ISRCTN51713388 https://doi.org/10.1186/ISRCTN51713388

Artificial intelligence project for improved sarcoma diagnoses for patient benefit

Submission date
06/12/2024

Recruitment status
Recruiting

Registration date
09/01/2025

Overall study status
Ongoing

Last Edited
17/12/2024

Condition category
Cancer

[ ] Prospectively registered
[ ] Protocol
[ ] Statistical analysis plan
[ ] Results
[ ] Individual participant data
[X] Record updated in last year

Plain English summary of protocol

Background and study aims
Cancer treatment is determined by how tumours are classified by pathologists, who provide diagnoses which encompass prognoses and predictions of responses to therapies.
Reaching a tissue diagnosis today requires more time than hitherto allocated to pathologists, driven by the rapidly increasing knowledge of important prognostic factors, an understanding of the molecular basis of disease and the availability of targeted and personalised treatments.
Diagnoses must be highly specific and accurate, distinguishing cancer subtypes from each other and from other diseases such as bone and soft tissue tumours versus regenerative /degenerative musculoskeletal (MSK) disease.
The increasing workload and complexity in reaching pathological diagnoses is compounded by a declining workforce: 29% of all UK-based pathologists are aged 55 and over. MSK pathology is specifically affected as there is a serious shortfall in pathologists in this subspecialty area.
Although primary MSK sarcomas represent about 2% of all cancers, there are over 100 subtypes described, as well as many common conditions that mimic sarcoma. The diagnosis has huge treatment implications and therefore crucially important to get right. Sarcomas can occur anywhere in the body (e.g. breast, lung) so they are frequently first encountered by non-expert MSK pathologists. This often leads to excessive consultant time in reaching diagnoses, and ordering inappropriate and excessive immunohistochemical and genetic tests. This protracts patients' journeys to treatment, resulting in unnecessary costs.
This study aims to find out whether artificial intelligence (AI) is a solution to these challenges and can help support (not replace) pathologists as already shown for common cancers (e.g. prostate and bowel).

Who can participate?
Patients who had a diagnosis of sarcoma or a mimic of sarcoma and whose tissue has been processed as part of their clinical care. The study is fully inclusive of all genders, age ranges, ethnicities and members of all socio-economic groups.

What does the study involve?
As AI requires large numbers of cases and pathology slides must be digitised to generate whole-slide images (WSIs). The availability of digital scanners across the UK has made this possible. This study will gather large numbers of sarcoma cases using archived slides/data from the last 40

years and prospectively. Members of direct clinical care teams will gather pathology images /data from up to 50,000 patients. The project will run for 10 years (collecting images/data for 5 years and 5 years follow-up).

What are the possible benefits and risks of participating?
The study aims to produce an AI algorithm to aid sarcoma diagnosis in the future. There are no direct benefits or risks to current patients.

Where is the study run from?
University College London (UK)

When is the study starting and how long is it expected to run for?
March 2023 to December 2033

Who is funding the study?
The majority of the funding has been secured from UK Research and Innovation (UKRI) along with smaller grants from Tom Prince Cancer Trust, Skeletal Cancer Trust, Sarcoma UK, Chordoma UK and Bone Cancer Research Trust (UK)

Who is the main contact?
Prof. Adrienne M Flanagan, a.flanagan@ucl.ac.uk

Contact information

Type(s)
Public, Scientific, Principal Investigator

Contact name
Prof Adrienne Flanagan

ORCID ID
http://orcid.org/0000-0002-2832-1303

Contact details
University College London
Cancer Institute
Huntley Street
London
United Kingdom
WC1E 6BT
+44 (0)7980290621
a.flanagan@ucl.ac.uk

Additional identifiers

EudraCT/CTIS number
Nil known

IRAS number
328987

ClinicalTrials.gov number
Nil known

Secondary identifying numbers
EDGE 161548

Study information

Scientific Title
An artificial intelligence solution for diagnosing, prognosticating as well as predicting the outcome of sarcomas and their mimics: a multi-centre study

Acronym
AI-SCOPE

Study objectives
It is hypothesised that artificial intelligence (AI) can be used to support pathologists in classifying sarcomas (rare cancers of bone and soft tissue) and their mimics and provide improved or similar classification performance compared to pathologists.

It is hypothesised that the AI Classifier will save pathologists' time and improve the patients' diagnostic pathways and that cost efficiencies can be made.

Ethics approval required
Ethics approval required

Ethics approval(s)
Approved 12/12/2023, Health and Social Care Research Ethics Committee B (HSC REC B) (Office for Research Ethics Committee Northern Ireland (ORECNI), Lissue Industrial Estate West, 5 Rathdown Walk, Lisburn, BT28 2RF, United Kingdom; +44 (0)28 95 361400; info.orecni@hscni.net), ref: 23/NI/0166

Study design
Multi-centre observational pseudonymised cohort study

Primary study design
Observational

Secondary study design
Cohort study

Study setting(s)
Hospital, University/medical school/dental school

Study type(s)
Diagnostic

Participant information sheet
Not available in web format, please use the contact details to request a participant information sheet

Health condition(s) or problem(s) studied
Bone and soft tissue tumours and their mimics

Interventions
This study involves a large-scale collection of digitised whole slide images (WSI), together with related demographic and clinical data, from a cohort of 35,000 - 50,000 patients which will be obtained from multiple hospital sites.

There is no recruitment to the study and no intervention.

Because of the rarity of sarcoma, this can only be achieved using retrospective cases archived over many years. The identification and preparation of the WSI and related clinical data, including its deidentification and pseudonymisation, will be undertaken at each collaborating pathology site by the direct clinical care team. Only de-identified and pseudonymised data will be accessible to researchers. The diagnoses given by the resulting algorithm will be compared with the diagnosis given by a panel of pathologists.

Intervention Type
Other

Primary outcome measure
The primary outcome measure is whether the algorithm has predicted the diagnosis correctly. The algorithm ranks the diagnosis in order of likelihood with the highest ranking being compared to the diagnosis agreed on by a panel of pathologists and/or additional molecular tests and is categorised as being correct or incorrect and a confusion table will be constructed. Measured at a single timepoint.

Secondary outcome measures
Measured at a single timepoint:
1. Diagnostic pathway efficiency and speed, measured as the rate at which a patient receives a diagnosis
2. The number of ancillary tests required, measured as the numbers requested by pathologists prior to reaching a diagnosis
3. Pathologist diagnostic efficiency, measured as the number of pathologists able to make accurate diagnoses without the need for excessive tests and referrals
4. Pathological and epidemiological insights into sarcomas, measured through the review of a large number of retrospective cases along with the development of an algorithm to improve diagnosis prospectively.
5. Prognosis and prediction of response to treatment, assessed by linking the algorithm's predictions with demographic and clinical outcome data ranging from the patient's initial date of diagnosis to either their date of death or date last seen.

Following exploratory data analysis with correlation plots, histograms and frequency tables, statistical modelling will be performed using survival analysis with Cox proportional hazard estimates and log-rank test as appropriate. Statistical significance will be set at 5%. Further statistical methods and machine learning techniques such as the random forest algorithm may be used to improve prediction of prognosis.

Overall study start date
01/03/2023

Completion date
31/12/2033

Eligibility

Key inclusion criteria
Patients who had a diagnosis of sarcoma or a mimic of sarcoma and whose tissue has been processed as part of their clinical care. The study is fully inclusive of all genders, age ranges, ethnicities and members of all socio-economic groups.

Participant type(s)
Patient

Age group
All

Lower age limit
1 Days

Upper age limit
100 Years

Sex
Both

Target number of participants
50000

Total final enrolment
50000

Key exclusion criteria
Does not meet the inclusion criteria

Date of first enrolment
01/10/2023

Date of final enrolment
31/12/2029

Locations

Countries of recruitment
England

United Kingdom

Wales

Study participating centre
Royal National Orthopaedic Hospital NHS Trust
Brockley Hill
Stanmore
United Kingdom
HA7 4LP

Study participating centre
Manchester University NHS Foundation Trust
Cobbett House
Oxford Road
Manchester
United Kingdom
M13 9WL

Study participating centre
The Newcastle upon Tyne Hospitals NHS Foundation Trust
Freeman Hospital
Freeman Road
High Heaton
Newcastle upon Tyne
United Kingdom
NE7 7DN

Study participating centre
The Royal Marsden NHS Foundation Trust
Fulham Road
London
United Kingdom
SW3 6JJ

Study participating centre
Nottingham University Hospitals NHS Trust - Queen's Medical Centre Campus
Nottingham University Hospital
Derby Road
Nottingham
United Kingdom
NG7 2UH

Study participating centre

Oxford University Hospitals NHS Foundation Trust
John Radcliffe Hospital
Headley Way
Headington
Oxford
United Kingdom
OX3 9DU

Study participating centre
Great Ormond Street Hospital Central London Site
Great Ormond Street
London
United Kingdom
WC1N 3JH

Study participating centre
Swansea Bay University Local Health Board
One Talbot Gateway
Seaway Drive
Seaway Parade Industrial Estate
Baglan Port Talbot
West Glamorgan
United Kingdom
SA12 7BR

Study participating centre
University College London Hospitals NHS Foundation Trust
250 Euston Road
London
United Kingdom
NW1 2PG

Study participating centre
The Robert Jones and Agnes Hunt Orthopaedic Hospital NHS Foundation Trust
Gobowen
Oswestry
United Kingdom
SY10 7AG

Study participating centre

Sheffield Teaching Hospitals NHS Foundation Trust
Northern General Hospital
Herries Road
Sheffield
United Kingdom
S5 7AU

Study participating centre
Cambridge University Hospitals NHS Foundation Trust
Cambridge Biomedical Campus
Hills Road
Cambridge
United Kingdom
CB2 0QQ

Sponsor information

Organisation
University College London

Sponsor details
4th Floor, West
250 Euston Road
London
England
United Kingdom
NW1 2PG
+44 (0)20 3447 9928
uclh.randd@nhs.net

Sponsor type
University/education

Website
http://www.ucl.ac.uk/

ROR
https://ror.org/02jx3x895

Funder(s)

Funder type
Government

Funder Name
UK Research and Innovation

Alternative Name(s)
UKRI

Funding Body Type
Government organisation

Funding Body Subtype
National government

Location
United Kingdom

Funder Name
Tom Prince Cancer Trust

Alternative Name(s)
UKRI

Funding Body Type
Government organisation

Funding Body Subtype
National government

Location
United Kingdom

Funder Name
Skeletal Cancer Trust

Funder Name
Sarcoma UK

Alternative Name(s)
SUK

Funding Body Type
Government organisation

Funding Body Subtype

Trusts, charities, foundations (both public and private)

Location
United Kingdom

Funder Name
Chordoma UK

Funder Name
Bone Cancer Research Trust

Alternative Name(s)
The Bone Cancer Research Trust, BCRT

Funding Body Type
Government organisation

Funding Body Subtype
Trusts, charities, foundations (both public and private)

Location
United Kingdom

Results and Publications

Publication and dissemination plan
Planned publication in a peer-reviewed journal, internal report, conference presentation and publication on website.

Intention to publish date
31/12/2025

Individual participant data (IPD) sharing plan
The datasets generated and/or analysed during the current study will be published as a supplement to the results publication

IPD sharing plan summary
Published as a supplement to the results publication