



# BMJ Open Multicancer early detection in a cohort of patients with confirmed and suspected cancer by measuring plasma amino acid cross sections with the Enlighten test: MODERNISED protocol

Sam Wilding <sup>1</sup>, Victoria Goss,<sup>1</sup> Wesley Sukdao,<sup>2</sup> Zaed Hamady,<sup>3</sup> Joanne Lord,<sup>4</sup> Adam Coleman,<sup>5</sup> Catherine Pointer,<sup>5</sup> Jocelyn Walters,<sup>1</sup> Will Herbert,<sup>1</sup> Katy McLaughlin,<sup>1</sup> Robert Waugh,<sup>1</sup> Nicola Irvine,<sup>5</sup> Tom Oliver,<sup>1</sup> Irene Soulsby <sup>6</sup>, Julie Hooper,<sup>6</sup> Simon J Crabb,<sup>1</sup> Gareth Griffiths,<sup>1</sup> Emma Yates,<sup>2</sup> Andrew Davies<sup>5</sup>

**To cite:** Wilding S, Goss V, Sukdao W, *et al.* Multicancer early detection in a cohort of patients with confirmed and suspected cancer by measuring plasma amino acid cross sections with the Enlighten test: MODERNISED protocol. *BMJ Open* 2025;15:e108220. doi:10.1136/bmjopen-2025-108220

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<https://doi.org/10.1136/bmjopen-2025-108220>).

Received 21 July 2025

Accepted 06 October 2025



© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY. Published by BMJ Group.

For numbered affiliations see end of article.

## Correspondence to

Professor Andrew Davies;  
a.davies@soton.ac.uk

## ABSTRACT

**Introduction** Detecting cancer earlier improves treatment options and long-term survival. A multicancer early detection test that reliably picks up early-stage cancer would potentially save lives and reduce the cost of treating cancer. One promising candidate is the *Enlighten* test, which applies machine learning to plasma amino acid concentrations to detect cancer. In a cohort of 77 patients recently diagnosed with breast, colorectal, pancreatic or prostate cancer, 60 (78%) were detected by the test (sensitivity), with no false positives in 20 healthy controls. The MODERNISED study will further develop the *Enlighten* test to detect 10 different cancers by adding bladder, lung, melanoma, oesophageal, ovarian and renal cancer to the test.

**Methods and analysis** MODERNISED (ISRCTN17299125) is a multicentre prospective, non-interventional, case-control study. We aim to recruit 1000 adult participants with a recent cancer diagnosis, 250 adult participants with symptoms of cancer where a cancer diagnosis was ruled out by the National Health Service (NHS) standard of care and 100 healthy adult volunteers. Cancer tissue of origin (ToO) will include bladder, breast, colorectal, lung, melanoma, oesophageal, ovarian, pancreatic, prostate and renal. Participants in the two non-cancer cohorts who are later diagnosed with cancer will be moved to the cancer cases cohort. The primary aim is to train and validate a machine learning algorithm to detect cancer, which will be evaluated by AUROC. Secondary aims include training and validating an algorithm to predict ToO and stage of cancer, exploring differences in performance by demographics and estimating how sensitivity varies across specificity cut-offs of 95%, 99% and 99.9%. These results will provide a statistically powered estimate of how well the *Enlighten* test can discriminate between individuals with and without cancer, which can then be validated for clinical use in further research.

**Ethics and dissemination** This study is sponsored by University Hospital Southampton NHS Foundation Trust and has been approved by the Health Research Authority and Health and Care Research West Midlands (24/

## STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ Multicancer early detection (MCED) tests are proposed as tools for improving the detection of cancer in the early stages, but most use cell-free DNA technology, which is expensive and has demonstrated poor sensitivity for stage I and stage II cancers.
- ⇒ Amino acids may be more successful in detecting early-stage cancer because immune responses are stronger during the early stages of cancer.
- ⇒ This study will use the *Enlighten* test, a new candidate MCED that applies machine learning to amino acid level patterns and has a sensitivity of 76% and specificity of 100% in a cohort of 72 cancer patients and 20 healthy controls.
- ⇒ This study will retrain and evaluate the *Enlighten* test in newly diagnosed cancer patients, symptomatic patients who had cancer ruled out and healthy volunteers across southern England.
- ⇒ Reliable and cheap MCED tests would potentially save lives and reduce the cost of treating cancer, but further validation in future research will be required before clinical implementation.

WM/0234). Results will be presented at scientific meetings and published in international peer-reviewed journals. Lay summaries of study progress and findings will be published on the Southampton Clinical Trial Unit's website. **Trial registration number** ISRCTN17299125.

## INTRODUCTION

Cancer is one of the leading causes of death worldwide, with 149 022 deaths attributed to cancer in England and Wales across 2023.<sup>1</sup> Detecting cancer earlier improves the number of treatment options available and long-term survival.<sup>2</sup> In 2019, the National Health Service (NHS) in the UK committed to increasing the proportion of cancers that



are diagnosed at stage I or II.<sup>3</sup> There has been minimal change in the proportion between 2019 (53.4%) and 2022 (54.8%), despite an increasing number of people referred for cancer diagnostics.<sup>4 5</sup> Additional tools are required to detect cancer at an early stage, which would improve the proportion of cancers diagnosed at an earlier stage, improve patient outcomes and reduce the cost burden on the system.

### Multicancer early detection (MCED) tests

MCEDs, tests that analyse a 'liquid-biopsy' such as blood, urine, saliva or stool and test for the presence of multiple cancers in a singular sample, are being proposed as a complement to traditional cancer diagnosis pathways to improve early detection.<sup>6</sup> A 2024 systematic review found 36 studies evaluating six different MCEDs, with most having a high risk of bias.<sup>7</sup> There are no randomised studies demonstrating that using MCEDs leads to an increase in the proportion of cancers diagnosed at stage I or II. The majority of MCEDs that have published data available are using cell-free DNA (cfDNA) or circulating-tumour DNA (ctDNA), which are fragments of DNA released into the bloodstream that may possess alterations indicative of cancer.<sup>8</sup>

The degree of mutations present within cfDNA and ctDNA increases as cancer progresses, so this approach may not be successful if the aim is to detect cancers at stage I or II, where the amount of material shed into the bloodstream is below the limit of detection.<sup>9</sup> The largest evaluation of an MCED in the UK, the SYMPLIFY study, used the GRAIL test on 5461 symptomatic patients referred for urgent investigation for a possible gynaecological, lung or gastrointestinal cancer. The sensitivity for cancer detection was much lower at stage I (24.2%) than at stage IV (95.3%).<sup>10</sup> This is a common trend among cfDNA MCEDs. One meta-analysis found a pooled sensitivity of 50.3% for stage I or II versus 77.4% for stage III or IV.<sup>11</sup> To increase early cancer detection, alternative techniques are required.

### The Proteotype Diagnostics Enlighten test

Proteotype Diagnostics (PD) have developed an alternative approach to MCED testing. Instead of measuring material released from tumours (cfDNA or ctDNA), amino acid concentrations are measured in plasma in the Enlighten test, which are affected by the immune response. Immune responses are stronger during the early stages of cancer<sup>12</sup> and therefore may be used to improve the detection of stage I and II cancers. Machine learning (an ensemble subspace discriminant classifier) was applied to classify tested samples to cancer or non-cancer in a retrospective observational study of 77 cancer patients (breast, colorectal, pancreatic and prostate) and 20 healthy volunteers. A sensitivity of 78% at 100% specificity was achieved using Enlighten, and concentrations were further differentiated in early-stage cases.<sup>13</sup>

Southampton Experimental Cancer Medicine Centre (ECMC) responded to an open call for partnerships with

PD to develop a study further evaluating the Enlighten test. Southampton Clinical Trials Unit (SCTU) then developed a research proposal with PD and ECMC as partners, which is academically sponsored by University Hospital Southampton Trust and jointly funded by the National Institute for Health and Care Research and the Office for Life Sciences (ref NIHR207538).

The primary aim of this study is to train a machine learning classifier to use amino acid concentrations from the Enlighten test to detect cancer (bladder, breast, colorectal, lung, melanoma, oesophageal, ovarian, pancreatic, prostate and renal) in comparison to adult participants with symptoms of cancer where a cancer diagnosis was ruled out and healthy volunteers. Participants will be randomised, with 75% used to train the classifier and 25% used to validate the classifier. Secondary aims include:

1. Training and validating a classifier to distinguish cancer patients from adult participants with symptoms of cancer, where a cancer diagnosis was ruled out.
2. Training and validating a classifier to distinguish cancer patients from healthy controls.
3. Evaluating discrimination and sensitivity by tissue of origin (ToO), stage, age, sex at birth and availability of a UK cancer screening programme.
4. Training and validating a classifier to predict ToO.
5. Training and validating a classifier to predict cancer stage.

## METHODS AND ANALYSIS

### Study design

MODERNISED (ISRCTN17299125) is a multicentre prospective, non-interventional, case-control study recruiting 1000 participants with known cancer, 250 under investigation for cancer where cancer was ruled out, and 100 healthy volunteers. Participants will be recruited from five secondary care sites and healthy volunteer register banks, or hospital or university staff in England. Recruitment began in January 2025. Recruitment is underway, and version 2 (01 April 2025) is the latest version of the study protocol at the time of submitting this manuscript. Important protocol modifications are reported to relevant parties as part of the SCTU's standard operating procedures.

### Study participants

The first cohort, 'cancer cases', will consist of 1000 adults recently diagnosed with one of the 10 cancers, recruited from secondary care sites prior to treatment. The second cohort, 'symptomatic controls', will consist of 250 adults either currently or previously under investigation for one of the 10 cancers, recruited from secondary care sites, who are found to be cancer-free following normal diagnostic pathways. Participants in this cohort who are found to have cancer during follow-up will be moved to the cancer cases cohort. These participants will be reached via screening of clinic lists and approached during routine

**Table 1** Inclusion and exclusion criteria

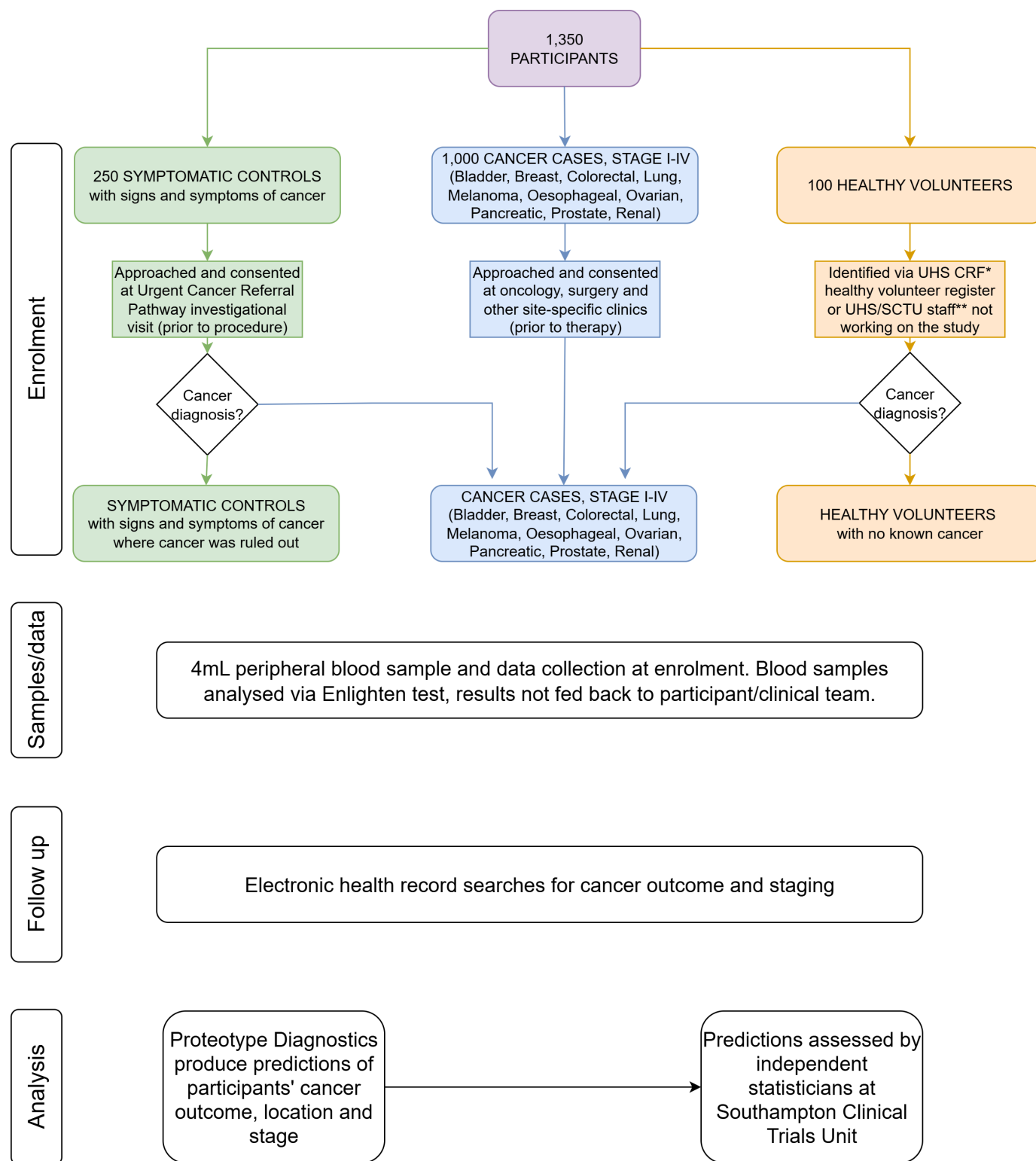
Entry arm	Inclusion criteria	Exclusion criteria
All	<ol style="list-style-type: none"> <li>1. Age 18 years or older.</li> <li>2. Able to provide a written informed consent.</li> <li>3. Willing to provide a 4 mL blood sample at the time of enrolment.</li> </ol>	<ol style="list-style-type: none"> <li>6. Pregnancy (by self-report of pregnancy status), or pregnancy within the last 12 months.</li> <li>7. Current febrile illness.</li> <li>8. Infection requiring hospitalisation within 30 days prior to blood draw.</li> <li>9. Acute exacerbation or flare of an inflammatory condition requiring escalation in medical therapy within 14 days prior to blood draw.</li> <li>10. Undergone a surgical procedure (including diagnostic biopsy) within 14 days prior to blood draw.</li> <li>11. Undergone a blood transfusion within 30 days prior to blood draw.</li> <li>12. Recipient of organ transplant or prior allogeneic stem cell transplant.</li> <li>13. Poor health status or unfit to tolerate a blood draw.</li> <li>14. Oral or intravenous corticosteroid use in the past 14 days prior to blood draw.</li> </ol>
Cancer cases	<ol style="list-style-type: none"> <li>4. Have either of the following: <ul style="list-style-type: none"> <li>► <b>Confirmed cancer diagnosis</b> of breast, melanoma, colorectal, lung, prostate, pancreatic, ovarian, oesophageal, renal or bladder cancer (any stage I–IV) <b>within 90 days prior to or up to 90 days after study blood draw</b>, based on assessment of a pathological specimen (including, but not limited to, biopsy from primary tumour site, lymph node or metastatic lesion, or cytology specimen).</li> <li>► <b>Or a high suspicion for a cancer diagnosis</b> of breast, melanoma, colorectal, lung, prostate, pancreatic, ovarian, oesophageal, renal or bladder cancer by clinical and/or radiological assessment, with <i>planned</i> biopsy or surgical resection to establish a definitive diagnosis within 3 months (90 days) <i>after</i> study blood draw.</li> </ul> </li> </ol>	<ol style="list-style-type: none"> <li>15. Currently receiving, or ever received, any of the following therapies to treat their current cancer: surgical management of the cancer beyond that required to establish the cancer diagnosis; local, regional or systemic chemotherapy, including chemoembolisation; targeted therapy, immunotherapy, including cancer vaccines; hormone therapy or radiation therapy.</li> <li>16. Known prior diagnosis of cancer, per guidance in protocol, separate from the confirmed or suspected cancer diagnosis associated with study enrolment. <b>Exception:</b> subjects with a history of non-melanoma skin cancer that has been effectively and exclusively managed by local or focal therapies, such as surgical resection, radiation therapy, cryotherapy or topical therapy, are eligible to enrol.</li> <li>17. Cancer recurrence.</li> <li>18. Synchronous malignancies.</li> </ol>
Symptomatic control	<ol style="list-style-type: none"> <li>5. Have either of the following: <ul style="list-style-type: none"> <li>► <b>Symptomatic</b> of breast, melanoma, colorectal, lung, prostate, pancreatic, ovarian, oesophageal, renal or bladder cancer, <b>but determined to be cancer-free</b> by clinical and/or radiological assessment and/or assessment of a pathological specimen (including, but not limited to, biopsy from suspected primary tumour site, lymph node or suspected metastatic lesion, or cytology specimen)—for a list of symptoms please refer to online supplemental file 1.</li> <li>► <b>Or an initial high suspicion for a cancer diagnosis</b> of breast, melanoma, colorectal, lung, prostate, pancreatic, ovarian, oesophageal, renal or bladder that was <b>ruled out</b> 90 days after study blood draw.</li> </ul> </li> </ol>	<ol style="list-style-type: none"> <li>19. Known current or prior diagnosis of cancer, per guidance in protocol. <b>Exception:</b> subjects with a history of non-melanoma skin cancer (eg, Basal Cell Carcinoma or Squamous Cell Carcinoma) that has been effectively and exclusively managed by local or focal therapies, such as surgical resection, radiation therapy, cryotherapy or topical therapy, are eligible to enrol.</li> </ol>
Asymptomatic control		<ol style="list-style-type: none"> <li>20. Known current or prior diagnosis of cancer, per guidance in protocol. <b>Exception:</b> subjects with a history of non-melanoma skin cancer that has been effectively and exclusively managed by local or focal therapies, such as surgical resection, radiation therapy, cryotherapy or topical therapy, are eligible to enrol.</li> </ol>

hospital visits for symptom investigation, or by phone call to invite them to attend an enrolment visit. The third cohort, 'healthy volunteers', will consist of 100 adults with no known cancer and no active symptoms indicative of cancer, recruited from volunteer groups at secondary care sites, or University and hospital staff if recruitment is slower than anticipated. These participants will be reached by advertisement of the study in email newsletters and invited to attend recruitment clinics. Participants in this cohort who are found to have cancer during

follow-up will be moved to the cancer cases cohort. Full inclusion and exclusion criteria can be found in [table 1](#). The study workflow is outlined in [figure 1](#).

### Baseline data collection

Clinic and MDT lists will be screened regularly (minimum two times per day) to identify cancer cases referred to relevant clinics. Clinic and 28-day faster diagnosis pathway lists will also be used to identify individuals referred to secondary care by their general practitioner



\* University Hospital Southampton Clinical Research Facility

\*\* Southampton Clinical Trials Unit

**Figure 1** MODERNISED study workflow. CRF, clinical research facility; SCTU, Southampton Clinical Trials Unit; UHS, University Hospital Southampton.

for symptoms, which indicate that they may have one of the 10 cancers being investigated. Healthy volunteers are recruited from registers where they have consented to be approached for involvement in research and will

be contacted by the research team. If recruitment of healthy volunteers is challenging, the protocol allows for members of staff at the University Hospital Southampton and the University of Southampton to be approached.



Those providing informed consent will be invited to a baseline visit ( $T_0$ ).

At the baseline visit, eligibility will be confirmed with a trained and delegated healthcare professional, who will also collect medical history, concomitant medications, demographics and the Index of Multiple Deprivation. Pseudonymised data will then be entered into Medidata's Rave Electronic Data Capture, which can then only be accessed by allocated staff at SCTU. Participants will then have 4mL of blood drawn into an EDTA tube. Any adverse events related to blood collection will be reported to the SCTU and summarised in the final study publication. All blood samples will be posted overnight Monday to Thursday at ambient temperature using Royal Mail Safeboxes (or equivalent) to the University Hospital Southampton NHS Foundation Trust, Wessex Investigational Sciences Hub (WISH). The Enlighten assay will be performed by Southampton ECMC-trained laboratory personnel housed in the WISH laboratory, and pseudonymised assay results (participant identifier and amino acid readings) will be transferred securely to PD to be converted to amino acid concentrations.

### Outcome measurement

Cancer outcome and staging will be retrieved from the Electronic Health Record (EHR) via local hospital research staff. This will be triggered for each participant once 3 months have passed since informed consent ( $T_1$ ), but data may be entered earlier where appropriate. Cancer diagnosis will be coded using ICD-10 to capture ToO, morphology and staging according to TNM (tumour, node, metastases) and cancer-specific staging system, as appropriate. Consent will be gained to collect further cancer diagnoses up to 2 years from participants' enrolment.

### Model training and validation

Participants will be randomised in a 3:1 ratio with varying permuted blocks, balanced by cancer ToO and stage, into a model training (75% of participants) or model validation (25% of participants) set. The randomisation list will be pregenerated by a statistician at SCTU. Those in the model training set will have their cohort (cancer case, symptomatic control and asymptomatic control), age and sex shared with PD to train machine learning algorithms in combination with amino acid concentrations. PD will be blinded to the diagnostic status of participants in the model validation set.

Symptomatic controls and healthy volunteers may later be diagnosed with cancer (and therefore move into the cancer cases cohort). Randomisation will therefore only occur on three occasions during the study to reduce the impact of imbalance between building and validation sets. All participants and care providers will remain blinded. Participants will not be randomised more than once. The occasions are as follows: (1) when the amino acid concentrations have been measured for the first batch of 20 participants' samples with complete data (to stress

test processes). (2) When the  $T_1$  visit has occurred for 50% of the target participants in the breast, colorectal, prostate and lung cancer cohorts. (3) When the  $T_1$  visit has occurred for all participants, and all data have been returned.

### Machine learning training and validation

Machine learning techniques considered will include: random forests, support vector machines, ensemble subspace discriminant classifiers, XGBoost, neural networks and deep learning algorithms. The final modelling strategy will be informed by comparative predictive performance and computational cost. Triplicate values of the five amino acids (Cysteine, Reduced Cysteine, Lysine, Tryptophan and Tyrosine) will be used as input parameters (or combined as average readings), with potential additions of participant age and sex. Feature importance will be measured by Shapley values. Stratified cross-validation will be used to reduce overfitting.

One algorithm will predict diagnostic status ('cancer' or 'no cancer'), a second will predict cancer ToO (bladder, breast, colorectal, lung, melanoma, oesophageal, ovarian, pancreatic, prostate and renal) and a third will predict cancer stage (I, II, III, IV). The trained algorithms will use amino acid concentrations (and potentially age and sex) to predict diagnostic status, ToO and cancer stage in the development and validation cohorts. These predictions will be provided to statisticians at SCTU to independently validate predictions against true status.

### Sample size calculation

A prediction model for a binary outcome (cancer yes/no status, prevalence 74%), 15 prediction parameters (five amino acids Cysteine, Reduced Cysteine, Lysine, Tryptophan and Tyrosine, potentially age and sex, and allowing for additional interactions in models), Nagelkerke's r-squared of 20% and target shrinkage of 10% requires a minimum sample size of 916 (678 events) to reduce overfitting.<sup>14</sup> Rounded up to 750 events and allowing a random 3:1 split between data available for model building and validation, the total required sample size is 1000 events, with 250 symptomatic controls, and 100 healthy volunteers, a total of 1350.

Of the 1350, the total sample size for validation would be 337, with 250 cancer cases, 62 symptomatic controls and 25 healthy volunteers. Because the distribution of the linear predictor derived from the prediction model is not known in advance, the anticipated Area Under the Receiver Operator Curve (AUROC) can be used to determine the required sample size for validation.<sup>15</sup> Using this simulation approach with a target CI width of 0.1, the prediction model's AUROC must be at least 0.84 to validate the model with sufficient accuracy with 337 individuals. As the pilot data demonstrated an AUROC of 0.98, it seems reasonable to assume an AUROC of 0.84 or higher can be achieved within this study.



## Analysis plan

Baseline medical history, concomitant medications, demographics and Index of Multiple Deprivation will be summarised using descriptive statistics by cohort. The number of symptomatic controls and healthy volunteers who are later diagnosed with cancer and moved into the cancer cases cohort will be presented.

PD will train machine learning algorithms for diagnostic status, ToO and cancer stage. The output for diagnostic status will be a continuous risk score. The output for ToO will be one of the 10 cancers involved in the study, and may include the percentage chance for each ToO. The output for cancer stage will be either stage I, stage II, stage III or stage IV, and may include the percentage chance for each stage. Statisticians at SCTU will produce an AUROC plot in the model training and validation groups for diagnostic status. The apparent c-statistic will be calculated and adjusted for optimism using bootstrapping<sup>16</sup> with accompanying 95% CIs. Sensitivity will be calculated in each group at cut-offs of 95%, 99% and 99.9% specificity with accompanying 95% CIs. Accuracy of the ToO and cancer stage predictions will be calculated in the model training and validation groups with accompanying 95% CIs.

These analyses will be repeated, stratified by early stage (stage I or II) and late stage (stage III or IV) cancer; ToO; symptomatic controls and healthy volunteers; age band at recruitment; biological sex at birth; availability of screening programme for cancer (screening test available (breast, cervical, colorectal and lung) or not available). All analyses will be conducted in R V.4.4.2 (or higher) or STATA V.17 (or higher).

## Ethics and dissemination

This study is sponsored by University Hospitals Southampton NHS Foundation Trust and has been approved by the Health Research Authority and Health and Care Research West Midlands (24/WM/0234). This study is funded jointly by the National Institute for Health Research and the Office for Life Sciences (ref NIHR207538). SCTU is committed to the responsible sharing of clinical study data and samples with the wider research community. Data and sample access are administered through the SCTU Data and Sample Release Committee, which will consider requests once the final analysis has been published.

Results will be presented at scientific meetings and published in international peer-reviewed journals. Lay summaries of study progress and findings will be published on SCTU's website.

## Patient and public involvement

This study benefits from two funded PPI coapplicants who have lived experience of cancer, have extensive experience in providing a patient's perspective in cancer research and an interest in equality, diversity and inclusion. They have contributed to the conception and development of the study and will be advising on the content

and tone of patient-facing materials. They will also lead the public dissemination strategy and participate in all regular study management group meetings.

## DISCUSSION

This study will determine the diagnostic accuracy of machine learning techniques using the Enlighten test to detect cancer. This will be achieved by running the assay in a group of recently diagnosed cancer patients, patients who are symptomatic but cancer was ruled out, and participants with no symptoms or cancer diagnosis. Utility in guiding cancer diagnosis will also be assessed by determining the accuracy of the predicted ToO and stage of cancer. Cancer diagnosis will be collected and updated via local EHRs across five secondary care sites in the south of England. The potential applicability of the machine learning algorithms in future research will be assessed through a held-out dataset.

While this study focuses on 10 sites of cancer (bladder, breast, colorectal, lung, melanoma, oesophageal, ovarian, pancreatic, prostate and renal), additional ToO may be added, which will require model retraining in future research. This will reduce the risk of false results where a tested individual has a cancer that is not covered by the algorithm, which can delay the time to diagnosis. To improve the proportion of cancers detected early, a substantial number of cancer patients diagnosed at stage I or II must be recruited into the study. The research team have reviewed the throughput of recruitment sites in previous years, and recruitment will be monitored and potentially tailored to ensure that these requirements are met.

Theoretically, amino acids will improve the sensitivity for MCEDs to detect early-stage cancers in comparison to the more commonly used cfDNA and ctDNA techniques. If the Enlighten test does perform in this regard, then it could be used as a triage tool prior to expensive diagnostic imaging for individuals suspected of cancer. A reliable rule-in tool for cancer detection would reduce the overall demand for diagnostic imaging. This would, in turn, allow for individuals with a lower pretest probability of cancer to be tested, therefore increasing the rate of cancers detected early.

If the results from this study are promising, then the aim of future research will be to demonstrate an improvement in early-stage cancer detection balanced against the harms of overdiagnosis and cost in a randomised controlled trial. The feasibility of progressing to this stage depends on the recruitment success for all combinations of cancer ToO and stages, and the clinical acceptability of a 10 cancer MCED versus developing the MCED across further cancer ToO.

## Author affiliations

<sup>1</sup>Cancer Research UK Southampton Clinical Trials Unit, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, UK

<sup>2</sup>Proteotype Diagnostics, Cambridge, UK

<sup>3</sup>University Hospital Southampton NHS Foundation Trust, Southampton, UK

<sup>4</sup>Southampton Health Technology Assessments Centre, University of Southampton, Southampton, UK

<sup>5</sup>Cancer Research UK/National Institute for Health and Care Research, Southampton Experimental Cancer Medicine Centre, University of Southampton, Southampton, UK

<sup>6</sup>Patient and Public Representative, Southampton, UK

**Acknowledgements** Preparation of this manuscript was supported by SW attending the University of Southampton Faculty of Medicine/Faculty of Environmental and Life Sciences Writing Retreat (January 2025).

**Contributors** The study concept and design were conceived by SW, VG, JW, WS, EY and AD. The study protocol and patient-facing documents were written by SW, VG, ZH, AC, JW, WH, KM, RW, IS, JH, GG and AD. The lab manual for sample processing was written by AC, CP, NI, WS and EY. SW wrote and revised the manuscript. All authors have read and agreed to the submitted version of this manuscript. SW is a guarantor.

**Funding** This work is jointly supported by the National Institute for Health Research and the Office for Life Sciences (ref NIHR207538).

**Competing interests** WS is CEO and cofounder of Proteotype Diagnostics, declares shares in Proteotype Diagnostics, and that Proteotype Diagnostics owns a relevant European patent (EP 4 196 797 B1). ZH declares payment for providing expert testimony for Medico and is a former president of BASO. AC declares that Proteotype Diagnostics has purchased a Tecan plate reader, calibration plate and transilluminator for the Experimental Cancer Medicine Centre to carry out work for this study. SJC declares payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Janssen, Ipsen, Merck, Astellas, AstraZeneca, BMS and Pfizer; support for attending meetings and/or travel from Bayer, Janssen, Merck and Novartis; participation on a Data Safety Monitoring Board or Advisory Board for Bayer, Merck, Astellas, Pfizer, MSD and AstraZeneca; institutional research support from AstraZeneca, Roche, Veracyte. EY is CSO and cofounder of Proteotype Diagnostics, declares shares in Proteotype Diagnostics, and that Proteotype Diagnostics owns a relevant European patent (EP 4 196 797 B1). AD declares research funding to the host organisation from AstraZeneca, Roche, Kite/Gilead, MSD, Cellcentric, BMS and Regeneron. He has received support to attend conferences from Roche and BMS. He has attended paid advisory boards for Roche, BMS, AbbVie, Sobi and AstraZeneca. Speaker honoraria from Roche, BMS, Sobi and Kite/Gilead. Participation of the Data Safety and Monitoring Board for BMS. All other authors declare no competing interests.

**Patient and public involvement** Patients and/or the public were involved in the design, conduct, reporting or dissemination plans of this research. Refer to the Methods section for further details.

**Patient consent for publication** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

## ORCID iDs

Sam Wilding <https://orcid.org/0000-0003-4184-2821>

Irene Soulsby <https://orcid.org/0000-0002-2291-831X>

## REFERENCES

- Office for National Statistics. Deaths registered summary statistics, England and Wales. Available: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/deathsregisteredsummarystatisticsenglandandwales> [Accessed 16 Jan 2025].
- Etzioni R, Urban N, Ramsey S, *et al.* The case for early detection. *Nat Rev Cancer* 2003;3:243–52.
- NHS. The NHS long term plan. 2019. Available: <https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf> [Accessed 16 Jan 2025].
- NHS England. CWT CRS – national time series Oct 2009 – Nov 2024 with revisions. Available: <https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/> [Accessed 16 Jan 2025].
- Cancer Research UK. Early cancer diagnosis data hub. Available: <https://crukcanerintelligence.shinyapps.io/EarlyDiagnosis/> [Accessed 16 Jan 2025].
- Guerra CE, Sharma PV, Castillo BS. Multi-cancer early detection: the new frontier in cancer early detection. *Annu Rev Med* 2024;75:67–81.
- Wade R, Nevitt S, Liu Y, *et al.* Multi-cancer early detection tests for general population screening: a systematic literature review. *Oncology (Williston Park, NY)* 2024.
- Stejskal P, Goodarzi H, Srovnal J, *et al.* Circulating tumor nucleic acids: biology, release mechanisms, and clinical relevance. *Mol Cancer* 2023;22:1–21.
- Medina JE, Dracopoli NC, Bach PB, *et al.* Cell-free DNA approaches for cancer early detection and interception. *J Immunother Cancer* 2023;11.
- Nicholson BD, Oke J, Virdee PS, *et al.* Multi-cancer early detection test in symptomatic patients referred for cancer investigation in England and Wales (SYMPHONY): a large-scale, observational cohort study. *Lancet Oncol* 2023;24:733–43.
- Park JH, Oh Y, Chung L-Y, *et al.* Systematic review and meta-analysis of the accuracy and applicability of blood-based multi-cancer early detection (MCED) in the general population. *JCO* 2023;41:3069.
- Adam JK, Odhav B, Bhoola KD. Immune responses in cancer. *Pharmacol Ther* 2003;99:113–32.
- Tang C, Correia P, Casimiro S, *et al.* Immunodiagnostic plasma amino acid residue biomarkers detect cancer early and predict treatment response. *Nat Commun* 2025;16.
- Riley RD, Ensor J, Snell KIE, *et al.* Calculating the sample size required for developing a clinical prediction model. *BMJ* 2020;368.
- Riley RD, Debray TPA, Collins GS, *et al.* Minimum sample size for external validation of a clinical prediction model with a binary outcome. *Stat Med* 2021;40:4230–51.
- Collins GS, Dhiman P, Ma J, *et al.* Evaluation of clinical prediction models (part 1): from development to external validation. *BMJ* 2024;384.