because of the derivatives. Lastly, we can examine the similar effect of preprocessing 2 in combination with background correction which has introduced some noise to the spectra.

There was no statistically significant difference between samples as determined by a one-way ANOVA ($F(19,4480) = 1.098$, $p < 0.345$) for no preprocessing. Similarly for Preprocessing 1 ($F(19,4280) < 0.001$, $p = 1.000$) and Preprocessing 2 ($F(19,4280) < 0.001$, $p = 1.000$) there was also no statistically significant difference. However, the F value indicates there is less between-group variance for Preprocessing 1 and Preprocessing 2. This is also supported by the p value. This further illustrates how an improved training set for the classification algorithm has been produced after preprocessing. We found similar results for all the other pharmaceuticals.

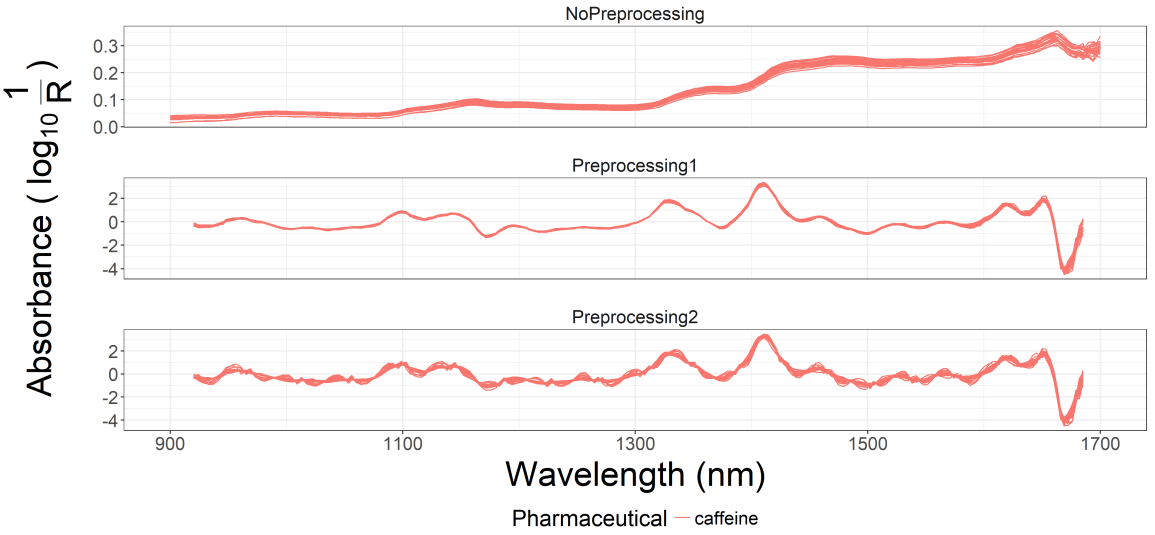


Fig. 4. Comparison between the three different preprocessing algorithms effect on 20 samples of caffeine tablets.

4.3 Results

In this section, we explore the accuracy of different classification algorithms and conduct an in-depth analysis of the processed data. To determine the appropriateness for miniaturised NIRS in a medical context, it is vital to establish how well the device performs. In a scenario where patient medicine has to be verified (*e.g.*, through classification), it is important that the output is accurate to avoid medication mismanagement.

4.3.1 Classification. We calculate the machine learning performance using 10-fold cross validation in 'scikit-learn' (a machine learning library for Python). Each object was sampled with a digital resolution of 225, which means that its absorbance of the illumination from the NIRS was measured at 225 different wavelengths between 900 – 1700nm. Each measured wavelength is considered a feature, and thus we have 225 unique features ($\lambda_1, \lambda_2, \ldots \lambda_{255}$) prior to preprocessing. As described in Section 4.2.1, the smoothing algorithm removes 5 values at the start and the end of the spectrum. As a result, our 'No Preprocessing' dataset has 225 features while the reminding two datasets each use a total of 215 features. In a later section we evaluate the opportunities for feature reduction to improve our understanding of the data, but it was not included as a part of the processing prior to classification.

In Table 4 we can observe the performance of three different classification models for each dataset. As expected, the performance is lowest for the 'No Preprocessing' dataset, as it still contains all the noise from sampling. When comparing 'Preprocessing 1' and 'Preprocessing 2' we notice that the former is performing better or equal for all three models. The background correction may be a contributing factor; if either background correction variable Xo or Y contains unwanted noise this can degrade the spectrum. Furthermore, the references should always be updated to minimize this effect, and in a deployment this can prove challenging. The derivatives could in turn amplify the noise, resulting in lower performance for 'Preprocessing 2'.

Table 4. Performance evaluation of three different types of preprocessing and models. (NB: Naive Bayes; KNN: K-Nearest Neighbor; RF: Random Forest)

Classification									
Metric (%)	No Preprocessing			Preprocessing 1			Preprocessing 2		
	NB	KNN	RF	NB	KNN	RF	NB	KNN	RF
$Accuracy_{(SD)}$	97.80 (0.031)	99.30 (0.011)	97.50 (0.032)	100.0 (0.000)	99.00 (0.012)	100.0 (0.000)	100.0 (0.000)	99.00 (0.012)	99.70 (0.008)
Precision	97.90	99.30	97.80	100.0	99.10	100.0	100.0	99.10	99.80
Recall	97.70	99.20	97.80	100.0	99.00	100.0	100.0	99.00	99.70
F1	97.80	99.20	97.70	100.0	99.00	100.0	100.0	99.00	99.70

Naive Bayes (NB) has overall the best performance out of all the classifiers, when considering Preprocessing 1 and Preprocessing 2, with 100% accuracy, precision and recall. Naive Bayes has been proven to produce reliable results in real world applications, despite the conditional independence assumption it makes [27]. Random Forest (RF) is the second best with an accuracy, precision and recall of 100% for Preprocessing 1 and 99.7% Preprocessing 2. In situations where a probability estimation is required, RF may be a more suitable choice [12]. K-Nearest Neighbor (k = 5) yielded the lowest performance in our test. The relative simple nature of the algorithm could not correctly distinguish some of the closely related classes, causing every 4 out of 400 sample to be misclassified. Interestingly, KNN performed best for No Preprocessing, indicating that it exceeded at describing the underlying relationship between samples in the somewhat noisy data. In general, the high accuracy across all methods underline how accurate and suitable NIRS is for analysis of pharmaceuticals.

4.3.2 Principal Component Analysis. Principal component analysis (PCA) was performed to visually describe the different classes in relationship to the preprocessing. The results are shown in Fig. 5 and can also be used to understand some of the errors made by the classification algorithms. We see that 'No preprocessing' has several overlapping clusters and there is little between-cluster variance. This describes the lower classification performance achieved on this dataset. For the two other datasets noise has been removed and multiple distinct clusters have formed. However, there is some overlap between certain clusters, specifically omega 3 and ranixal is spread out for both Preprocessing 1 and Preprocessing 2. While it is hard to notice, Preprocessing 1 data points are also slightly more condensed. One contributing factor as to why some clusters are closer together then others is that they might share similar absorbance features in specific parts of the spectrum. This may also explain why the relative simple nature of KNN struggles with achieving 100% accuracy. While a plot of the PCA analysis provides a visual understanding of the data, it is a simplified representation. We emphasise that with clusters we mean how the pharmaceutical data points from the two principal components group together, we did not apply a clustering method to the PCA components. The ellipses are drawn on the plot to emphasise how certain clusters have a higher variance than others.

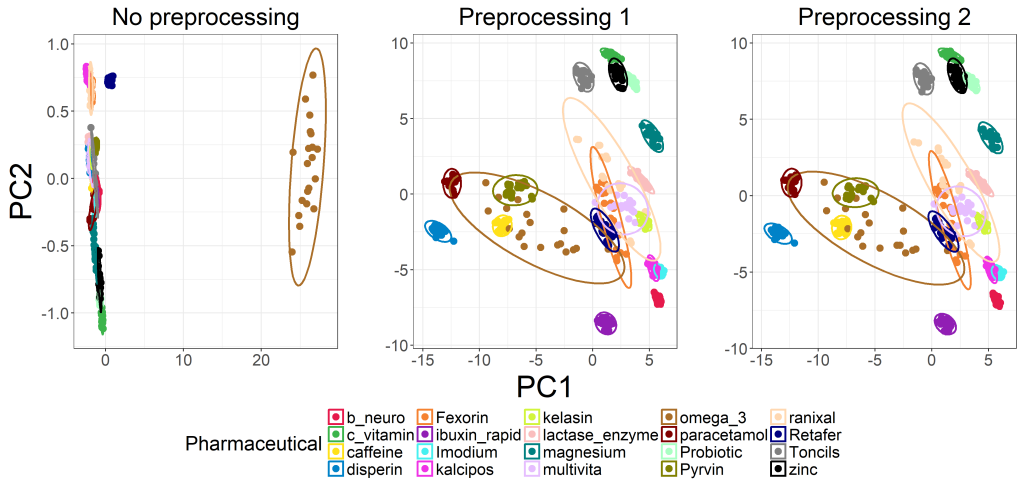


Fig. 5. Principal component analysis of the three datasets with ellipses drawn around each pharmaceutical type to indicate the variance in that group. The proportion of variance captured by PC1 + PC2 for the three plots are: 0.9898 + 0.00347 (No pre), 0.5764 + 0.2164 (Pre 1), 0.5764 + 0.2164 (Pre 2).

4.3.3 Feature Evaluation. Feature evaluation was also conducted in order to assess the importance of each wavelength. The results are visualized in Fig. 6. The top graphs indicate that each method successfully labelled most of the peaks as important. These commonly contain useful information to identify the sample, as its underlying chemical composition is causing the absorbance in the region and thus creating the peak. However, the device utilised in the sampling was over-sensitive in the wavelengths from ~1600 to 1700 nm. This part was recognised as less important by the feature evaluation methods. The bottom graphs illustrate the importance of each individual wavelength. The Gini method is most aggressive, labelling a majority of the points in the upper importance range. It is followed by MDL (Minimum Description Length), which has more dispersion and RF which is spread out more evenly.

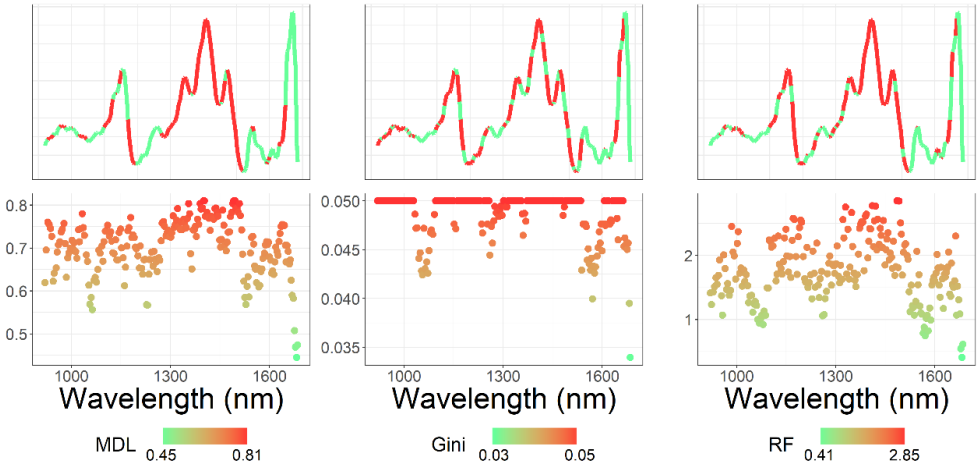


Fig. 6. Bottom: Feature evaluation using MDL, Gini and RF showcasing the relative importance of each wavelength. Top: All features (*i.e.*, wavelengths) which score ≥ to the 50th percentile of the respective feature evaluation method are marked with red (caffeine spectra used for showcasing).

Feature evaluation and selection can help to reduce model complexity and increase performance by removing insignificant parts of the spectrum which contain noise or little information [52]. This may lead to lower computation times as less data has to be processed. However, absorbance peaks may shift depending on scan conditions and too much reduction can lead to a loss of data. This tradeoff can be observed in Fig. 7, which displays recursive feature elimination (RFE) with 10-fold cross validation (CV) for RF and NB (not supported for KNN). The accuracy of both RF and NB increases rapidly until it stabilises at 100% with 10 variables for RF and 38 variables for NB. We calculated the RFECV using the caret R library, as the function in scikit-learn for Python only supports linear models. The results indicate that both NB and RF is suitable for feature reduction and can have its complexity greatly reduced. While we do not utilise feature reduction prior to training our classifiers, we included it to provide a more in-depth exploratory analysis of the data.

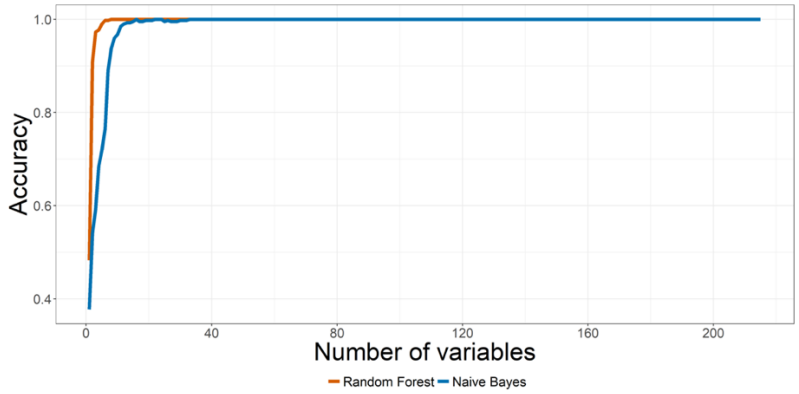


Fig. 7. Change in prediction accuracy with number of features.

4.3.4 Training Scores. Another important aspect of working with machine learning and miniaturised NIRS is knowing how much training data is required to achieve satisfactory accuracy. Fig. 8 shows how the accuracy increases when the models (*i.e.*, KNN, NB, RF) have more training data available. The graphs have a steep initial curve and begin to saturate at 140 training examples for all algorithms. RF experiences a modest

dip in accuracy at 220 training examples, however all of them gradually achieve the accuracy values listed in Table 4. The training scores are higher than the CV score, as this involves training and testing the model on the same data. We noticed that KNN initially had a low training score as well, which is caused by the numbers of neighbours (k = 5). This parameter was tuned for using the entire training set and causes underfitting when used with less training data.

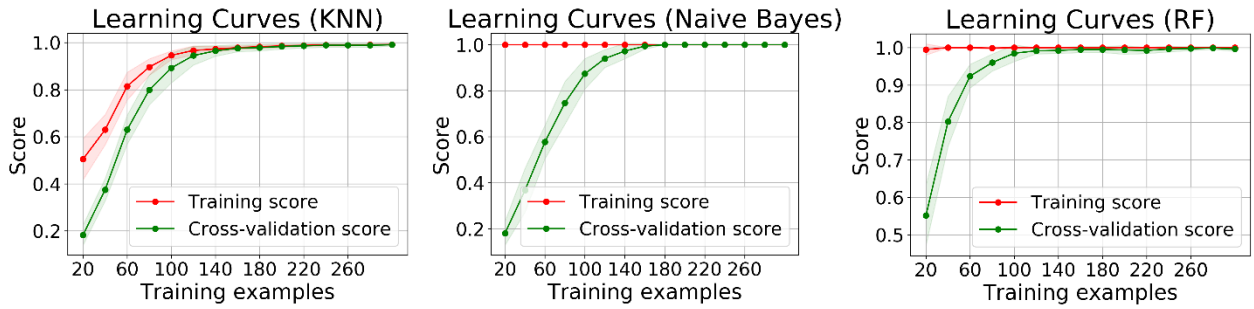


Fig. 8. Change in accuracy with size of training data.

4.4 Evaluation

We conducted a final user study to evaluate our theoretical findings in an applied setting and to compare our solution to existing methods currently in use inside the care centre. As much of the qualitative data had already been collected regarding the nurses' opinion on the device and the opportunities it presents, we focused our evaluation on benchmarking the device's performance.

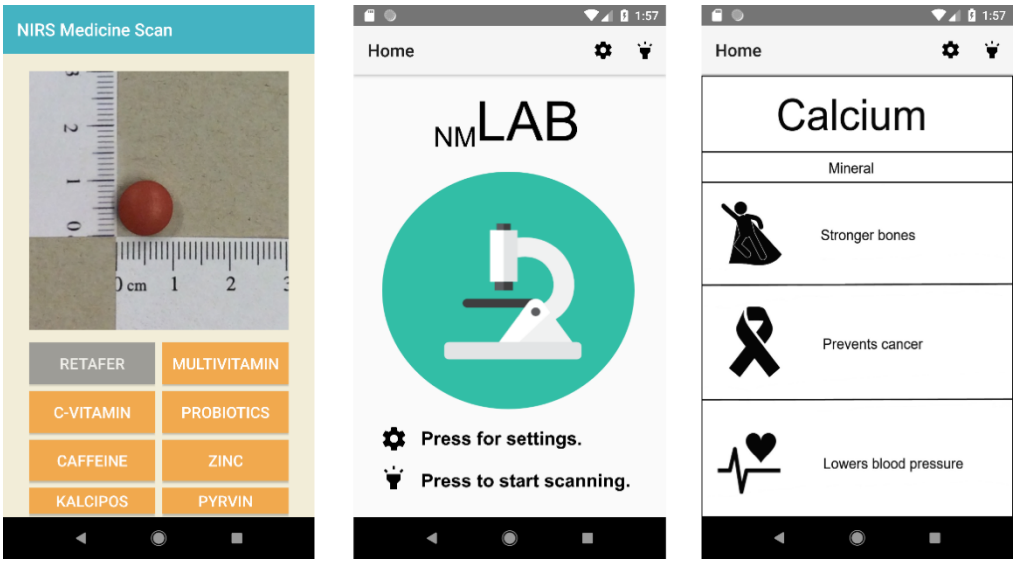
4.4.1 Protocol. The study was conducted with eight care workers at a local care centre (1 male, 7 females; ages: 26 - 59 years old, M = 31, SD = 13.42; work experience: 1 - 30 years, M = 9.5, SD = 11.38). Participants occupied different positions (1 student, 5 nurses, 2 registered nurses). The participants were asked to complete two different pill-sorting tasks. Both tasks consisted of sorting 10 pharmaceuticals (number 1-10 in Table 3 and Fig. 2) where every two pharmaceuticals formed a 'pair of visual similarity'. The following five pairs were formed (each number refers to an individual pharmaceutical in Table 3): 1-6, 2-7, 3-8, 4-9, and 5-10. Judging from Fig. 2 we can observe that each pair, starting from the leftmost, has increasingly more similar visual appearance. This was to facilitate a varying degree of difficulty in our test, as only giving the participants similar looking tablets would be highly challenging (for the first task).

Even though pills were initially selected for their similar visual appearance, all pills were mixed together for the sorting tasks. This accurately replicates how pills arrive in prescription bags to the care centre. We opted for this approach in order to simulate a situation in which a nurse needs to identify a specific pharmaceutical while there are two or more similar looking pills in the bag. While few pill bags come with 10 pharmaceuticals with this particular level of visual similarity at the same time, we aimed to challenge their current practices. As mentioned earlier, when a care worker needs to identify a certain pill (*e.g.*, when sorting prescription bags), they scan the QR code of the bag with a smartphone app. The application will then list the names of the pills with its corresponding picture. It is then up to the care worker to pick a pill from the bag and establish a visual match with the pictures shown on the phone.

1. The first task was to sort the 10 unique pharmaceuticals using a custom Android application (shown in Fig. 9. 1a), replicating current practice. With this application the user can select different pills based on their name, after which a picture and size information of the selected pill appear on the screen. This approach is used to establish a baseline of the current method for identifying pharmaceuticals. The participants also had a ruler available for use, however this solution was not always feasible as some tablets had the same dimensions.

2. The second task consisted of using the NIRS device (Fig. 3) in combination with a custom Android application (Fig. 9. 1b and 1c) to scan the pills. The app utilised the RF model built in section 4.3 using Preprocessing 1 data to classify each pill. Using this application, participants simply had to position the tablet on the NIRS sample holder, press scan in the application, and wait for the scan result.

The participants received a short introduction to both apps before conducting the respective task. We kept a pseudo anonymous reference symbol of each pill for ground truth and asked the participants to place the pill in one of ten cups (each cup was labelled with one of the ten names of the tablets chosen for the study) after they had identified the pill. We never revealed the ground truth of the pills and started with task 1 before moving on to task 2. The participants could sort the tablets in any order using the tools available to that respective task. The participants were informed that each pill was unique, and thus in the first task the participants would sort one tablet per cup. However, if the application in the second task would classify two pharmaceuticals as zinc, the participant would have to place both in the zinc cup, regardless of their knowledge regarding uniqueness of the pills.



(1a) (1b) (1c)

Fig. 9. (1a): The app for first sorting tasks in the user study. (1b): Main menu of the app for second sorting task. (1c): Result screen of app for the second sorting task, shown to the participant after scanning a calcium tablet.

4.4.2 Results. Sorting manually in task 1 proved to be difficult, and the participants struggled even with the more distinctive pills as summarized in Table 5. We observe that the accuracy is only 50% for manual identification, meaning half of the pharmaceuticals (*i.e.*, 40) were wrongly classified. Given the nature of the manual sorting task (*i.e.*, the row sum and column sum of the confusion matrix will always be 8), the precision, recall and F1 score is also 50%. For the second sorting task, where participants used the NIRS application, the accuracy is much higher. A total of 77 pills were correctly classified, resulting in 96.25% accuracy and precision, recall, and F1 score in the same range. We observed that most of the errors using the NIRS came from suboptimal placement of the tablets on the NIRS scanner.

Table 5. Performance evaluation of the classification results from the user study for sorting tablets manually and with the use of NIRS.

	Classification	
Metric (%)	Manual	NIRS
Accuracy	50.00	96.25
Precision	50.00	96.88
Recall	50.00	96.25
F1	50.00	96.56

Judging from the manual confusion matrix shown in Fig. 10, the classifications seem randomly spread out besides the 40 correct ones along the diagonal. By comparing with the overview of the tablets in Fig. 2. we see that the more visually different ones, such as Multivitamin, Pyrvin and Retafer were the easiest to classify manually. In contrast, the more visually similar ones such as C-vitamin, Caffeine, Lactase Enzyme and Zinc were more challenging. The confusion matrix for NIRS shows a clear trend along the diagonal. Multivitamin was classified as Retafer once, and probiotic as zinc twice. The data points from multivitamin is in close proximity to Retafer as seen in the PCA analysis in Fig. 5., and also probiotic and zinc are closely located. This indicates that poor object placement (e.g., object does not cover the entire lens) should be addressed carefully before a real-world implementation of the technology.

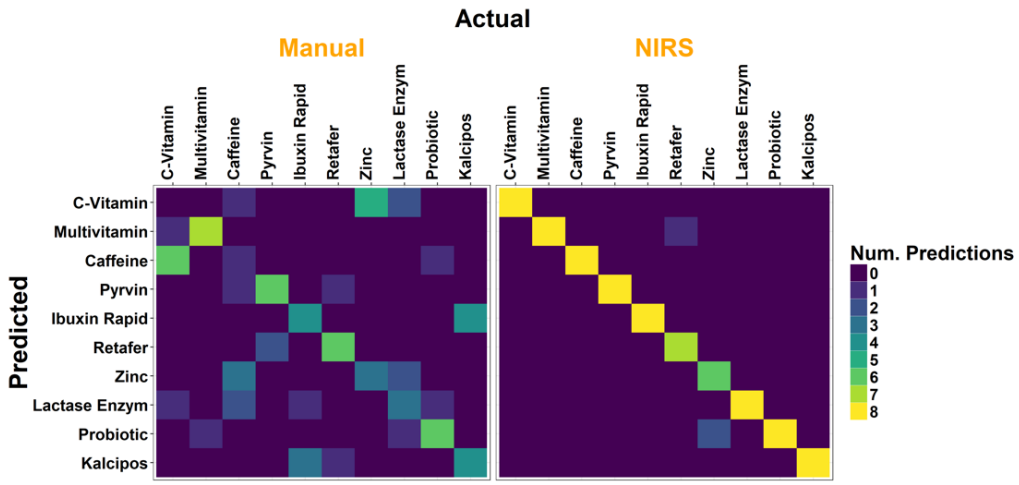


Fig. 10. Confusion matrix of the classification results from the user study for sorting tablets manually and with the use of NIRS.

5 DISCUSSION

The issues related to mismanagement of medication is evident in our society and leads to more than just economic repercussions [15]. In fact, errors occurring during administration of pharmaceuticals have been the source of many fatal medical reports in the past [20]. Structural changes have been requested in different sections of the pharmaceutical supply chain to improve communications and information flow [41]. In this paper, we focus on the final stage of the supply chain, which entails pill verification prior to administration. As this stage acts as the "last line of defence", it is important to develop reliable mechanisms to reduce the potential for medication errors. Our findings identified a significant need for tools that enable nurses to accurately identify pharmaceuticals in situ and in real-time. Furthermore, we highlight how NIRS is a highly

accurate instrument for pharmaceutical analysis, even in the miniaturised form factor. As the hardware continues to develop, new exciting use cases can open up. In the future, we envision that NIRS can be embedded in pill dispensers and other suitable care tools in need of accurate ground truth.

5.1 Addressing Medication Mismanagement

Most of the currently available systems for medication management help the user with scheduling and notification. In the context of an elderly care centre there are further protocols in place aimed at reducing the likelihood of medication errors. In comparison to the issues found in the hospital environment in which Pérez *et al. [48]* evaluated their RFID system, matching between a patient's medication bag and user was not the main problem identified in the elderly care centre. There is already a functional barcode system in place, however this did not allow for the identification of individual pharmaceuticals found in the bags. Specifically, it was noticeable during our user study that the pill identification methods presently utilised in the elderly care facility were not meticulous enough considering the potential consequences of a false positive. As it stands, nurses frequently have to rely on their own judgement to interpret the identity of a pill. This practice is susceptible to human error due to pill similarity, inhibited cognition caused by stress and work overload [41], and medication misplacement.

During our discussion with the nurses we received positive feedback on the potential of miniaturised NIRS' on improving their workflow as well as reducing the likelihood of medication errors, thus improving a patient's quality of life. A list of four main impact areas of miniaturised NIRS in elderly care was derived from the user study. Furthermore, previous work has already demonstrated that it is feasible for novice end users to utilise this advanced instrument via a 3D printed casing for the miniaturised NIRS and an accompanying smartphone application aimed at guiding the scanning process [29]. This system could be modified to fit a care environment, with targeted functionality, such as built in patient medication lists, automated drug interaction checks and data encryption [36]. At first the device could be introduced to the nursing home as an optional tool, which is consistently accessible when the care workers need to confirm or identify pharmaceuticals. As they grow more accustomed to the device, it could be integrated as part of their workflow. The small form factor and light weight of the device, allows it to easily be carried around for in situ scanning. In a homecare setting, NIRS modules could be embedded in currently available medicine scheduling and notification systems. This would enable the system to have an accurate ground truth of which pills are taken.

5.2 Accurately Identifying Pharmaceuticals using Miniaturised NIRS

The results from our experiments show how miniaturised NIRS can identify pharmaceuticals with high accuracy. The preparation of the data was a vital step to achieve the desired outcome. 'Preprocessing 1' was the leading option in terms of performance, and 100% classification accuracy was achieved utilising NB and RF. This is partially caused by the fact 'Preprocessing 1' does not involve background correction and has less risk of noise amplification caused by derivatives. The user study conducted to evaluate our theoretical findings in a realistic scenario, further underlined NIRS suitability for pill identification in an elderly care centre. In the sorting task, nurses achieved 96.25% accuracy using NIRS, compared to 50% when manually identifying with a smartphone application. This indicates that the existing practices and tools for correctly identifying prescription medicine is not adequate, as nurses struggle to separate medication that is not visually significantly different. Equipping the nurses with the NIRS device could potentially reduce the magnitude of medication mismanagement drastically. While we built the scan library in controlled conditions with stable temperature and humidity, in practice these factors would only impact the result in extreme conditions, such as a tropical climate. Another aspect that may influence the accuracy in real world usage is poor sample positioning, however previous work has shown how this can be alleviated through good design of sample holders and assistive software [29]. Yet, we established that the sample holder in our user study still had room for improvements, as small number of pills where placed in suboptimal positions. This lead to 3 pills being misclassified out of 80. A number that should be reduced before an actual long-term deployment of the technology could start in the care centre.

The choice of classification algorithm is not a straightforward decision and is affected by multiple factors. The PCA analysis highlights how well miniaturised NIRS captures information about the pharmaceuticals, as predominantly clear clusters form after preprocessing. Yet, it also emphasizes the potential risk scenarios, as some data points from different classes are in short proximity of each other. It is therefore crucial that the chosen classification algorithm appropriately determines appropriate class boundaries, especially in these circumstances. Opting for a model that reports probability estimation alongside the predictions [12], would enable the nurses to interpret the results with greater confidence. In the confined set of cases where the result might be indecisive, the system can suggest to the user to discard that specific pill, thus reducing the risk associate with misclassifications. The threshold would need to be calibrated to achieve a minimum of FPs, as patient safety is the top priority. Results could also consist of multiple options alongside confidence intervals rather than one result if the system is unsure, allowing the user to filter the options based on their own perception. Working with medication, it is critical that the precision is high in order to give the most accurate information to the nurses. Yet, recall is also important to make sure each medication is identified. Based on the F1 score in Table 4, NIRS has adequate performance for pharmaceutical classification. When introducing such a tool to a healthcare environment, this is exceedingly important, as nurses will rely on the tool to make the right choices for the patients.

Complex models might increase the total time taken before the nurse receives a response. While we considered it beyond the scope of this paper to go into details on processing load and time of each individual model, we evaluated the features through multiple methods. The results show that the RF model could be simplified significantly, while retaining accuracy. A simplified model is faster, easier to interpret (a widely discussed trait in the UbiComp community [34]), can run more efficiently on smart devices [37], and can potentially contain less noise. In a setting such as a care centre, where the health of patients is at stake, understanding the topology of the algorithm is crucial.

Finally, before miniaturised NIRS could be deployed in a care facility, a broad knowledge base containing all relevant medications would be necessary. This library could be built using crowdsourcing methods [23]. It would also require periodical updates when new pills are prescribed to patients. To determine guidelines for data collection we explored the training scores of KNN, NB, and RF. The results are in line with literature on benchtop instruments with close to 20 samples of each item being sufficient to achieve high accuracy. These databases could initially be built locally and eventually centralised databases with scans from both manufacturers, pharmacies, and other partners could emerge. In cases where nurses come across pills with no prior training data, clustering could potentially be a powerful tool to understand the attributes of the pill. For example, if the pill has different levels or similar combinations of active pharmaceutical ingredients to a pill in the training set, they will be clustered together in close proximity.

5.3 Next Generation Healthcare Technology

We argue that many of the challenges related to medication mismanagement can be alleviated by introducing miniaturised NIRS to the care personnel. We initially considered other identification methods, such as machine vision techniques [16]. However, they still have limitations due to their inability to analyse the chemical composition of the pills [56]. For example, computer vision techniques are constrained to using visual features such as size, shape, and colour in their classification algorithms. A mobile application utilising machine vision [19] achieved less than 80% accuracy using these three features when identifying pharmaceuticals with seemingly significantly different characteristics. However, the number of samples in their training set was significantly larger (N = 95) then ours and the execution time of their method was slightly lower. Similarly, the HelpmePills application [10] is also confined to the low number of visual features for pharmaceutical classification. In addition, their setup relies on setting up the smartphone in the correct position and angle using seemingly random objects (*e.g.*, books) which seem cumbersome and inappropriate for a healthcare context. Furthermore, it is challenging to evaluate their approach as it appears that training and testing was completed on the same dataset. A limitation affecting current machine learning methods is the ability to only analyse an object based on its surface area. For example, the same capsule can hold two

vastly different drugs, however this will not be apparent based on visual inspection. In comparison, NIRS techniques uses up to several hundred features (depending on hardware and scan configuration) to precisely infer the inner molecular composition of an object.

The elderly care facility is a hectic environment and the tools designed for nurses must efficiently integrate with their workflow. Time is always a constraint and care workers quickly change between various tasks [30]. The sample scan time was set to 2 seconds to maintain the flow of interaction [11]. It is the nurses' decision when they deem it necessary to use the NIRS for identification of a tablet. Initially, it is meant as an optional tool that can be used when they are unsure about a particular pill. We designed the scanning application to be simple to use, as we do not want to add complexity to the nurses' workflow (*e.g.*, force additional labour). We received positive feedback on the application when the nurses used it to identify pharmaceuticals. They did not seem to mind the scan time and expressed how practical and useful the application seemed. Yet, to make a definite conclusion the use would have to be evaluated through a long-term deployment at a care centre. Like other healthcare technology, it is vital that the users find value in the tool to achieve a high user adoption rate [65].

We consider our work an important first step towards enabling this type of technology in an elderly care setting. As mentioned, there exists a large body of work related to the work of traditional benchtop and less portables NIRS instruments for pharmaceutical analysis [13,26]. However, these are at an entirely different level when it comes to price, size, weight, complexity, and intended use. Additionally, previous work generally focuses on conducting the analysis in a lab setting, and the procedure involves much more complex sample preparation (*e.g.*, grinding tablets to powder form). While it illustrates some of the potential in this technology, it is vital to evaluate this new generation of NIRS devices in the new environment and use cases it enables (*e.g.*, practical applications close to the end user).

5.4 Limitations

We recognize several limitations in the presented work. Our user study was conducted at a single care centre, limiting the generalizability of the results. Yet, we expect similar issues to arise in other facilities as well, at least here in Finland, given how the treatment and protocols in such facilities is governed and supervised by the same governmental institutions. Furthermore, we limited our study to a set of 10 unique pharmaceuticals. In reality, the set of pharmaceuticals is likely to be considerably larger. This raises new questions on construction of the knowledge base and the risk of misclassification, and the set of 10 was chosen to reliably demonstrate the feasibility of our approach.

5.5 Future work

We have the opportunity to complete similar studies in a subacute hospital specialising in elderly care. This is important because different care centres operate with different number of patients and varying staffing procedures, affecting the level of individual care offered to patients. It is expected that the explored solution would be most useful to those care centres in which nurses are responsible for a high number of patients or where there is a limited availability of experienced staff. This yields us a great opportunity to explore in our future work, together with domain experts using our solutions in their authentic work environments. Enabling rapid and accurate in-situ identification of pharmaceuticals can benefit a plethora of scenarios besides the one presented in our work. As highlighted previously, the miniaturised NIRS could potentially be a useful tool in a homecare setting. However, we must emphasize that it would require significant work on the usability before elderly, or even other users (*e.g.*, visually, situationally impaired [53]) would be able to use it as an independent solution. The recent advancements in hardware development (decrease size, weight, and price) also opens new opportunities for distribution of the technology at a larger scale. Drug consumption rooms [63] could use the device for detection of dangerous chemicals in narcotics. In developing countries, the device could be deployed in central location bundled with a tablet, allowing users to identify counterfeit medication (*e.g.*, antimalarial tablets [44]).

Further work should also explore and suggest strategies on how pharmaceutical companies, care centres, or other interested individuals could all contribute to the construction of knowledge bases on pharmaceuticals that the use of NIRS requires. To this end, care centres embracing this technology could also operate with "hyperlocal" knowledge bases containing only pharmaceuticals that are relevant in their context, effectively reducing the risk of misclassification.

6 CONCLUSION

Our work has systematically investigated the applicability for assistive medication management in elderly care using miniaturised NIRS. We conducted a user study with nurses working at a local care centre to analyse the medication procedures and explore use cases which could benefit from the technology. The findings are summarised in a list containing the four main impact areas, indicating a significant demand for a tool which enables rigorous identification of pills. Subsequently, we benchmark the performance of different preprocessing and machine learning algorithms, to establish proof of miniaturised NIRS capacity to accurately identify pharmaceuticals. The results show that miniaturised NIRS can identify pills with high precision, making it exceedingly relevant for the main impact areas discovered in our qualitative study. In our ongoing work, we intend to deploy an in-situ scanner at a subacute hospital specialising in elderly care.

REFERENCES

[1] A.P. Popov, A.V. Bykov and I.V. Meglinski. 2017. Influence of local probe pressure on diffuse reflectance spectra of human skin measured in vivo. *J. Biomed. Opt.*, 22.

[2] Manel Alcalà, Marcelo Blanco, Daniel Moyano, Neville W Broad, Nada O'Brien, Don Friedrich, Frank Pfeifer and Heinz W Siesler. 2013. Qualitative and quantitative pharmaceutical analysis with a novel hand-held miniature near infrared spectrometer. *Journal of Near Infrared Spectroscopy*, 21 (6). 445-457.

[3] Roman M Balabin and Sergey V Smirnov. 2011. Variable selection in near-infrared spectroscopy: benchmarking of feature selection methods on biodiesel data. *Analytica chimica acta*, 692 (1). 63-72.

[4] RJ Barnes, Mewa Singh Dhanoa and Susan J Lister. 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied spectroscopy*, 43 (5). 772-777.

[5] Regina M Benjamin. 2012. Medication adherence: helping patients take their medicines as directed. *Public health reports*, 127 (1). 2.

[6] Craig F Bohren and Donald R Huffman. 2008. *Absorption and scattering of light by small particles*. John Wiley & Sons.

[7] Donald A Burns and Emil W Ciurczak. 2007. *Handbook of near-infrared analysis*. CRC press.

[8] A Candolfi, R De Maesschalck, DL Massart, PA Hailey and ACE Harrington. 1999. Identification of pharmaceutical excipients using NIR spectroscopy and SIMCA. *Journal of pharmaceutical and biomedical analysis*, 19 (6). 923-935.

[9] Wen-Wei Chang, Tung-Jung Sung, Heng-Wei Huang, Wei-Chih Hsu, Chi-Wei Kuo, Jhe-Jhao Chang, Yi-Ting Hou, Yi-Chung Lan, Wen-Cheng Kuo and Yu-Yen Lin. 2011. A smart medication system using wireless sensor network technologies. *Sensors and Actuators A: Physical*, 172 (1). 315-321.

[10] António Cunha, Telmo Adão and Paula Trigueiros. 2014. Helpmepills: A mobile pill recognition tool for elderly persons. *Procedia Technology*, 16. 1523-1532.

[11] Jim Dabrowski and Ethan V Munson. 2011. 40 years of searching for the best computer system response time. *Interacting with Computers*, 23 (5). 555-564.

[12] Theresa Dankowski and Andreas Ziegler. 2016. Calibrating random forests for probability estimation. *Statistics in medicine*, 35 (22). 3949-3960.

[13] Klara Dégardin, Aurélie Guillemain, Nicole Viegas Guerreiro and Yves Roggo. 2016. Near infrared spectroscopy for counterfeit detection using a large database of pharmaceutical tablets. *Journal of pharmaceutical and biomedical analysis*, 128. 89-97.

[14] M Dyrby, S Balling Engelsen, L Nørgaard, M Bruhn and L Lundsberg-Nielsen. 2002. Chemometric quantitation of the active substance (containing C≡ N) in a pharmaceutical tablet using near-infrared (NIR) transmittance and NIR FT-Raman spectra. *Applied Spectroscopy*, 56 (5). 579-585.

[15] Frontier Economics. 2014. Exploring the costs of unsafe care in the NHS: a report prepared for the Department of Health. *London, Frontier Economics*.

[16] Juergen Gall, Angela Yao, Nima Razavi, Luc Van Gool and Victor Lempitsky. 2011. Hough forests for object detection, tracking, and action recognition. *IEEE transactions on pattern analysis and machine intelligence*, 33 (11). 2188-2202.

[17] Robert Geirhos, David HJ Janssen, Heiko H Schütt, Jonas Rauber, Matthias Bethge and Felix A Wichmann. 2017. Comparing deep neural networks against humans: object recognition when the signal gets weaker. *arXiv preprint arXiv:1706.06969*.

[18] Ma Inmaculada González-Martín, Guillermo Wells Moncada, Claudio González-Pérez, Nelson Zapata San Martín, Fernando López-González, Iris Lobos Ortega and Jose-Miguel Hernández-Hierro. 2014. Chilean flour and wheat grain: tracing their origin using near infrared spectroscopy and chemometrics. *Food chemistry*, 145. 802-806.

[19] Andreas Hartl, Clemens Arth and Dieter Schmalstieg. 2011. Instant medical pill recognition on mobile phones. In *proc. of IASTED International Conference on Computer Vision*, 188-195.

[20] Nicole Hartnell, Neil MacKinnon, Ingrid Sketris and Mark Fleming. 2012. Identifying, understanding and overcoming barriers to medication error reporting in hospitals: a focus group study. *BMJ Qual Saf*, 21 (5). 361-368.
[21] M Hayakawa, Y Uchimura, K Omae, K Waki, H Fujita and K Ohe. 2013. A smartphone-based medication self-management system with realtime medication monitoring. *Applied clinical informatics*, 4 (1). 37-52.
[22] GM Hieftje. 1972. Signal-to-noise enhancement through instrumental techniques. *Analytical Chemistry*, 44 (6). 81A-88A.
[23] Simo Hosio, Jorge Goncalves, Niels van Berkel and Simon Klakegg. 2016. Crowdsourcing Situated & Subjective Knowledge for Decision Support. In *International Joint Conference on Pervasive and Ubiquitous Computing Adjunct*, ACM, 1478-1483. DOI: https://doi.org/10.1145/2968219.2968591
[24] Hanna-Leena Huttunen, Simon Klakegg, Nils Van Berkel, Aku Visuri, Denzil Ferreira and Raija Halonen. 2017. Understanding elderly care: a field-study for designing future homes. In *Proceedings of the 19th International Conference on Information Integration and Web-based Applications & Services*, ACM, 390-394.
[25] Texas Instrument, DLP® NIRscan™ Nano EVM User's Guide. Accessed from http://www.ti.com/tool/DLPNIRNANOEVM
[26] Marzena Jamrógiewicz. 2012. Application of the near-infrared spectroscopy in the pharmaceutical technology. *Journal of pharmaceutical and biomedical analysis*, 66. 1-10.
[27] Liangxiao Jiang, Chaoqun Li, Shasha Wang and Lungan Zhang. 2016. Deep feature weighting for naive Bayes and its application to text classification. *Engineering Applications of Artificial Intelligence*, 52. 26-39.
[28] Matti T Kinnunen, Alexey P Popov, Jerzy Plucinski, Risto A Myllyla and Alexander V Priezzhev. 2004. Measurements of glucose content in scattering media with time-of-flight technique: comparison with Monte Carlo simulations. In *Saratov Fall Meeting 2003: Optical Technologies in Biophysics and Medicine V*, International Society for Optics and Photonics, 181-192.
[29] S. Klakegg, J. Goncalves, N. van Berkel, C. Luo, S. Hosio and V. Kostakos. 2017. Towards Commoditised Near Infrared Spectroscopy. In *Conference on Designing Interactive Systems*.
[30] S. Klakegg, N. van Berkel, A. Visuri, C. Luo, J. Goncalves, S. Hosio, H. Huttunen and D. Ferreira. 2017. Informing Caregivers Through an Assistive Tool: An Investigation of Elderly Care Metrics. In *British HCI*, Sunderland, UK.
[31] Simon Klakegg, Niels van Berkel, Aku Visuri, Hanna-Leena Huttunen, Simo Hosio, Chu Luo, Jorge Goncalves and Denzil Ferreira. 2017. Designing a context-aware assistive infrastructure for elderly care. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, Maui, Hawaii, ACM, 563-568. DOI: https://doi.org/10.1145/3123024.3124403
[32] Simon Klakegg, Chu Luo, Jorge Goncalves, Simo Hosio and Vassilis Kostakos. 2016. Instrumenting smartphones with portable NIRS. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, ACM, 618-623.
[33] V Korhonen, T Myllylä, M Yu Kirillin, AP Popov, AV Bykov, AV Gorshkov, EA Sergeeva, M Kinnunen and V Kiviniemi. 2014. Studies of light propagation in optical sensing of the cerebral cortex of the human brain. *IEEE Journal of Selected Topics in Quantum Electronics*, 20 (2). 7100310.
[34] V. Kostakos, D. Ferreira, J. Goncalves and S. Hosio. 2016. Modelling Smartphone Usage: A Markov State Transition Model. In *International Joint Conference on Pervasive and Ubiquitous Computing*, 486-497. DOI: https://doi.org/10.1145/2971648.2971669
[35] Jiangbo Li, Wenqian Huang, Xi Tian, Chaopeng Wang, Shuxiang Fan and Chunjiang Zhao. 2016. Fast detection and visualization of early decay in citrus using Vis-NIR hyperspectral imaging. *Computers and Electronics in Agriculture*, 127. 582-592.
[36] Chu Luo, Angelos Fylakis, Juha Partala, Simon Klakegg, Jorge Goncalves, Kaitai Liang, Tapio Seppänen and Vassilis Kostakos. 2016. A data hiding approach for sensitive smartphone data. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, 557-568.
[37] Chu Luo, Miikka Kuutila, Simon Klakegg, Denzil Ferreira, Huber Flores, Jorge Goncalves, Mika Mäntylä and Vassilis Kostakos. 2017. TestAWARE: A Laboratory-Oriented Testing Tool for Mobile Context-Aware Applications. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1 (3). 80. DOI: https://doi.org/10.1145/3130945
[38] Marena Manley. 2014. Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. *Chemical Society Reviews*, 43 (24). 8200-8214.
[39] Harald Martens and Edward Stark. 1991. Extended multiplicative signal correction and spectral interference subtraction: new preprocessing methods for near infrared spectroscopy. *Journal of pharmaceutical and biomedical analysis*, 9 (8). 625-635.
[40] Medisafe, Research. Accessed from https://medisafe.com/research/
[41] Eija Metsälä and Ulla Vaherkoski. 2014. Medication errors in elderly acute care–a systematic review. *Scandinavian journal of caring sciences*, 28 (1). 12-28.
[42] MicroNIR, MicroNIR. Accessed from https://www.viavisolutions.com/en-us/osp/products/micronir-spectrometers
[43] Jørgen Alexander Mølmann, Olavi Junttila, Øystein Johnsen and Jorunn Elisabeth Olsen. 2006. Effects of red, far-red and blue light in maintaining growth in latitudinal populations of Norway spruce (Picea abies). *Plant, cell & environment*, 29 (2). 166-172.
[44] Gaurvika ML Nayyar, Joel G Breman, Paul N Newton and James Herrington. 2012. Poor-quality antimalarial drugs in southeast Asia and sub-Saharan Africa. *The Lancet infectious diseases*, 12 (6). 488-496.
[45] NIRScan, DLP® NIRscan™ Nano Evaluation Module. Accessed from http://www.ti.com/tool/DLPNIRNANOEVM
[46] Alan V Oppenheim. 1999. *Discrete-time signal processing*. Pearson Education India.
[47] Celio Pasquini. 2003. Near infrared spectroscopy: fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, 14 (2). 198-219.
[48] María Martínez Pérez, Mariano Cabrero-Canosa, José Vizoso Hermida, Lino Carrajo García, Daniel Llamas Gómez, Guillermo Vázquez González and Isabel Martín Herranz. 2012. Application of RFID technology in patient tracking and medication traceability in emergency care. *Journal of medical systems*, 36 (6). 3983-3993.

[49] Alexey P Popov, Alexander V Bykov, Sami Toppari, Matti Kinnunen, Alexander V Priezzhev and Risto Myllylä. 2012. Glucose sensing in flowing blood and intralipid by laser pulse time-of-flight and optical coherence tomography techniques. *IEEE Journal of Selected Topics in Quantum Electronics*, 18 (4). 1335-1342.
[50] Åsmund Rinnan, Frans van den Berg and Søren Balling Engelsen. 2009. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry*, 28 (10). 1201-1222.
[51] Cyril Ruckebusch, F Orhan, A Durand, T Boubellouta and JP Huvenne. 2006. Quantitative analysis of Cotton–Polyester textile blends from near-infrared spectra. *Applied spectroscopy*, 60 (5). 539-544.
[52] Claude Sammut and Geoffrey I Webb. 2011. *Encyclopedia of machine learning*. Springer Science & Business Media.
[53] Z. Sarsenbayeva, J. Goncalves, J. García, S. Klakegg, S. Rissanen, H. Rintamäki, J. Hannu and V. Kostakos. 2016. Situational Impairments to Mobile Interaction in Cold Environments. In *International Joint Conference on Pervasive and Ubiquitous Computing*, 85-96. DOI: https://doi.org/10.1145/2971648.2971734
[54] Abraham Savitzky and Marcel JE Golay. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36 (8). 1627-1639.
[55] SCiO, SCiO. Accessed from https://www.consumerphysics.com/
[56] Heinz W Siesler, Yukihiro Ozaki, Satoshi Kawata and H Michael Heise. 2008. *Near-infrared spectroscopy: principles, instruments, applications*. John Wiley & Sons.
[57] Michael C Sokol, Kimberly A McGuigan, Robert R Verbrugge and Robert S Epstein. 2005. Impact of medication adherence on hospitalization risk and healthcare cost. *Medical care*, 43 (6). 521-530.
[58] Milan Sonka, Vaclav Hlavac and Roger Boyle. 2014. *Image processing, analysis, and machine vision*. Cengage Learning.
[59] Lan Sun, Chang Hsiung, Christopher G Pederson, Peng Zou, Valton Smith, Marc von Gunten and Nada A O'Brien. 2016. Pharmaceutical raw material identification using miniature near-infrared (MicroNIR) spectroscopy and supervised pattern recognition using support vector machine. *Applied spectroscopy*, 70 (5). 816-825.
[60] Tellspec, Tellspec. Accessed from http://tellspec.com/
[61] Paul Trueman, DG Taylor, K Lowson, A Bligh, A Meszaros, D Wright, J Glanville, Jennifer Newbould, Mike Bury and Nick Barber. 2010. Evaluation of the scale, causes and costs of waste medicines. Report of DH funded national project.
[62] Joe Wolf, Paul Michael Carter, Simon Geoffrey Parker, Robert Fearn and Niki Frampton. 2006. Multirate cochlear stimulation strategy and apparatus, Google Patents.
[63] Sara Woods. 2014. Drug Consumption Rooms in Europe. *Organisational overview. Regenboog Group, Amsterdam*.
[64] Maciej S Wróbel, Alexey P Popov, Alexander V Bykov, Matti Kinnunen, Malgorzata JÄ and Valery V Tuchin. 2015. Measurements of fundamental properties of homogeneous tissue phantoms. *Journal of biomedical optics*, 20 (4). 045004-045004.
[65] Long Wu, Jhao-Yin Li and Chu-Ying Fu. 2011. The adoption of mobile healthcare by hospital's professionals: An integrative perspective. *Decision Support Systems*, 51 (3). 587-596.
[66] Danfei Xu, Gerald E Loeb and Jeremy A Fishel. 2013. Tactile identification of objects using Bayesian exploration. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, IEEE, 3056-3061.
[67] Jimmy Zumba, James Rodgers and Matthew Indest. 2017. Impact of temperature and relative humidity on the near infrared spectroscopy measurements of cotton fiber micronaire. *Textile Research Journal*. 0040517517720499.

Received November 2017; revised February 2018; accepted April 2018