

# Earlier detection of breast cancer in blood using Raman spectroscopy

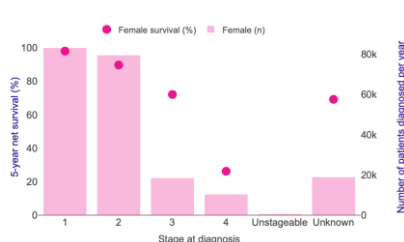
Imran Mohamed<sup>1,2</sup>, Zoe Barber<sup>4</sup>, Kelvin Gomez<sup>3</sup>, Peter Dunstan<sup>1</sup>, Tom Wilkinson<sup>1</sup>, Dean Harris<sup>1,2</sup>

<sup>1</sup>Swansea University, <sup>2</sup>Swansea Bay University Health Board, <sup>3</sup>Aneurin Bevan University Health Board, <sup>4</sup>Cwm Taf Morgannwg University Health Board

## Background

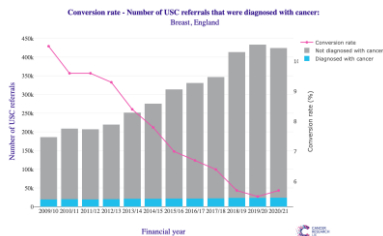
- Breast cancer is the most common cancer in the UK
- Management is through gold standard triple assessment
- Early diagnosis is key to improving survival

Breast cancer five-year net survival by stage, with incidence by stage (all data: adults diagnosed 2013-2017, followed up to 2018)



### Clinical challenge

- Conversion rate of urgent suspected cancers is 5.7%
- Conversion rate of urgent breast symptom referrals is 1.3%
- Bottleneck for early detection and management



## Aim

An exploratory multi-site prospective non-interventional observational cohort study in symptomatic patients referred to tertiary breast services with suspected cancer was conducted to determine Raman spectral differences between invasive cancer, DCIS and control samples

## Inclusion Criteria

- Female patients aged 18 years or over
- Confirmed histopathological diagnosis of breast cancer OR benign outcome after assessment from triple assessment clinic
- Participants had fasted serum sample (15ml) taken

## Methods

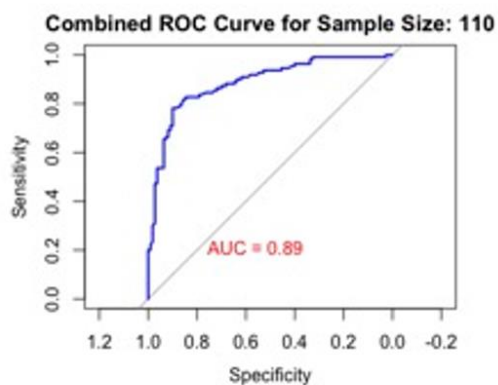
Prospective cohort pilot observation study across 3 NHS Wales Sites: SBUHB, ABUHB, CTMUHB

Serum samples were analysed using Raman spectroscopy

A random forest model was built to explore the ability of machine learning to accurately classify cancer or control cases.

## Results

- Data presented are taken from two groups: cancer and control.
- Cancer group: 176 patients (152 invasive cancer, 23 DCIS); median age 64.7years (range 38.6–92.6years), median BMI 30 (range 13–82).
- Control group: 116 patients; median age 44.3years (range 18.6–83.4years), median BMI 26 (range 18–60).
- The best performing random forest model included 110 samples (AUC 0.89, accuracy 0.81, sensitivity 0.84, specificity 0.78).



sample size	accuracy	sensitivity	specificity	AUC
110	0.81	0.84	0.78	0.89

## Outcomes

- There is good evidence to support that Raman spectroscopy can be used to interrogate biological samples to provide spectra for interpretation.
- Building the random forest model involves a complex interplay between sample composition, model variability, and the inherent randomness in the training process.
- A larger sample size may improve the model's performance.