

Objective determination of peripheral edema in heart failure patients using short-wave infrared molecular chemical imaging

Aaron G. Smith,^a Reina Perez,^b Aaron Thomas,^b Shona Stewart,^{a,*}
Arash Samiei,^a Arjun Bangalore,^a Heather Gomer,^a,^{ORCID} Marlina B. Darr,^a
Robert C. Schweitzer,^a Sandhya Vasudevan,^a Jeffrey Cohen,^{a,b},^{ORCID}
J. Christopher Post,^a Srinivas Murali,^b and Patrick J. Treado,^a^{ORCID}

^aChemImage Corporation, Pittsburgh, Pennsylvania, United States

^bAllegheny General Hospital, Pittsburgh, Pennsylvania, United States

Abstract

Significance: Peripheral pitting edema is a clinician-administered measure for grading edema. Peripheral edema is graded 0, 1+, 2+, 3+, or 4+, but subjectivity is a major limitation of this technique. A pilot clinical study for short-wave infrared (SWIR) molecular chemical imaging (MCI) effectiveness as an objective, non-contact quantitative peripheral edema measure is underway.

Aim: We explore if SWIR MCI can differentiate populations with and without peripheral edema. Further, we evaluate the technology for correctly stratifying subjects with peripheral edema.

Approach: SWIR MCI of shins from healthy subjects and heart failure (HF) patients was performed. Partial least squares discriminant analysis (PLS-DA) was used to discriminate the two populations. PLS regression (PLSR) was applied to assess the ability of MCI to grade edema.

Results: Average spectra from edema exhibited higher water absorption than non-edema spectra. SWIR MCI differentiated healthy volunteers from a population representing all pitting edema grades with 97.1% accuracy ($N = 103$ shins). Additionally, SWIR MCI correctly classified shin pitting edema levels in patients with 81.6% accuracy.

Conclusions: Our study successfully achieved the two primary endpoints. Application of SWIR MCI to monitor patients while actively receiving HF treatment is necessary to validate SWIR MCI as an HF monitoring technology.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.26.10.105002](https://doi.org/10.1117/1.JBO.26.10.105002)]

Keywords: molecular chemical imaging; short-wave infrared; heart failure; artificial intelligence; peripheral edema.

Paper 210090RR received Mar. 22, 2021; accepted for publication Sep. 16, 2021; published online Oct. 23, 2021.

1 Introduction

Heart failure (HF) can result from any structural or functional abnormality that impairs the ability of the ventricle to eject blood, referred to as systolic HF, or to fill with blood, termed diastolic HF. Prevalence of HF is widespread in the USA with 5.7 million people diagnosed; projections estimate an increase past 8 million by 2030.¹ The economic cost of HF was estimated to be \$31 billion in 2012.¹ More than 40% of the cost is attributed to hospitalization, with volume retention and overload being the most common reason for HF hospitalization. According to a study of more than 1.3 million HF patients, 24.8% of patients were readmitted within 30 days of

*Address all correspondence to Shona Stewart, stewart@chemimage.com

discharge.² The American Heart Association has launched a program called Rise Above Heart Failure that assists healthcare professionals in educating patients concerning heart failure.³ In addition to patient education, there is a need to better understand the disease state and monitor patients both inside and outside a clinical setting in order to reduce hospital length of stay and patient readmission rate, improve quality of life for HF patients, and reduce the economic burden of HF.

HF is diagnosed based on a patient's signs and symptoms, past medical history, family history, and test results, including blood tests and imaging. Fluid retention is one of the cardinal manifestations of HF and may accumulate in the lungs, abdomen, and extremities. Peripheral edema describes interstitial fluid volume overload in the lower extremities and is a sign of hypervolemia and progression of disease. Edema occurs when an excessive volume of fluid accumulates in the tissues, either within cells (cellular edema) or within the collagen-mucopolysaccharide matrix distributed in the interstitial spaces (interstitial edema).

The traditional clinical assessment of peripheral edema in HF patients, which is still widely used, is subjective. In this examination, a clinician applies pressure with his or her index finger to an anatomical point, usually the patient's ankle or mid-tibia. The anatomical point is subjectively chosen to represent the location of greatest edema. This technique captures pit depth and the time needed for the skin to return to its original appearance (recovery time) as a single edema score ranging from 0 to 4+.⁴ This measurement will vary upon the experience of the clinician, time constraints, and individual patient factors.⁵ Alternative, non-invasive measurements have been investigated such as ankle circumference, but these methods lack reliability, feasibility, or correlation with the classic clinical assessment of pitting edema.⁵ Assessment of edema in the upper extremities (forearm) is not practiced clinically. Measurement of daily weight as a surrogate for volume retention is fraught with inaccuracies and is clinically unreliable.⁶

Additional methods to track patient progress including thoracic impedance measurements, pulmonary lung fluid tracking, pulse oximetry, and blood testing for natriuretic peptides are in various stages of commercial development.⁷⁻¹⁰ A non-invasive and non-contact monitoring option, such as imaging, would lead to reduced costs, reduced hospital readmission rate, and greater patient compliance. Invasive implantable devices, such as CardioMEMS, have shown a reduction in hospitalization rate compared to control groups.¹¹ The implantable devices routinely monitor pulmonary artery pressure and edema but not peripheral edema. There is currently no commercially available imaging modality that could accurately measure and monitor peripheral edema in HF patients.

This study explores the use of short-wave infrared (SWIR) molecular chemical imaging (MCI) to non-invasively and objectively measure the level of edema in a patient's limb. MCI has been explored for several medical applications due to the non-contact, non-destructive nature of the technique and use of non-damaging illumination that can be used to obtain multiple measurements from a given patient as needed.¹² These characteristics are useful in medical applications in which invasive tests can lead to patient discomfort or non-compliance with physician testing schedules. SWIR MCI is potentially suitable for quantitatively measuring subcutaneous interstitial fluid due to water being a strong absorber in the SWIR spectral range. We expect to see evidence of higher water absorption in SWIR spectra from subjects with edema than from subjects without edema. Additionally, the SWIR wavelength range exhibits decreased scattering when compared with the more conventionally exploited near-infrared (NIR) range. Furthermore, the longer wavelengths in the SWIR allow deeper penetration of light into tissues.^{13,14} The penetration depth with SWIR will allow the measurement of patients with varying skin thickness, such as obese patients, and has relative insensitivity to skin pigmentation.¹⁵ The epidermis layer varies by body location, averaging 0.1 to 0.15 mm in thickness; the dermis layer can be 1.5 to 4 mm; and the subcutaneous layer is thicker than the dermis.¹⁶ The study described here serves as an opportunity to establish technical feasibility and inform future directions for the potential development of a clinical instrument. The objectives were to investigate a primary endpoint of discrimination between healthy and HF patients with edema and a second primary endpoint of whether a SWIR-based index [the CardioVerification Index (CVI)] could be conceived and used to predict the pitting edema level using the ground truth provided by the traditional pitting edema test.

2 Methods

2.1 Study Design

Patient enrollment was conducted in two cohorts, with each cohort centered at a single site. Twenty healthy volunteers were enrolled in the first cohort with no presence of edema and no other excluding criteria present, as determined by clinical personnel. Patients at Allegheny General Hospital (Pittsburgh, PA) who were under care of the principal investigator and had a documented HF diagnosis were informed of the study at an inpatient or outpatient setting. Patients were consented to the second cohort if they agreed to participate in the study. This observational study included one data collection for each enrolled patient. Forty-seven HF patients were enrolled. Data sets that did not meet the quality standards were not included, and 28 HF patients are represented in the reported results.

The edema patient population that met the quality standards was 58.3% male, 89% Caucasian, and 50% diabetic with an average age of 70.2 years. Demographics in the healthy population were closely matched in gender and ethnicity with 63% male and 95% Caucasian. The healthy population analysis pool had an average age of 39.3 years.

2.2 Instrument

A tower-based sensor was used to facilitate measurements of patients with limited or no mobility. Figure 1 shows the imaging system and major components. The sensor head unit (SHU) used in this study has been previously described in a benchtop configuration that was used to detect illegal drugs in correctional facility mail.¹⁷ The benchtop configuration was impractical for imaging limbs in the present study, therefore, the sensor was incorporated into the tower configuration to allow for movement of the instrument between data collection locations. Briefly, this sensor collects hyperspectral images via a SWIR multiconjugate filter (Gen 5 SWIR multiconjugate filter, ChemImage Corp., Pittsburgh, PA) that filters light, which is subsequently captured by the detector, a SWIR InGaAs focal plane array (TAU SWIR camera, FLIR, Santa Barbara, CA). This SHU is attached to an articulating arm, which enables it to be positioned over patient limbs, normal to each limb, at an approximate working distance of 70 cm (Fig. 2). A light housing has been attached to the SHU and is equipped with six 20 W halogen light bulbs (MR-16 EuroStar-Reflecko #1000010, Ushio, Tokyo, Japan) for illumination of the 163 mm × 124 mm field of view. Motorized focus optics were used by the instrument operator to maintain focus between limbs and patients. Once positioned, the operator controls data collection via Spectral Kitchen[®] software (ChemImage Corp., Pittsburgh, PA), which runs on the workstation that has been integrated onto the tower.

Daily operational qualification testing was completed to confirm that the spectral and imaging quality of the instrument was suitable for collection. Instrumentation used for this study will be useful in future studies; however, a smaller, handheld sensor will allow for increased ease

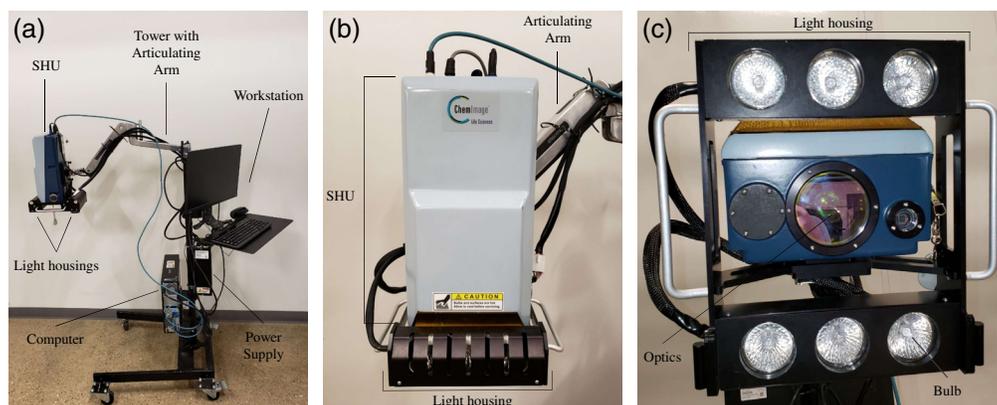


Fig. 1 (a) Tower-based SWIR MCI instrument with main components. (b) Close-up view of the SHU attached to the articulating arm. (c) View of the lighting and optical components on the SHU.



Fig. 2 SWIR MCI of a healthy volunteer shin. Future form factors are envisioned to not require a tower and will be handheld.

of use among intended use demographics. The utility of SWIR sensors is highlighted by effective use in biological and non-biological applications.

2.3 Depth Penetration

Pitting edema is typically graded up to 10 mm using the conventional clinical method, so measurement of edema in deep tissue is ideal. We have modeled the noise equivalent depth of detection (NEDoD) of pitting edema using SWIR imaging, and estimate that at 1160 nm, depth of detection in normal tissue can be as deep as 9.8 mm, whereas in tissue containing edema, the depth of detection is slightly lower at 8.1 mm. In order to quantify depth sensitivity in SWIR imaging, we modeled sensitivity matrices ($\Delta R/\Delta\mu_a$, where R is reflectance and μ_a is the absorption coefficient) in healthy and edema-containing shin tissue using an optical tomography software program, NIRFAST.¹⁸ NEDoD values were quantified by determining a cut off depth in the sensitivity matrices with respect to the instrument noise. Realistic noise was utilized to calculate NEDoD values by extracting noise from 1 h drift measurements acquired on our instrument. Increased optical absorption was modeled in edema-containing shin tissue with $\sim 70\%$ increased water content versus healthy tissue.^{19,20} Two-fold lower optical scattering was modeled in edema-containing tissue to represent dilution of the number density of scatterers in the presence of excess blood and water.¹⁹

2.4 Data Collection

Protocols and consent forms were developed by the investigators with review and approval by the Western Institutional Review Board. Prospective patients were screened against criteria defined in Table 1. Enrolled patients were assigned a deidentified study ID and metadata was collected. Pitting edema level in each shin was measured by clinical staff ~ 10 min prior to measurements. Compression stockings, if applicable, were removed followed by a minimum 5-min recumbent period before data collection. Data collection involved an RGB image and a series of SWIR images covering 1000 to 1700 nm. For a given field of view, individual wavelength frames were captured at 5 nm intervals across this spectral range, with an exposure time of 32 ms. The set of wavelength frames is referred to as a hypercube, a three dimensional stack of images, where X and Y are spatial dimensions and the Z dimension represents the spectral dimension. Therefore, each pixel location in a hypercube has an individual spectrum. Collection of each hypercube takes about 50 s.

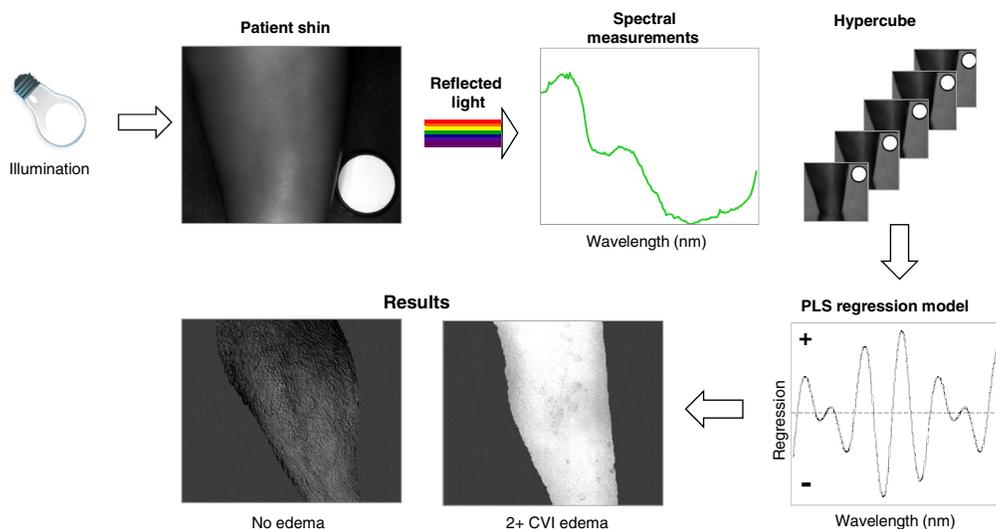
Table 1 Criteria used to screen patients for enrollment.

No.	Item	Criteria status
1	Males or females age 18 to 90 years	Inclusion (all cohorts)
2	Evidence from history, physical examination, imaging studies, and laboratory tests that support the diagnosis of HF	<ul style="list-style-type: none"> • Inclusion for HF cohort • Exclusion for healthy volunteers
3	Subjects who are incapable of providing consent may be non-compliant or whose care could be compromised	Exclusion (all cohorts)
4	Subjects who have a confirmed or suspected pregnancy	Exclusion (all cohorts)
5	Subjects with overt skin lesions	Exclusion (all cohorts)
6	Patients who have peripheral edema caused by conditions unrelated to HF	Exclusion (all cohorts)
7	Patients with a diagnosis of diabetes (type I or type II)	<ul style="list-style-type: none"> • Exclusion for healthy volunteers • Not an exclusion for HF cohort

Each patient had both shins and forearms imaged, and all data that met quality standards were included in the analysis. A typical data collection for all four patient limbs was 15 min or less. Data were collected using ChemImage-developed Spectral Kitchen[®] software. Periodic reviews of enrollment were conducted by clinical staff, and de-identified metadata were shared with the sponsor. MCI technology is used for diffuse reflectance measurements where reflected light is measured from a patient's limb, and data are recorded at specified wavelengths in a hypercube as described in Fig. 3.

2.5 Analysis

The data collection and analysis steps are illustrated in Fig. 3. The data collection starts with the illumination of the patient's shin or forearm. A molecular chemical image or hypercube, comprising an image frame at each wavelength interrogated is subsequently collected. Spectra corresponding to soft tissue are extracted from the hypercubes and applied to multivariate statistical and classification methods, such as partial least squares regression (PLSR). The PLSR model generates, from the spectra, regression vectors that are used to predict CVI values for any given spectrum (pixel). The two output images ("no edema" and "2+ CVI edema") illustrate the fact

**Fig. 3** Overview of MCI data collection and processing.

that the PLSR model can be used to calculate CVI values for every spectrum (pixel) in an input hypercube, thereby outputting a PLS score image.

2.5.1 Preprocessing

The impact of non-uniform illumination was reduced by dividing patient hypercubes by an image of a 99% reflectance standard (Gigahertz-Optik, Amesbury, MA) filling the field of view. Instrument response was additionally accounted for by performing a local background division of a reflectance standard placed in the field of view with the patient limb. Processing steps were carried out using ChemImage Xpert[®] software (ChemImage, Corp., Pittsburgh, PA).

2.5.2 Data extraction

Ten average spectra were extracted from SWIR hypercubes by selecting regions of interest (ROIs) comprising 462 pixels per ROI. In the shin hypercubes, five ROIs on either side of the tibia were sampled. The tibia was avoided because the bone is more reflective than the soft tissue and thereby distorts the signal associated with edema. Forearm average spectra were sampled with five ROIs on each side of the limb. The spectral extraction was enabled by a graphical tool in ChemImage Xpert that extracts the average spectrum associated with a given ROI drawn by a user on an image frame associated with a hypercube in the GUI. The software supports exporting the spectra to a spectral file that can then be used in downstream processing.

2.5.3 Supervised classification models for discrimination analysis to support endpoint 1

Partial least squares discriminant analysis (PLS-DA) is a statistical method of analysis employed for data reduction, model creation, and data classification.²¹⁻²³ This supervised classification method was performed in MATLAB with cross validation and was used to create a two-class model discriminating between spectra from healthy volunteers and those from HF patients exhibiting peripheral edema. PLS-DA is a binary classification method and the results presented consider shins and forearms to be separate populations for analysis. All shins are used in the development of the shin model. However, leave-one-out cross validation is used so the prediction of CVI values for a given shin does not include the spectra in that shin. In essence, a separate model is generated and used for each patient shin. The same process is followed for the forearm models.

The performance of a PLS-DA model is assessed through a receiver operating characteristic (ROC) curve analysis. A ROC curve is a plot showing sensitivity versus 1-specificity of a test for a binary system. The area under the ROC curve (AUC) is a measure that is often used to compare the performance of ROC curves using a single value. An AUC of 1 indicates 100% classification accuracy of data.

2.5.4 Supervised quantitative analysis to support endpoint 2

PLSR is closely related to PLS-DA and was also performed in MATLAB. Both algorithms share the same PLS approach that reduces the dimensionality of the data space from n dimensions (the number of wavelengths) to k dimensions (the number of selected basis vectors) in a manner that preserves the maximum amount of variance from both the X -block (spectral data) and the Y -block (edema score information).²⁴ PLSR, like PLS-DA, is a supervised method that requires a set of training data consisting of spectra and edema scores to result in a mathematical model that allows the prediction of edema scores for data not included in the training process. PLSR is a regression method that generates real values corresponding to non-integer assignments of edema scores. The values obtained from PLSR analysis are termed the CVI, which are compared to the pitting edema ground truth by rounding a CVI score to the nearest whole integer.

The PLSR approach used in this work utilizes a hierarchy of models. The first model discriminates between pitting edema levels of 0 (no evidence of edema) and all other levels of pitting edema. The second model discriminates between edema levels of 1+ versus the

combination of 2+, 3+, and 4+. The third model discriminates 4+ from a combination of 2+ and 3+. The final model discriminates between pitting edema levels of 2+ and 3+. The forearm population available for this study did not have 4+ data sets, so the third model was not needed for that population. The PLSR models were built using cross validation and were evaluated using accuracy as the figure of merit.

3 Results

Figure 4(a) displays example shin and forearm RGB images from a patient with edema and a volunteer without edema, whereas Fig. 4(b) shows a spectral image at 1200 nm indicating ROIs from which spectra were sampled. The images in Fig. 4(b) are displayed at a consistent wavelength between limbs using 1200 nm; however, any of the evaluated SWIR wavelengths could be used to illustrate the data. Figure 4(c) shows example processed images that may reveal features, such as bone or vascular locations not obvious in RGB or raw spectral images. Spectra were extracted from soft tissue regions to eliminate influence from the tibia, which is more reflective. The extracted regions represented several areas of the shin to account for local edema changes and variability of molecular species. Figure 4(d) displays representative reflectance spectra from the ROIs for both shins and forearms.

The principal macromolecules represented in the SWIR spectra are water, lipids, and collagen.²⁵ Lipids, and to a lesser extent, collagen and water, absorb between ~1120 and 1230 nm.^{25,26} The peaks in this spectral region are associated with the second overtone of C–H stretching in lipids and collagen and a vibrational overtone of the O–H bond in water. The larger envelope from 1300 to 1650 nm encompasses mostly collagen and water, peaking at ~1480 nm, and some lipid, absorbing closer to 1400 nm.^{25,26} The presence of collagen is reflected in the absorption peaking at ~1500 nm, representing a combination of CH₂ stretching and non-stretching modes. Water and lipids are represented by the first overtone of OH stretching, closer to 1430 nm.

A portion of collected data was not suitable for analysis due to incorrect wavelength step size (wavelength distance between collected images in a hypercube), file save failures, file corruptions, or protocol issues such as the absence of the flatfielding correction standard (99% reflectance standard disc) or incorrect standoff distance. In total, 27 of 94 shins were removed from the HF population, corresponding to 11 subjects in which both shins were excluded, and 5 subjects in which one shin was excluded. Similarly, 12 forearm data sets were found to be not evaluable. The removal of data due to quality included one volunteer in the healthy population who had an edema level other than 0 and another with a diagnosis of diabetes.

The primary analysis endpoints for the study were to (1) resolve if SWIR MCI can accurately discriminate between healthy and edema patients and (2) investigate if SWIR MCI can accurately discriminate amongst all pitting edema levels. Acceptance criteria for each endpoint were established as an accuracy of 80% or greater. Two-class PLS-DA models were built to discriminate between healthy patients without pitting edema and HF patients with all levels of pitting edema. The distribution of edema in evaluable limbs is summarized in Table 2. Pitting edema ground truth level was measured in the shin area as part of the routine care of HF patients. Forearm pitting edema was not collected as part of this pilot study, and the applicable shin ground truth was applied to the forearm data. The forearm population was selected using corresponding forearm data from the final shin population. Some data sets from the forearm population did not meet the data quality requirements outlined above and were removed from the evaluated data.

The discrimination plots for the shin and forearm models are shown in Figs. 5(a) and 5(b). The ROC curves, shown in Fig. 5(c), correspond to an AUC of 0.993, accuracy of 97.1%, sensitivity of 97.0%, and specificity of 97.2% for the shin population. The test for forearm edema had an AUC of 0.993, accuracy of 96.7%, sensitivity of 98.2%, and specificity of 94.4%. Figure 5(d) displays the average spectra from the edema and healthy population shins and forearms. These results are encouraging, and suggested additional analysis be undertaken to understand if discrimination between levels of pitting edema is feasible.

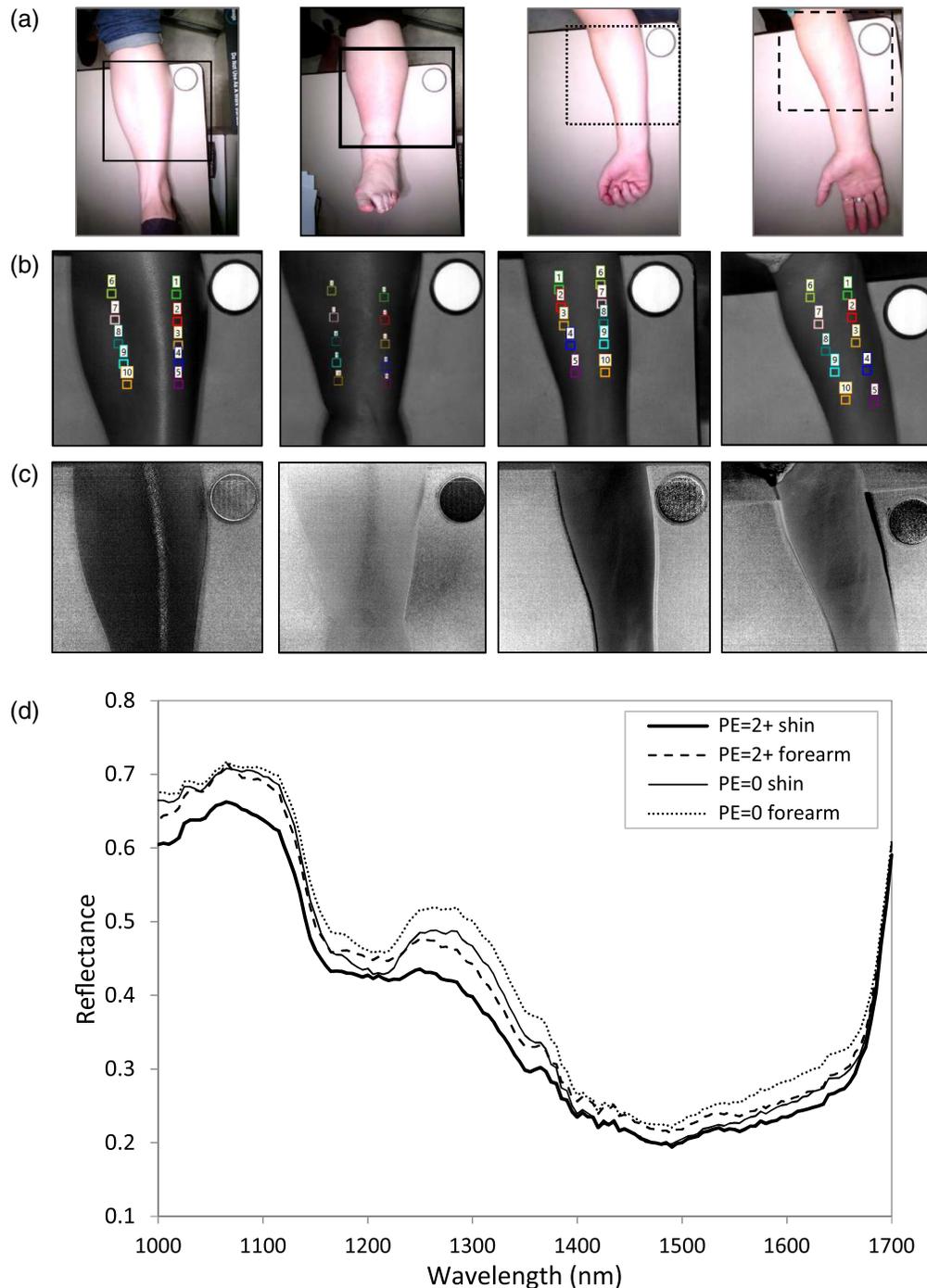


Fig. 4 Example data from a healthy volunteer and an edema patient. (a) RGBs of a healthy PE = 0 volunteer shin (left), 2+ edema shin (middle left), healthy PE = 0 volunteer forearm (middle right) and a 2+ edema forearm (right). Areas sampled for SWIR analysis indicated by blue boundaries. (b) Spectral images at 1200 nm and ROIs were used for spectral extraction. (c) Processed images displaying bone and soft tissue regions in shin and vasculature in forearm images. (d) Average spectra from the ROIs.

The second primary endpoint was investigated using hierarchical PLSR to build a model using extracted spectra and the ground truth edema level as the dependent variable. The shin model generated CVI scores that predict healthy and HF patient edema levels with an accuracy of 81.6% when compared to the ground truth. This model was then applied to shin hypercubes to generate score images that enabled visualization of edema in participant shins. Figure 6(a) shows

Table 2 Distribution of edema in patients.

No.	Pitting edema level	Number of evaluable shins	Number of evaluable forearms
1	0/absence of edema	36	36
2	1+	10	8
3	2+	31	27
4	3+	22	20
5	4+	4	0

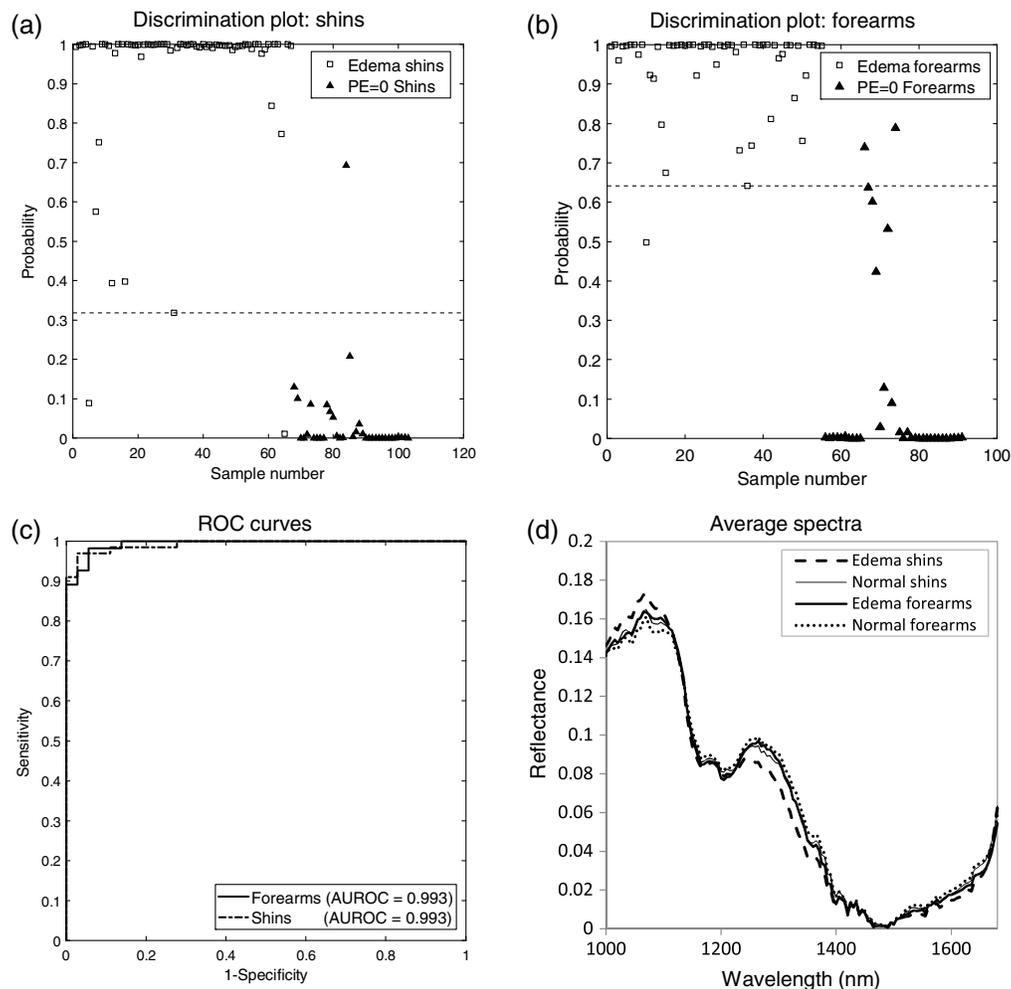


Fig. 5 Differentiation of healthy and edema populations using PLS-DA. (a) Discrimination plot for shin data. Shin data are classified as edema for values above the dotted threshold. (b) Discrimination plot for forearm data. Forearm data are classified as edema for values above the dotted threshold. (c) ROC curves for two class PLS-DA models discriminating healthy patients from HF patients with peripheral edema. (d) Average reflectance spectra for edema and healthy shins and forearms used in models.

box plots of predicted pitting edema level versus the provided ground truth with the standard deviation represented as a box. The red dot indicates the median, the center line in the box is the mean, and the bottom and top edges of the box represent the 25th and 75th percentiles, respectively. The “whiskers” around the box plot represent the extremes of the distribution that are not

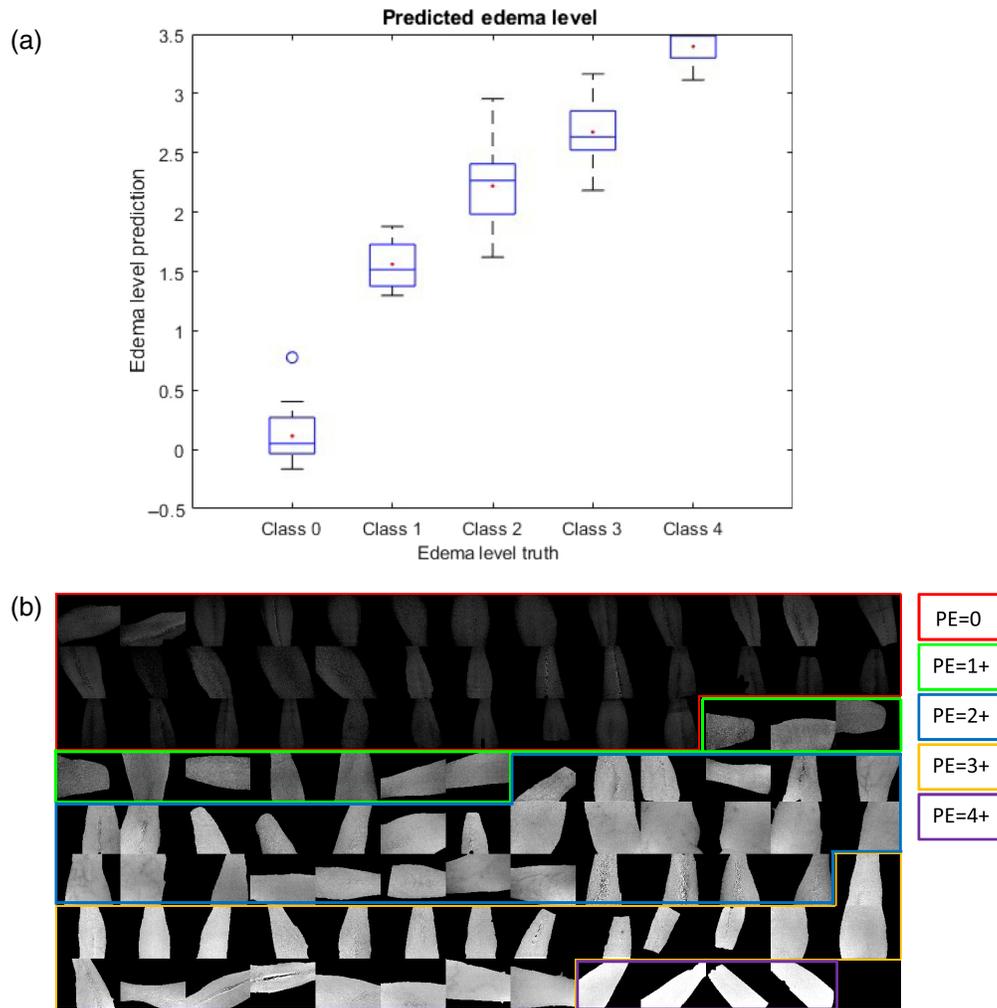


Fig. 6 (a) Box plot for hierarchical models comparing the predicted shin edema level using CVI scores and ground truth provided by clinical staff. (b) Output images for the hierarchical PLSR model. Shins are displayed in row-order of increasing ground truth edema level. Shins 1 to 36 starting in upper left corner (PE = 0 ground truth; red box); shins 37 to 46 (PE = 1+ ground truth; green box); shins 47 to 60 (PE = 2+ ground truth; blue box); shins 61 to 78 (PE = 3+ ground truth; yellow box); and shins 79 to 80 (PE = 4+ ground truth; purple box). The overall classification accuracy for this model is 86%.

considered outliers and the outliers are represented by open and closed circles, where open circles are outliers and closed circles are extreme outliers. Figure 6(b) comprises the PLSR score images generated when this model is applied to subject hypercubes. Each pixel in the image is generated by multiplying the spectrum for that pixel by the regression vector that is the output of the PLSR model building exercise. By processing every pixel, a score image for that model is assembled. The average intensity of the pixels, in general, increases as the predicted edema level increases. The same methodology was applied to the forearm population as shown in Fig. 7. A hierarchical PLSR model using forearm data discriminated edema levels with an accuracy of 63.8%.

4 Discussion

In this pilot study, the primary endpoint of a non-invasive and non-contact SWIR MCI-based approach to discriminate between limbs with and without edema was achieved as judged by the high accuracy of the two-class PLS-DA model. The second primary analysis endpoint of

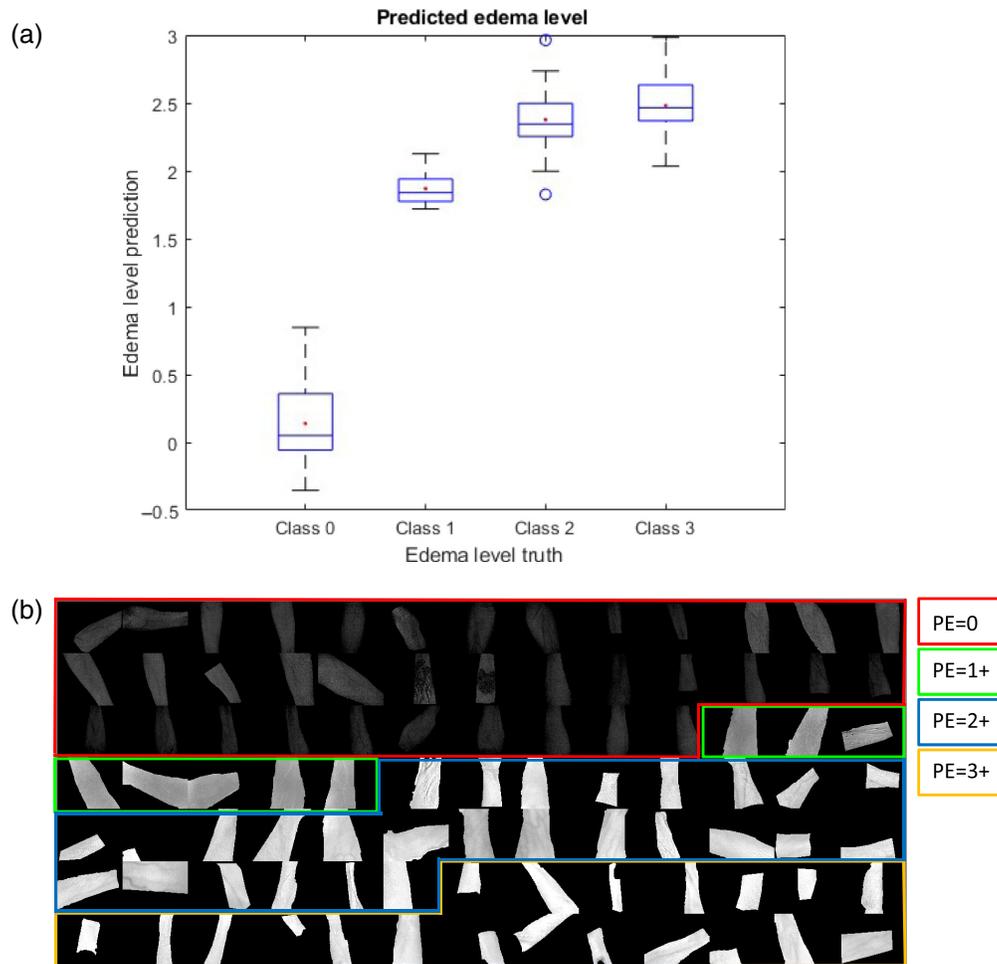


Fig. 7 (a) Box plots for hierarchical models comparing the predicted edema level using CVI scores and ground truth provided by clinical staff. (b) Example output images for the hierarchical PLSR model. Forearms are displayed in row-order of increasing ground truth edema level. Forearms 1 to 36 starting in upper left corner (PE = 0 ground truth; red box); forearms 37 to 46 (PE = 1+ ground truth; green box); forearms 47 to 58 (PE = 2+ ground truth; blue box); forearms 59 to 76 (PE = 3+ ground truth; yellow box). The overall classification accuracy for this model is 71.1%.

predicting edema level in patients gave encouraging results, but a larger study is needed to obtain and train on greater populations for all edema levels. The majority of biological tissue-based spectroscopic studies have focused on the measurement of chromophores such as hemoglobin, which absorb in the visible wavelength range.^{27–29} However, the SWIR region offers benefits that have not been exploited in high volume until more recently, due to recent developments in SWIR-based optics, e.g., InGaAs sensors.¹⁵ First, the SWIR spectral region is highly favorable for thicker samples due to the longer wavelengths and transparency of the region to blood. Additionally, the absorption coefficients of key chromophores: water, lipids, and collagen are higher in the SWIR region compared with the visible and NIR regions. For example, the absorption coefficient of water is more than 60 times higher at 1440 nm than at the 975-nm NIR peak.³⁰ In a similar trend, lipid absorbs many magnitudes more at 1210 nm than at 930 nm.^{28,29}

The basis of discrimination among the populations can be explored by comparing spectra from the healthy and edema populations. Figures 8(a) and 8(b) show average normal and edema reflectance spectra representing the shin and forearm data, respectively, used to generate the two PLS-DA models. In the key region of 1120 to 1230 nm, it is clear that absorption is higher in spectra from subjects with edema in both shins and forearms. We can exploit spectral analysis further by evaluating the variable importance in projection (VIP) scores from the model. Shin and forearm VIP scores are shown alongside average model spectra in Figs. 8(a) and 8(b),

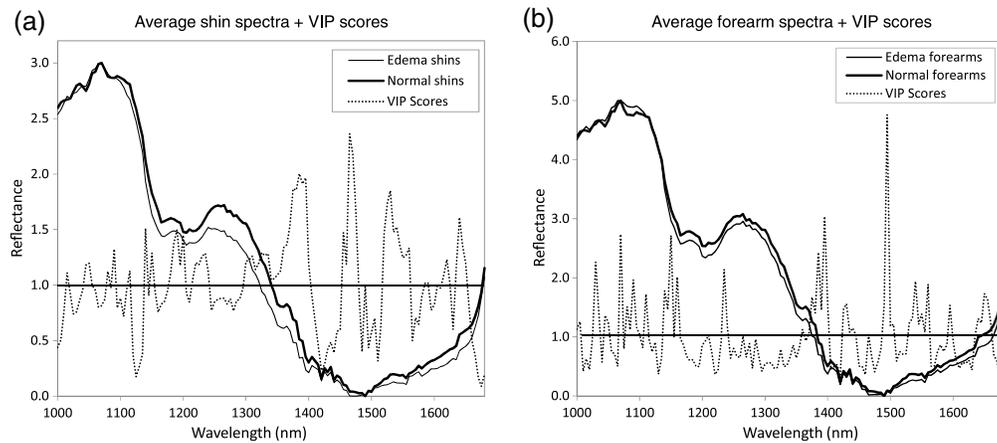


Fig. 8 (a) Average shin reflectance spectra and VIP scores calculated from the PLS-DA model for edema in shins. (b) Average forearm reflectance spectra and VIP scores calculated from the PLS-DA model for edema in forearms.

respectively. A VIP score is a measure of a variable's importance in a PLS model. VIP scores can be used for variable selection and for identifying the spectral regions of importance in a model.³¹ A variable (in this case, wavelength) with a VIP score >1 can be considered important in a given model. The VIP scores associated with the PLS-DA shin model indicate the spectral regions that play a significant part in discriminating between healthy subjects and those with edema, regardless of grade. Several spectral locations with VIP scores >1 are evident. As expected, these locations represent all three macromolecules—lipids, collagen, and water. VIP scores higher than 1 at 1140 and 1150 nm reflect water content; 1200 and 1235 nm correspond to lipid and collagen absorption. However, the highest VIP scores are at 1385, 1465, and 1530 nm. This broad spectral envelope corresponds to the absorption of mostly water and collagen, and lipids to a lesser extent. At 1385 and 1465 nm, OH stretching likely causes the observed absorption, and at 1530 nm, we expect there to be water influence as well as contribution from CH₂ stretching and non-stretching vibrations in collagen. The forearm VIP scores are very similar to the shin VIP scores.

What these analyses tell us is that not only are edema and normal shins and forearms distinguishable using SWIR MCI, the basis of discrimination is water content in addition to collagen and lipid differences. Because subjects with high levels of edema can have as high as a 70% increase in water content,^{19,20} we estimate that water differences impart a high bearing on differentiating between edema and normal tissue spectra.

Among other collected metadata, the influence of diet and medications will be monitored to determine effects on results during treatment. Alternate prediction strategies, such as artificial intelligence (AI), may be employed when a larger data set is available with greater patient populations. Currently, PLSR is a good choice for model building, given the limited and the spectral nature of the data. Deep neural networks are often chosen for machine learning and AI but are a better choice when the data consists of large training sets of images. Certain AI tools, such as data augmentation, will be investigated via spectral additions of water to current samples to increase the training set coverage and provide more robust models. Greatly increased data collections for an individual patient when used in a monitoring mode could allow for adaptive background estimation which could better model the spectral background. Next generation handheld sensors will incorporate more advanced computer vision techniques for standoff distance and angle calculations as well as for motion correction.

The prediction accuracy of the shin model is greater than that of the forearm model. The ground truth for a given shin was used as ground truth for the corresponding forearm. This may have influenced the degradation of forearm accuracy, as shin edema may be more severe due to gravity. The effect of sampling location will continue to be investigated, as flexibility in this area will increase clinical utility for patients with physical impairments, morbidly obese populations, and limited mobility patients. Evaluating the forearm, for example, which is not currently used to

measure pitting edema, allows us to assess the upper extremity as an option for edema measurement, which would give us flexibility when designing future form factors.

Subsequent clinical trials will, when possible, include a more diverse population in terms of ethnicity, pitting edema level, and age to better capture biological variability and increase the robustness of models.

In an additional experiment, individual two-class PLS-DA models were built from subjects' right and left shin measurements. Using individual shin pitting edema ground truth, the accuracy of the right shin model was higher than that of the left side shin model (100% versus 96.1%, respectively; data not shown). Although not currently statistically different, we will continue to monitor this trend in future clinical trials to understand if sampling from multiple limbs is necessary, and to understand optimal sampling location. Any metadata in a patient's history that may influence the results, such as a heart bypass operation or which veins were harvested from a patient's leg for bypass surgery, will be monitored.

The feasibility of monitoring dynamic peripheral edema levels in patients over time, including patients who are receiving treatment, will be investigated in upcoming clinical trials. Ultimately, these trials will help understand if MCI technology used consistently on a periodic basis can reduce readmission rates for HF patients. Efforts are underway to develop a hand-held form factor that will automate features where possible to reduce the training burden. These efforts are being conducted by CardioVere, a spinout of ChemImage. The eventual goal is to develop strategies to automatically integrate MCI results into the electronic health record to alert clinicians when a patient has a condition change and requires clinical intervention.

We acknowledge some of the limitations of the study, including underrepresentation of non-Caucasian ethnicities and patients with 4+ pitting edema. Data sets from four enrolled non-Caucasian patients met data quality standards and were included in the reported models. Although we currently do not have evidence suggesting that a change in performance would occur when including a larger number of non-Caucasian patients, this will be investigated in future studies with a larger patient population. Though we only included HF patients with pitting edema, it is possible that some of the patients had non-pitting edema, which may have influenced the results. The cohort without HF had a lower average age than the cohort with HF. Skin thickness may decrease with increased age and this will be considered in future protocols. Finally, this study included forearm assessment for edema, which is typically not part of routine evaluation of HF patients. Understanding the body locations useful in the evaluation of edema will be explored in additional studies. Although this study only obtained ground truth from the shin area, the protocols for future studies are anticipated to collect ground truth from each location sampled. Importantly, the evaluation of sampling locations is not limited only to the extremities used in this study.

5 Clinical Perspectives

Tracking volume status in HF patient care is vital to prevent the progression of the disease. Accurate tracking helps to manage symptoms better by making simple lifestyle changes and modifying therapeutics. This technology may provide a reliable, objective, non-invasive, and non-contact methodology for quantitative peripheral edema measurement. As the technology matures, it is envisioned that patient self-monitoring, with wireless transmission of edema levels to physicians from home, can aid clinicians in monitoring HF patients for improved patient outcomes and to improve the future of individualized medicine. This increases the chance of modifying treatment plans appropriately based on accurate monitoring of edema level. This ultimately may reduce HF hospital readmission rates and unscheduled, emergent visits to the clinic or emergency room.

Disclosures

Authors identified as affiliated with ChemImage Corporation are employees and receive a salary from ChemImage Corporation. ChemImage employees disclose the following: ChemImage employment; ownership interest, including stock and stock options; patents; and other intellectual property.

Acknowledgments

This work was supported financially by ChemImage research and development funding, which included program development and sponsoring the clinical trial described in this report. The authors would like to thank the volunteers and patients for participating in the study and colleagues at ChemImage for instrument development and engaging discussion of project goals. The authors also acknowledge the clinical staff at Allegheny General Hospital for aid in developing regulatory documents, patient enrollment, and data collection. Enrollment and data collection support were provided by Jessica Burkardt and Laurie Machen. Regulatory support from Allegheny General Hospital was provided by Heather McDonald and Meghana Bohra. The concept for this work and data interpretation is attributed to J. Christopher Post, Jeffrey Cohen, and Patrick J. Treado. Additional study design and data interpretation input were provided by Arash Samiei and Srinivas Murali, the latter also serving as principal investigator for the clinical study. Reina Perez and Aaron Thomas provided protocol refinement and served as the primary coordinators for the clinical study. Shona Stewart, Arjun Bangalore, Heather Gomer, Marlena B. Darr, and Robert C. Schweitzer provided data analysis and interpretation; Heather Gomer also provided algorithm development. Aaron G. Smith contributed to study design, data collection, data interpretation, and initial drafting of the manuscript. Sandhya Vasudevan developed SWIR depth penetration models.

References

1. D. Mozaffarian et al., "Heart disease and stroke statistics - 2016 update: a report from the American Heart Association," *Circulation* **133**(4), e38-360 (2016).
2. R. Cheng et al., "Outcomes in patients with heart failure with preserved, borderline, and reduced ejection fraction in the Medicare population," *Am. Heart J.* **168**(5), 721-730 (2014).
3. American Heart Association (editorial staff), "Rise above heart failure toolkit for healthcare professionals," AHA, <https://www.heart.org/en/health-topics/heart-failure/heart-failure-tools-resources/rise-above-heart-failure-toolkit> (accessed 4 May 2020).
4. "Grading of edema," Med-Health, <http://www.med-health.net/Edema-Grading.html> (accessed 4 May 2020).
5. K. Brodovicz et al., "Reliability and feasibility of methods to quantitatively assess peripheral edema," *Clin. Med. Res.* **7**(1-2), 21-31 (2009).
6. P. Lynga et al., "Weight monitoring in patients with severe heart failure (WISH). A randomized controlled trial," *Eur. J. Heart Fail* **14**(4), 438-444 (2012).
7. M. Sato et al., "Bioelectrical impedance analysis in the management of heart failure in adult patients with congenital heart disease," *Congenit. Heart Dis.* **14**(2), 167-175 (2019).
8. H. Brunner-La Rocca and S. Sanders-van Wijk, "Natriuretic peptides in chronic heart failure," *Card Fail Rev.* **5**(1), 44-49 (2019).
9. E. Watanabe et al., "Prognostic importance of novel oxygen desaturation metrics in patients with heart failure and central sleep apnea," *J. Card Fail.* **23**(2), 131-137 (2017).
10. N. Uriel et al., "Relationship between noninvasive assessment of lung fluid volume and invasively measured cardiac hemodynamics," *J. Am. Heart Assoc.* **7**(22), e009175 (2018).
11. M. Givertz et al., "Pulmonary artery pressure-guided management of patients with heart failure and reduced ejection fraction," *J. Am. Coll. Cardiol.* **70**(15), 1875-1886 (2017).
12. G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *J. Biomed. Opt.* **19**(1), 010901 (2014).
13. D. C. Sordillo et al., "Short wavelength infrared optical windows for evaluation of benign and malignant tissues," *J. Biomed. Opt.* **22**(4), 045002 (2017).
14. E. Thimsen, B. Sadtler, and M. Y. Berezin, "Shortwave-infrared (SWIR) emitters for biological imaging: a review of challenges and opportunities," *Nanophotonics* **6**(5), 1043-1054 (2017).
15. H. Zhang et al., "Penetration depth of photons in biological tissues from hyperspectral imaging in shortwave infrared in transmission and reflection geometries," *J. Biomed. Opt.* **21**(12), 126006 (2016).

16. P. Kolarsick, M. Kolarsick, and C. Goodwin, "Anatomy and physiology of the skin," *J. Dermatol. Nurses Assoc.* **3**(4), 203–213 (2011).
17. R. Schweitzer et al., "Automated chemical imaging identification of illegal drugs in correctional facilities mail," *J. Chemometr.* **32**(10), e3038 (2018).
18. H. Dehghani et al., "Near infrared optical tomography using NIRFAST: algorithm for numerical model and image reconstruction," *Commun. Numer. Methods Eng.* **25**(6), 711–732 (2008).
19. J. R. Weber et al., "Multispectral imaging of tissue absorption and scattering using spatial frequency domain imaging and a computed-tomography imaging spectrometer," *J. Biomed. Opt.* **16**(1), 011015 (2011).
20. A. J. Lin et al., "Spatial frequency domain imaging of intrinsic optical property contrast in a mouse model of Alzheimer's disease," *Ann. Biomed. Eng.* **39**(4), 1349–1357 (2011).
21. H. Martens and T. Naes, *Multivariate Calibration*, John Wiley & Sons, New York (1992).
22. M. Barker and W. Rayens, "Partial least squares for discrimination," *J. Chemom.* **17**(3), 166–173 (2003).
23. S. Stewart et al., "Distinguishing between renal oncocytoma and chromophobe renal cell carcinoma using Raman molecular imaging," *J. Raman Spectrosc.* **45**(4), 274–280 (2014).
24. K. Beebe, R. Pell, and S. Seasholtz, *Chemometrics: A Practical Guide*, 1st ed., pp. 286–287, John Wiley & Sons, New York (1998).
25. R. H. Wilson et al., "Review of short-wave infrared spectroscopy and imaging methods for biological tissue characterization," *J. Biomed. Opt.* **20**(3), 030901 (2015).
26. T. J. Allen et al., "Spectroscopic photoacoustic imaging of lipid-rich plaques in the human aorta in the 740 to 1400 nm wavelength range," *J. Biomed. Opt.* **17**(6), 061209 (2012).
27. K. J. Zuzak et al., "Visible reflectance hyperspectral imaging: characterization of a noninvasive, in vivo system for determining tissue perfusion," *Anal. Chem.* **74**, 2021–2028 (2002).
28. R. Nachabe et al., "Diagnosis of breast cancer using diffuse optical spectroscopy from 500 to 1600 nm: comparison of classification methods," *J. Biomed. Opt.* **16**(8), 087010 (2011).
29. C. Tsai, J. Chen, and W. Wang, "Near-infrared absorption property of biological soft tissue constituents," *J. Med. Biol. Eng.* **21**, 7–14 (2001).
30. G. M. Hale and M. R. Querry, "Optical constants of water in the 200-nm to 200-m wavelength region," *Appl. Opt.* **12**(3), 555–563 (1973).
31. I. Chong and C. Jun, "Performance of some variable selection methods when multicollinearity is present," *Chemometr. Intell. Lab Syst.* **78**(1–2), 103–112 (2014).

Aaron G. Smith received his BS degree from Oakland University and his PhD from Texas A&M University and completed postdoctoral work at the University of Pittsburgh prior to joining ChemImage. He is a biochemist with a background in molecular biology and electron microscopy. His publications include the detection of tumors using molecular chemical imaging, stem cell development, chloroplast division, and the development of a protocol that reduces the required time for a common assay, the Western Blot.

Reina Perez worked as a tissue collection associate at AHN's Cardiovascular Research Institute. Her background is in pathology, phlebotomy, and bench research.

Aaron Thomas received his MS degree in healthcare informatics from Chatham University. He worked as a clinical research coordinator at AHN's Cardiovascular Research Institute. Now, he works as a clinical systems analyst in AHN's Division of Infectious Disease. His focus was data analytics in heart failure and pulmonary hypertension clinical trials.

Shona Stewart received her BS and PhD degrees from Queensland University of Technology and was a postdoctoral researcher at the University of Michigan. In 2002, she joined the ChemImage Biomedical R&D team. She is an analytical chemist with a background in spectroscopy. She has extensive experience in techniques including Raman, surface-enhanced Raman scattering, and diffuse reflectance spectroscopy. For more than 15 years, she has been employing hyperspectral imaging technologies to address unmet clinical needs.

Arash Samiei earned his medical degree from Tehran University of Medical Sciences and completed fellowships in the Department of Pediatrics at the University of Chicago and at the Northwestern University's Lurie Cancer Center focusing on the biology of cancer. Currently, he is a clinical and translational researcher at ChemImage, supporting development of new imaging technology aimed at real-time detection of cancer and anatomic structures.

Arjun Bangalore received his MSc degree from the Central College, Bangalore, India, his PhD from Ohio University, and he was a postdoctoral researcher at the Argonne National Laboratory. He is an analytical chemist with a background in data analysis in remote sensing, and his formal training is in chemometrics. He is experienced in R&D, evaluation, and troubleshooting of instrument components, development of application solutions, prototyping innovative analysis pathways, developing custom software for specific analysis problems, and standalone software solutions.

Heather Gomer earned her BS and MS degrees in electrical and computer engineering from Carnegie Mellon University. She expanded her signal and image processing experience while at Raytheon BBN Technologies by developing detection and classification algorithms. She is a principal algorithm scientist with a background in signal and image processing. Her focus, more than the last 11 years at ChemImage, has been algorithm development and data analysis of spectral- and imaging-based research efforts.

Marlena B. Darr attended Carnegie Mellon University where she received her BS degree in biological sciences from Mellon College of Science. She is a biologist with a background in cell biology. Currently, she is a scientist II in ChemImage's Life Sciences group where she supports several ongoing research efforts, including those related to the development of an intra-operative imaging tool that will provide information to surgeons in real time, without the use of reagents.

Robert C. Schweitzer received his undergraduate degrees in chemistry and computer science and his PhD in analytical chemistry from Ohio University. He is a lead algorithm scientist with a background in chemometrics and programming. His work at ChemImage over the last 21 years has included chemical image analysis, empirical model building, and a variety of additional algorithm topics.

Sandhya Vasudevan received her PhD in electrical engineering from the University of Notre Dame, Indiana. She is a senior algorithm scientist at ChemImage with a background in computational tissue optics, study of light tissue interactions, spectroscopy, as well as modeling of light propagation in tissue. Her doctoral research focused on the role of broadband near-infrared spectroscopy in breast cancer diagnosis.

Jeffrey Cohen is the chief executive officer at ChemImage, the chief physician executive, Community Health and Innovation at Allegheny Health Network, president of Triangle Urology Group, and associate professor of surgery at Drexel University/Allegheny General Hospital.

J. Christopher Post received his medical degree from the University of Florida in 1983 and his PhD in human genetics from the Graduate School of Public Health at the University of Pittsburgh. He completed two years of a pediatric residency at the Massachusetts General Hospital in Boston and a fellowship in pediatric otolaryngology at the Children's Hospital of Pittsburgh. He is the chief medical officer at ChemImage and responsible for the clinical implementation of ChemImage technologies.

Srinivas Murali is an internationally recognized cardiologist and system chairman of the Department of Cardiovascular Medicine and the co-chairman of the Cardiovascular Institute at Allegheny Health Network. He is a professor of medicine at Drexel University College of Medicine. He has authored more than 130 peer-reviewed publications, given around 240 presentations, and served as a clinical investigator in 120 clinical trials in the field of heart failure, pulmonary hypertension, mechanical circulatory support, and cardiac transplantation.

Patrick J. Treado is an analytical chemist who has pioneered the field of chemical imaging. He founded ChemImage in 1994 while an assistant professor at the University of Pittsburgh. His research interests involve the development of vis-NIR, IR, and fluorescence (i.e., molecular) chemical imaging technology, and its application to materials analysis, clinical diagnostics, and process monitoring. He is the author of more than 150 publications and 118 patents on chemical imaging technology, methodology, and applications.